

La revolución del genoma del arroz: de un grano antiguo a Green Super Rice

Rod A. Wing^{1,2*}, Michael D. Purugganan^{3,4} y Qifa Zhang^{5*}

Resumen | El arroz es un cultivo básico para la mitad de la población mundial, que se espera que crezca en 3 mil millones en los próximos 30 años. También es un modelo clave para estudiar la genómica de los agroecosistemas. Esta doble función coloca al arroz en el centro de un enorme desafío que enfrenta la agricultura: cómo aprovechar la genómica para producir suficientes alimentos para alimentar a una población mundial en expansión. Científicos de todo el mundo están investigando la variación genética entre las especies de arroz domesticado y sus parientes silvestres con el objetivo de identificar loci que puedan explotarse para generar una nueva generación de cultivos sostenibles conocidos como Green Super Rice.

Resistencia al alojamiento
La capacidad de las plantas para soportar vientos de alta velocidad, como los de los tifones anuales en los trópicos.
Por lo general, la resistencia al acame ocurre al reproducirse para obtener tallos más rígidos, baja estatura o ambos.

heterosis
También conocido como vigor híbrido. Un fenómeno por el cual el híbrido producido al cruzar dos líneas genéticamente distintas (normalmente endogámicas) supera agrónomicamente a cada uno de sus padres (por ejemplo, en términos de mayor rendimiento y crecimiento más rápido).

¹Universidad de Arizona, Tucson, AZ, EE. UU.
²Instituto Internacional de Investigación del Arroz, Los Baños, Filipinas.

³Universidad de Nueva York, Nueva York, NY, EE. UU.

⁴Universidad de Nueva York Abu Dhabi, Abu Dhabi, Emiratos Árabes Unidos.

⁵Universidad Agrícola de Huazhong, Wuhan, China.

*correo electrónico: rwing@arizona.edu; qifazh@mail.hzau.edu.cn

<https://doi.org/10.1038/s41576-018-0024-z>

Como uno de los principales cultivos básicos del mundo, el arroz es un componente esencial de la dieta y los medios de subsistencia de más de 3500 millones de personas. A su tasa de crecimiento actual, se prevé que la población mundial alcance los ~10y000 millones de personas para 2050 (ref.1). Gran parte de este aumento ocurrirá en regiones pobres y densamente pobladas del mundo que ya dependen en gran medida del arroz (como África y el sur de Asia), en que subsiste una población que se desempeñará

Los programas tradicionales de mejoramiento de cultivos se basan en identificar y cruzar plantas con fenotipos deseables desde el punto de vista agronómico. Estos enfoques han dado como resultado la adopción de variedades semienanas para una mejor resistencia al acame y la explotación de la heterosis, lo que ha hecho que los rendimientos del arroz aumenten sustancialmente durante el último medio siglo. Sin embargo, estos aumentos en la productividad han sido costosos en términos de recursos y sus efectos ambientales adversos en muchas áreas productoras de arroz; tales efectos se derivan del uso excesivo de pesticidas, fertilizantes y agua, entre otros insumos. El concepto de Green Super Rice (GSR) se propuso por primera vez hace 10 años2 como un medio para satisfacer las demandas futuras y al mismo tiempo reducir los costos y la huella ecológica asociada con una mayor productividad. Los atributos clave de las variedades GSR incluyen requisitos reducidos de fertilizantes, pesticidas y agua; mayor rendimiento; granos más sabrosos y nutritivos; la capacidad de crecer en tierras marginales; y reducción de las emisiones de gases de efecto invernadero (Fig. 1a).

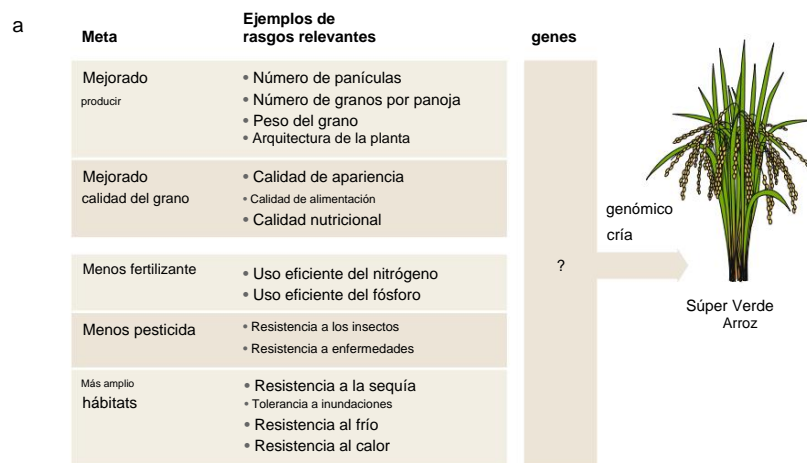
Para lograr este objetivo, será necesario comprender, manejar y utilizar mejor la variación genética existente presente en el arroz domesticado y sus parientes silvestres. En este sentido, los estudios de genómica funcional y comparativa serán clave para identificar y comprender los componentes genéticos responsables de los rasgos agrónomicamente beneficiosos (Fig. 1b). Rice está bien posicionada para liderar el camino en estos nuevos enfoques de mejoramiento genómico por varias razones. A ~400Mb, el genoma del arroz diploide

es el más pequeño entre los cereales domesticados, lo que lo hace particularmente apto para estudios genómicos. Además, el arroz fue la primera planta de cultivo en ser completamente secuenciada3, y la disponibilidad de un genoma de alta calidad permite estudios genómicos evolutivos, funcionales y de población. Además, el género *Oryza* comprende especies silvestres y domesticadas con una larga trayectoria evolutiva que contiene un reservorio de genes y rasgos prácticamente sin explotar que pueden usarse para mejorar los cultivos (Fig. 2a,b). De hecho, el arroz se ha cultivado en el Viejo Mundo durante miles de años y en el Nuevo Mundo durante cientos de años, tiempo durante el cual se ha adaptado a una variedad de ubicaciones geográficas y una plétora de condiciones ambientales (Fig. 2c); los genes seleccionados durante la domesticación y la adaptación pueden potencialmente aprovecharse sistemáticamente para mejorar los cultivos.

De hecho, aunque es difícil de cruzar, se han producido con éxito híbridos entre todos los tipos de genoma no AA, lo que ha dado como resultado la introgresión de muchas características valiosas en el arroz cultivado4 .

Aquí, revisamos cómo los estudios genómicos comparativos y funcionales de las colecciones de germoplasma de arroz domesticado y silvestre pueden usarse para informar los programas de mejoramiento, con énfasis en cómo están contribuyendo al desarrollo de variedades GSR. Discutimos los conocimientos adquiridos al estudiar la domesticación del arroz asiático (*Oryza sativa*) y el arroz africano (*Oryza glaberrima*), que ocurrieron en dos continentes diferentes con una diferencia de aproximadamente 6000 años. Describimos cómo las colecciones enteras de germoplasma ahora se secuencian y fenotizan para identificar las regiones genómicas que son importantes para la productividad y la adaptación de los cultivos y cómo este proceso se ha visto facilitado por los avances en las tecnologías de fenotipado de campo y la disponibilidad de un número cada vez mayor de muestras casi libres de lagunas. genomas de referencia que cubren la amplitud de la diversidad de *Oryza* cultivada y silvestre.

Finalmente, abordamos la necesidad de una genómica funcional y mejoramiento



de estos eventos pasados puede identificar características agronómicas beneficiosas que han sido seleccionadas en el transcurso de la domesticación y diversificación de cultivos. Los enfoques comparativos, de población y de genómica funcional se pueden usar para determinar qué genes son responsables de los rasgos seleccionados, y esta información genética se puede usar potencialmente para dar forma a futuros esfuerzos de mejoramiento.

Introgresión de genes durante la domesticación y diversificación del arroz asiático. *O. sativa* muestra diferenciación genética en subespecies o grupos varietales, que han sido validados por estudios genómicos⁶. Estos grupos varietales incluyen japónica tropical, japónica templada e indica, aromáticas y el grupo aus menos conocido⁷.

(Figura 2b). Sin embargo, no ha quedado claro si estos grupos varietales surgieron de uno o múltiples eventos de domesticación de novo^{7–10}. Recientes modelos genéticos de población y análisis filogenómicos basados en secuencias genómicas completas de alta calidad de los diferentes grupos varietales y de las especies silvestres de *Oryza* sugieren una imagen compleja de la evolución del arroz, en la que las poblaciones de arroz existentes se originaron a partir de diferentes poblaciones ancestrales de *O. rufipogon* y/o *Oryza nivara* que divergió hace ~300 000–400 000 años¹⁰ (Fig. 3a). En este modelo, la domesticación de novo parece haber ocurrido solo una vez, en japonica, y la posterior introgresión de alelos de japonica en el arroz silvestre o en una población proto-indica y/o proto-aus dio lugar a otras poblaciones de arroz asiático que hoy en día hacen los diferentes grupos varietales de arroz^{9,10}.

El origen de indica (y posiblemente de aus)¹⁰ ilustra el papel de la introgresión en el movimiento de alelos de domesticación seleccionados entre poblaciones distintas y representa un ejemplo del Neolítico temprano de mejoramiento del arroz por hibridación entre poblaciones divergentes. El análisis de las regiones genómicas introgresadas de japonica a indica ha demostrado que contienen una serie de genes que son responsables de los rasgos deseables desde el punto de vista agronómico, algunos de los cuales todavía se utilizan en los programas de mejoramiento en la actualidad. Estos genes incluyen el alelo que no se rompe de SH4 (ref.¹¹), que mejora la retención de semillas maduras en la panoja de arroz y que llegó al subcontinente del sur de Asia a través del movimiento del arroz japonica temprano, posiblemente a través de la antigua Ruta de la Seda hacia noroeste de la India¹²; el gen de color RC, que conduce al color de grano blanco culturalmente deseable¹³; y el gen del crecimiento erecto PROG1, que evita la pérdida de rendimiento por encamado¹⁴.

Eventos de domesticación independientes en Asia y África seleccionados por genes similares.

Las domesticaciones independientes dentro del género *Oryza* proporcionan un contexto comparativo para el estudio de la evolución de rasgos agronómicos clave. Mientras que *O. sativa* evolucionó en Asia, se produjo una domesticación separada en África occidental que dio origen a *Oryza glaberrima* a partir de su ancestro silvestre *Oryza barthii*¹⁵ (Figura 3b). Los análisis genómicos y filogenéticos de la población de las secuencias del genoma completo sugieren un evento primario de domesticación tierra adentro para esta especie hace unos 3500 años, seguido de una diversificación entre las poblaciones costeras y del interior, y entre los grupos del norte y del sur^{15,16}.

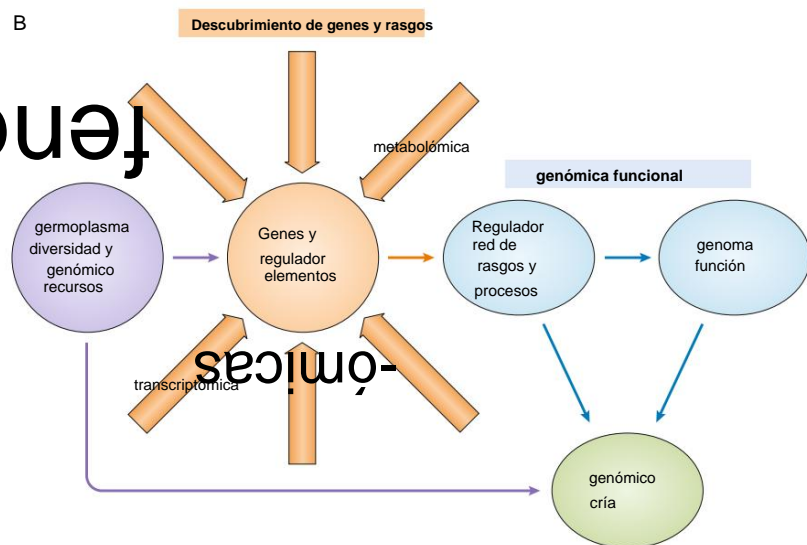


Figura 1 | **Estrategias basadas en la genómica para el desarrollo de Green Super Rice.** **un** | Comprender qué genes regulan los rasgos "verdes" deseables puede ayudar a generar cultivos mejorados, como nuevas variedades de Green Super Rice, a través de enfoques de mejoramiento genómico. **segundo** | Un resumen de la investigación genómica en la comunidad mundial del arroz y cómo se relaciona con el mejoramiento genómico. Las flechas indican el flujo de trabajo y de información. El objetivo final es identificar y comprender la función del complemento completo de genes en el genoma del arroz.

El progreso realizado en cualquier etapa del flujo de trabajo se puede aplicar al mejoramiento genómico.

tubería para acelerar el ritmo de la mejora de cultivos a través de tecnologías de mejoramiento genómico una vez que se hayan identificado los genes candidatos.

Genómica de la domesticación

Durante el período Neolítico, que comenzó hace unos 11 500 años, los primeros agricultores transformaron las plantas silvestres en especies domesticadas con mayor rendimiento y estabilidad del rendimiento⁵. Por ejemplo, antes de la Revolución Verde, el rendimiento del arroz asiático domesticado (*O. sativa*) promediaba 4,1 toneladas por hectárea, que es 3,7 veces más que el rendimiento promedio de 1,12 toneladas por hectárea de su ancestro silvestre, *Oryza rufipogon*⁵. Además, las poblaciones de cultivos sufrieron una diversificación adaptativa a medida que se expandieron más allá de los rangos de sus especies originales y encontraron tanto nuevos entornos como nuevos contextos culturales humanos. El estudio

mejoramiento genómico
Enfoques que utilizan los datos, conocimientos, recursos, genes y tecnologías generados por la investigación genómica para mejorar los programas de mejoramiento.

