

54. M. Steiniger-White and W. S. Reznikoff, *J. Biol. Chem.*, in press.
55. T. W. Wiegand and W. S. Reznikoff, *J. Bacteriol.* **174**, 1229 (1992).
56. D. York and W. S. Reznikoff, *Nucleic Acids Res.* **24**, 3790 (1996).
57. R. R. Isberg and M. Syvanen, *J. Biol. Chem.* **260**, 3645 (1985).
58. Z. Otwinowski and W. Minor, *Methods Enzymol.* **276**, 307 (1997).
59. A. T. Brunger *et al.*, *Acta Crystallogr. D* **54**, 905 (1998).
60. A. Roussel and C. Cambilau, in *Silicon Graphics Geometry Partners Directory* (Silicon Graphics, Mountain View, CA, 1991), vol. 86.
61. N. S. Pannu and R. J. Read, *Acta Crystallogr. A* **52**, 659 (1996).
62. P. D. Adams, N. S. Pannu, R. J. Read, A. T. Brünger, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5018 (1997).
63. M. Carson, *Methods Enzymol.* **277**, 493 (1997).
64. R. J. Read, *Acta Crystallogr. A* **42**, 140 (1986).
65. G. H. Cohen, *J. Mol. Biol.* **190**, 593 (1986).
66. ———, *J. Appl. Crystallogr.* **30**, 1160 (1997).
67. We thank L. Mahnke for providing the first samples of purified transposase, J. B. Thoden for help in collecting the x-ray data, and T. Naumann, A. Bhasin, M. Steiniger-White and R. Saecker for helpful discussions. We gratefully acknowledge the help of N. Duke, F. Rotella, and A. Joachimak at the Structural Biology Center Beamline, Argonne National Laboratory, in collecting the data. This research was supported in part by NIH grants AR35186 (I.R.) and GM50692 (W.S.R.). W.S.R. is a recipient of a Vilas Associates Award and is the Evelyn Mercer Professor of Biochemistry and Molecular Biology. D.R.D. was supported by the NIH Biotechnology Training Grant. The use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Energy Research, Contract W-31-109-Eng-38. The PDB accession number for the coordinates and structure factors is 1F3I for the transposase/DNA complex.

26 April 2000; accepted 5 June 2000

## fw2.2: A Quantitative Trait Locus Key to the Evolution of Tomato Fruit Size

Anne Frary,<sup>1\*</sup> T. Clint Nesbitt,<sup>1\*</sup> Amy Frary,<sup>1†</sup>  
Silvana Grandillo,<sup>1‡</sup> Esther van der Knaap,<sup>1</sup> Bin Cong,<sup>1</sup>  
Jiping Liu,<sup>1</sup> Jaroslaw Meller,<sup>2</sup> Ron Elber,<sup>2</sup> Kevin B. Alpert,<sup>1</sup>  
Steven D. Tanksley<sup>1§</sup>

Domestication of many plants has correlated with dramatic increases in fruit size. In tomato, one quantitative trait locus (QTL), *fw2.2*, was responsible for a large step in this process. When transformed into large-fruited cultivars, a cosmid derived from the *fw2.2* region of a small-fruited wild species reduced fruit size by the predicted amount and had the gene action expected for *fw2.2*. The cause of the QTL effect is a single gene, *ORFX*, that is expressed early in floral development, controls carpel cell number, and has a sequence suggesting structural similarity to the human oncogene *c-H-ras* p21. Alterations in fruit size, imparted by *fw2.2* alleles, are most likely due to changes in regulation rather than in the sequence and structure of the encoded protein.

cDNA library (constructed from the small-fruited genotype, *L. pennellii* LA716). About 100 positive cDNA clones were identified that represent four unique transcripts (cDNA27, cDNA38, cDNA44, and cDNA70) that were derived from genes in the *fw2.2* YAC contig. A high-resolution map was created of the four transcripts on 3472 F<sub>2</sub> individuals derived from a cross between two nearly isogenic lines (NILs) differing for alleles at *fw2.2* (Fig. 2A) (8). The four cDNAs were then used to screen a cosmid library of *L. pennellii* genomic DNA (9). Four positive, nonoverlapping cosmids (cos50, cos62, cos69, and cos84) were identified, one corresponding to each unique transcript. These four cosmid clones were assembled into a physical contig of the *fw2.2* region (10) (Fig. 2B) and were used for genetic complementation analysis in transgenic plants.

