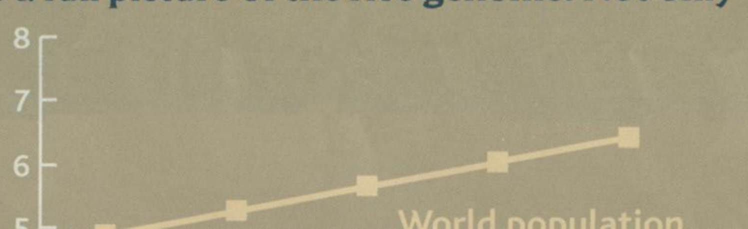
RICE

Rice has been a staple of the human diet for many thousands of years. Rice was originally domesticated in China. Excavations from the Hemudu Relics in Zhejiang Province, China, have found fossilized rice estimated at 7,000 years old (inset), roughly reflecting the time when agriculture developed. Approximately one-third of the world's population relies on rice for a significant portion of their food. About 11% of the world's arable land is used for growing rice, and about 29% of total cereal production in 2000 was due to rice.

Now, research teams from both academic and private-sector groups are assembling the knowledge necessary to have a full picture of the rice genome. Not only will this support understanding gene function in rice, but it will also open the door to elucidating crops such as corn and wheat that have larger and

zed rice grains



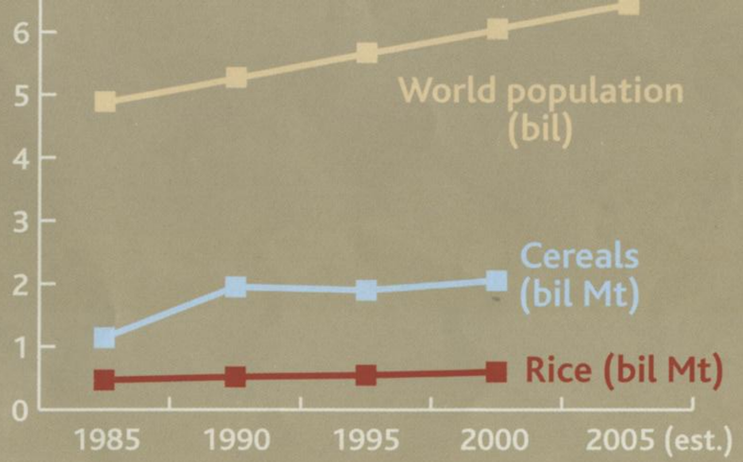
T

rs pro-
in the
ks, and
traits.
grains
to cer-
nother
saline

grown
he rice
left its
parallel
ated as
in the
nd in a
rticular
os that

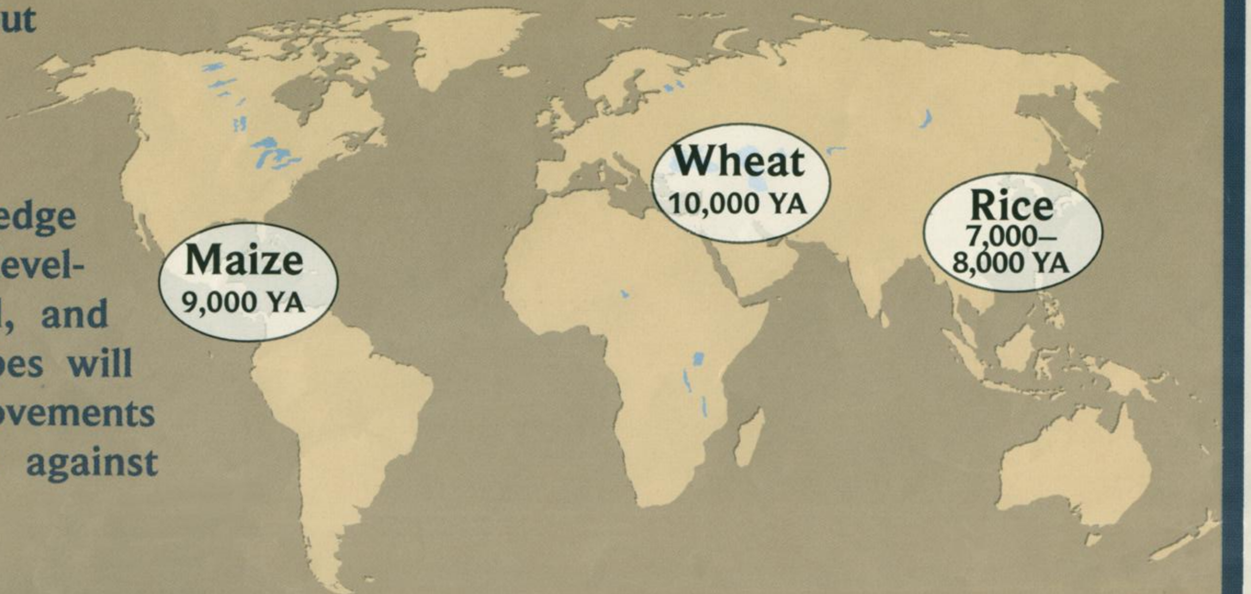
the door to elucidating crops such as corn and wheat that have larger and more complex genomes. The groups have used a variety of methods, including whole-genome and map-based sequencing and analysis of physical maps, to generate draft sequences of the entire rice genome. Rice, *Oryza sativa*, includes two main strains that likely diverged from one another as long as 2 to 3 million years ago (YA): *indica*, which is generally adapted to growth in the tropics, and *japonica*, which is generally adapted to growth in temperate regions. Both strains are represented in the rough drafts of the rice genome that are published in this issue, *Science* **296**, 79–100 (2002).

With cereal production roughly holding steady in recent years, continued world population growth will put added pressure on food production mechanisms that will need to be met. Genetic knowledge being uncovered about developmental, physiological, and morphological phenotypes will contribute to crop improvements important in the fight against world hunger.



Recent and future estimates of world population growth contrast with agricultural productivity of total cereals and rice. [Statistics of FAOSTAT {http://apps.fao.org/}, the online database from the Food and Agriculture Organization of the United Nations]