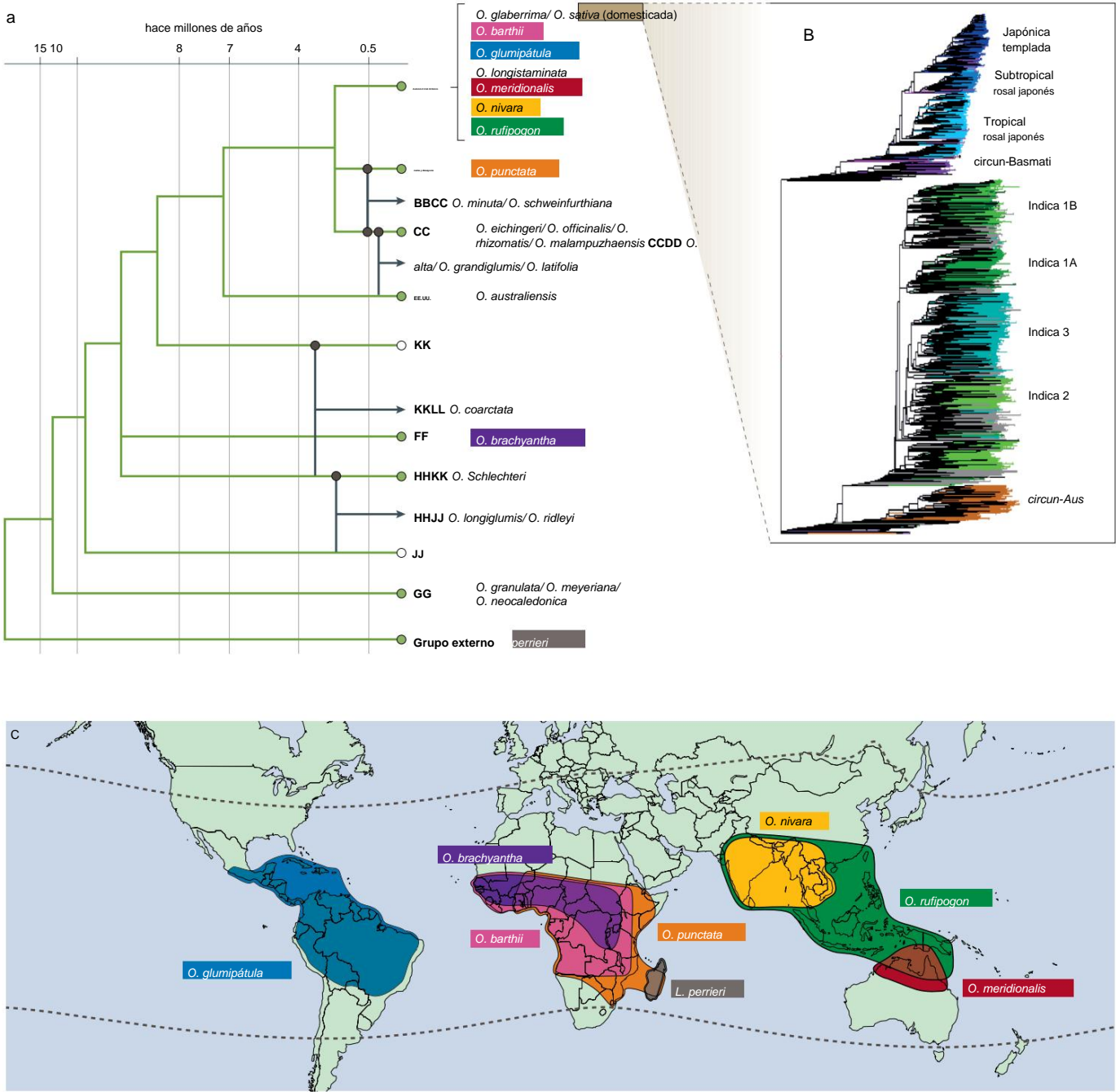


Figura 2 | **Filogenia y distribución del género *Oryza*** . un | Árbol filogenético del género *Oryza* (modificado de refs45,111). El género *Oryza* contiene 27 especies, dos de las cuales están domesticadas. El género alberga 11 tipos de genoma existentes (6 diploides y 5 poliploides) que tienen una variación de 3,6 veces en el tamaño del genoma (que va desde ~350 a 1283Mb). Las flechas indican el origen de los poliploides, los círculos cerrados representan a los padres maternos y los círculos abiertos representan genomas diploides no identificados. **segundo** | El análisis genómico confirma que *Oryza sativa* se puede subdividir en nueve subpoblaciones30; se muestra un análisis del conjunto de datos de acceso del Proyecto de genomas de arroz 3000 (3K RGP) usando el programa de software ADMIXTURE112 **do** | Mapa mundial que muestra que el género *Oryza* está ampliamente distribuido. En el mapa se indican los límites de crecimiento del arroz cultivado (líneas discontinuas) y las ubicaciones endémicas de siete especies secuenciadas recientemente, incluido el grupo externo *Leersia perrieri*. La parte **a** está adaptada con permiso de ref.45, Elsevier y ref.111, PNAS. La parte **c** está adaptada de la referencia 42, Macmillan Publishers Limited, CC-BY-4.0.

Tipos de genoma
El género *Oryza* está compuesto por ~27 especies existentes que albergan 11 tipos de genoma (GT) distintos, 6 de los cuales son diploides (n=12; GT: AA, BB, CC, EE, FF y GG) y 5 de los cuales son poliploides (n=24; GT: BBCC, CCDD, HHJJ, HHKK y KKLL). Estos GT se definieron en función de la citogenética (es decir, el número, el tamaño y la forma de los cromosomas), la hibridación fluorescente in situ (FISH) y la hibridación genética.

La secuenciación del genoma del arroz africano ha proporcionado información clave sobre los genes que subyacen a la evolución paralela de los rasgos de domesticación15. El análisis genómico comparativo, por ejemplo, confirma que HD1 (también conocido como SE1), un gen que controla el tiempo de floración tardío

perdido en el arroz africano en comparación con su progenitor silvestre *O. barthii* y puede explicar la insensibilidad al fotoperíodo y el tiempo de floración sincronizado de esta especie de cultivo15. Los escaneos en busca de firmas de selección en el genoma del arroz tardío implican muchos loci que también parecen ser

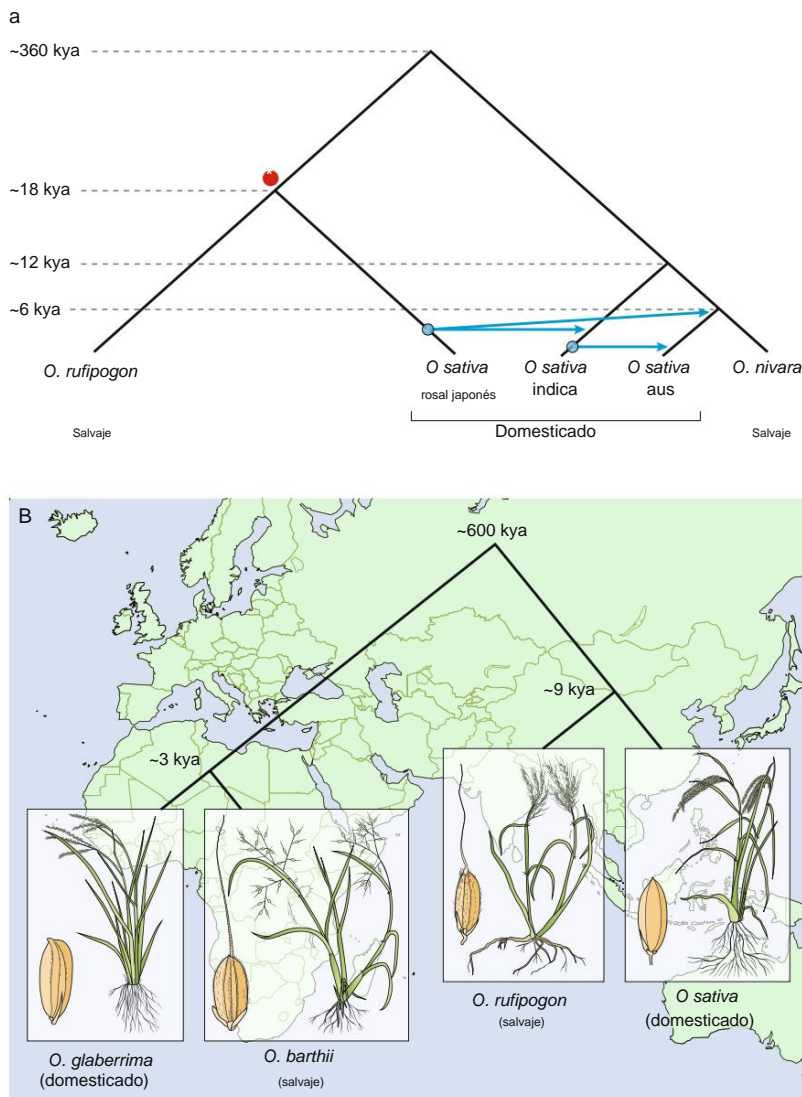


Figura 3 | Domesticación de *Oryza sativa* y *Oryza glaberrima*. un | Relaciones entre varias subespecies de *Oryza sativa* y sus ancestros, *Oryza rufipogon* y *Oryza nivara*. El asterisco rojo indica el momento de la divergencia de japonica de los linajes existentes de *O. rufipogon* (~18kya). Esta fecha temprana puede reflejar la divergencia entre las poblaciones actuales de *O. rufipogon* y las poblaciones extintas de arroz silvestre que fueron los progenitores de *O. sativa* variegata japonica. La evidencia sugiere que la hibridación resultó en la introgresión de genes de japonica a indica y aus (indicado por las flechas azules)¹⁰. b| El arroz asiático y africano se domesticaron de forma independiente con una diferencia de aproximadamente 6000 años. kya, hace mil años. La parte **a está adaptada de la ref. 10, Choi, JY La paradoja del arroz: orígenes múltiples pero domesticación única en el arroz asiático. *mol. Biol. Evol.* 2017, 34(4), 969–379, con autorización de Oxford University Press. La parte **b** está adaptada de la ref. 113, Macmillan Publishers Limited.**

importante para las características agronómicas clave en el arroz asiático, incluido el gen semienano SD1, el gen de respuesta al estrés NAC6 y el locus 15 SH4 de destrucción. Curiosamente, al menos dos genes involucrados en el rasgo de no fragmentación en el arroz asiático (SH1 y SH4) están asociados con mutaciones que son potencialmente responsables del fenotipo de no fragmentación en *O. glaberrima*¹⁵. El gen SH1 se elimina en el arroz africano, mientras que el ortólogo SH4 de *O. glaberrima* muestra una expresión reducida y se encuentra cerca de un barrido selectivo, lo que es consistente con la selección para no desgranar en esta especie de cultivo domesticado. Estos estudios sugieren que, para ciertos rasgos clave, solo un número limitado

se seleccionan varios genes durante y después de la domesticación. Por lo tanto, los genes identificados como reguladores de rasgos agronómicos importantes en una especie (aquí, el arroz africano) podrían usarse para desarrollar esos rasgos en GSR.

Evolución adaptativa: genes implicados en la propagación del arroz asiático y africano. Después del inicio de la domesticación, una especie cultivada comienza a expandirse desde su centro de origen y se traslada a regiones con distintas características ambientales; por lo tanto, necesita evolucionar adaptativamente a estos nuevos entornos. Además, algunos loci en una población ancestral dada pueden seleccionarse para nuevos entornos porque generan rasgos clave que siguen siendo deseables en distintas culturas y lugares. Por lo tanto, algunos de estos fenotipos (como la respuesta al estrés abiótico y la calidad del grano) podrían ser extremadamente importantes para la reproducción de GSR.

Por ejemplo, estudios de la historia evolutiva del arroz revelan la acción de selección positiva para el gen *WAXY17,18*. Los alelos de este gen tienen efectos importantes sobre los principales determinantes de la calidad culinaria y comestible del arroz: contenido de amilosa, temperatura de gelificación y longitud de gel del grano de arroz¹⁹. La variación en estos rasgos ha contribuido en gran medida a la diversidad culinaria del arroz y la forma en que se consume. Los estudios evolutivos moleculares revelan que el arroz glutinoso (pegajoso) lleva un alelo de *WAXY* con un sitio de empalme mutante, que probablemente se originó en el sudeste de Asia continental y se extendió por el este de Asia y finalmente se incorporó a la japónica templada; la dispersión geográfica del alelo está asociada con la conveniencia cultural del *arroz glutinoso*^{17,18}.

Otra asociación clara de un rasgo agronómico importante y la adaptación geográfica es la tolerancia a la salinidad en *O. glaberrima*¹⁶. Los cultivares de todo el rango geográfico del arroz africano poseen cierto nivel de tolerancia a la sal, con la excepción de las poblaciones costeras del sur¹⁶. Los estudios de asociación del genoma completo (GWAS) han identificado 11 loci asociados con la tolerancia a la salinidad¹⁶, uno de los cuales se superpone a *HAK5*, un gen transportador de potasio de alta afinidad asociado con la tolerancia a la sal del arroz²⁰ que se induce en *O. glaberrima* tras el estrés salino¹⁶. Además, dos de los loci GWAS abarcan regiones genómicas que han sufrido una selección asociada con la adaptación geográfica en el arroz africano; una de estas regiones contiene 41 genes¹⁶, incluido el locus *PPL*, que es miembro de una familia de genes asociados con la tolerancia a la sal del arroz²¹.

También se han realizado GWAS en los dos principales grupos de arroz indica que han sido generados por actividades de mejoramiento independientes en China (indica I) y el sudeste asiático (indica II, que ha sido generado en gran medida por el *Instituto Internacional de Investigación del Arroz, IRRI*)²². Este estudio identificó ~200 regiones que contenían firmas de domesticación o selección artificial y que abarcaban el 7,8 % del genoma del arroz. Estas regiones albergan alrededor de 4000 genes de elementos no transponibles, incluidos muchos con funciones asociadas con características agronómicas importantes. Los ejemplos incluyen *GN1A23*, que afecta el número de granos por panoja; *SD1* (ref.24), que afecta la altura de la planta; *AMT1;1* (ref.25), que regula la absorción de nitrógeno; *XA4* (ref.26) y *XA26* (también conocido como *XA3*) (ref.27), que están involucrados en la resistencia a enfermedades; y *RF1* (ref.28), que restaura la fertilidad del arroz híbrido.

introgresión

La transferencia de genes y segmentos genómicos de una especie o población a otra a través de la hibridación.

fenotipado de campo
El uso de sensores y sistemas de cámaras de última generación, montados en tractores, puentes y drones, para medir los rasgos fenotípicos de las plantas (como la altura, el ángulo de la hoja, el peso de 1000 granos, la presión de enfermedades y la temperatura del dosel, entre otros). en el transcurso de una temporada de crecimiento.