The constructs (11) were transformed into two tomato cultivars, Mogeor (fresh market-type) and TA496 (processing-type) (12). Both tomato lines carry the partially recessive large-fruit allele of *fw2.2*. Because *fw2.2* is a QTL and the *L. pennellii* allele is only partially dominant, the primary transformants (R0), which are hemizygous for the transgene, were self-pollinated to obtain segregating R1 progeny. In plants containing the transgene (13), a statistically significant reduction in fruit weight indicated that the plants were carrying the small-fruit allele of *fw2.2* and that complementation had been achieved. This result was only observed in the R1 progeny of primary transformants fw71 and fw107, both of which carried cos50 (Fig. 1B and Table 1) (14). That the two complementing transformation events are independent and in different tomato lines (TA496 and Mogeor) indicates that the cos50 transgene functions similarly in different genetic backgrounds and genomic locations. Thus, the progeny of plants fw71 and fw107 show that *fw2.2* is contained within cos50.

Most QTL alleles are not fully dominant or recessive (5). The small-fruit *L. pennellii* allele for *fw2.2* is semidominant to the large-fruit *L. esculentum* allele (7). R2 progeny of fw71 were used to calculate the gene action [*d/a* = dominance deviation/additivity; calculated as described in (5)] of cos50 in the transgenic plants.

In natural populations, most phenotypic variation is continuous and is effected by alleles at multiple loci. Although this quantitative variation fuels evolutionary change and has been exploited in the domestication and genetic improvement of plants and animals, the identification and isolation of the genes underlying this variation have been difficult.

Conspicuous and important quantitative traits in plant agriculture are associated with domestication (1). Dramatic, relatively rapid evolution of fruit size has accompanied the domestication of virtually all fruit-bearing crop species (2). For example, the progenitor of the domesticated tomato (*Lycopersicon esculen-*

*tum*) most likely had fruit less than 1 cm in diameter and only a few grams in weight (3). Such fruit was large enough to contain hundreds of seeds and yet small enough to be dispersed by small rodents or birds. In contrast, modern tomatoes can weigh as much as 1000 grams and can exceed 15 cm in diameter (Fig. 1A). Tomato fruit size is quantitatively controlled [for example, (4)]; however, the molecular basis of this transition has been unknown.

Most of the loci involved in the evolution and domestication of tomato from small berries to large fruit have been genetically mapped (5, 6). One of these QTLs, *fw2.2*, changes fruit weight by up to 30% and appears to have been responsible for a key transition during domestication: All wild *Lycopersicon* species examined thus far contain small-fruit alleles at this locus, whereas modern cultivars have large-fruit alleles (7). By applying a map-based approach, we have cloned and sequenced a 19-kb segment of DNA containing this QTL and have identified the gene responsible for the QTL effect.

**Genetic complementation with *fw2.2*.** A yeast artificial chromosome (YAC) containing *fw2.2* was isolated (8) and used to screen a

<sup>1</sup>Department of Plant Breeding and Department of Plant Biology, 252 Emerson Hall, Cornell University, Ithaca, NY 14853, USA. <sup>2</sup>Department of Computer Science, Cornell University, Ithaca, NY 14853, USA.

\*These authors contributed equally to this work.

†Present address: Department of Biological Sciences, Clapp Laboratory, Mount Holyoke College, South Hadley, MA 01075, USA.

‡Present address: Research Institute for Vegetable and Ornamental Plant Breeding, IMOF-CNR, Via Università 133, 80055 Portici, Italy.

§To whom correspondence should be addressed.

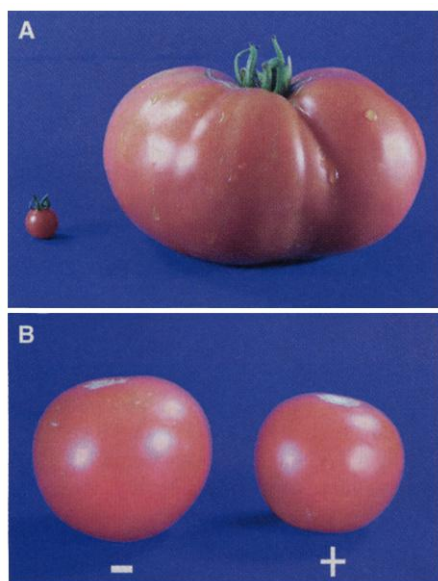
The transgene had a *d/a* of 0.51; in previous work with nearly isogenic lines (NILs), *fw2.2* had a *d/a* of 0.44. This similarity of gene action is consistent with the conclusion that the *cos50* transgene carries *fw2.2*.

