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

Plant sequence completed

ArticleInfo		
ArticleID	:	3863
ArticleDOI	:	10.1186/gb-spotlight-20001214-02
ArticleCitationID	:	spotlight-20001214-02
ArticleSequenceNumber	:	300
ArticleCategory	:	Research news
ArticleFirstPage	:	1
ArticleLastPage	:	3
ArticleHistory	:	RegistrationDate : 2000–12–14 OnlineDate : 2000–12–14
ArticleCopyright	:	BioMed Central Ltd2000
ArticleGrants	:	
ArticleContext	:	130591111

William Wells
Email: wells@biotext.com

In the 14 December Nature the Arabidopsis Genome Initiative (AGI) announces the completion of the first plant genome sequence. The published sequence covers 115.4 megabases of the estimated 125-megabase genome of the thale cress *Arabidopsis thaliana*, a model plant. The sequence contains 25,498 genes encoding proteins from 11,000 families. The gene number is high for two reasons: an ancient diploidization, and the presence of 1,528 tandem arrays (containing 4,140 individual genes), which probably arose after unequal crossing-over events during recombination.

After publication of the sequence of *Arabidopsis* chromosome 2 and chromosome 4 last year, the new reports cover the remaining sequences of chromosomes 1 (Theologis *et al*., *Nature* 2000, **408:**816-820), 3 (Salanoubat *et al*., *Nature* 2000, **408:**820-822) and 5 (Tabata *et al*., *Nature* 2000, **408:**823-826). There is also a lengthy analysis of the whole sequence by the entire AGI team (*Nature* 2000, **408:**796-815).

The analysis includes descriptions of sequence polymorphisms, membrane transport systems, chromatin components, and genes encoded within centromeres. Specific findings include:

- A large number of peptide transporters.- A total number of gene regulators similar to that found in other organisms, but many transcription factors (16 of the 29 classes found in *Arabidopsis*) appear to be unique to plants.- No genes encoding intermediate filament proteins or structures linking actin to extracellular proteins, but 420 proteins are presumed to function in the synthesis or modification of the cell wall.- An absence of genes for G-proteins of the Ras, Rho, Rac and Cdc42 families, and of genes encoding proteins of several other major signal-transduction pathways (including Wingless/Wnt, Hedgehog, Notch, and JAK/STAT).- Different families of receptor kinases have been expanded to cover developmental processes similar to those seen in other organisms. Receptor tyrosine kinases are lacking in *Arabidopsis*, although the plant has at least 340 genes for receptor Ser/Thr kinases.- A substantial dose of genes similar to cyanobacterial genes, presumably acquired via the endosymbiont that became the plastid. Many of these genes encode proteins involved in basic metabolic processes.

References

1. *Nature*, [http://www.nature.com/nature/]

2. *Arabidopsis thaliana*: a model plant for genome analysis.

3. Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

4. Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*.

5. Cereon *Arabidopsis* polymorphism collection, [http://www.arabidopsis.org/cereon/]

6. Genomic comparisons of membrane transport systems, [http://www-biology.ucsd.edu/~ipaulsen/transport/]

7. The plant chromatin database, [http://Ag.Arizona.Edu/chromatin/chromatin.html]

8. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*: rDNA, telomeres and centromeres , [http://preuss.bsd.uchicago.edu/arabidopsis.genome.html]