Genomas de referencia
También conocida como secuencia de referencia (RefSeq). Ensamblaje del genoma que se utiliza para representar la secuencia completa del genoma de un organismo dado. Idealmente, RefSeq estará libre de espacios y tendrá errores de secuencia cero. Sin embargo, los ensamblajes del genoma pueden perder hasta el 50 % de la secuencia completa del genoma, principalmente debido a la tecnología de secuenciación utilizada (por ejemplo, secuenciación de lectura corta) y las herramientas de ensamblaje disponibles.

Revolución verde
El aumento sustancial en la producción de granos que comenzó a fines de la década de 1960 y principios de la de 1970. fue el resultado de adopción generalizada de variedades de trigo y arroz de alto rendimiento mejoradas para incorporar genes seminales y un uso más sistemático de fertilizantes nitrogenados y pesticidas.

Hibridación
El proceso o resultado de realizar cruces genéticos entre individuos de distintas especies o poblaciones muy divergentes.

Barrio selectivo
Una región genómica que parece estar bajo selección natural o artificial. En el contexto de esta Revisión, consideramos que es una región del genoma que incluye y rodea un rasgo de domesticación (por ejemplo, rendimiento o rotura de granos).

Estrés abiótico
Un estrés considerado de origen no biológico, como el calor, la sal, la sequía, los nutrientes, la luz y la oscuridad, entre otros.

Estudios de asociación del genoma completo
(GWAS). Un enfoque de mapeo que se basa en una correlación estadística observada entre las variantes genómicas individuales (como polimorfismos de un solo nucleótido (SNP)) y fenotipos específicos en una población natural.

Además, el rendimiento de grano se correlacionó positivamente con el número de firmas de selección, lo que indica que las firmas podrían ser útiles para predecir el potencial agronómico e implica a los loci seleccionados como objetivos potenciales para los programas de mejoramiento de cultivos.

Maximizar la utilidad de los bancos de germoplasma
Una extensión lógica de los estudios evolutivos discutidos anteriormente es realizar análisis genómicos comparativos sistemáticos de las poblaciones de arroz de Asia y África para identificar genes asociados con la adaptación local al estrés abiótico y otras presiones ambientales y culturales, proporcionando así nuevos objetivos para los esfuerzos de mejoramiento. Dichos análisis requieren un genotipado extensivo (es decir, resecuenciación) combinado con un fenotipado de laboratorio y de campo de alto rendimiento en múltiples condiciones ambientales con el objetivo de traducir genotipo a fenotipo en un formato amigable para el criador.

Generación de bancos de germoplasma digitales. Prácticamente todos los cultivos tienen colecciones de germoplasma, llamadas bancos de germoplasma, que se utilizan para almacenar, preservar y mantener el germoplasma como semilla y/o plantas vivas con fines de reproducción y conservación. Dichos bancos son esenciales para garantizar que la diversidad de cultivos se conserve a largo plazo y que se pueda acceder a ellos fácilmente, generalmente a través de un acuerdo universal de transferencia de material. La mayoría de los bancos de germoplasma de cultivos son nacionales, y la comunidad internacional no siempre tiene fácil acceso a sus semillas. Sin embargo, el [Banco de Germoplasma de Arroz](#) (IRRI, Filipinas), y la [Asociación Internacional de Investigación Agrícola \(CGIAR\)](#) mantiene una serie de bancos de germoplasma internacionales donde los investigadores de todo el mundo pueden acceder y depositar germoplasma valioso con fines de investigación. Hay tres bancos de genes de arroz internacionales: el [Banco de genes de arroz internacional](#) (IRRI, Filipinas), que posee ~128 000 accesiones de arroz y 4 464 parientes silvestres; el banco de germoplasma de [arroz de África](#) (Centro Africano del Arroz (ARC), Costa de Marfil), que posee ~20 000 accesiones; y [Oryzabase](#) (Japón), que posee ~22 000 accesiones, incluidas poblaciones y poblaciones cultivadas, silvestres, mutantes y genéticas.

La resecuenciación de grandes poblaciones es una herramienta importante que se utiliza para desentrañar la estructura de la población, detectar firmas de selección y mapear loci de rasgos cuantitativos (QTL). A medida que los costos de secuenciación se desploman y el rendimiento de las plataformas tecnológicas continúa aumentando, las comunidades de cultivos ahora contemplan la posibilidad de volver a secuenciar colecciones completas de germoplasma para crear bancos de genes digitales. En 2014, la Academia de Ciencias de China (CAS), el Instituto de Genómica de Beijing (BGI) y el IRRI realizaron un experimento piloto de banco genético digital (denominado Proyecto 3000 Genomas de Arroz (3K RGP)) en el que >3000 accesiones de arroz diversas de 89 países fueron resecuenciadas, con una edad de cobertura promedio de 14x (ref.29). Los >17 Tb resultantes de datos sin procesar se filtraron hasta 18,9 millones de polimorfismos de un solo nucleótido (SNP) cuando se asignaron a la secuencia de referencia de Nipponbare (RefSeq)3 solo. En su artículo, propusieron el uso de 'xian' y 'geng' como nombres de grupos varietales, en lugar de indica y japonica respectivamente, en reconocimiento del hecho de que estos nombres se han usado en China durante varios miles de años, lo que corroboramos. Análisis filogenético inicial de 200 000 conjuntos aleatorios de SNP

las variantes permitieron subdividir las 3.000 accesiones en 5 grupos varietales distintos: 2 grupos principales (geng (o japonica) y xian (o indica)); 2 grupos pequeños (aus/boro y basmati/sadri); y un pequeño grupo intermedio (mixto). Estos cinco grupos ahora se han resuelto en nueve grupos mediante el análisis de la estructura de la población30 (Fig. 2b). En total, estimamos que hasta la fecha se han vuelto a secuenciar más de 10y000 accesiones de arroz silvestre y domesticado, y los consorcios se están formando para asegurar el financiamiento para hacer realidad los bancos de germoplasma digitales para la mayoría de los cultivos principales durante los próximos 5 años.

Hacia genomas de referencia estándar de platino. Los esfuerzos para generar bancos de genes digitales serán más poderosos en combinación con genomas de referencia de alta calidad y casi libres de brechas (conocidos como Platinum Standard Reference Sequences, PS-RefSeqs) que abarquen el espectro del arroz cultivado y silvestre. La asignación de datos de resecuenciación a estos genomas de referencia permitirá detectar la gran mayoría de la diversidad de alelos y haplotipos en el genoma del género Oryza pan. Este enfoque también ayudaría a identificar la redundancia, lo que permitiría seleccionar las accesiones más diversas para su uso futuro. PS-RefSeqs están disponibles para el arroz africano O. glaberrima15 y para los dos principales grupos varietales de arroz asiático cultivado: O. sativa japonica3 y O. sativa indica31,32. Los borradores de genomas se publicaron por primera vez para japonica33 e indica (conocidos como 93–11)34 en 2002, pero el primer genoma de referencia verdadero para el arroz, y de hecho para cualquier especie de cultivo, fue el del cultivar japonica Nipponbare, publicado en 2005 (ref. .3). Este Nipponbare RefSeq3 siguió siendo el RefSeq de mayor calidad de cualquier genoma de cultivo hasta 2016, cuando se lanzaron dos genomas de referencia de calidad similar para indica: Minghui 63 y Zhenshan 97 (ref.31).

Los análisis comparativos de estos dos genomas indica descubrieron amplias diferencias estructurales, especialmente con respecto a inversiones, translocaciones, inserciones y delecciones, y duplicaciones segmentarias. Más recientemente, se informó un ensamblaje de novo del cultivar indica Shuhui498 con solo cinco lagunas y siete cromosomas completos de extremo a extremo32. Estas tres secuencias indica de referencia han reemplazado esencialmente la secuencia indica original 93–11 lanzada en 2002.

Ahora se están realizando esfuerzos para generar ensamblajes de genomas de alta calidad para representantes de las seis especies de Oryza con genoma AA salvaje, así como accesiones representativas de especies que albergan otros diez tipos de genomas. De hecho, durante los próximos 2 años, la comunidad de Oryza puede esperar tener acceso a PS-RefSeqs para 2 a 4 accesiones representativas de cada una de las 9 subpoblaciones de arroz cultivado, según lo define el 3K RGP30, y un conjunto completo de PS-RefSeqs para las 25 especies de Oryza silvestres definidas, incluidas 10 especies poliploides (Fig. 2a,b). Hasta la fecha, se han publicado ensamblajes para 35 especies cultivadas y silvestres15,31,32,35–44 (Fig. 2c; consulte la Tabla complementaria 1), y la secuenciación y las actualizaciones de novo están en curso para al menos 30 genomas. Estos proyectos incluyen la secuenciación de novo de tres especies del genoma CC (N. Kurata y M. Shenton, comunicación personal, y RAW, datos no publicados) y la primera especie poliploide de Oryza, Oryza coarctata (tipo de genoma KKLL), un

Selección artificial
Selección de rasgos deseables
que los seres humanos llevan a
cabo consciente y deliberadamente.

resecuenciación
Técnica utilizada para muestrear un
genoma individual sin necesidad de
generar una secuencia genómica
completa. Los datos de resecuenciación
generalmente consisten en lecturas
de secuencias cortas (250 pb) con
una cobertura genómica baja (0,1 a
10 veces) que se mapean por
complementariedad de secuencias a
una secuencia de referencia (RefSeq)
para detectar variaciones genéticas
(como polimorfismos de un solo
nucleótido (SNP) e indels).) entre el
individuo resecuenciado y el RefSeq.

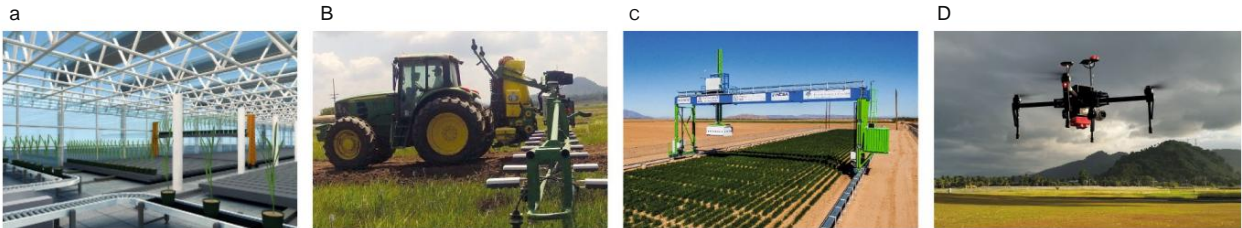
Loci de rasgos cuantitativos
(QTL). Loci genéticos que
contribuyen (positiva o
negativamente) a rasgos no
discretos, como el rendimiento,
la calidad del grano, el estrés hídrico y térmico

especies altamente tolerantes a la sal (RAW, datos no publicados).
La mayoría de los esfuerzos de secuenciación del genoma de *Oryza*
salvaje han sido encabezados por el Proyecto Internacional de
Alineación de Mapas de *Oryza* (I-OMAP), y en el IRRI se mantiene una
colección permanente y públicamente disponible de especímenes
vivos^{42,45}.

Generación de datos fenómicos. Si bien los bancos de genes
digitales son valiosos por derecho propio, la capacidad de vincular la
información genotípica con los fenotipos en un formato de alto
rendimiento es de suma importancia para la identificación rápida de
nuevas combinaciones de alelos adecuadas para el desarrollo de
cultivares GSR. Las plataformas de fenotipado de campo e invernadero
de alto rendimiento, tanto en entornos académicos como industriales,
han avanzado rápidamente en los últimos 10 años y están mejorando
continuamente. Por ejemplo, se ha desarrollado en China una instalación
de fenotipado de arroz de alto rendimiento basada en un invernadero,
que incluye un marcador de rasgos de rendimiento novedoso que
puede medir automáticamente el número total de espiguillas, el número
de granos llenos, la fertilidad de las espiguillas, el rendimiento por
planta, el peso de 1000 granos. , y forma y tamaño de grano a razón
de una planta por minuto⁴⁶ (Fig. 4a).
Recientemente, se utilizaron sensores ultrasónicos y de reflectancia
multiespectrales montados en tractores para generar datos de tipo
fenotípico para un experimento de campo de alto rendimiento en una
población de arroz que comprende 1.516 células endogámicas recombinantes.

líneas⁴⁷ (Fig. 4b). En comparación con las mediciones manuales, el
sistema de tractor de alto rendimiento fue igual de preciso para
características como la altura de la planta, el tiempo de floración, el
rendimiento del grano y el índice de cosecha, pero tomó una fracción
del tiempo para recopilar datos. Otros sistemas de alto rendimiento
basados en el campo incluyen plataformas de pórtico, como el [TERRA-
REF Field Scanalyzer en Arizona](#)
(Fig. 4c) y la plataforma de fenotipado LeasyScan en el Instituto
Internacional de Investigación de Cultivos para los Trópicos Semiáridos
(ICRASAT)⁴⁸ en India, y vehículos aéreos no tripulados⁴⁹ (Fig. 4d).

