The objection to the use of cut twigs of plants and lack of pre-starvation of the beetles in my experiments is not sustainable. The reduction in attractiveness of cut twigs and the lack of pre-starvation of beetles may have influenced the feeding rate and resulted in underestimates. If I had taken care of these two factors, the beetles would have certainly fed more! And, anyway, it was not my intention to force-feed the beetles.

The authors argue that there were no mass releases of the beetles and yet go on to claim that the Government of India is trying to stop it[2].

Thus, every claim made by the authors[2] seems to be on loose grounds, it is a pity that they made little effort to critically examine my findings. In addition, what they miss out is the implication of such a finding on which Ganeshaiah and Uma Shaanker[15] have elaborated. Therefore, to me, more than the 'Mexican beetle', the purpose of the rejoinder[2] seems elusive.

It is widely accepted that wherever it was released, *Z bicolorata* has significantly reduced the population of *P hysterophorus*[5, 6], but a rigorous proof is lacking. There is an urgent need now to establish quantitatively its impact on *P. hysterophorus* and to investigate the resultant changes in the composition of flora and fauna. A proper analysis of the changes that might have cascaded through the plant community, due to introduction of *Z bicolorata*, should not only serve as an excellent guide for any such biological control programme in future[15] but also tell us more about the functioning of biological communities[16].

1. Kumar, A. R V, *Curr. Sci*, 1992, **63**, 729–730.
2. Chakravarthy, A. K and Bhat, N. S, 1993, **65**, 905–906.
3. Mayr, E, *Principles of Systematic Zoology*, Tata-McGraw Hill, New Delhi, 1977.
4. *ESA, Common Names of Insects and Related Organisms*, Entomological Society of America, Hyatsville, MD, USA, 1989
5. Jayanth, K. P., Status paper on the reported feeding by the Mexican beetle *Zygogramma bicolorata* on sunflower in Bangalore, Mimeographed report, IIHR, Bangalore, 1992, p 10.
6. Manjunath, T. M., On the feeding habits of the Mexican beetle, *Zygogramma bicolorata* – the present status, Mimeographed report of the committee set up to investigate into the Mexican beetle and its host plants, 1992, p. 5.
7. Jayanth, K. P., Report on the studies carried out at IIHR on the feeding by the Mexican beetle on sunflower, Mimeographed report, IIHR, 1992, p. 4
8. Thyagrajan, K. S, *Nature*, 1993, **361**, 387
9. Schoonhoven, L. M, *Entomol Exp Appl*, 1967, **10**, 270–272
10. Jermy, T, Hanson, F. E and Dethier, V G, *Entomol. Exp Appl*, 1968, **11**, 223–229
11. Kovalev, O V., Proc. 7th Session, All-Union Entomol Soc, Leningrad, 1974, pp 31–32
12. Futuyma, D. J, in *Herbivorous Insects: Host Searching Behaviour and Mechanisms* (ed Ahamad, S), Academic Press, New York, 1983
13. Sridhar, S, *Frontline*, 9–22 November, 1991, p 100
14. Chakravarthy, A K and Bhat, N S, *J. Oilseed Res.*, 1993 (in press). As cited vide ref 2
15. Ganeshaiah, K. N and Uma Shaanker, R., *Curr. Sci.*, 1992, **63**, 732–734
16. Pimm, S L., in *Conservation Biology* (ed. Soule, M F), Sinauer Associates, Mass, USA, 1986, pp 309–329.

A. R. V. KUMAR

Department of Agricultural Entomology
University of Agricultural Sciences
GKVK, Bangalore 560 065, India

After split genes it is now split proteins

Protein splicing or 'protein carpentry[1, 2]', a term used by some, is the formation of a functional and mature protein by the removal of an intervening segment of a polypeptide from a precursor followed by joining of the flanking regions. It differs from the well-documented proteolytic cleavage in that the latter involves the removal of the polypeptide segment(s) either from the carboxy- or the amino-terminus to create the functional protein. Protein splicing has been reported from both prokaryotic and eukaryotic systems, viz. yeast[3, 4], *Thermococcus litoralis*[5] and *Mycobacterium tuberculosis*[6, 7] which implies that this phenomenon though very widely distributed shares remarkable similarities.