***fw2.2* corresponds to *ORFX* and is expressed in pre-anthesis floral organs.** Sequence analysis of *cos50* (15) revealed two open reading frames (ORFs) (Fig. 2C): one corresponding to cDNA44, which was used to isolate *cos50*, and another 663-nucleotide (nt) gene, *ORFX*, for which no corresponding transcript was detected in the initial cDNA library screen. The insert also contains a highly repetitive, AT-rich (80%) region of 1.4 kb (Fig. 2C). Previous mapping of *fw2.2* had identified a single recombination event that delimited the "rightmost" end of the *fw2.2* candidate region

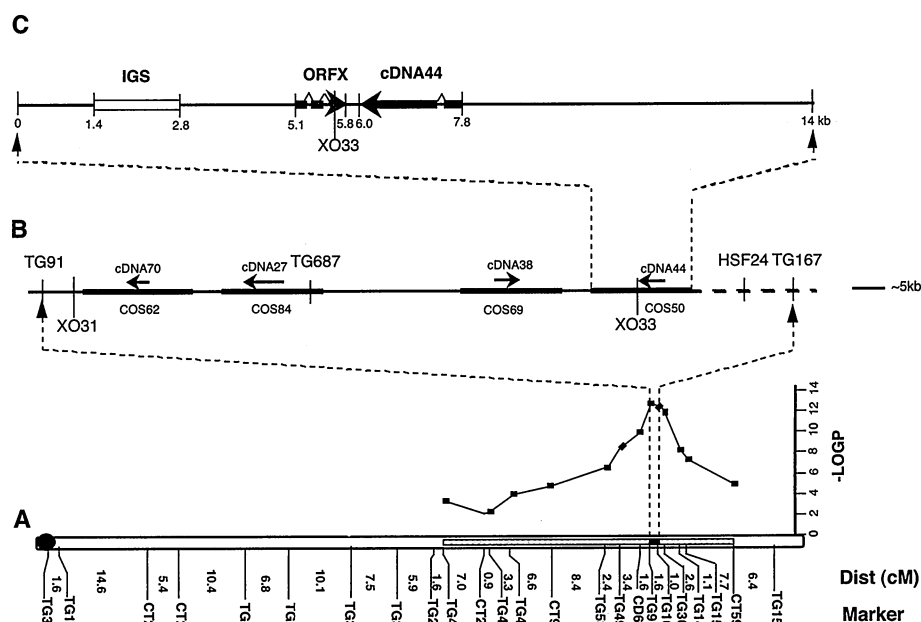
[XO33 in (8)]. Comparison of genomic DNA sequence from this recombinant plant with that of the two parental lines indicated that XO33 is within 43 to 80 nt 5' from the end of *ORFX* (Fig. 2C). Because genetic mutation(s) causing change in fruit size must be to the left of XO33, cDNA44 cannot be involved, and *ORFX* or an upstream region is the likely cause of the *fw2.2* QTL phenotype.

*ORFX* is transcribed at levels too low to be detected through standard Northern hybridization protocols in all pre-anthesis floral organs (petal, carpels, sepals, and stamen) of both large- and small-fruited NILs; however, semi-quantitative reverse transcriptase-polymerase chain reaction (RT-PCR) analysis indicated that the highest levels were expressed in carpels (16)

(Fig. 3A). In addition, comparison of the relative levels of *ORFX* transcript in the carpels of the NILs showed significantly higher levels in the small-fruited NIL (TA1144) than in the large-fruited NIL (TA1143) (TA1143/TA1144 carpel transcript levels, mean ratio = 0.51; for the null hypothesis mean = 1,  $P = 0.02$ ). The observation of *ORFX* transcription in pre-anthesis carpels suggests that *fw2.2* exerts its effect early in development. To test this hypothesis, we compared the floral organs from the small- and large-fruited NILs. Carpels (which ultimately develop into fruit), styles, and sepals of the large-fruited NIL were already significantly heavier at anthesis ( $P = 0.0007$ , 0.001, and 0.001, respectively) than their counterparts in the small-fruited NIL. Stamen and petals



