Research publications

The research productivity of 234 SRAs who availed full tenure of 36 months in the pool was analysed during the period of study. A total of 290 research papers had been published, 92 papers had been accepted for publication and 163 papers were communicated to various national and international journals by these SRAs. Further, 11 patents were filed by these SRAs during this period.

Conclusion

In the last 40 years, the Scientists' Pool scheme has offered opportunity to over 19,000 highly qualified Indian S&T personnel and over 10,000 had joined the pool to utilize their research potential and simultaneously look for regular positions. The scheme has been made flexible in nature for allowing the SRAs to join the institutions of their choice and pursue a research proposal conceived by them. Majority of the SRAs succeeded in getting absorbed in the sectors of their placement. The study indicates that the SRAs are to be encouraged to join the industrial sector in order to improve their employment prospects. The research output of the SRAs has been appreciable. Over the years many SRAs[1] acquired eminent positions and attained the level of Directors in various R&D organizations and Professors or equivalent positions in various universities and other institutions of higher learning. The scheme has proved to be an effective instrument in developing and providing continuously a band of trained R&D scientists and technologists to the nation.

1. Pal, A. K., Ahuja, I. M. L., Das, S. R., Ranglani, G., Chhiber, O. K., Nandy, R. K., Sehgal, V. P., Pail, S. and Rajagopal, N. R., *J. Sci. Ind. Res.*, 1996, 55, 765–771.

Received 21 November 1997; revised accepted 20 January 1998.

The origin of adaptive mutations: Explaining the mutational spectra of Lac$^+$ revertants of the *Escherichia coli* strain FC40

S. K. Bhattacharjee and S. K. Mahajan*

Reversion of lacI33-lacZ frameshift mutation to produce Lac$^+$ colonies takes place by demonstrably different mechanisms in growing and starving cultures of strains (such as FC40) harbouring this mutation. The revertants appear to arise in starving cells only under conditions where they can immediately promote growth. This is reminescent of Lamarckism. Here we propose a mutagenic mechanism which is Darwinian inasmuch as it is blind to the adaptive fitness of the mutants but can explain the mutational spectra seen in FC40 and certain other strains. This mechanism can produce mutations only in the neighbourhood of a methylatable cytosine.

SINCE the Darwinian theory of evolution freed biology from all transcendence and did not require living organisms to have any mysterious properties unexplained by the chemical properties of their constituents, it found ready scientific acceptance (in spite of considerable oppostion from certain non-scientific quarters) and by 1930s it became almost the universally accepted theory of evolution, at least as far as the evolution of higher organisms was concerned. However, microorganisms seemed to be somewhat exceptional because they very rapidly adapted to changing environments, suggesting a possible role for environment in producing specific (directed) mutations which could increase the fitness of the organisms. In 1940s the fluctuation analysis of Luria and Delbruck and other experiments based on sib selection and replica plating showed that bacterial mutants which survived a lethal environmental change were largely produced before the population was exposed to the lethal challenge (see ref.1 for a detailed discussion). These mutations, therefore, could not have been directed by the new environment. Since these results were consistent with the Darwinian theory, bacteria were admitted into the fold of neo-Darwinism, in spite of a warning

The authors are in the Molecular Biology and Agriculture Division, Bhabha Atomic Research Centre, Mumbai 400 085, India.
*For correspondence. (e-mail: mbad@magnum.barc.ernet.in)

by Delbruck[2] that the above experiments had not ruled out the possibility that bacteria exposed to prolonged nonlethal stress might produce specifically beneficial mutations directed by the environment. Subsequent discovery of the double helical structure of DNA, which suggested mutation production due to random errors during replication, and other developments in molecular biology, e.g. the universality of the genetic code, the existence of silent and neutral as well as harmful mutations, were all consistent with neo-Darwinism and led to near universal rejection of the Lamarckian idea of organisms being able to produce exclusively useful mutations. However, certain recent reports from several laboratories have reopened this debate.

Between 1954 and 1963 Ryan and his coworkers published several papers showing that when a population of His⁻ bacteria was incubated in a medium lacking histidine, apparently adaptive mutations arose in a time-dependent manner after several days of incubation[3]. These mutations appeared to be directed by the future advantage they conferred on the host. However, due to the universal appeal and acceptance of Darwinism by that time their possible Lamarckian implications were ignored. In 1988, Cairns and his coworkers[4] reported that when *E. coli* cells of a Lac⁻ strain were plated on minimal medium plates containing lactose as the sole source of carbon, apparently adaptive Lac⁺ mutations arose continuously for several days thereafter, in addition to some pre-existing Lac⁺ mutations which form colonies within the first two days. These mutations seemed to have Lamarckian characteristics in that they were produced only if they conferred a growth advantage, i.e. if lactose was present on the plates and there was no other nutritional deficiency. Since then, there have been several other reports of production of similar Lamarckian-looking mutations in several bacterial and yeast systems[5].

There have been three types of general responses to these reports:

(i) To rigorously examine the design and interpretation of the concerned experiments to ascertain the validity of their claim that the selective conditions produce only immediately beneficial mutations. This has resulted in detection of significant flaws in several of the experiments making their claims of directed mutagenesis untenable (see refs 6 and 7 for review). However, in a few cases the evidence for the preferential, if not exclusive, recovery of the beneficial mutations under nonlethal selection has not been seriously challenged. These cases include the generation of Lac⁺ revertants in the strain FC40 and its derivatives (see below).