Protein splicing was first reported[3, 4] in 1990 by two groups working independently on the gene encoding the catalytic subunit of vacuolar type proton-translocating adenosine triphosphatase (TFPI or VMA1) in the yeast *Saccharomyces cerevisiae*. Surprisingly, the gene contained a single open reading frame (ORF) encoding a putative protein of 1071 amino acids (119 kDa) which displayed an actual molecular mass of 67 kDa on SDS-polyacrylamide gels. Analysis of the predicted amino-acid sequence revealed a high degree of homology to the catalytic subunits of H^+-ATPases from several different species. Alignment of the deduced sequence with that of carrot and *Neurospora* revealed that the regions of homology mapped to the amino-terminal and the carboxy-terminal while the middle spacer region encoding a putative 50 kDa protein of 454 amino acids did not exhibit any homology. Scientists at the New England Biolabs[5] while trying to clone the Vent DNA polymerase gene from *Thermococcus* found the single ORF apparently coding for a protein of approximately 180 kDa actually coded for one with a molecular mass of 93 kDa. Again comparison of the deduced amino-acid sequence with other DNA polymerase sequences showed that the Vent DNA polymerase contains two intervening protein sequences (IVS1 and IVS2) that interrupt the conserved motifs. In a very similar situation, Davis *et al.*[6] identified a *recA*-like gene from *Mycobacterium tuberculosis* with a single continuous ORF and deduced that the protein encoded by it would have a predicted molecular mass of 85 kDa which, however, turned out to

be 38 kDa on a Western blot. Comparisons with numerous other recA proteins revealed homologies towards the amino- and the carboxy-terminals with an additional 440 unrelated amino acids in the spacer region In summary, in all the above examples, the genes had a much larger coding capacity compared to the actual size of the protein. Was it then a novel case of protein processing or was it merely another instance of RNA processing?

Although RNA splicing is not widespread in prokaryotes and so is the presence of a single long ORF even for intron-containing genes, it was the simple and natural explanation for the above phenomena. Northern blots and reverse transcriptase PCR experiments[3,4,7] confirmed the presence of RNA species equivalent to the size of the complete ORFs. Neither processed or shortened RNA products could be seen nor conserved sequences or secondary structure representation of intron classes could be found. Moreover, since two proteins equivalent to the predicted sizes of the flanking regions and the spacer region were detected, the other possibility of some kind of protein-processing mechanism, proteolytic cleavage or protein splicing, was explored. Homology between the spacer regions of TFP1 and recA[6,7] is limited to two short motifs of six amino acids each (C, F/L, A, K/E, G, T) and (V, V, V, H, N, C) at their N- and C-termini which were immediately suspected to be the protein-processing junctions. Silent mutations in all the six amino acids[7], i.e. changes at the third wobble position did not affect processing or generation of mature protein, whereas change in a codon at the first position (proposed junction) abolished both. This clearly indicated the involvement of a process other than RNA splicing. The major problem with this protein splicing hypothesis, however, was the inability to detect and establish a precursor–product relationship with the normal wild type genes. The relief came with the analysis of a mutant version of TFP1 gene[4] where an unprocessed species equal to the full length precursor could indeed be detected, albeit in very low amounts Possibly the rate of processing was reduced in the mutated versions which allowed the detection of precursors. *Subsequently, Davis et al*[7] *further* demonstrated in the recA system, conversion of the precursor protein to the processed product (a functional recA protein and the spacer protein) in pulse-chase experiments. This was again using a construct where a short in-frame deletion in the spacer region was introduced in the *recA* locus. This was another strong evidence in support of a protein-processing mechanism. The observation that in a mutated *recA* locus of *Mycobacterium tuberculosis*[7] the spacer protein is derived from the centre of the *recA* locus and not from a terminal end further eliminated the possibility of *proteolytic cleavage in such events.*

The role of the spacer protein remains something of an enigma. Both in yeast[4] and *Thermococcus litoralis*[5], deletion of the entire spacer region still gave rise to a functional protein product indistinguishable from the wild type product. On the other hand, in-frame deletions and frameshift mutations in the spacer region of the *Mycobacterium tuberculosis recA* locus abolished the processing mechanism which in turn abolished the recA protein activity[7]; restoration of the original ORF by a second mutation in the spacer abolished the action of the first mutation. This then raises the question whether certain sequences within the central region are involved in processing in some cases and not in others? If this is so, then how does processing occur in those where these sequences do not play a role? The spacer domain region of TFP1 encodes a fairly abundant and stable protein which has recently been shown[8] to exhibit a site-specific endonuclease activity having extensive homology with a site-specific DNA endonuclease (termed as HO) involved in mating-type switching in yeast. *Thermococcus litoralis* DNA polymerase intervening sequence (IVS2) also has a site-specific endonuclease activity[5] which, however, has not yet been experimentally demonstrated for the recA spacer region. The spacer homology with HO was concentrated within two dodecapeptide motifs that HO shares with a large family of DNA endonucleases These dodecapeptides are encoded in group I introns and also as independent genes. Endonucleases encoded by group I introns are involved in a process called 'intron homing'[9]. Do these spacer proteins comprise a new class of homing endonucleases?