Cereal Domestication History



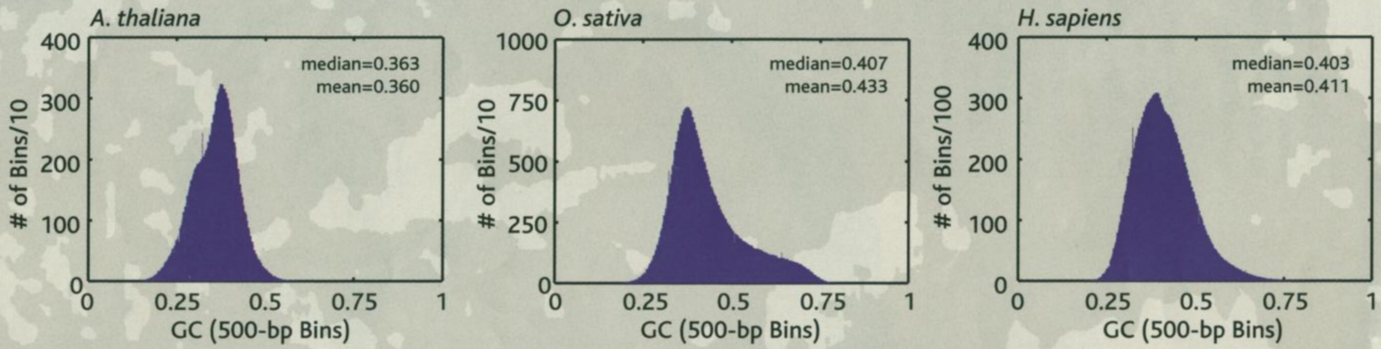
Genomic Information

1
3

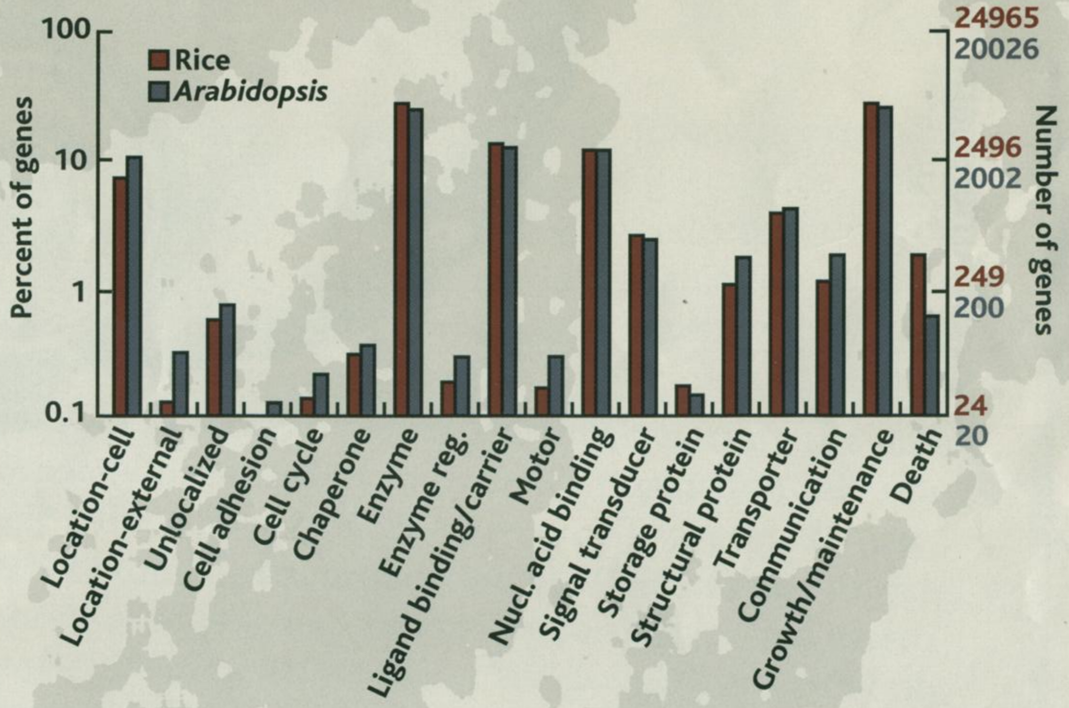
Genomic Information

The evolutionary lineages of rice and *Arabidopsis* likely diverged roughly 200 million years ago. Although their genomes do not show extensive synteny, there are similarities among many of the proteins encoded. Rice genes diverge from *Arabidopsis* genes in GC content, codon usage, and amino acid usage, in ways that complicate efforts at gene annotation. There is a gradient in the GC content of rice genes that does not appear in *Arabidopsis* genes, with the 5'- end being up to 25% richer in GC than the 3'- end (*1*). For the 