Independientemente de la plataforma de alto rendimiento, existe
una necesidad apremiante de coordinar actividades intensivas de
fenotipado en un conjunto común de accesiones bajo múltiples
condiciones de estrés biótico y estrés abiótico. En 2011, 15 instituciones
asociadas establecieron el [Global Rice
Phenotyping Network](#) para fenotipar subpaneles bien caracterizados
(como los paneles ORYTAGE y Phenómica de la Adaptación del Arroz
y el Potencial de Rendimiento (PRAY)) que representan subpoblaciones
de indica, japónica tropical y aus y que pueden genotiparse usando
una variedad de métodos^{50–54} , incluida una matriz común de alta
densidad de 700y000 SNP⁵⁵. Recientemente, el interrogatorio de dos
paneles PRAY indica y aus para loci asociados con respuestas
tempranas a la tolerancia a la salinidad identificó un locus previamente
no detectado en el cromosoma 11 que



	Invernadero	Tractor	Portal	Zumbido
Rasgos	Altura y temperatura del dosel, tiempo de floración, estructura 3D, contenido de nitrógeno y clorofila, componentes del rendimiento y biomasa	Días a la floración, altura del dosel, porcentaje de cobertura y nivel de nitrógeno	Altura y temperatura del dosel, tiempo de floración, estructura 3D, contenido de nitrógeno y clorofila, componentes del rendimiento y biomasa	Días a la floración, altura del dosel, porcentaje de cobertura y nivel de nitrógeno 1 a 2 cm de píxel por resolución espacial
Capacidad (Ha/h)*	0.003	1.0	0.1	1,0 (frecuencia de muestreo de 10–20 Ha/h)
ventajas	Completamente automatizado, de alta precisión, opciones ilimitadas de sensores y ambiente controlado	Plataforma móvil multipropósito, facilidad de uso, manejo eficiente de datos y ambiente de campo	Completamente automatizado, de alta precisión, muchas opciones de sensor y campo ambiente	Muestreo rápido, bajo costo, flexible diseños, desplegables en cualquier lugar y entorno de campo
Contras	Costoso, de bajo rendimiento y entorno sin campo	Semiautomatizado, inestable plataforma, diseños fijos y dependientes del clima	Ubicación costosa, lenta y fija y dependiente del clima	Semiautomatizado, sensor limitaciones, preocupaciones regulatorias y dependientes del clima

*La capacidad incluye el tiempo para el procesamiento básico de datos

Figura 4 | **Plataformas de fenotipado de alto rendimiento.** **un** | Instalación de invernadero de fenotipado de arroz de alto rendimiento (HRPF) en Wuhan, China. El HRPF puede fenotipar ~4000 plantas por día. **segundo** | Escáner de campo de alto rendimiento (HTFS) que determina el fenotipo de un arrozal en el Instituto Internacional de Investigación del Arroz (IRRI), Los Baños, Filipinas. El HTFS puede fenotipar 3000 parcelas de 25 plantas (de genotipo idéntico) por hora. **do** | Escáner de campo TERRA-REF para fenotipado de trigo. El Lemnatec Field Scanalyzer del Centro Agrícola Maricopa de la Universidad de Arizona y la Estación de Investigación de Tierras Áridas del Departamento de Agricultura de EE. UU. en Maricopa, Arizona, es el robot de análisis de cultivos de campo más grande del mundo. Este robot de escaneo de campo de fenotipado de alto rendimiento tiene un pórtico de acero de 30 toneladas que se mueve de forma autónoma a lo largo de dos rieles de acero de 200 metros mientras captura continuamente imágenes de los cultivos que crecen debajo de él con una diversa gama de cámaras y sensores. **re** | Dron de vuelo autónomo utilizado para fenotipado en el IRRI. El procesamiento automatizado de imágenes produce un modelo 3D y ortomosaicos georreferenciados de alta resolución que se utilizan para el análisis de datos a nivel de parcela. Ha/hr, hectáreas por hora. La parte **a** es cortesía de Wanneng Yang, Universidad Agrícola de Huazhong, China. Las partes **b** y **d** son cortesía de Stephen Klassen, IRRI, Filipinas. La parte **c** es cortesía de Mark Yori, Phoenix Drone Services, EE. UU.

xian
También conocido como indica, un grupo importante de arroz cultivado en Asia que se cultiva ampliamente en las regiones tropicales y subtropicales de Asia y está parcialmente aislado reproductivamente del arroz geng.
Geng
También conocido como japonica, un grupo importante de arroz cultivado en Asia que se cultiva ampliamente en las regiones templadas de Asia y otras áreas y está aislado reproductivamente parcialmente del arroz xian.
Ejemplares de vales vivos
Accesiones de una sola planta seleccionadas para ser representativas de una especie en particular. En el caso de la <i>Oryza</i> más salvaje especie, los vales se pueden propagar clonalmente indefinidamente.
Estrés biótico
Un estrés considerado de origen biológico, como los patógenos vegetales (bacterias, hongos y virus) y plagas animales (insectos y nematodos), entre otros.
Sitio de unión de nucleótidos y proteínas repetidas ricas en leucina (NBS-LRR)
Miembros de una gran clase de proteínas codificadas por muchos genes de resistencia a enfermedades o insectos (<i>genes R</i>) de las plantas. Una proteína NBS-LRR contiene un dominio del sitio de unión de nucleótidos (NBS) y un dominio repetido rico en leucina (LRR), que se cree que confieren la especificidad de la resistencia.
Inmunidad desencadenada por efectores
Una respuesta de defensa que se inicia cuando una molécula efectora de un patógeno es reconocida por un sitio de unión de nucleótidos del hospedador localizado en el citoplasma y una proteína repetida rica en leucina (NBS-LRR).
Inmunidad activada por patrón
Una respuesta de defensa que se inicia cuando un patrón molecular asociado a un patógeno es reconocido por el receptor de reconocimiento de patrón del huésped correspondiente.

promueve la transpiración eficiente⁵⁵. El IRR1 ha anunciado recientemente otro esfuerzo mucho más grande llamado Global Rice Array (GRA), que propone fenotipar el conjunto de datos de accesión 3K RGP en múltiples ubicaciones alrededor del mundo y bajo numerosas condiciones de campo, invernadero y laboratorio. Por ejemplo, se están realizando esfuerzos para importar el conjunto de datos de acceso 3K RGP a los EE. UU. para el fenotipado no solo en las regiones de cultivo de arroz de los EE. UU., sino también en otras regiones como el TERRA-REF Field Scanalyzer en Maricopa, Arizona (Fig. 4c), para detectar y mapear fenotipos en condiciones ambientales semiáridas.

Acceso e integración de datos genómicos y fenómicos. El fácil acceso a los datos genómicos y fenotípicos es de vital importancia para la investigación aplicada y básica en arroz. Hay varios recursos disponibles para acceder y explorar genomas de referencia de arroz individuales. Por ejemplo, el [Proyecto de Anotación de Arroz \(RAP\)](#)⁵⁶ y [Michigan](#) Los portales de datos de [State University DB \(MSU-DB\)](#) se consideran los principales sitios de acceso para acceder a Nipponbare RefSeq y todas las anotaciones asociadas. Estos sitios se crearon de forma independiente al inicio del Proyecto Internacional de Secuenciación del Genoma del Arroz (IRGSP, por sus siglas en inglés) original y culminaron en una unificación de ensamblaje en 2013 (ref.35), que desde entonces se ha utilizado para todas las actualizaciones de anotaciones posteriores de Nipponbare RefSeq. Se puede encontrar una base de datos para PS-RefSeqs para cultivares de arroz indica Minghui 63 y Zhenshan 97 en [Rice Information Plant Pathogen](#) (RIGW)^{31,57,58} y [R498 en MBKBase](#)³². Alternativamente, se puede acceder a estos y muchos otros ensamblajes del genoma del arroz a través de [Genbank](#).

Aunque estas bases de datos y ensamblajes son extremadamente útiles, tienden a centrarse en genomas únicos y no están configurados para integrar o interrogar otros genomas de arroz cultivado y arroz silvestre a medida que se ponen en línea. Para satisfacer esta necesidad, [Gramene](#) y [Ensemble plantas](#) contienen información sobre el genoma de las accesiones de *Oryza* tanto cultivadas como silvestres, así como múltiples herramientas analíticas para el análisis de datos. Por ejemplo, estos sitios permiten a los investigadores comparar y descargar fácilmente grandes regiones genómicas de interés en múltiples ensamblajes de genomas. Gramene también alberga la [base de datos Rice Diversity](#), que contiene datos de genotipo y fenotipo para la longitud del grano y la arquitectura de la panícula de 1568 variedades diversas de arroz⁵⁹.

A medida que los bancos de germoplasma digitales continúan creciendo, se necesitan bases de datos que puedan almacenar, mapear y entregar SNP, variación estructural y datos fenotípicos a pedido. Una de esas bases de datos para el arroz se llama [SNP-Seek60](#) (Figura 5). SNP-Seek puede mostrar millones de SNP en cientos o miles de accesiones en tiempo real y contiene puntajes de fenotipo para 72 características. En una actualización de 2016, SNP-Seek asignó el conjunto de datos de acceso 3K RGP a cuatro ensamblajes de genoma preliminares adicionales (IR 64 (indica), 93–11 (indica), DJ 123 (aus) y Kasalath (aus)), lo que resultó en el descubrimiento de ~11 millones de nuevos SNP y ~0,5 millones de nuevas inserciones y deleciones³⁰. Estos resultados argumentan firmemente que se necesita una resecuenciación continua y un mapeo de múltiples RefSeqs si queremos obtener una comprensión completa de la variación genética que existe en el arroz cultivado y silvestre.

Del locus candidato al gen

Hasta ahora, hemos revisado el uso de estudios evolutivos y de asociación para identificar regiones del genoma que están bajo selección. Sin embargo, muchos genes importantes para la agricultura relevantes para la reproducción de GSR ya se han aislado y caracterizado funcionalmente utilizando enfoques más tradicionales, como la clonación basada en mapas de mutaciones anulares o inducidas de forma natural o el etiquetado de transposones. A fines de 2017, se analizaron funciones biológicas de un total de 2996 genes utilizando varios enfoques (datos de [funRiceGenes61](#)) y se pueden clasificar en categorías funcionales relacionadas con los objetivos de GSR2 (Fig. 6).

Resistencia a enfermedades o plagas. [La utilización de genes de resistencia a enfermedades e insectos proporciona el enfoque más económico y efectivo para reducir los pesticidas y, por lo tanto, es uno de los principales objetivos en el mejoramiento de GSR.](#) Aproximadamente el 8% de los genes de arroz que se han caracterizado funcionalmente hasta la fecha tienen funciones en la resistencia a enfermedades o plagas (Fig. 6b). Las dos enfermedades principales que con frecuencia conducen a grandes pérdidas de rendimiento en la mayoría de las áreas de cultivo de arroz son el tizón bacteriano y el tizón fúngico, que son causados por *Xanthomonas oryzae* pv. y *Magnaporthe oryzae*, respectivamente. En la última década, se han caracterizado 11 genes de resistencia al tizón bacteriano y 27 genes de resistencia al añublo fúngico y están disponibles con fines de mejoramiento^{62–65}. La comparación de los productos génicos predichos reveló características muy diferentes de las interacciones planta-patógeno entre estas dos enfermedades. Todos menos tres de los loci de resistencia a la explosión fúngica codifican sitios de unión de nucleótidos y proteínas repetidas ricas en leucina (NBS-LRR), que se supone que proporcionan inmunidad desencadenada por efectores y funcionan individualmente o en pares para conferir resistencia específica de raza. Por el contrario, sólo uno de los genes de resistencia al tizón bacteriano codifica una proteína NBS-LRR. Tres de los diez genes restantes activan la expresión de otros genes de resistencia del arroz en lugar de conferir resistencia directamente; cuatro son genes de resistencia recesiva que previenen la invasión del patógeno a través de varios procesos; se supone que dos confieren inmunidad activada por patrón; y el gen final codifica una quinasa asociada a la pared que mejora la resistencia al alojamiento además de proporcionar resistencia al tizón bacteriano⁶⁵. Claramente, los mecanismos moleculares subyacentes a la resistencia al tizón bacteriano son mucho más diversos que los de la resistencia al añublo fúngico, lo que sugiere que se deben adoptar diferentes estrategias para lograr una resistencia duradera a estas dos enfermedades. La formación de pirámides de genes que pertenecen a diferentes categorías funcionales puede proporcionar una solución viable para mejorar la resistencia al tizón bacteriano, mientras que el desarrollo de una serie multilínea de líneas casi isogénicas, cada una con un gen de resistencia diferente, puede ser la solución para una resistencia duradera al añublo del arroz. Se han identificado 30 genes que proporcionan resistencia al saltahoja pardo, una de las plagas del arroz más dañinas, y 12 de ellos han sido [clonados](#)^{66,67}. Las características moleculares de estos genes son sorprendentemente parecidas a las que confieren resistencia a la explosión fúngica: 8 de los 12 genes codifican proteínas NBS-LRR, que presumiblemente proporcionan resistencia desencadenada por efectores como se observa en la interacción planta-patógeno. Por lo tanto, adoptar un enfoque multilínea similar también podría ser una estrategia eficaz para lograr una resistencia duradera a este insecto. Muchos de

