

PublisherInfo		
PublisherName	:	BioMed Central
PublisherLocation	:	London
PublisherImprintName	:	BioMed Central

Rice genome revealed I: the *indica* sequence

ArticleInfo		
ArticleID	:	4349
ArticleDOI	:	10.1186/gb-2002-3-7-reports0035
ArticleCitationID	:	reports0035
ArticleSequenceNumber	:	23
ArticleCategory	:	Paper report
ArticleFirstPage	:	1
ArticleLastPage	:	4
ArticleHistory	:	RegistrationDate : 2002-4-26 Received : 2002-4-26 OnlineDate : 2002-6-28
ArticleCopyright	:	BioMed Central Ltd2002
ArticleGrants	:	

Jean-Nicolas Volff

Abstract

Together with the *japonicarice* genome sequence, this publicly available draft sequence of the *indica* rice genome should boost the genetic improvement of rice and other cereal crops

Significance and context

Rice (*Oryza sativa*) is the staple crop for more than half of the world's population. As well as its immense economic value, rice's intrinsic genomic characteristics predestined it to become, after the thale cress *Arabidopsis thaliana*, one of the first plants with a completely sequenced genome. Sequencing the 'small' gene-rich part of the rice genome (400-450 million base pairs, Mbp) should boost studies on other cereals, including wheat (3,000 Mbp) and maize (16,000 Mbp), both of which have much larger genomes but similar gene content and order. The importance of the rice genome is underlined by the existence of several parallel, sometimes collaborating, sequencing projects being carried out by different academic institutions and biotechnology companies. Two *O. sativa* subspecies - *japonica* and *indica* - are the main subspecies grown agriculturally. Yu *et al.* report the draft sequence of the genome of *O. sativa*ssp. *indica*, the major rice subspecies grown in China and other parts of the Asia-Pacific region. The sequencing team is a collaboration between scientists from the Beijing Genomics Institute and other Chinese academic institutions. In the same issue of *Science*, the sequence of a variety of the *japonica* subspecies is reported by a team from the Swiss agricultural company Syngenta and the US biotech company Myriad Genetics (see related report - [Genome Biology 3\(7\):reports0036](#)).

Key results

Using the whole-genome shotgun strategy, a method based mainly on random sequencing that has already been successfully applied to the fly and human genomes, Yu *et al.* estimated the size of the rice genome at 466 Mbp and succeeded in determining 362 Mbp of genomic sequence disrupted by about 100,000 gaps. The predicted number of genes (46,022-55,615) indicates that rice has more genes than *Arabidopsis* (estimated at 25,498 genes). This can be explained by additional duplication events in the rice lineage after its separation from the *Arabidopsis* lineage about 150-200 million years ago. The extreme divergence of some of the resulting duplicates might explain why 50% of the rice genes predicted by Yu *et al.* have no obvious homologs in *Arabidopsis* and other organisms. The mean gene size in rice was 4.5 kb, as compared to 2.4 kb in *Arabidopsis* and 72 kb in humans. In contrast to the

situation in the human genome, transposon sequences in rice were mainly found in intergenic regions rather than in introns. Yu *et al.* observed a gradient in both GC content and codon usage in protein-coding rice sequences, with the 5' end being typically up to 25% richer in GC than the 3' end.

Links

Some commentaries on the sequencing of the rice genome are available free of charge from [The rice genome](#) index page at the *Science* website. The draft sequence of the genome of *O. sativa*ssp. *indica* reported by Yu *et al.* can be downloaded freely from [Rice GD](#) and is also available through the National Center for Biotechnology Information (NCBI) [Oryza sativa\(rice\) genome view](#), along with other rice sequences in the public domain. The Myriad Genetics/Syngenta draft sequence of *O. sativa*ssp. *japonica* Nipponbare is available to academic researchers under certain conditions from the [Torrey Mesa Research Institute - the rice genome](#) page. Information about the Japanese-led public International Rice Genome Sequencing Project (IRGSP) with links to the different institutions involved and to sequences deposited in public databases can be found at the [Academia Sinica Plant Genome Center](#).

Reporter's comments

There is no doubt that the draft sequences of the rice genome by Yu *et al.* and by Goff *et al.* (see related report - [Genome Biology 3\(7\):reports0036](#)) are milestones in rice research that should help us to understand the physiology, developmental biology, genetics and evolution of cereals and other plants and to boost the improvement of species important for agriculture. The draft sequences, of course, still contain numerous gaps. The generation of a highly accurate, mainly gap-free genome sequence should be accelerated through collaboration with the IRGSP, which is sequencing mapped clones, and to which the agricultural company Monsanto has donated its rice sequence data. In the future, functional genomics will be needed to determine the exact number of genes and their function(s) in rice, particularly those genes that are apparently specific to plants, and genes present in rice but not in *Arabidopsis*. Comparative genomics, particularly between different cereal species, should allow the molecular characterization of loci already identified by classical genetics as potential targets for improvement of agricultural varieties.

Table of links

[Science](#)

[Genome Biology 3\(7\):reports0036](#)

[The rice genome](#)

Rice GD

Oryza sativa(rice) genome view

Torrey Mesa Research Institute - the rice genome

Academia Sinica Plant Genome Center

References

1. Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, et al: A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science. 2002, 296: 79-92. 0036-8075