**Fig. 1. (A)** Fruit size extremes in the genus *Lycopersicon*. On the left is a fruit from the wild tomato species *L. pimpinellifolium*, which like all other wild tomato species, bears very small fruit. On the right is a fruit from *L. esculentum* cv Giant Red, bred to produce extremely large tomatoes. **(B)** Phenotypic effect of the *fw2.2* transgene in the cultivar Mogeor. Fruit are from R1 progeny of *fw107* segregating for the presence (+) or absence (–) of *cos50* containing the small-fruit allele.



**Fig. 2. High-resolution mapping of the *fw2.2* QTL. (A)** The location of *fw2.2* on tomato chromosome 2 in a cross between *L. esculentum* and a NIL containing a small introgression (gray area) from *L. pennellii* [from (8)]. **(B)** Contig of the *fw2.2* candidate region, delimited by recombination events at XO31 and XO33 [from (8)]. Arrows represent the four original candidate cDNAs (70, 27, 38, and 44), and heavy horizontal bars are the four cosmids (cos62, 84, 69, and 50) isolated with these cDNAs as probes. The vertical lines are positions of restriction fragment length polymorphism or cleaved amplified polymorphism (CAPs) markers. **(C)** Sequence analysis of *cos50*, including the positions of cDNA44, *ORFX*, the A-T-rich repeat region, and the "rightmost" recombination event, XO33.

**Table 1. Average fruit weights and seed numbers (23) for R1 progeny of several primary transformants. Unless otherwise noted, progeny are from independent R0 plants. Numbers in parentheses are the numbers of R1 individuals tested.**

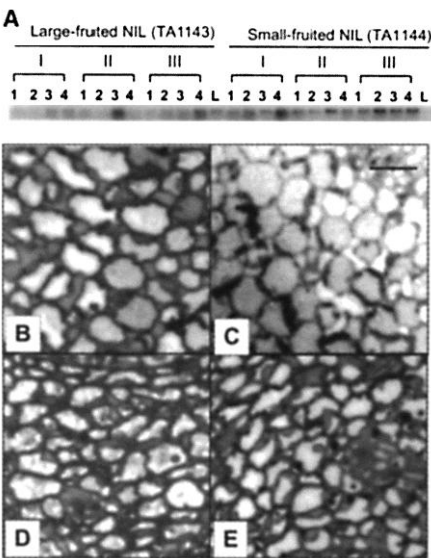
Cosmid	Cultivar	R0 plant no.	Average fruit weight (g)		<i>P</i> value	Average seed number		<i>P</i> value
			+Transgene	–Transgene		+Transgene	–Transgene	
50*	TA496	fw71	41.6 (18)	56.4 (7)	<0.0001	32.6 (18)	28.3 (7)	0.40
50*	TA496	fw71	47.7 (23)	68.1 (12)	<0.0001	31.4 (23)	27.4 (12)	0.44
50	Mogeor	fw107	25.4 (21)	40.9 (7)	<0.0001	24.1 (21)	28.2 (7)	0.34
62	Mogeor	fw59	46.5 (18)	48.0 (9)	0.70	36.1 (18)	36.5 (9)	0.94
62	TA496	fw70	51.0 (21)	51.3 (3)	0.94	28.3 (21)	39.8 (3)	0.04
69	Mogeor	fw51	50.0 (14)	51.7 (10)	0.58	29.8 (14)	34.8 (10)	0.15
84	Mogeor	fw95	49.4 (18)	47.9 (5)	0.71	33.0 (18)	35.5 (5)	0.62

\*R1 progeny of the same primary transformant.