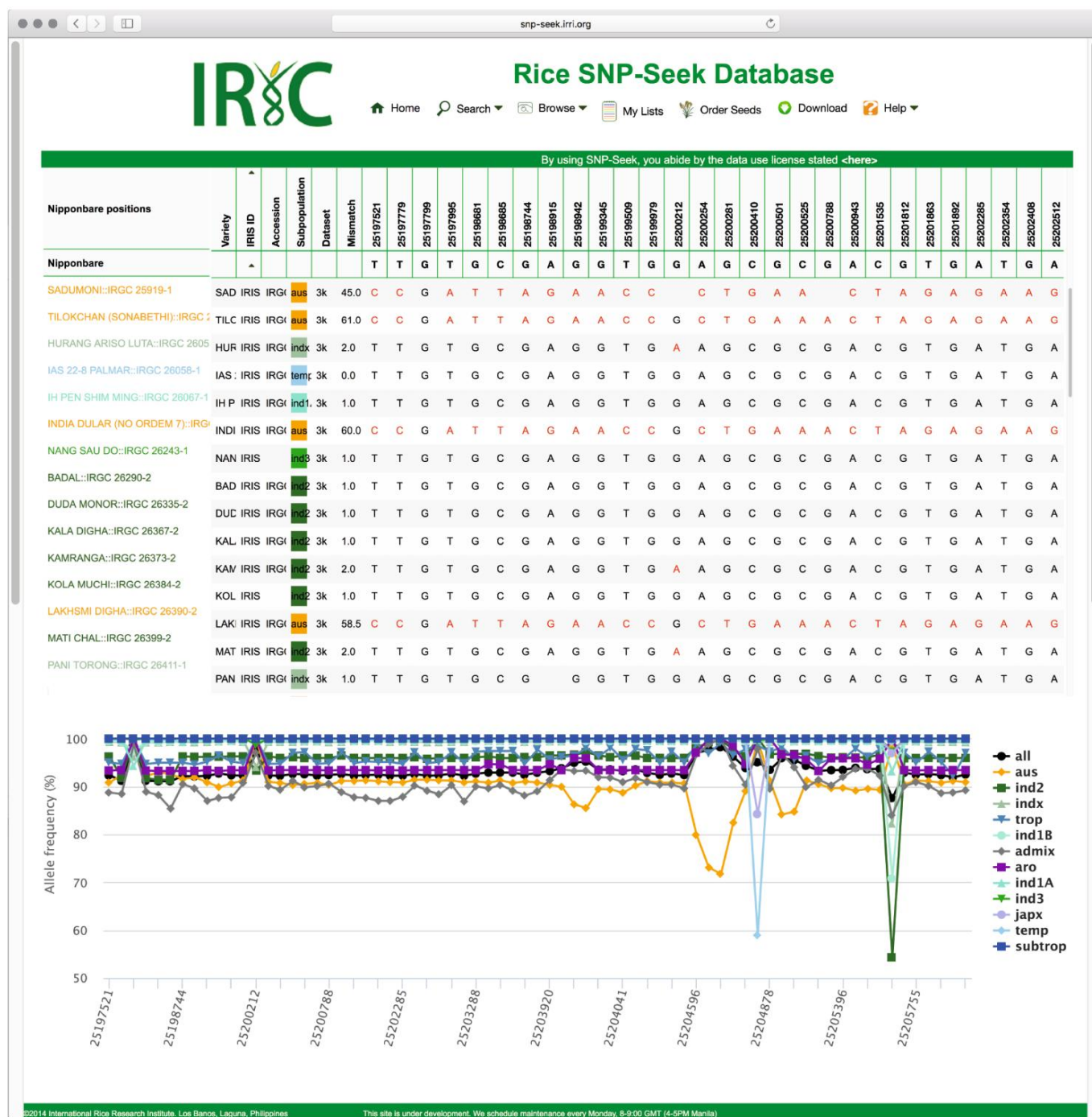


Figura 5 | **caracterización de los genomas de arroz 3K a través de la base de datos snP-seek.** Salida de la base de datos Rice SNP-Seek que perfila llamadas de polimorfismo de un solo nucleótido (SNP) y frecuencias alélicas en una región del gen de destrucción *SH1*. El panel superior muestra SNP específicos (en rojo) en relación con la secuencia de referencia de Nipponbare (RefSeq) en 28 posiciones de nucleótidos en 15 accesiones en el gen *SH1*. El panel inferior muestra las frecuencias alélicas en el gen *SH1* (región de ~8,6 kb) en el cromosoma 3 para el conjunto de datos de acceso completo del Proyecto del Genoma del Arroz 3000 (3K RGP). Tenga en cuenta la baja diversidad alélica de SNP 25197521 a 25204041 seguida de un gran aumento en la diversidad, especialmente en templado japonica en SNP 25204788 e indica-2, indica-1B e indica-x en SNP 25205648. Estos datos indican que estos dos SNP están bajo fuerte selección artificial en esas subpoblaciones.

estos genes han sido ampliamente utilizados en programas de mejoramiento de arroz. Por ejemplo, BPH1, BPH2, BPH3 y BPH4 se han incorporado a una serie de variedades de arroz liberadas por el IRRI, y BPH14 y BPH15 se han incorporado a varios cultivos superiores en China⁶⁷.

También cabe mencionar que se han identificado numerosos genes que proporcionan resistencias cuantitativas, duraderas y no específicas de raza o de patógeno, y que también podrían ser de utilidad para mejorar la resistencia a plagas o patógenos⁶⁸.

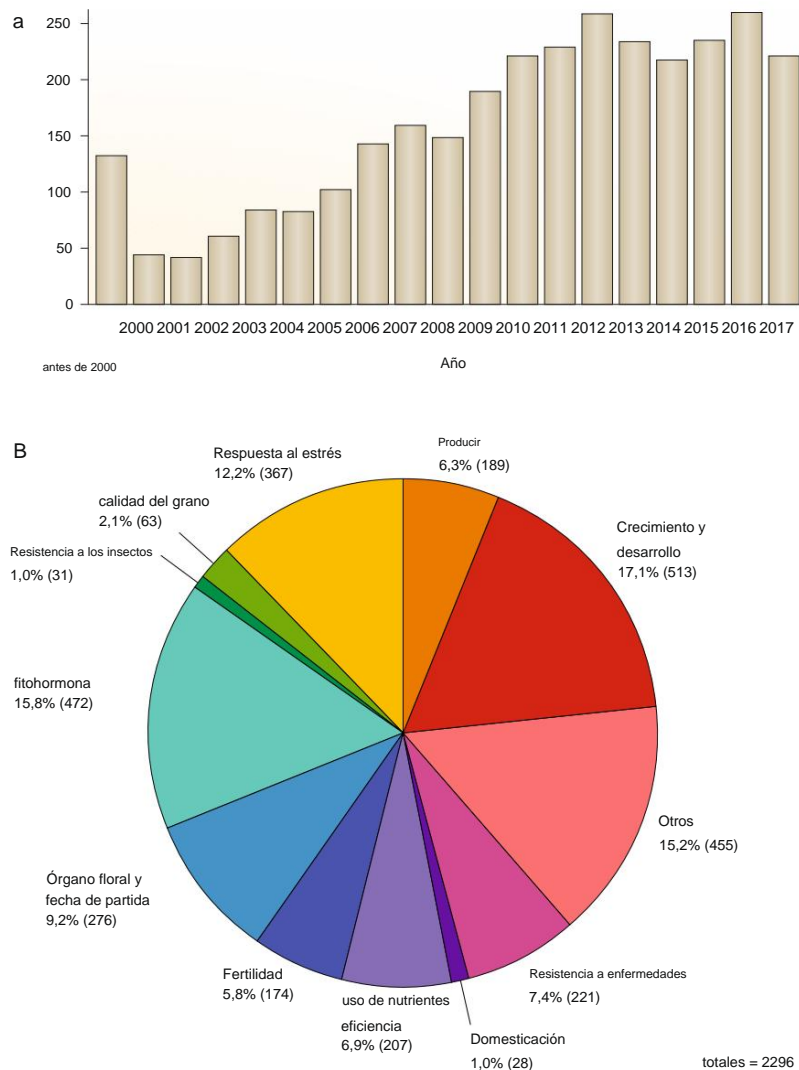


Figura 6 | **Muchos genes de arroz han sido caracterizados y clasificados por función.** un | Gráfico que muestra la cantidad de genes de arroz analizados funcionalmente en la literatura publicada por año según los datos de PubMed a los que se accedió y analizó usando funRiceGenes61. **segundo** | Clasificación funcional de esos genes de acuerdo a las características deseables para el desarrollo del Súper Arroz Verde.

Resistencia al estrés abiótico. Brindar resistencia al estrés abiótico, como la sequía, las inundaciones, la temperatura y la salinidad, es un objetivo intrínseco de GSR, y se han identificado más de 360 genes que están involucrados en las respuestas al estrés (Fig. 6b).

La sequía en particular ha sido identificada como una limitación importante para la producción de arroz⁶⁹, pero a pesar de los tremendos esfuerzos durante las últimas 2 décadas, se han identificado muy pocos genes que puedan ser de uso práctico para mejorar el arroz resistente a la sequía⁷⁰. Sin embargo, se ha logrado un progreso importante en la identificación de genes que confieren resistencia a una variedad de estreses abióticos, incluida la evitación de la sequía (DRO1)⁷¹, la sumersión o la tolerancia a las inundaciones (SUB1 (ref.72), SNORKEL1 (ref.73) y SNORKEL2 (ref.73)), tolerancia a baja temperatura (qLTG3-1 (ref.74), LTG1 (ref.75) (también conocido como HBD2) y COLD1 (ref.76)), tolerancia al calor (TT1 (ref.77) y HTAS78) y tolerancia a la sal (SKC1, también conocido como HKT8)⁷⁹. Varios de estos genes han sido incorporados en programas de mejoramiento⁷⁰, y el ejemplo más conocido es la introgresión de SUB1 de

la variedad autóctona FR13A en una serie de cultivares para mejorar en gran medida la tolerancia a la inmersión⁷².

Uso eficiente de nutrientes. Se han identificado más de 200 genes que mejoran la eficiencia del uso de nutrientes en las plantas de cultivo (Fig. 6b), y el desarrollo de plantas que contengan dichos genes puede proporcionar el enfoque más efectivo para reducir la aplicación de fertilizantes. Los mecanismos moleculares subyacentes al uso de nitrógeno y fósforo se comprenden mejor que los de otros nutrientes y, por lo tanto, son el foco principal de los esfuerzos de mejoramiento de GSR en la actualidad. Se han identificado varios genes potencialmente útiles implicados en el uso de nitrógeno. Por ejemplo, se ha demostrado que un alelo natural del gen transportador de nitrato NRT1.1B del arroz indica mejora la absorción de nitrato en el arroz japónica ica en condiciones de campo⁸⁰. Un mutante natural de DEP1-1, que codifica una proteína Gy, exhibe un crecimiento vegetativo insensible al nitrógeno con una mayor absorción y asimilación de nitrógeno⁸¹. La sobreexpresión de NRT2.3b (también conocido como NRT2.3), un gen transportador de nitrato sensible al pH, aumenta la absorción no solo de nitrógeno sino también de hierro y fósforo⁸². Otros genes involucrados en el uso eficiente del fósforo incluyen el gen PSTOL1 que codifica la proteína quinasa, que aumenta los rendimientos en suelos con deficiencia de fósforo al conferir tolerancia a la deficiencia de fósforo⁸³; y el gen que codifica el transportador de distribución de fósforo similar a SULTR, que cuando se daña en los nudos de las plantas de arroz reduce la asignación de fósforo al grano y reduce su contenido de fósforo sin incurrir en una penalización del rendimiento⁸⁴.

Rendimiento mejorado. El rendimiento de una planta de arroz individual está determinado por tres rasgos componentes: el número de panículas, el número de granos por panícula y el peso del grano⁸⁵. A nivel de población, la arquitectura de la planta también es un rasgo componente importante del rendimiento, ya que determina el número de panículas por unidad de área⁸⁶. Los genes que regulan el rendimiento tienden a ser altamente pleiotrópicos⁸⁵, es decir, afectan muchos rasgos diferentes. Por lo tanto, obtener un rendimiento óptimo a menudo implica equilibrar los diferentes efectos fenotípicos. Por ejemplo, el gen GHD7 regula el número de granos por panícula pero también tiene un gran efecto sobre el tamaño de la planta y el tiempo de floración⁸⁷. También se han encontrado efectos similares para DTH8 (refs 88,89) (también conocido como GHD8 y HD5) GHD7.1 (ref.90) (también conocido como HD2) y HD1 (ref.91), y se ha demostrado que la manipulación del tiempo de floración usando diferentes combinaciones de estos cuatro genes predice el rendimiento de las plantas⁹². También se ha demostrado que la regulación a la baja del gen SGDP7 (también conocido como FZP) por variación natural disminuye el tamaño de grano pero aumenta el número de granos por panícula⁹³. Además, se necesitan niveles óptimos de expresión del gen IPA1 (también conocido como WFP) para lograr una arquitectura de planta ideal que tenga menos macollos pero más productivos, lo que resulta en un mayor rendimiento⁹⁴. La pleiotropía asociada con los genes que afectan el rendimiento significa que puede parecer que están subrepresentados en las clasificaciones funcionales (Fig. 6b); sin embargo, los genes clasificados como que afectan el crecimiento y el desarrollo⁹⁵ y el órgano floral y la fecha de espiga también afectan el rendimiento directa o indirectamente. Por lo tanto, es importante caracterizar redes reguladoras completas, en lugar de rasgos componentes o genes individuales, cuando se considera el rendimiento. Arroz híbrido

ha contribuido en gran medida al aumento de la producción de granos a nivel mundial, y la explotación de la heterosis entre las subespecies y/o grupos varietales de indica y japonica se ha considerado durante mucho tiempo como el siguiente paso para aumentar aún más los niveles de rendimiento. Desafortunadamente, los avances se han visto limitados por la esterilidad híbrida intersubespecífica. Sin embargo, ahora se han identificado y caracterizado varios loci de esterilidad híbrida, lo que promete estimular el progreso en esta [área](#)^{96–99}.

Mejora de la calidad del grano. La calidad del grano consta de un grupo de características que son importantes para los consumidores y, en cierta medida, para los productores de arroz (es decir, agricultores y molineros). Aunque las preferencias varían en diferentes partes del mundo, la calidad del grano de arroz generalmente se caracteriza por las siguientes características: apariencia (forma y translucidez); calidad de molienda (tiza e integridad del arroz molido); calidad para cocinar y comer (la combinación de contenido de amilosa, consistencia del gel y temperatura de gelificación); y calidad nutricional (contenido de macronutrientes y micronutrientes). Amplios estudios genéticos han identificado muchos loci que regulan los rasgos de calidad del grano¹⁰⁰

(Fig. 6b), incluyendo WAXY, que tiene efectos importantes en la calidad culinaria¹⁹; ALK, que regula la temperatura de gelificación¹⁰¹; y CHALK5, que afecta la translucidez de los granos¹⁰². Además, los genes identificados por sus efectos sobre el tamaño del grano (largo y ancho), como GS3 (ref. 103) (también conocido como LK3) y GW5 (refs. 104,105), también afectan muchos aspectos de la calidad del grano porque la apariencia y el molino La calidad de los granos está estrechamente relacionada con el tamaño y la forma del grano. Con la mayor conciencia de los efectos de los productos vegetales que promueven la salud y la demanda de un arroz más nutritivo, la investigación genómica funcional también debería abordar la necesidad de granos de arroz enriquecidos en macronutrientes y micronutrientes, como hierro y zinc, así como clases especiales de metabolitos como el γ -carotenoide y la antocianina. Se han logrado avances importantes en el aumento de los niveles de algunos de los nutrientes utilizando enfoques transgénicos^{106,107}; sin embargo, la aceptación pública de los alimentos modificados genéticamente sigue siendo baja. Por lo tanto, la disponibilidad de variedades de arroz mejoradas con nutrientes puede depender de la identificación e incorporación de alelos naturales utilizando enfoques de mejoramiento genómico.