20% and 36% of predicted rice and *Arabidopsis* genes that can be functionally classified, the distributions are highly similar.

GC content for *Arabidopsis*, rice, and human genomes (*1*).



Functional classification of rice genes, according to the Gene Ontology Consortium, and assigned by homology to categorized *Arabidopsis* genes (*1*).



RICE GENOME STATISTICS

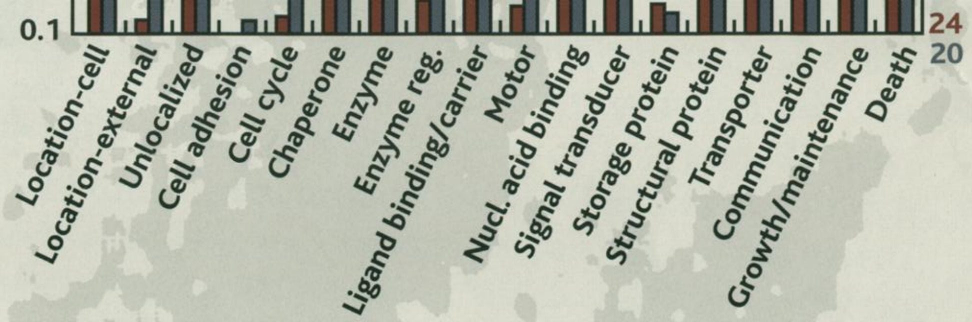
	Yu *et al.* (*1*)*	Goff *et al.* (*2*)†
Rice strain sequenced	*indica*	*japonica*
Genome size (Mb)	466	420
Number of genes	46–56K	32-50K
Duplicated genes (%)	74	77
Number of transposable elements	>24.9%‡	4,220
Number of simple sequence repeats	1.7%‡	46,666
Draft quality assessment:		