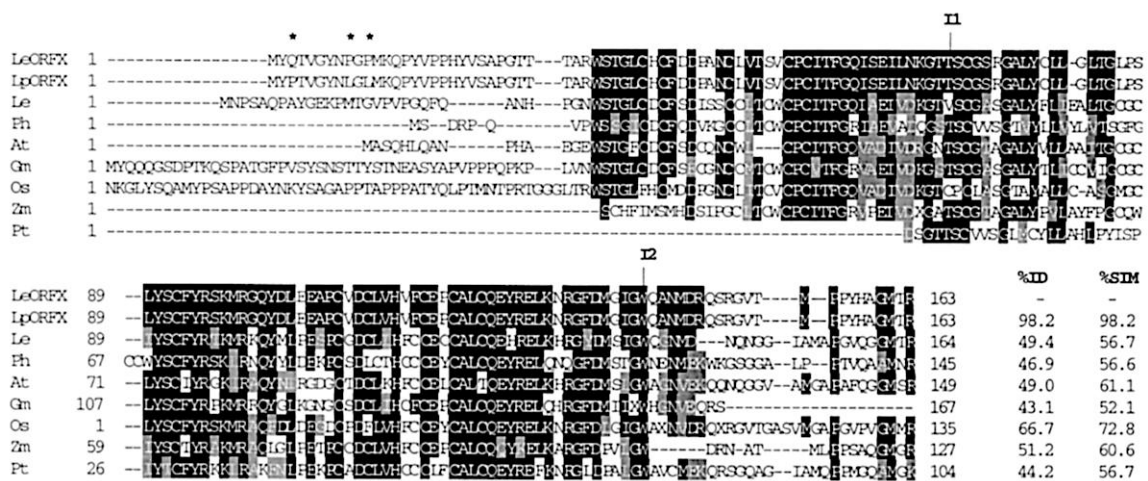
of the *ORFX* cDNA against sequences in the GenBank expressed sequence tag (EST) database found matches only with plant genes. Matches (up to 70% similarity) were found with ESTs in both monocotyledonous and dicotyledonous species. In addition, a weaker match (56.7% similarity) was found with a gymnosperm, *Pinus* (Fig. 4). In tomato, at least four additional paralogs of *ORFX* were identified in the EST database. Although only one *Arabidopsis* EST is represented in the database, eight additional homologs of *ORFX* appear in *Arabidopsis* genomic sequence, often in two or three-gene clusters and having intron-exon arrangements similar to those of

**ORFX.** None of the putative homologs of *ORFX* has a known function. Thus, *ORFX* appears to represent a previously uncharacterized plant-specific multigene family.

**The basis for allelic differences at *fw2.2*.** In order to understand the basis for allelic differences at *fw2.2*, we compared the *L. pennellii* and *L. esculentum* *ORFX* alleles by amplifying and sequencing an 830-nt fragment containing *ORFX* [including 55 nt from the 3' untranslated region (UTR) and 95 nt from the 5' UTR] from both NILs (Fig. 4). Of the 42 nt differences between the two alleles, 35 fell within the two predicted introns, 4 represent silent mutations, and only 3 cause amino acid changes. All three of the substitutions occurred within the first nine residues of the ORF (asterisks in Fig. 4). Although the start methionine cannot be determined with certainty, if the second methionine in the ORF (M12 in Fig. 4) were used, this would place all three potential substitutions in the 5' UTR. Conservation between the alleles



**Fig. 4.** The results of CLUSTALW alignment of LpORFX (*L. pennellii*, AF261775) and LeORFX (*L. esculentum*, AF261774) with 7 representatives of 26 matched from the GenBank EST and nucleotide databases and the contigs assembled from the TIGR (The Institute for Genomic Research) tomato EST database (24). LpORFX and LeORFX residues are shaded black when identical to at least 73% of all the genes included in the analysis. Shading in the other genes represents the black residues in the optimize alignment. Peptide acid residues over the ESTs may be only part of *tomato hybrida*, AF049922 (AU068795), Zm (Zea mays), AF049922 (AU068795).



*L. esculentum* EST is contig TC3457 from the TIGR EST database. At represents a predicted protein from *Arabidopsis* genomic sequence (AB015477.1). The positions of the introns in ORFX are indicated as I1 and I2, and the three residue differences between LpORFX and LeORFX are denoted by asterisks. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

suggests that the *fw2.2* phenotype is probably not caused by differences within the coding region of *ORFX*, but by one or more changes upstream in the promoter region of *ORFX*. Variation in upstream regulatory regions of the *teosinte branched1* gene has also been implicated in the domestication of maize (21). However, differences in fruit size imparted by the different *fw2.2* alleles may be modulated by a combination of sequence changes in the coding and upstream regions of *ORFX* (22).