Mejoramiento genómico de Green Super Rice

Cada año durante el último medio siglo, los enfoques de mejoramiento convencionales han generado cientos de cultivares de arroz con rendimiento y calidad mejorados para mantener el ritmo de las crecientes demandas de alimentos de una población mundial en crecimiento. Además de aumentar el rendimiento y la calidad, el desarrollo de variedades GSR también se enfoca en brindar resistencia a múltiples estreses bióticos y abióticos y en el uso eficiente de nutrientes. Se requiere una gran cantidad de genes de varias fuentes para conferir todos estos rasgos, y combinarlos en un solo cultivo sería muy difícil de lograr con tecnologías convencionales. Sin embargo, nuestra comprensión cada vez mayor de la evolución, función y regulación de los genes y rasgos clave del arroz, y el desarrollo de plataformas técnicas sofisticadas para el fenotipado y el mejoramiento molecular, significa que cada vez es más factible científicamente y técnicamente práctico generar GSR. por mejoramiento genómico (Fig. 1).

Conceptualmente, el mejoramiento genómico consta de dos componentes principales: el diseño genómico y la selección del genoma completo. Usando un cultivar élite específico como punto de partida, la fase de diseño implica hacer coincidir una lista de mejoras de características requeridas (como áreas de cultivo específicas, alto rendimiento, alta calidad de grano, uso eficiente de nutrientes, uso eficiente del agua y resistencia a las principales enfermedades e insectos, entre otros) a una lista de genes que pueden generar los fenotipos deseados. Deben identificarse los recursos de germoplasma para estos genes objetivo y deben determinarse las estrategias para ensamblar los genes e introducirlos en el cultivar élite. Claramente, las diferentes regiones productoras de arroz enfrentan diferentes desafíos en términos de recursos y medio ambiente y, por lo tanto, cada región requerirá su propia combinación de rasgos verdes y el desarrollo de su propia variedad GSR. El diseño de GSR también se verá afectado por el rápido desarrollo socioeconómico que se está produciendo en muchos países productores de arroz, como China, donde se están produciendo cambios drásticos en los sistemas de producción de arroz. Por lo tanto, los programas de mejoramiento de GSR deben actualizarse periódicamente para incluir rasgos que aumenten la eficiencia de la producción de arroz en los nuevos sistemas agrícolas.

El sistema de selección tiene dos componentes: una plataforma para el genotipado de las plantas con respecto a los genes diana y los antecedentes genómicos y un sistema de selección de genes específicos. La plataforma de genotipado toma la forma de un chip de reproducción. Se han desarrollado varios chips de reproducción^{108–110} basados en datos de variación genética obtenidos de proyectos de resecuenciación a gran escala y en datos de función genética obtenidos de la bibliografía que informa sobre la caracterización de genes individuales. El sistema de selección de genes específicos debe comprender un marcador funcional para seleccionar cada gen objetivo y marcadores de ADN estrechamente vinculados que flanquean cada gen objetivo (Fig. 7) para detectar la recombinación entre los genomas del donante y el receptor y para facilitar la selección de líneas en las que el gen diana se ha incorporado con precisión. La literatura científica contiene una gran cantidad de información funcional sobre los genes del arroz, que se puede utilizar para desarrollar sistemas de selección específicos de genes.

Este sistema de selección proporciona un medio eficaz para la mejora específica de cualquier característica del arroz para la que se haya identificado un locus causal. Por ejemplo, se ha utilizado el mejoramiento genómico para generar una serie multilínea resistente al anfibio fúngico (QZ, datos no publicados y F. Zhou, comunicación personal): cuatro genes de resistencia al anfibio (PI1, PI2 (también conocido como PI2), PI9 y PIGM) se incorporaron individualmente al genoma de Kongyu131, el cultivar de arroz más cultivado en la provincia de Heilongjiang, en el noreste de China. Todo el proceso involucró un cruce, cuatro retrocruzamientos y una generación de autopolinización (Fig. 7).

Durante el proceso de reproducción, el chip de reproducción 6K108 y el sistema de selección específico de genes desarrollado para cada gen facilitaron la selección de los genes objetivo, los eventos de recombinación y los antecedentes genómicos. Se obtuvieron cuatro líneas casi isogénicas (NIL), cada una de las cuales contenía un fragmento de ADN muy corto (<200 kb) que contenía uno de los genes de resistencia a la explosión relevantes de la línea donante correspondiente. Una prueba de campo a gran escala (~500Ha) mostró que estas líneas eran fenotípicamente muy uniformes con los niveles requeridos de resistencia, tanto en rodales puros de líneas individuales como en una mezcla de las cuatro líneas. Esta prometedora estrategia para generar multilíneas con resistencia duradera

chip de cría

Una micromatriz que permite el genotipado de alto rendimiento y la selección de descendientes en programas de mejoramiento.

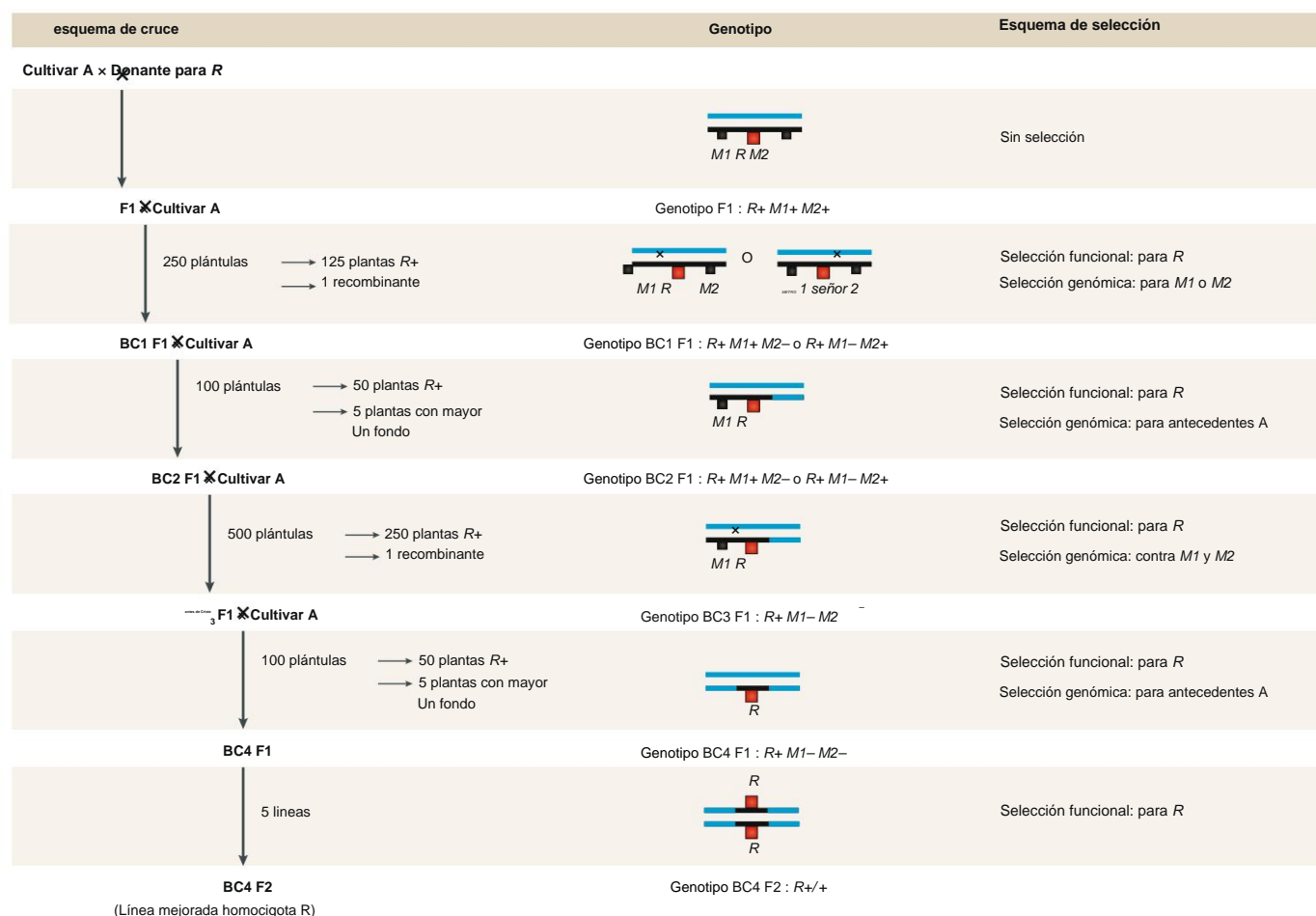


Fig. 7 |Un esquema de mejoramiento genómico para introducir con precisión un gen en el trasfondo de un cultivar élite. Se introduce un gen diana (R) en el fondo del Cultivar A mediante retrocruzamientos sucesivos. La selección se realiza utilizando un sistema de selección de genes específicos que consta de un marcador funcional para seleccionar el gen objetivo (selección directa) y marcadores de ADN en cada lado del gen objetivo (M1 y M2) para identificar la recombinación entre el gen objetivo y el fondo genético (selección hacia atrás). Cada uno de los marcadores de ADN debe ser <100 kb del gen objetivo. Los números de plantas indicados para cada generación son expectativas basadas en una proporción promedio de distancia física a genética de 250 kb/cM en todo el genoma del arroz.

en condiciones de producción a gran escala ilustra dos claras ventajas del mejoramiento genómico en comparación con el mejoramiento convencional. El primero es la velocidad: se obtuvieron líneas homocigóticas mejoradas en 2,5 años (con tres generaciones por año), mientras que con métodos de mejoramiento convencionales se requieren 5 a 6 años. La segunda ventaja, y la más clara, es la incorporación precisa del gen objetivo en el genoma de la línea receptora. Esta característica permite generar una serie de NIL, cada uno de los cuales se diferencia de los demás por un fragmento genómico muy pequeño, lo que proporciona una resistencia genéticamente diversa pero fenotípicamente uniforme en un campo de arroz. Se ha clonado en arroz una gran cantidad de genes que confieren resistencia al anillo que podrían usarse para generar una gran cantidad de NIL a través de este enfoque multilínea. La rotación espacial y/o temporal de diferentes mezclas que comprenden de cuatro a cinco NIL del grupo más grande podría proporcionar una estrategia para una resistencia duradera en la producción de arroz. Además, los genes para diferentes rasgos se pueden combinar fácilmente en un trasfondo genético homogéneo con un solo cruce. Por lo tanto, las variedades GSR podrían desarrollarse incorporando múltiples genes, cada uno para un rasgo

Conclusiones y perspectivas de futuro

Ya se han desarrollado y plantado en grandes áreas de China y otros países asiáticos, y en ensayos de demostración en varios países africanos, cultivares con características individuales de GSR (como resistencia a la explosión, resistencia al saltahoja pardo, resistencia a la sequía o un uso más eficiente del nitrógeno). Además, se han iniciado importantes proyectos para desarrollar GSR en China y por la Fundación Bill y Melinda Gates para la Cooperación Internacional.

Los avances tecnológicos recientes en la biología del genoma (como la secuenciación, la edición del genoma y el mejoramiento molecular) han abierto la puerta para que los biólogos de plantas exploren y prueben rápidamente un suministro ilimitado de variación natural en todo el género *Oryza* a nivel de secuencia de ADN. La tarea más urgente y desafiante que enfrenta la comunidad ahora es cómo asociar esta variación de secuencia natural con fenotipos en ecosistemas variados. Para cerrar esta brecha de genotipo a fenotipo, proponemos que los investigadores de arroz de todo el mundo se asocien con centros internacionales de arroz (como IRRI, AfricaRice y el Centro Internacional de Agricultura Tropical) para desarrollar una base de alto rendimiento

fenotipado en grandes paneles de accesiones de arroz resecuenciadas (como el conjunto de datos de accesiones 3K RGP) en diversas condiciones de campo. Los equipos internacionales de temas específicos interrogarían a estos paneles por su conjunto específico de rasgos, como el rendimiento agronómico, la arquitectura de la raíz, la calidad del grano y los microbiomas, entre otros. Todos los datos se depositarían en una base de datos central (como SNP-Seek en <https://www.snpseek.org/>) y se utiliza para guiar y acelerar el proceso de mejoramiento genómico para desarrollar cultivares GSR adaptados a cualquier parte del mundo. El trabajo de fenotipado de línea de base podría ser financiado por los donantes del centro CGIAR y los interrogatorios de temas específicos por parte de las agencias de subvenciones nacionales. Este enfoque, si se adopta, podría servir como modelo para otros sistemas de cultivo.

De hecho, el progreso hacia el desarrollo de GSR ya está estableciendo un paradigma para los objetivos de mejoramiento en otros cultivos. Por ejemplo, en 2017, el Ministerio de Agricultura de China implementó Green Variedades, un nuevo sistema de evaluación y certificación de variedades para el lanzamiento de las principales variedades de cultivos que requieren menos fertilizantes, pesticidas o agua. Este esquema, por primera vez, ha abierto la puerta para el reconocimiento de variedades verdes en un sistema oficial de certificación de variedades. Tales iniciativas fomentarán el desarrollo de sistemas de agricultura verde y contribuirán al objetivo final de garantizar la producción sostenible de alimentos para los 10,000 millones de personas que se prevé que habitarán el mundo para 2050.

Published online: 05 June 2018

1. Departamento de Asuntos Económicos y Sociales de las Naciones Unidas/División de Población. Perspectivas de la población mundial: hallazgos clave y tablas de avance. ESA https://esa.un.org/unpd/wpp/publications/Files/WPPP2017_Resultados clave.pdf (2017).

2. Zhang, Q. Estrategias para desarrollar un súper arroz verde. *proc. Academia Nacional. ciencia EE. UU.* **104**, 16402–16409 (2007).

Este artículo introduce el concepto de RSG.