RICE GENOME WEB SITES

Beijing Genomics Institute http://btn.genomics.org.cn/rice

Clemson University Genomics Institute http://www.genome.clemson.edu/

Cold Spring Harbor Laboratory, U.S. http://www.cshl.org/



and
nics
ute,

jing
ute,
nap,
02);
MO,
ize,
nce
esa
n.

RICE GENOME STATISTICS

	Yu *et al.* (1)*	Goff *et al.* (2)†
Rice strain sequenced	*indica*	*japonica*
Genome size (Mb)	466	420
Number of genes	46–56K	32-50K
Duplicated genes (%)	74	77
Number of transposable elements	>24.9%‡	4,220
Number of simple sequence repeats	1.7%‡	46,666
Draft quality assessment:		
Shotgun coverage	4.2x	>6x
Number of contigs	127,550	42,109

Genome analysis produces a huge variety of additional statistics, including for example:

Yu *et al.* estimate that the average gene size is 4.5 kb, and that 49.4% of predicted rice genes have homologs in *Arabidopsis*.

Goff *et al.* estimate that there are 513 ribosomal RNA gene repeats and 688 centromeric sequence repeats.

Citations:
(1) J. Yu *et al.*, *Science* 296, 79 (2002).
(2) S. A. Goff *et al.*, *Science* 296, 92 (2002).

Notes:
* Yu *et al.* estimate that 90.8% of the sequence has a Phrap-estimated error rate of better than 1 in 1000.
† Goff *et al.* estimate error rates as 1 bp difference between Syd and the "public" data per kb, and 1 InDel difference per 2 kb.
‡ Percent of genome.