A reduction in cell division in carpels of the small-fruited NIL is correlated with overall higher levels of *ORFX* transcript, suggesting that *ORFX* may be a negative regulator of cell division. Whether the *ORFX* and *RAX* proteins share common properties other than predicted three-dimensional structure and control of cell division awaits future experimentation. An affirmative result may reflect an ancient and common origin in the processes of cell cycle regulation in plants and animals.

#### References and Notes

1. J. Doebley, A. Stec, J. Wendel, M. Edwards, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9888 (1990).
2. J. Smartt and N. W. Simmonds, *Evolution of Crop Plants* (Longman, London, 1995).
3. C. M. Rick, R. W. Zobel, J. F. Fobes, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 835 (1974).
4. A. H. Paterson et al., *Genetics* **127**, 181 (1991).
5. S. D. Tanksley, *Annu. Rev. Genet.* **27**, 205 (1993).
6. S. Grandillo, H. M. Ku, S. D. Tanksley, *Theor. Appl. Genet.* **99**, 978 (1999).
7. K. B. Alpert, S. Grandillo, S. D. Tanksley, *Theor. Appl. Genet.* **91**, 994 (1995).
8. K. B. Alpert and S. D. Tanksley, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 15503 (1996).
9. Details are available at *Science Online* at [www.sciencemag.org/feature/data/1050401.shl](http://www.sciencemag.org/feature/data/1050401.shl).
10. The Expand Long Template PCR System (Boehringer Mannheim) was used.
11. Constructs were electroporated into *Agrobacterium tumefaciens* strain ABI-AZ08 (Monsanto, St. Louis, MO).
12. A. Frary and E. D. Earle, *Plant Cell Rep.* **16**, 235 (1996).
13. The presence of the transgene was assayed by PCR and Southern hybridization analyses.
14. A total of 11 primary transformants were generated for *cos50*. Although all of these plants carried *np11*, only two individuals (*fw71* and *fw107*) contained the *L. pennellii* portion of the transferred DNA, as determined by PCR analysis with primers designed from the *L. pennellii* sequence of *cos50*.
15. Cosmid sequencing is described at *Science Online* at [www.sciencemag.org/feature/data/1050401.shl](http://www.sciencemag.org/feature/data/1050401.shl).
16. RT-PCR is described at *Science Online* at [www.sciencemag.org/feature/data/1050401.shl](http://www.sciencemag.org/feature/data/1050401.shl).
17. 5' and 3' rapid amplification of cDNA ends (RACE) is described at *Science Online* at [www.sciencemag.org/feature/data/1050401.shl](http://www.sciencemag.org/feature/data/1050401.shl).
18. The predicted *ORFX* protein was compared to a training set of 594 structures (chosen from the Protein Data Base to eliminate redundancy) by using the LOOP algorithm (J. Meller and R. Elber, in preparation). See also [www.tc.cornell.edu/reports/NIH/resource/CompBiologyTools/looppl/](http://www.tc.cornell.edu/reports/NIH/resource/CompBiologyTools/looppl/).
19. The three-dimensional structure of c-H-ras p21 (6q21) is shown at *Science Online* at [www.sciencemag.org/feature/data/1050401.shl](http://www.sciencemag.org/feature/data/1050401.shl).
20. Reviewed in S. R. Sprang, *Curr. Opin. Struct. Biol.* **7**, 849 (1997).
21. R. Wang, A. Stec, J. Hey, L. Lukens, J. Doebley, *Nature* **398**, 236 (1999).
22. P. C. Phillips, *Trends Genet.* **15**, 6 (1999).
23. Seed number is included in the analysis because reduced fertility, as evidenced by reduced seed per fruit, can decrease fruit size. Thus, these data show that the change in fruit size associated with *cos50* is not a byproduct of reduced fertility.
24. The alignment of *LpORFX* and *LeORFX* with a total of 26 genes is shown at *Science Online* at [www.sciencemag.org/feature/data/1050401.shl](http://www.sciencemag.org/feature/data/1050401.shl).
25. We thank J. Nasrallah, C. Aquadro, J. Doebley, K. Schmid, and W. Swanson for critical review of the manuscript. We also thank C. Lewis and N. van Eck for technical assistance. Supported by grants to S.D.T. from the National Research Initiative Cooperative Grants Program, U.S. Department of Agriculture Plant Genome Program (No. 97-35300-4384); the National Science Foundation (No. DBI-9872617); and the Binational Agricultural Research and Development Fund (No. US 2427-94) and by a grant from the NIH NCRR (National Center for Research Resources) to R.E. for development of LOOP at Cornell Theory Center. We dedicate this paper to the memory of Dr. Kevin Alpert whose research inspired this work.