3. Matsumoto, T. et al. La secuencia basada en mapas del genoma del arroz. *Naturaleza* **436**, 793–800 (2005).

Este artículo describe la secuencia final del genoma del arroz, el primer genoma de cultivo en ser secuenciado.

4. Yang, C., Yang, Z., Hu, J., He, G. & Shu, L. Estudio sobre la resistencia a la chicharrita parda en líneas introgresivas de arroz silvestre. *Acta Phytophylac. Pecado.* **26**, 197–202 (1999).

5. Buckler, ES, Thomsberry, JM y Kresovich, S. Diversidad molecular, estructura y domesticación de gramíneas. *Gineta. Res.* **77**, 213–218 (2001).

6. Calcedo, AL et al. Patrones de polimorfismo de nucleótidos en todo el genoma en arroz domesticado. *PLoS Genet.* **3**, 1745–1756 (2007).

7. Cívaj, P., Craig, H., Cox, CJ & Brown, TA Three domesticaciones geográficamente separadas del arroz asiático. *Nat. Plantas* **1**, 15164 (2015).

8. Molina, J. et al. Evidencia molecular de un solo origen evolutivo del arroz domesticado. *proc. Academia Nacional. ciencia EE. UU.* **108**, 8351–8356 (2011).

9. Huang, X. et al. Un mapa de la variación del genoma del arroz revela el origen del arroz cultivado. *Naturaleza* **490**, 497–501 (2012).

10. Choi, JY et al. La paradoja del arroz: orígenes múltiples pero domesticación única en el arroz asiático. *mol. Biol. Evol.* **34**, 969–979 (2017).

Este estudio utiliza múltiples genomas de referencia y modelos demográficos para respaldar un solo evento de domesticación del arroz japónica seguido de una introgresión para producir variedades indica y aus.

11. Li, C., Zhou, A. y Sang, T. Domesticación del arroz por reduciendo la rotura. *Ciencia* **311**, 1936–1939 (2006).

12. Stevens, CJ et al. Entre China y el sur de Asia: un Corredor de Asia central de dispersión de cultivos e innovación agrícola en la Edad del Bronce. *Holocene* **26**, 1541–1555 (2016).

13. Sweeney, MT y col. Diseminación global de una sola mutación que confiere pericarpio blanco en arroz. *PLoS Genet.* **3**, e133 (2007).

14. Tan, L. et al. Control de una transición clave de crecimiento postrado a erecto en la domesticación del arroz. *Nat. Gineta.* **40**, 1360–1364 (2008).

15. Wang, M. et al. La secuencia del genoma del arroz africano (*Oryza glaberrima*) y evidencia de domesticación independiente. *Nat. Gineta.* **46**, 982–988 (2014).

Este documento informa sobre el primer genoma de referencia disponible para el arroz africano y proporciona evidencia de una selección convergente pero independiente de un conjunto común de genes durante dos procesos de domesticación geográfica y culturalmente distintos.

16. Meyer, RS et al. Historia de la domesticación y adaptación geográfica inferida de un mapa SNP de arroz africano. *Nat. Gineta.* **48**, 1083–1088 (2016).

17. Olsen, KM & Purugganan, MD Evidencia molecular sobre el origen y evolución del arroz glutinoso. *Genética* **162**, 941–950 (2002).

18. Olsen, KM et al. Selección bajo domesticación: evidencia de un barrido en la región genómica cerosa del arroz. *Genética* **173**, 975–983 (2006).

Este estudio proporciona una demostración temprana de la huella genómica de la selección asociada con un rasgo culturalmente significativo.

19. Tan, YF et al. Los tres rasgos importantes para cocinar y comer la calidad de los granos de arroz están controlados por un solo locus en un híbrido de arroz élite, Shanyou 63. *Theor. aplicación Gineta.* **99**, 642–648 (1999).

20. Yang, T. et al. El papel de un transportador de potasio OsHAK5 en la adquisición y el transporte de potasio desde las raíces hasta los brotes en arroz con bajos niveles de suministro de potasio. *Fisiol. vegetal.* **166**, 945–959 (2014).

21. Ruan, SL et al. Identificación proteómica de OsCYP2, una ciclofilina de arroz que confiere tolerancia a la sal en plántulas de arroz (*Oryza sativa* L.) cuando se sobreexpresa. *BMC Plant Biol.* **11**, 34 (2011).

22. Xie, W. et al. Firmas de mejoramiento de arroz reveladas por un mapa de variación genómica de una gran colección de germoplasma. *proc. Academia Nacional. ciencia EE. UU.* **112**, E5411–5419 (2015).

Este estudio identifica dos grupos principales de arroz indica, basados en firmas de mejoramiento, que resultaron de actividades de mejoramiento independientes en diferentes regiones de Asia.

23. Wang, J. et al. La selección artificial de Gn1a juega un papel importante en la mejora de los rendimientos de arroz en diferentes regiones ecológicas. *Arroz* **8**, 37 (2015).

24. Spielmeier, W., Ellis, MH y Chandler, PM Semidwarf (sd-1), arroz de la “revolución verde”, contiene un gen defectuoso de giberelina 20-oxidasa. *proc. Academia Nacional. ciencia EE. UU.* **99**, 9043–9048 (2002).

25. Hoque, MS, Masle, J., Udvardi, MK, Ryan, PR y Upadhyaya, NM La sobreexpresión del gen OsAMT1-1 del arroz aumenta la absorción y el contenido de amonio, pero perjudica el crecimiento y desarrollo de las plantas bajo una nutrición rica en amonio. *Función Biol. vegetal* **33**, 153–163 (2006).

26. Khush, GS, Mackill, DJ & Sidhu, GS en *Bacterial Blight of Rice* (eds Banta, SJ, Cervantes, E. & Mew, TW) 207–217 (International Rice Research Institute, 1989).

27. Sun, X. et al. Xa26, un gen que confiere resistencia a *Xanthomonas oryzae* pv. *oryzae* en el arroz, codifica una proteína similar a la quinasas del receptor LRR. *Planta* **J. 37**, 517–527 (2004).

28. Kazama, T. & Toriyama, K. Un gen que contiene una repetición de pentatricopéptido que promueve el procesamiento de ARN atp6 aberrante de arroz citoplasmático con esterilidad masculina. *FEBS Lett.* **544**, 99–102 (2003).

29. El Proyecto 3000 Genomas de Arroz. El proyecto 3.000 genomas de arroz. *GigaScience* **3**, 7 (2014).

30. Wang, W. et al. Variación genómica en 3.010 accesiones diversas de arroz asiático cultivado. *Naturaleza* **557**, 43–49 (2018).

Este artículo describe los análisis pangenómicos que descubrieron, de 29 millones de SNP, 2,4 millones de índices cortos, más de 90,000 variantes estructurales y más de 10,000 genes nuevos, todos los cuales contribuyen a la variación de la población en el arroz cultivado en Asia. Además, los patrones detectados de introgresión en varios genes de domesticación respaldan múltiples domesticaciones independientes en el arroz asiático.

31. Zhang, J. et al. Amplia divergencia de secuencias entre los genomas de referencia de dos variedades élite de arroz indica Zhenshan 97 y Minghui 63. *proc. Academia Nacional. ciencia EE. UU.* **113**, E5163–E5171 (2016).

32. Du, H. et al. Secuenciación y ensamblaje de novo de un genoma de arroz indica casi completo. *Nat. común* **8**, 15324 (2017).

33. Goff, SA et al. Un borrador de secuencia del genoma del arroz (*Oryza sativa* L. ssp. japonica). *Ciencia* **296**, 92–100 (2002).

34. Yu, J. et al. Un borrador de secuencia del genoma del arroz (*Oryza sativa* L. ssp indica). *Ciencia* **296**, 79–82 (2002).

35. Kawahara, Y. et al. Mejora de la *Oryza sativa* Genoma de referencia de Nipponbare utilizando secuencias de próxima generación y datos de mapas ópticos. *Arroz* **6**, 4 (2013).

36. Yamamoto, T. et al. Fina definición del pedigrí. haplotipos de cultivares de arroz estrechamente relacionados mediante el descubrimiento de polimorfismos de un solo nucleótido en todo el genoma. *BMC Genomics* **11**, 267 (2010).

37. Takagi, H. et al. MutMap-Gap: genoma completo la resecuenciación de la progenie masiva de F2 mutante combinada con el ensamblaje de novo de las regiones de brecha identifica el gen Pii de resistencia al ahüello del arroz. *Fitol nuevo.* **200**, 276–283 (2013).

38. Lu, L. et al. El seguimiento de los resultados de todo el genoma de un explosión de elementos transponibles durante décadas de amplificación. *proc. Academia Nacional. ciencia EE. UU.* **114**, E10550–E10559 (2017).

39. Yu, J. et al. Los Genomas de *Oryza sativa*: una historia de duplicaciones. *PLoS Biol.* **3**, e38 (2005).

40. Reddy, MM & Ulaganathan, K. Proyecto de secuencia del genoma de *Oryza sativa* elite indica cultivar RP Bio-226. *Parte delantera. ciencia de las plantas* **6**, 896 (2015).

41. Mahesh, HB et al. Ensamblaje del genoma del arroz indica, anotación y extracción de genes de resistencia a la enfermedad del ahüello. *BMC Genomics* **17**, 242 (2016).

42. Stein, JC et al. Los genomas de 13 parientes del arroz silvestre y domesticado destacan la conservación genética, el recambio y la innovación en todo el género *Oryza*. *Nat. Gineta.* **50**, 285–296 (2018).

43. Zhang, Y. et al. El genoma y la transcriptómica comparativa del arroz silvestre africano *Oryza longistaminata* brindan información sobre el mecanismo molecular de la rizomatosis y la autoincompatibilidad. *mol. Planta* **8**, 1683–1688 (2015).

44. Chen, JF et al. La secuenciación del genoma completo de *Oryza brachyantha* revela los mecanismos subyacentes a la evolución del genoma de *Oryza*. *Nat. común* **4**, 1595 (2013).

45. Jacquemin, J., Bhatia, D., Singh, K. y Wing, RA El proyecto internacional de alineación del mapa de *Oryza* : desarrollo de una plataforma de genómica comparativa de todo el género para ayudar a resolver la pregunta de los 9 mil millones de personas. *actual Opinión Biol. vegetal* **16**, 147–156 (2013).

46. Yang, W. et al. Combinando alto rendimiento fenotipado y estudios de asociación de todo el genoma para revelar la variación genética natural en el arroz. *Nat. común* **5**, 5087 (2014).

Este trabajo informa sobre una instalación de fenotipado de alto rendimiento y demuestra que los datos se pueden usar para GWAS de rasgos agronómicos.

47. Tänger, P. et al. Alto rendimiento basado en campo el fenotipado identifica rápidamente las regiones genómicas que controlan los componentes del rendimiento en el arroz. *ciencia Rep.* **7**, 42839 (2017).

48. Vadez, V. et al. LeasyScan: un concepto novedoso que combina imágenes 3D y lisimetría para el fenotipado de alto rendimiento de rasgos que controlan el presupuesto hídrico de la planta. *Exp. J. Bot.* **66**, 5581–5593 (2015).

49. Shi, Y. et al. Vehículos aéreos no tripulados para alta fenotipado de rendimiento e investigación agronómica. *PLoS ONE* **11**, e0159781 (2016).

50. Courtois, B. et al. Mapeo de asociación de todo el genoma de los rasgos de la raíz en un panel de arroz japónica. *PLoS ONE* **8**, e78037 (2013).

51. Rebolledo, MC et al. Fenotípica y genética
diseción de rasgos componentes para el vigor temprano en arroz
utilizando modelos de crecimiento de plantas, análisis de contenido de
azúcar y mapeo de asociaciones. *Exp. J. Bot.* **66**, 5555–5566 (2015).

52. Qiu, X. et al. Estudio de asociación de todo el genoma de la apariencia del
grano y la calidad de la molienda en una colección mundial de
germoplasma de arroz indica. *PLoS ONE* **10**, e0145577 (2015).

53. Rebolledo, MC et al. Combinando análisis de imágenes,
estudios de asociación del genoma completo y diferentes ensayos de
campo para revelar regiones genéticas estables relacionadas con la
arquitectura de la panícula y el número de espiguillas por panícula en el
arroz. *Parte delantera. ciencia de las plantas* **7**, 1384 (2016).

54. Kikuchi, S. et al. Mapeo de asociación de todo el genoma
plasticidad fenotípica en arroz. *Entorno de células vegetales*. **40**, 1565–
1575 (2017).

55. Al-Tamimi, N. et al. Locus de tolerancia a la salinidad revelados en arroz
utilizando fenotipado no invasivo de alto rendimiento.
Nat. común **7**, 13342 (2016).

56. Sakai, H. et al. Base de datos del proyecto de anotación de arroz
(RAP-DB): una base de datos integradora e interactiva para la genómica
del arroz. *Fisiol. de células vegetales*. **54**, e6 (2013).

57. Zhang, J. et al. Construyendo dos referencias de arroz indica
genomas con datos de secuenciación final emparejada de PacBio de
lectura larga e Illumina. *ciencia Datos* **3**, 160076 (2016).

58. Canción, JM et al. Portal de información sobre el arroz: un
plataforma bioinformática integral para genomas de arroz Indica. *mol.*
Planta **11**, 505–507 (2018).

59. McCouch, SR et al. Recursos de acceso abierto para el mapeo de
asociación de todo el genoma en arroz.
Nat. común **7**, 10532 (2016).

60. Mansueto, L. et al. Actualización de la base de datos de búsqueda de SNP
de Rice: nuevos SNP, índices y consultas. *Ácidos Nucleicos Res.* **45**,
D1075–D1081 (2017).

61. Yao, W., Li, G., Yu, Y. y Ouyang, Y. Conjunto de datos funRiceGenes para una
comprensión y aplicación integrales de los genes funcionales del arroz.
Gigasciencia **7**, 1–9 (2018).

62. Zhang, HT & Wang, SP Progreso en funcional
Estudios genómicos de resistencia a enfermedades del arroz. *Boletín
chino de ciencias de la vida*. **28**, 1189–1199 (2016).