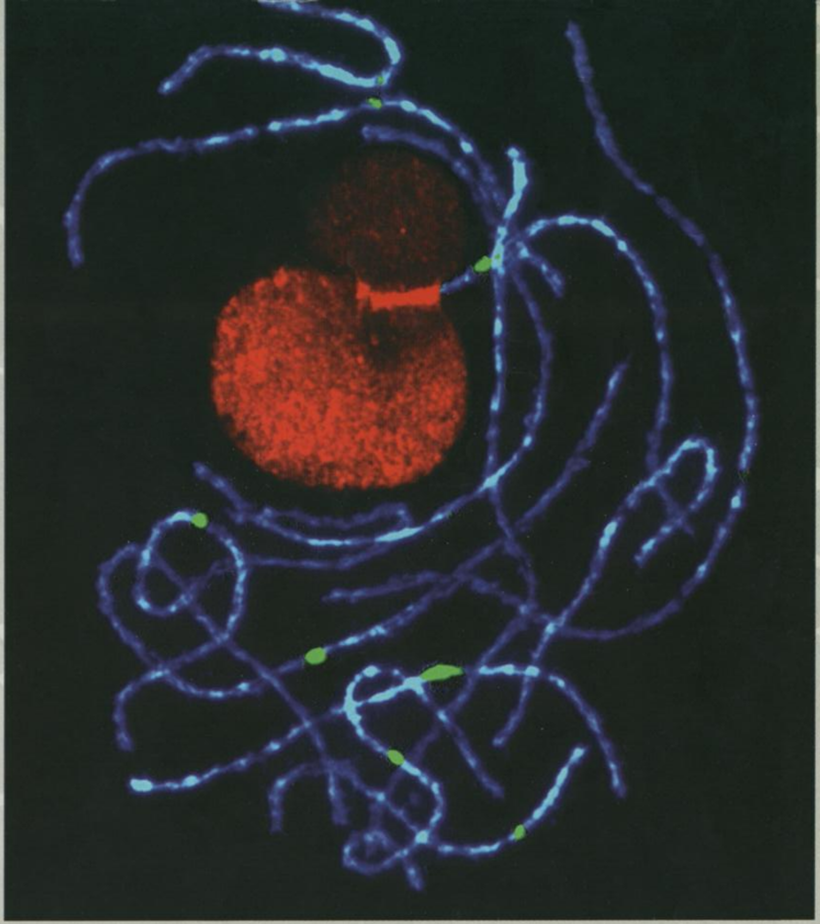
RICE GENOME WEB SITES

Beijing Genomics Institute	http://btn.genomics.org.cn/rice
Clemson University Genomics Institute	http://www.genome.clemson.edu/
Cold Spring Harbor Laboratory, U.S.	http://www.cshl.org/
Genoscope, France	http://www.genoscope.cns.fr/
Gramene, a comparative mapping resource for grains	http://www.gramene.org/
International Rice Research Institute, Phillipines	http://www.irri.org/
IRD Rice Genetics Group, France	http://www.mpl.ird.fr/rice/
Korea Rice Genome Program	http://biogen.niast.go.kr/english/egframe.htm
Monsanto Rice Genome Project	http://www.rice-research.org/
National Center for Gene Research, Chinese Academy of Sciences	http://www.ncgr.ac.cn/index.html
National Center for Genetic Engineering and Biotechnology, Thailand	http://www.cs.ait.ac.th
Oryzabase, Japan	http://www.shigen.nig.ac.jp/rice/oryzabase/
Plant Genome Initiative, Rutgers, U.S.	http://pgir.rutgers.edu/News.html
Rice Genome Project in Taiwan, China	http://genome.sinica.edu.tw/index_n.html
Rice Genome Program, Japan	http://rgp.dna.affrc.go.jp/
Syngenta/TMRI	http://www.tmri.org
TIGR Rice Gene Index	http://www.tigr.org/tdb/ogi/
U.S. Rice Genome Site	http://www.usricegenome.org/
Washington University School of Medicine	http://genome.wustl.edu/gsc/
Wisconsin Rice Genome Project	http://www.gcow.wisc.edu/Rice/index.htm

(YEAR 2050): ONE PERSON WILL HAVE LESS THAN 1/3 ACRE

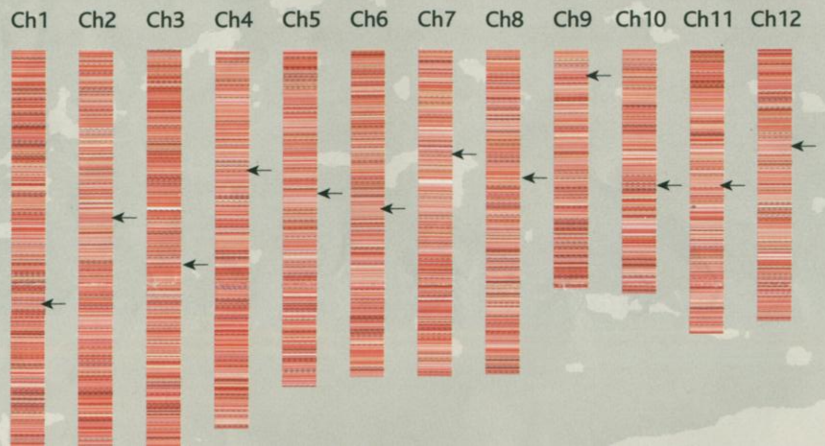
The Rice Chromosomes

Rice chromosomes taken from a meiotic pachytene cell of the *japonica* rice variety show the major features of the genome. The 12 chromosomes, ranging in size from roughly 50 to 25 Mb, are stained by 4',6-diamidino-2-phenylindole (DAPI) in blue. The nucleolus and the nucleolar organizing region, containing the ribosomal RNA gene repeats in a cluster on chromosome 9, are labeled in red. Centromeres are in green, labeled by a probe specific to the centromeric satellite repeat RCS2.