14 March 2000; accepted 4 May 2000

## REPORTS

# Stellar Production Rates of Carbon and Its Abundance in the Universe

H. Oberhummer,<sup>1\*</sup> A. Cs    ,<sup>2</sup> H. Schlattl<sup>3</sup>

The bulk of the carbon in our universe is produced in the triple-alpha process in helium-burning red giant stars. We calculated the change of the triple-alpha reaction rate in a microscopic 12-nucleon model of the <sup>12</sup>C nucleus and looked for the effects of minimal variations of the strengths of the underlying interactions. Stellar model calculations were performed with the alternative reaction rates. Here, we show that outside a narrow window of 0.5 and 4% of the values of the strong and Coulomb forces, respectively, the stellar production of carbon or oxygen is reduced by factors of 30 to 1000.

The formation of <sup>12</sup>C through the triple-alpha process takes place in two sequential steps in the He-burning phase of red giants. In the first step, the unstable <sup>8</sup>Be with a lifetime of only about 10<sup>-16</sup> s is formed in a reaction

equilibrium with the two alpha particles,  $\alpha + \alpha \rightleftharpoons {}^8\text{Be}$ . In the second step, an additional alpha particle is captured,  ${}^8\text{Be}(\alpha, \gamma){}^{12}\text{C}$ . Without a suitable resonance in <sup>12</sup>C, the triple-alpha rate would be much too small to account for the <sup>12</sup>C abundance in our universe. Hoyle (1) suggested that a resonance level in <sup>12</sup>C, at about 300 to 400 keV above the three-alpha threshold, would enhance the triple-alpha reaction rate and would explain the abundance of <sup>12</sup>C in our universe. Such a level was subsequently found experimentally when a resonance that possessed the required properties was discovered (2, 3). It is the

second 0<sup>+</sup> state in <sup>12</sup>C, denoted by 0<sub>2</sub><sup>+</sup>. Its modern parameters (4) are  $\epsilon = (379.47 \pm 0.18)$  keV,  $\Gamma = (8.3 \pm 1)$  eV, and  $\Gamma_\gamma = (3.7 \pm 0.5)$  meV, where  $\epsilon$  is the resonance energy in the center-of-mass frame relative to the three-alpha threshold, and  $\Gamma$  and  $\Gamma_\gamma$  are the total width and radiative width, respectively.

The isotope <sup>12</sup>C is synthesized further in the He burning in red giants by alpha capture to the O isotope <sup>16</sup>O, leading to an abundance ratio in the universe of <sup>12</sup>C:<sup>16</sup>O  $\approx$  1:2 (5). If the carbon abundance in the universe were suppressed by orders of magnitude, no carbon-based life could have developed in the universe. But the production of O is also necessary because no spontaneous development of carbon-based life is possible without the existence of water.

Here, we investigated the abundance ratios of C and O by starting from slight variations of the strength of the nucleon-nucleon (N-N) interaction with a microscopic 12-nucleon model. In previous studies, only hypothetical ad hoc shifts of the resonance energy of the 0<sub>2</sub><sup>+</sup> state were considered (6). Some preliminary results of our calculations are reported elsewhere (7).

The resonant reaction rate for the triple-alpha process ( $r_{3\alpha}$ ) proceeding via the ground state of <sup>8</sup>Be and the 0<sub>2</sub><sup>+</sup> resonance in <sup>12</sup>C is given approximately by (5)

<sup>1</sup>Institute of Nuclear Physics, Vienna University of Technology, Wiedner Hauptstrasse 8-10, A-1040 Vienna, Austria. <sup>2</sup>Department of Atomic Physics, E  tv  s University, P  zm  ny P  ter S  t  ny 1/A, H-1117 Budapest, Hungary. <sup>3</sup>Max-Planck-Institut f  r Astrophysik, Karl-Schwarzschild-Str. 1, D-85741 Garching, Germany.

\*To whom correspondence should be addressed. E-mail: [ohu@kph.tuwien.ac.at](mailto:ohu@kph.tuwien.ac.at)