Is protein splicing operating through an autocatalytic mechanism involving the spacer region? Though compelling, the experimental evidences in support of such a mechanism are entirely indirect. Atleast in some cases it seems likely, since processing has been demonstrated to occur in an *in vitro* transcription/translation system[4,7]. Heterologous *in vivo* expression studies also indicate on the least, that splicing factor(s), if required, are also abundantly present to interact and process the intervening protein from distantly related organisms. There is no evidence to show that the separated protein products are ligated to give rise to a mature functional protein but since *no small sized proteins corresponding to* the flanking regions have been detected, it is assumed that they must be combining to form a mature functional molecule The nature of the joining reaction is also not known. There is, however, a similar example from studies on the biosynthesis of jackbean concanavalin A[10,11] where also a central sequence is removed from the mature protein and the flanking sequences are joined in a 'turnaround' topology[3]. There are indications that splicing has occurred by transpeptidation. These studies could provide clues for determining the nature of the linkage in the present cases.

All evidences although circumstantial and indirect do point to a novel protein-processing mechanism. Nevertheless, a lot many issues require further thought and analyses. Consensus sequences at the 'splicing junctions' as yet based only on homology need a deeper look. Since experiments have suggested that sequences within the spacer of the *recA* locus are somehow involved in processing, interplay of the 'junction sequences' and these has to be elucidated. If protein splicing is an autocatalytic mechanism, as speculated, then it should be possible to demonstrate the above interplay of sequences *in vitro*. It would also be interesting to determine if the reaction is wholly autocatalytic or it requires certain factors. This would indeed have profound implications in biotechnology as polypeptides for human use could then be produced in its mature form *in vitro* at much lower costs But, the ushering in of 'protein introns'[2] analogous to the classical introns poses nomenclature problems Should these stretches of peptide be termed as 'introns' or 'intervening sequences'? These terms are traditionally used at the gene level, do we then need a totally new nomenclature for such elements at the protein level? The more fundamental question is the phenomenon of protein splicing *per se* What is the additional benefit for the

organism to harbour split proteins? It is highly unlikely that this could provide mutational advantage in a manner somewhat analogous to introns. Could this then be a mechanism to regulate the activity of a gene in a situation where the limited coding capacity of the prokaryotic or a lower eukaryotic genome (such as yeast) does not permit complex regulatory circuits. It was thought about three decades ago that there is a co-linearity of the gene with the polypeptide chain. This concept was shattered first with the discovery of RNA splicing and editing in eukaryotes and in a few prokaryotes. Protein carpenters are now giving yet another blow which renders inference of protein sequence solely through analysis of DNA a somewhat doubtful affair.

1 Hendrix, R W, *Curr Biol*, 1991, **1**, 71–73

2. Shub, D A. and Goodrich-Blair, *Cell*, 1992, **71**, 183–186.

3 Hirata, R, Ohsumi, Y, Nakano, A, Kawasaki, H, Suzuki, K. and Anraku, Y, *J Biol Chem*, 1990, **265**, 6726–6733

4 Kane, P M, Yamashiro, C. T, Wolczyk, D F., Neff, N, Goebl, M and Stevens, T H, *Science*, 1990, **250**, 651–657

5 Perler, F B, Comb, D G, Jack, W E, Moran, L S, Qiang, B, Kucera, R B., Benner, J, Slatko, B E, Nwankwo, D O, Hempstead, S. K, Carlow, C K S and Jannasch, H, *Proc Natl Acad Sci. USA*, 1992, **89**, 5577–5581

6. Davis, E O, Sedgwick, S. G and Colston, M J, *J Bacteriol*, 1991, **173**, 5653–5662

7. Davis, E. O., Jenner, P. J, Brooks, P C., Colston, M J and Sedgwick, S G., *Cell*, 1992, **71**, 201–210

8 Gimble, F S and Thorner, J, *Nature*, 1992, **357**, 301–306

9 Dujon, B, Belfort, M, Butow, R A, Jacq, C., Lemieux, C, Perlman, P. S. and Vogt, V M, *Gene*, 1989, **82**, 115–118

10 Carrington, D M, Auffret, A. and Hanke, D E, *Nature*, 1985, **313**, 64–67

11 Bowles, D J, Marcus, S E, Pappin, D J C, Findlay, J B C, Eliopoulos, E, Maycox, P R. and Burgess, J, *J Cell Biol*, 1986, **102**, 1284–1297

ACKNOWLEDGEMENT. We thank Dr S K Basu for critical comments

Received 16 August 1993, accepted 16 September 1993.

ANJALI JAIN
SEYED E. HASNAIN

Eukaryotic Gene Expression Laboratory
National Institute of Immunology
New Delhi 110 067, India