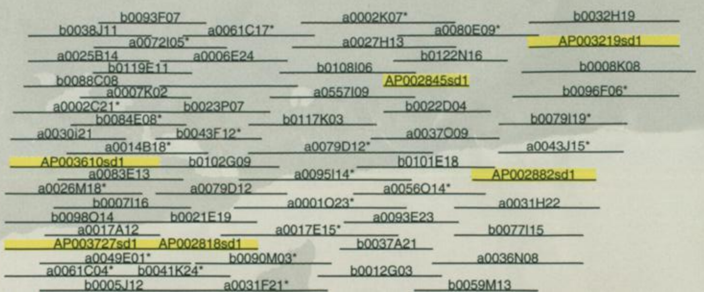


A diagrammatic representation (below) shows the distribution of genes across the rice genome. Although genes are generally evenly distributed, there are particular regions that show unusually dense or sparse collections of genes. The gene density is coded in color — darker red color indicates higher gene densities. Arrows indicate centromeres.

RICE GENE DENSITY ALONG INDIVIDUAL CHROMOSOMES



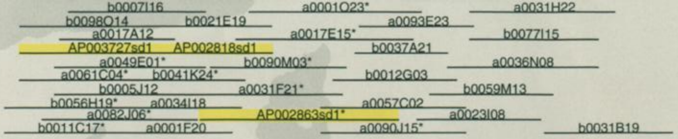
Shown below is a portion of the integrated physical map from one segment of chromosome 1. The figure represents integration of data from multiple private-sector and academic sources. The simulated digest clones in yellow are in silico fingerprints of sequenced P1-derived artificial chromosome (PAC) clones integrated into the bacterial artificial chromosome (BAC) physical map. BAC markers derived from the Monsanto draft sequence allowed this contig to be anchored to chromosome 1.




