with Ri*rolB* (Figs 4, 5). Thus, Ng*rolB* can potentially encode a polypeptide 48 amino acids shorter than Ri*rolB* (Fig. 5). The sequence of Ng*rolB* beyond the early stop is strongly homologous with Ri*rolB*. The above considerations reveal that *N. glauca rolB* and *rolC* genes are at least partially degenerate in that *rolB* is truncated and *rolC* right has both a frameshift and a stop codon. As Ng*rolC* left is intact and Ng*rolC* right is not, it is reasonable to assume their common ancestral sequence resembled Ng*rolC* left at these positions. All five reading frames are preceded by a sequence similar to a TATAA box and all show a concensus eukaryotic ribosome binding site of the form A/GXXATGG[9].

Since the divergence of these *A. rhizogenes* and *N. glauca* sequences, only a single frameshift has occurred in 1,857 bp of reading frame. In contrast, 25 frameshifts are evident in 1,240 bp of homologous intergenic DNA. This difference implies that selection has maintained the open reading frames in both sequences. The reading frames are 83% homologous in DNA sequence, whereas the homology in the intergenic regions averages only 75%, not counting frameshifts and not including obviously non-homologous regions. Comparing Ng*rolC* left and Ri*rolC*, the number of changes at the third base of the codon exceeds the sum of such changes at the first and second. The amino-acid sequences of the predicted polypeptides are 75% homologous (Fig. 5). Both sets of reading frames have been conserved in length and content for at least some of the period of time since their divergence.

The Ri-T_L homologous region found in *N. glauca* is a large, imperfect, inverted repeat. Similar structures can be generated on integration of T-DNAs[6]. None of the *A. rhizogenes* strains examined to date are duplicated in this region[2,3,10,11]. Therefore, the inverted repeat was probably generated in the genus *Nicotiana* either at integration or by a later duplication. The levels of sequence divergence between the left and the right arms can be used to set a minimum estimate on the divergence since the original integration event. The sequence divergence between the left and right arms is 1.5% and for the *rolC* region alone is 1.9%, whereas the divergence between Ri*rolC* and either Ng*rolC* is 17%. Most of the divergence preceded duplication of the *N. glauca* region and may have occurred in the *Agrobacterium* strain inciting the infection. If duplication occurred at infection, >85% of the divergence would have been in separate *Agrobacterium* strains. It seems possible that *A. rhizogenes* strains more related to the inciting bacterium than pRiA4b exist.

The left arm of the *N. glauca* inverted repeat contains an intact *rolC* reading frame and a *rolB* reading frame nearly 80% of the length of the *A. rhizogenes rolB* reading frame. The integrity of these reading frames might be taken as evidence for conservation, and thus function, in the genus *Nicotiana*. However, most of the divergence preceded duplication of this region and may have occurred in *Agrobacterium* where selective pressure would maintain these reading frames. If this is true, the presence of these reading frames in *N. glauca* may merely reflect the passage of insufficient time to allow them to become obviously degenerate.

We have presented evidence for a lateral interkingdom transfer of DNA from *Agrobacterium* to *Nicotiana* and its subsequent passage in evolutionary time. Plants transformed with the T_L DNA of present-day strains of *A. rhizogenes* show various developmental changes, including reduced apical dominance, wrinkled leaves and an increased width-to-length ratio in leaves[8]. These phenotypes are presumably caused by the expression of the integrated *rol* genes. The presence of comparable reading frames in the genome of *N. glauca* may have in evolution and may now significantly contribute to phenotype of plants containing them.

We thank Alice Montoya and Al Legro for help and advice, and Donna Akyioshi, Toby Bradshaw, Anath Das, Dean Rigeur and Steve Zeigler for critical reading of the manuscript. I.J.F. was supported initially by a Damon-Runyon postdoctoral fellowship and subsequently by an SERC (UK) postdoctoral fellowship.

Received 12 July; accepted 30 October 1985.