63. Deng, Y. et al. La regulación epigenética de los receptores antagonistas
confiere resistencia al añublo del arroz con equilibrio en el rendimiento.
Ciencia **355**, 962–965 (2017).

64. Li, W. et al. Un alelo natural de un factor de transcripción en
el arroz confiere una resistencia de amplio espectro al añublo. *Celda* **170**,
114–126 (2017).

65. Hu, K. et al. Mejora de múltiples rasgos agronómicos por un gen de resistencia
a enfermedades a través del refuerzo de la pared celular. *Nat. Plantas* **3**,
17009 (2017).

66. Zhao, Y. et al. La diversidad alélica en un gen NLR BPH9 permite que el
arroz combata la variación del saltohojas.
proc. Academia Nacional. ciencia EE . UU . **8**, 12850–12855 (2016).

67. Hu, J., Xiao, C. & He, Y. Avances recientes en la genética y el
mejoramiento molecular de la resistencia a la chicharrita parda
en el arroz. *Arroz* **9**, 30 (2016).

68. Ke, Y., Deng, H. & Wang, S. Avances en la comprensión de la resistencia de
amplio espectro a los patógenos en el arroz. *planta j.*
90, 738–748 (2017).

69. Lin, JY & Shen, M. *Restricciones de producción de arroz en China* (CAB
International, 1996).

70. Hu, H. & Xiong, L. Ingeniería genética y mejoramiento de cultivos resistentes
a la sequía. *año Rev. Plant Biol.* **65**, 715–741 (2014).

71. Uga, Y. et al. Control de la arquitectura del sistema raíz por
EL ENRAIZAMIENTO PROFUNDO 1 aumenta el rendimiento del arroz en
condiciones de sequía. *Nat. Gineta.* **45**, 1097–1102 (2013).

72. Xu, K. et al. Sub1A es un gen similar al factor de respuesta al etileno que
confiere al arroz tolerancia a la sumersión.
Naturaleza **442**, 705–708 (2006).

**Este artículo describe la clonación y caracterización del gen SUB1A .
Desde entonces, este gen se integro en las megavariiedades
asiáticas para ayudar a resistir inundaciones por inmersión de hasta 2
semanas.**

73. Hattori, Y. et al. Los factores de respuesta del etileno
SNORKEL1 y SNORKEL2 permiten que el arroz se adapte a aguas
profundas. *Naturaleza* **460**, 1026–1030 (2009).

74. Fujino, K. et al. Identificación molecular de un locus de rasgos cuantitativos
principales, qLTG3-1, que controla la germinabilidad a baja temperatura
en el arroz. *proc. Academia Nacional. ciencia
EE . UU .* **105**, 12623–12628 (2008).

75. Lu, G. et al. Rice LTG1 participa en el crecimiento adaptativo y la aptitud
física a baja temperatura ambiente. *planta j.*
78, 468–480 (2014).

76. Ma, Y. et al. COLD1 confiere tolerancia al frío en arroz.
Celda **160**, 1209–1221 (2015).

77. Li, XM et al. Los alelos naturales de un gen de la subunidad γ 2 del proteasoma
contribuyen a la termotolerancia y adaptación del arroz africano. *Nat. Gineta.*
47, 827–833 (2015).

78. Liu, J. et al. El dedo anular ubiquitina E3 ligasa OsHTAS mejora la
tolerancia al calor al promover el cierre estomático inducido por
H2O2 en el arroz. *Fisiol. vegetal.*
170, 429–443 (2016).

79. Ren, ZH et al. Un locus de rasgos cuantitativos de arroz para la tolerancia a
la sal codifica un transportador de sodio. *Nat. Gineta.*
37, 1141–1146 (2005).

80. Hu, B. et al. La variación en NRT1.1B contribuye a la divergencia en el
uso de nitratos entre las subespecies de arroz.
Nat. Gineta. **47**, 834–838 (2015).

81. Sol, H. et al. Las proteínas G heterotrímicas regulan la eficiencia del
uso de nitrógeno en el arroz. *Nat. Gineta.* **46**, 652–656 (2014).

82. Fan, X. et al. La sobreexpresión de un transportador de nitrato sensible al pH
en el arroz aumenta el rendimiento de los cultivos. *proc. Academia
Nacional. ciencia EE . UU .* **113**, 7118–7123 (2016).

83. Gamuyao, R. et al. La proteína quinas Pstol1 del arroz tradicional
confiere tolerancia a la deficiencia de fósforo. *Naturaleza* **488**, 535–
539 (2012).

84. Yamaji, N. et al. Reducción de la acumulación de fósforo en granos de arroz
con un transportador deteriorado en el nodo.
Naturaleza **541**, 92–95 (2017).

85. Xing, Y. & Zhang, Q. Bases genéticas y moleculares de
rendimiento de arroz. *año Rev. Plant Biol.* **61**, 421–442 (2010).

86. Wang, Y. & Li, J. Base molecular de la arquitectura vegetal.
año Rev. Plant Biol. **59**, 253–279 (2008).

87. Xue, W. et al. La variación natural en Ghd7 es una
regulador importante de la fecha de espiga y el potencial de
rendimiento en el arroz. *Nat. Gineta.* **40**, 761–767 (2008).

88. Wei, X. et al. DTH8 suprime la floración en el arroz, lo que influye
en la altura de la planta y el potencial de rendimiento
simultáneamente. *Fisiol. vegetal.* **153**, 1747–1758 (2010).

89. Yan, W.-H. et al. Un QTL importante. *Ghd8*, juega un papel pleiotrópico en la
regulación de la productividad del grano, la altura de la planta y la fecha de
espiga en el arroz. *mol. Planta* **4**, 319–330 (2011).

90. Yan, W. et al. La variación natural en Ghd7.1 juega un papel importante
en el rendimiento y la adaptación del grano en el arroz.
Resolución celular **23**, 969–971 (2013).

91. Zhang, Z.-H. et al. Pleiotropismo del fotoperíodo
alelo insensible de Hd1 en la fecha de espiga, la altura de la planta y las
características de rendimiento en arroz. *PLoS ONE* **7**, e25238 (2012).

92. Zhang, J. et al. Las combinaciones de los genes Ghd7, Ghd8 y Hd1 definen
en gran medida la adaptación eco geográfica y el potencial de rendimiento
del arroz cultivado. *Filol. nuevo*. **208**, 1056–1066 (2015).

93. Bai, X. et al. La duplicación de un silenciador aguas arriba de FZP aumenta
el rendimiento de grano en el arroz. *Nat. Plantas* **3**, 885–893 (2017).

94. Zhang, L. et al. Una matriz en tándem natural alivia
represión epigenética de IPA1 y conduce a un arroz de rendimiento
superior. *Nat. común* **8**, 14789 (2017).

95. Zhang, D. & Yuan, Z. Control molecular de la hierba
desarrollo de la inflorescencia. *año Rev. Plant Biol.* **65**, 553–578 (2014).

96. Chen, J. et al. Un sistema trialélico de S5 es un importante
regulador de la barrera reproductiva y compatibilidad de híbridos indica-
japónica en arroz. *proc. Academia Nacional. ciencia
EE . UU .* **105**, 11436–11441 (2008).

97. Largo, Y. et al. Esterilidad masculina híbrida en arroz controlada por interacción
entre alelos divergentes de dos genes adyacentes. *proc. Academia Nacional.
ciencia EE . UU .* **105**, 18871–18876 (2008).

98. Mizuta, Y., Harushima, Y. & Kurata, N. Polen de arroz
incompatibilidad híbrida causada por la pérdida recíproca de genes
duplicados. *proc. Academia Nacional. ciencia EE . UU .* **107**, 20417–
20422 (2010).

99. Yang, J. et al. Un sistema killer-protector regula tanto la esterilidad híbrida
como la distorsión de la segregación en el arroz.
Ciencia **337**, 1336–1340 (2012).

**Este estudio caracteriza el locus S5 para el aislamiento
reproductivo entre las subespecies indica y japónica, que consta de
tres genes adyacentes que forman un sistema asesino-protector. Este
gen se usa ampliamente en el mejoramiento de arroz híbrido
intersubespecífico.**

100. Yu, Y., Wing, RA y Li, J. en *Genética y genómica del arroz* (eds. Zhang,
Q. y Wing, RA) 237–254 (Springer-Verlag, 2013).

101. Gao, Z. et al. Clonación basada en mapas del gen ALK,
que controla la temperatura de gelatinización del arroz.
ciencia China C. Ciencias de la vida. **46**, 661–668 (2003).

102. Li, Y. et al. Chalk5 codifica una vacuolar H(+) -
translocación de pirofosfatasa que influye en la tiza del grano en el
arroz. *Nat. Gineta.* **46**, 398–404 (2014).

103. Ventllador, C. et al. GS3, un QTL principal para la longitud y el peso
del grano y un QTL menor para el ancho y el grosor del grano en
el arroz, codifica una supuesta proteína transmembrana. *teor. aplicación
Gineta.* **112**, 1164–1171 (2006).

104. Weng, J. et al. Aislamiento y caracterización inicial de GW5, un QTL principal
asociado con el ancho y el peso del grano de arroz. *Resolución celular*
18, 1199–1209 (2008).

105. Shomura, A. et al. La eliminación en un gen asociado con el tamaño del
grano aumentó los rendimientos durante la domesticación del arroz.
Nat. Gineta. **40**, 1023–1028 (2008).

106. Ye, X. et al. Ingeniería de la vía biosintética de la
provitamina A (betacaroteno) en endospermo de arroz
(sin carotenoides). *Ciencia* **287**, 303–305 (2000).

107. Zhu, Q. et al. Desarrollo de “Endospermo Púrpura
Rice” mediante la ingeniería de biosíntesis de antocianinas en el
endospermo con un sistema de apilamiento de transgenes de alta
eficiencia. *mol. Planta* **10**, 918–929 (2017).

108. Yu, H., Xie, W., Li, J., Zhou, F. y Zhang, Q.A.
matriz de SNP de genoma completo (RICE6K) para el
mejoramiento genómico en arroz. *Bioteconología vegetal.* **J. 12**, 28–37 (2014).

109. Chen, H. et al. Una matriz de genotipado SNP de alta densidad
para la biología del arroz y el mejoramiento molecular. *mol. Planta* **7**, 541–
553 (2014).

110. Singh, N. et al. Chip SNP de 50 K basado en un gen de copia única para
estudios genéticos y mejoramiento molecular en arroz.
ciencia Rep. **5**, 11600 (2015).

111. Ge, S., Sang, T., Lu, BR y Hong, DY Filogenia de los genomas del arroz
con énfasis en los orígenes de las especies alotetraploides. *proc.
Academia Nacional. ciencia EE . UU .* **96**, 14400–14405 (1999).

112. Alexander, DH, Novembre, J. & Lange, K. Estimación rápida basada
en modelos de ascendencia en individuos no relacionados. *Genoma
Res.* **19**, 1655–1664 (2009).

113. Purugganan, MD Una historia genómica evolutiva de dos especies de
arroz. *Nat. Gineta.* **46**, 931–932 (2014).

Agradecimientos

RAW recibió el apoyo de la Cátedra Bud Antle de Excelencia en Agricultura y Ciencias de la Vida, el Fondo de Investigación AXA y NIFA-HATCH AR27-1360510-H25-230. MDP fue apoyado por subvenciones del Programa de Investigación del Genoma Vegetal de la Fundación Nacional de Ciencias de EE. UU., la Fundación de la Familia Zegar y el Instituto de Investigación de Abu Dhabi de la Universidad de Nueva York. QZ recibió el apoyo de subvenciones del Programa Nacional 863 2104AA10A604, el Programa Nacional de Investigación y Desarrollo Clave 2016YFD0100903, el Fondo Asignado para el Sistema de Investigación Agrícola de China de China (CARS-01-05) y la Fundación Bill y Melinda Gates. Los autores también agradecen

K. McNally y S. Klassen por la lectura crítica del manuscrito antes de su publicación.

Contribuciones de autor

Todos los autores contribuyeron en todos los aspectos de la redacción de esta revisión.

Conflicto de intereses

Los autores declaran no tener conflictos de intereses.

nota del editor

Springer Nature se mantiene neutral con respecto a los reclamos jurisdiccionales en mapas publicados y afiliaciones institucionales.

Información del revisor

Nature Reviews Genetics agradece a Guo-Liang Wang, Jean Christophe Glaszmann y a los demás revisores anónimos por su contribución a la revisión por pares de este trabajo.

Información suplementaria

Hay información complementaria disponible para este artículo en <https://doi.org/10.1038/s41576-018-0024-z>.

Enlaces relacionados

Arroz de África: <http://www.africarice.org/warda/genebank.asp>

Grupo Consultivo e Investigación Agrícola Internacional (CGIAR): <https://www.cgiar.org/>

Conjunto de plantas: <http://plants.ensembl.org/index.html>

funRiceGenes: <http://funricegenes.ncpg.cn/>

GenBank: <https://www.ncbi.nlm.nih.gov/assembly/?term=Oryza>

Red Global de Fenotipado de Arroz: <http://ricephenonetwork.iri.org/>

Gramene: http://oge.gramene.org/genome_browser/index.html

Centro Internacional de Agricultura Tropical (CIAT): <http://ciat.cgiar.org/>

banco internacional de germoplasma de arroz: <http://iri.org/our-work/investigación/diversidad-genetica/banco-de-genes-de-arroz-internacional>

Consortios internacionales de informática del arroz (iRIC): <http://iric.iri.org/>

instituto internacional de investigación del arroz (IRRI): www.iri.org

Proyecto de anotación de arroz (RAP): <http://rapdb.dna.afrc.go.jp/>

DB de la Universidad Estatal de Michigan (MsU-DB): <http://rice.plantbiology.msu.edu/>

Orizabase: <https://shigen.nig.ac.jp/rice/orizabase/>

Portal de información sobre el arroz (RiGw): http://rice.hzau.edu.cn/cgi-bin/rice/download_ext

R498 en MBKBase: www.mbkbase.org/R498

Base de datos de Diversidad de Arroz: <http://www.ricediversity.org/data/index.cfm>

búsqueda de sNP: <http://snp-seek.iri.org/>

Scanalyzer TeRRa-ReF Field en Arizona: www.terraref.org