Fossilized rice grains

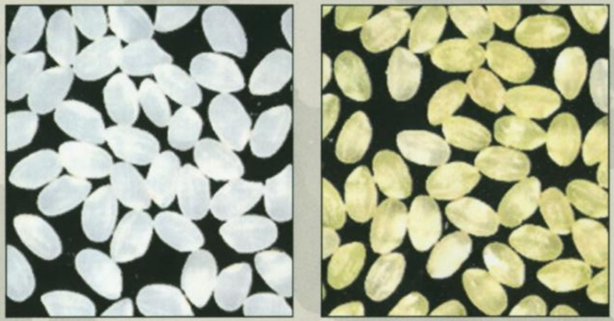
PREDICTED FUTURE LAND USE (YE

IRES APPROXIMATELY 1/2 ACRE



Applications

EVALUATION OF GENE FUNCTION



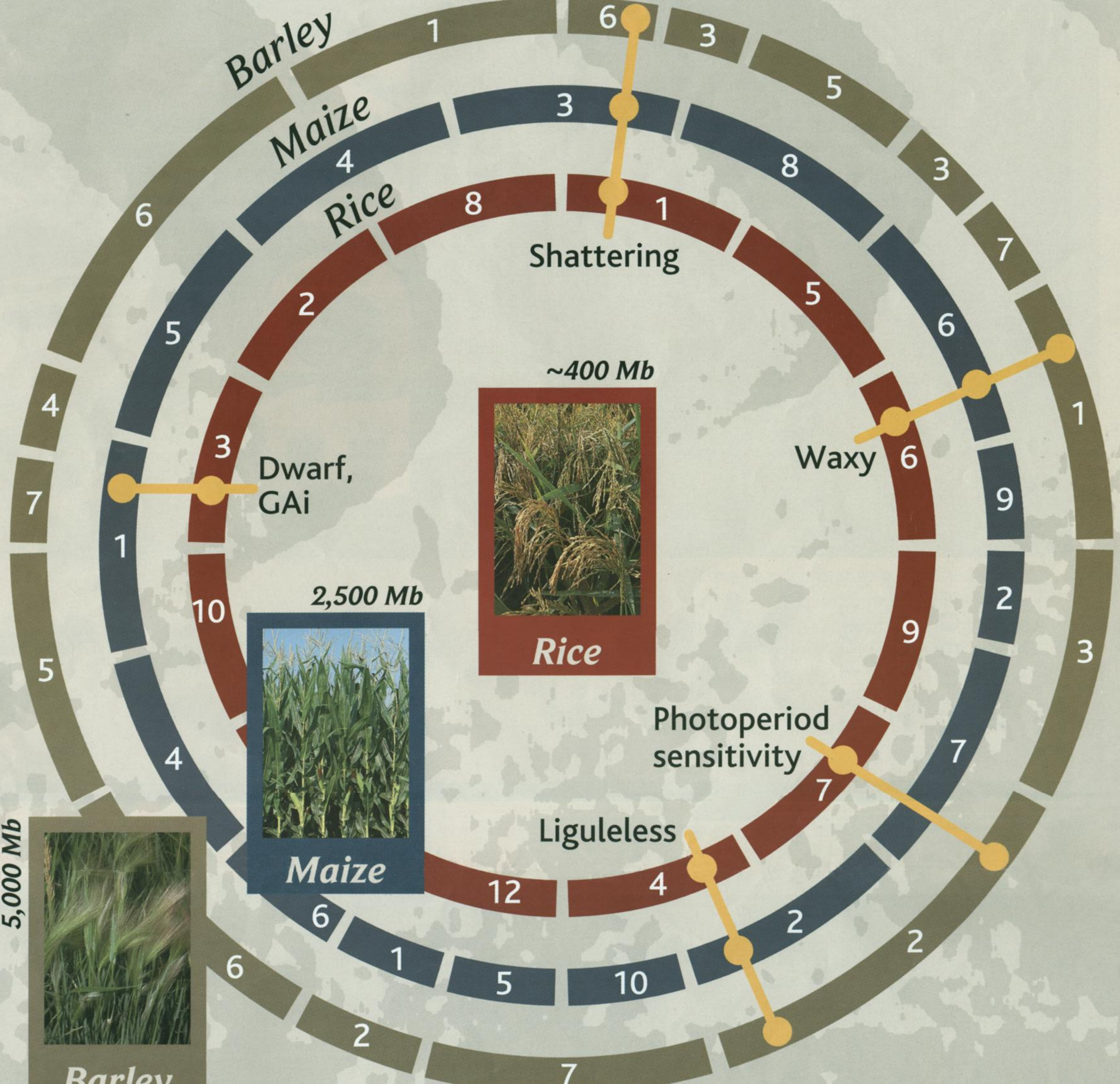
NUTRITIONAL QUALITIES

DISEASE MANAGEMENT

Crop Improvement. Through selective breeding, humans have for many thousands of years promoted development of desirable traits and discouraged development of undesirable traits in the crops we rely upon. More explicit knowledge of gene function, the genetic control networks, and the cellular physiology that cause these traits allows for more refined management of the traits. For example, nutritional qualities of the grain can be improved, as shown here with rice grains engineered to express precursors to vitamin A. Improving the plant's intrinsic resistance to certain diseases can lessen the ecological impact of pesticide use and limited crop yields. Another example lies in the need for continued food production in adverse conditions — plants more tolerant to drought or saline conditions can extend crop production capabilities.