1. White, F. F., Ghidossi, G., Gordon, M. P. & Nester, E. W. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3193–3197 (1982).
2. Huffman, G. A., White, F. F., Gordon, M. P. & Nester, E. W. *J. Bact.* **157**, 269–276 (1984).
3. White, F. F., Taylor, B. H., Huffman, G. A., Gordon, M. P. & Nester, E. W. *J. Bact.* **164**, 33–44 (1985).
4. White, F. F., Garfinkel, D. J., Huffman, G. A., Gordon, M. P. & Nester, E. W. *Nature* **310**, 348–350 (1983).
5. Taylor, B. H., White, F. F., Nester, E. W. & Gordon, M. P. *Molec. gen. Genet.* (in the press).
6. Kwok, W. W., Nester, E. W. & Gordon, M. P. *Nucleic Acids Res.* **13**, 459–471 (1985).
7. Costantino, P., Spano, Pomponi, M., Benvenuto, E. & Ancora, G. *J. Molec. appl. Genet.* **2**, 465–470 (1984).
8. Tepfer, D. *Cell* **37**, 959–967 (1984).
9. Barker, R. F., Idler, K. B., Thompson, D. V. & Kemp, J. D. *Pl. molec. Biol.* **2**, 335–350 (1983).
10. Lahners, K., Byrne, M. C. & Chilton, M.-D. *Plasmid* **11**, 130–140 (1984).
11. Spano, L., Pomponi, M., Costantino, P., Van Slogteren, G. M. S. & Tempe, J. *Pl. molec. Biol.* **1**, 291–300 (1982).
12. Thomashow, M. F., Nutter, R., Montoya, A. L., Gordon, M. P. & Nester, E. W. *Cell* **19**, 729–739 (1980).
13. Vieira, J. & Messing, J. *Gene* **19**, 259–268 (1982).
14. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
15. Guo, L.-H. & Wu, R. *Nucleic Acids Res.* **10**, 2065–2084 (1982).

Erratum

Mycobacterium leprae-specific protein antigens defined by cloned human helper T cells

T. H. M. Ottenhoff, P. R. Klatser, J. Ivani, D. G. Elferink, M. Y. L. de Wit & R. R. P. de Vries

Nature **319**, 66–68 (1986).

IN this letter there were errors in the Note Added in Proof, which which should read: Recent experiments show that also on the recombinant 36K *M. leprae*-specific protein, expressed in *Escherichia coli*, different *M. leprae*-specific protein determinants are detected by these *M. leprae*-specific TLC.

Corrigendum

Efficiency of petroleum expulsion from shale source rocks

D. Leythaeuser, M. Radke & R. G. Schaefer

Nature **311**, 745–748 (1984).

AS discovered only recently, the squalane product which we used as internal standard for quantitation of the gas chromatography data was contaminated by 5.4% pristane. Therefore, the data based on pristane concentrations shown in Figs 2 and 3, respectively, are in error. Values were corrected accordingly and selected rock samples re-analysed for verification. Listed are sample depth (m); the corrected pristane/*n*-heptadecane ratio; and, in parentheses, the pristane/phytane ratio: 1,126.7 = 0.60 (3.21); 1,131.3 = 0.44 (2.95); 1,132.6 = 0.58 (2.65); 1,133.1 = 0.48 (2.50); 1,135.3 = 0.29 (1.60); 1,142.8 = 0.21 (1.96); 1,148.1 = 0.14 (1.82); 1,152.65 = 0.19 (1.96); 1,155.5 = 0.11 (1.58); 1,159.5 = 0.06 (2.14); 1,161.0 = 0.06 (1.26); 1,162.5 = 0.05 (1.16).

As seen from these corrected data, our earlier conclusion of an increasing fractionation between pristane and *n*-heptadecane (expressed by the increase of the ratio between both compounds) with increasing degree of hydrocarbon expulsion from shale source rocks remains valid. However, the magnitude of this fractionation and the corresponding ratio changes are considerably less than we previously thought.