Genome Colinearity Among the Grasses. The rice genome is the smallest of the genomes from the grasses most commonly grown as crops. However, the overall organization of genes is preserved, such that, at a general level, information derived from the rice genome constructs a useful map for guidance through the larger genomes of maize and barley. Evolution has most certainly left its mark — inversions, translocations, transposable elements, and a slew of mutations all serve to add variation to otherwise parallel genome structures. Nonetheless, enough genomic similarity exists that the genomes of these grain species can be approximated as concentric circles, and the information from the smaller genome of rice provides insight into where to find genes of interest in the larger genomes of the other grasses. Each circle represents a single grain species, with its chromosomes collected end-to-end in a parsimonious manner to best match the structure of the rice genome. A few examples of genetic loci associated with particular traits are shown to illustrate that similar genes occur in similar portions of the genomes across these grasses. Other grain crops that show colinearity with rice include wheat, oats, pearl millet, sorghum, sugar cane, and foxtail millet.

D APPROXIMATELY 6,000 ACRES CURRENT: ONE PERSON REQUIRES



Rice strain sequ

Genome size (N

Number of gen

Duplicated gen

Number of tran

Number of simp

Draft quality as

PRE-AGRICULTURE: ONE PERSON REQUIRED APPROXIMATELY 6,

Science Coordinators: Pamela J. Hines and Barbara R. Jasny
Contributors: Gane Ka-Shu Wong and Jun Yu, Beijing Genomics Institute, Beijing 101300, China, Hangzhou Genomics Institute, Hangzhou 310007, China, and University of Washington Genome Center, Seattle, WA 98195, USA; Huanming Yang, Beijing Genomics Institute, Beijing 101300, China, Hangzhou Genomics Institute, Hangzhou 310007, China, Institute of Genetics, Chinese Academy of Sciences, Beijing 100101, China; Stephen A. Goff, Torrey Mesa Research Institute, Syngenta, San Diego, CA, 92121, USA.
Illustrations: Background art: Terry E. Smith; based on a photograph from the Honghe Hani Terraces in Yunnan Province (China), Liwen Ma and Baoxing Qiu, Beijing Genomics Institute, Beijing 101300, China; Fossilized rice from the Hemudu Relics in Zhejiang Province (China), Sheng Zhao, Hangzhou Genomics Institute, Hangzhou 310007, China; Fluorescent rice chromosomes, Zhukuan Cheng and Jiming Jiang, University of Wisconsin, Madison, WI, USA; Integrated physical map, Rod A. Wing, Carol Soderlund, Mingsheng Chen, Clemson University Genomics Institute, Clemson, SC 29634, USA; M. Chen *et al.*, *Plant Cell* **14**, 537 (2002); Evaluation of gene function in rice plants, Santanu Dasgupta, Monsanto Research Centre, Bangalore, Karnataka, India, and Gerard Barry, Monsanto, St. Louis, MO, USA. "Golden" rice grains, X. Ye, *et al., Science* **287**, 303 (2000); Diseased plants, W.Y. Song *et al., Science* **270**, 1804 (1995); Concentric circles showing rice, maize, barley genomes, adapted by Tien-Hung Lan, Torrey Mesa Research Institute, Syngenta, San Diego, CA, 92121, USA, on the basis of M. D. Gale and K. M. Devos, *Science* **282**, 656 (1998) and in silico mapping data (ref. 2); Rice gene density, cereal domestication history, Tien-Hung Lan, Molly Dunn, and Stephen A. Goff, Torrey Mesa Research Institute, Syngenta, San Diego, CA, 92121, USA; Barley crop, Gladys Lucille Smith/California Academy of Sciences; Corn and Rice crops, Grant Heilman.
Reviewer: Chris Somerville, Carnegie Institution, Stanford CA 94305, USA.
Design and Art Direction: C. Faber Smith; **Production Assistance:** Debra Morgenegg, Alan Stonebraker; **Copyediting:** Harry Jach.

Rice strain sequ

Genome size (N

Number of gen

Duplicated gen

Number of tran

Number of simp

Draft quality as

Shotgun cove

Number of co

Genome analysis
including for exa

Yu *et al.* estimat
of predicted rice

Goff *et al.* estim
and 688 centrom

Citations:
(1) J. Yu *et al.*, Sc
(2) S. A. Goff *et a*

Notes:
* Yu *et al.* estima
error rate of bett
† Goff *et al.* esti
the "public" data
‡ Percent of gen