(ii) To propose, and in some cases also to test, molecular models that could explain the occurrence of these mutations in terms of mechanisms which are blind to the adaptive fitness of mutants they produce. Several such models have been proposed. These include, reverse transcription of specific mutant mRNAs[4], mutagenic transcription[8], continuous generation by an error-prone, low level DNA repair synthesis of mutations which are fixed only if they promote cell growth but are otherwise reverted by a slow acting mismatch repair system[9,10], amplification of mutant genes[11], RecA-dependent stable DNA replication in non-dividing cells[12], and the hypermutable state model, according to which small proportion of a stressed population enters a hypermutable state which leads either to acquisition of a growth-promoting mutation or to eventual cell death[13] by presumably lethal mutations. Empirical testing[14] and careful theoretical examination has shown most of these models to be untenable. However, there is some evidence in partial support of three of these models, namely the-slow-repair-during-starvation model, the stable-DNA-replication in non-dividing cells model (also called the 'toe in water' model), and the hypermutable state model. These models are not mutually exclusive.

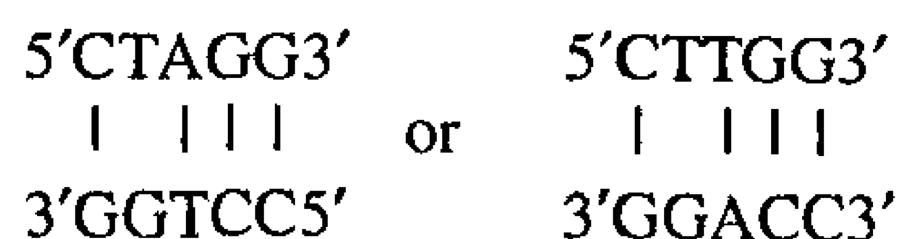
(iii) To empirically examine the nature of the late-arising beneficial mutations and the genetic mechinery responsible for their production with a view to ascertaining whether they are produced by a special mechanism (distinct from the mechanism(s) responsible for the production of mutations with similar phenotypes in growing cells) and whether this mechanism can be understood in terms of the known processes of molecular biology. This approach has largely concentrated on the reversion of *lacI33* : : *lacZ* fusion in FC40 (refs 14–22). Based on the spectrum of the mutational sites and the enzymatic machinery required for producing these mutations, it appears that in this strain the mechanism that produces the post-plating Lac⁺ revertants is distinct from the ones(s) responsible for producing Lac⁺ revertants in growing cells. Inasmuch as the former mechanism seemed to produce only the beneficial mutations after exposure to the selection pressure in the absence of cell division, without producing other mutations (e.g. trpE972 → Trp⁺) which would not be immediately advantageous on the selection plates, it would appear to be a non-Darwinian (Lamarckian?) mechanism of mutagenesis.

However, the crucial question to distinguish a Lamarckian from a Darwinian mechanism should be whether it can foresee the phenotypic consequences of specific mutations before producing them or is random in the sense of being blind to these consequences. Other kinds of apparent non-randomness, e.g. an ability to produce mutations only in certain sequence contexts, would not make the mechanism non-Darwinian, merely because it produces mutations in some genes and not in others. Here we propose that the mechanism responsible for the late-arising Lac⁺ mutants in FC40 is a Darwinian mechanism in the above sense and that the failure to recover non-adaptive mutants in the few attempts made so far can be understood in terms of the biochemical

properties of this mechanism. The mechanism proposed by us can also explain the recovery, of the (apparently) adaptive mutations in several other cases as well as nonrecovery of non-adaptive mutations.

The rationale for the model

The most likely cause of mutations produced in a non-replicating cell not exposed to any external environmental mutagen (such as FC40 plated on lactose M9 plates) should be some intracellular enzymatic processes. In *E. coli* the only known process which can produce deletions in mononucleotide tracts is polymerase slippage during DNA synthesis or unequal homologous recombination. If one such process is indeed responsible for the late arising Lac$^+$ in FC40, then its range must be very short because no mutations are seen in other mononucleotide repeats present only a few bp away from the affected tracts (Figure 1). One known process involving very short stretches (≈ 10 nt) of repair synthesis is the so-called VSPMR (very short patch of mismatch repair) system in which the DNA synthesis is initiated at a nick made by a strand specific, sequence specific endonuclease, Vsr (refs 23, 24). This endonuclease recognizes T : G mismatches present in certain specific sequence contexts, namely,



and nicks the strand containing the mismatched T immediately 5′ to this T. T : G mismatches in some other sequence contexts, namely, CTAGG, CTTGG, CTAGC, TTAGG, GTAGG, CTAGT, CTAGA, TTTGG, CTCGG, ATAGG, CTTGC, TTAGA, TTAGC and CTGGG (with the second T as the mispaired base in each case), which may differ from the above canonical at one or two bases also act as substrates of this enzyme though with variable efficiencies[25,26]. The 3′OH end generated by Vsr nicking serves as the primer for the VSP repair synthesis which does not appear to extend beyond 10–12 bp (ref. 25). If during the course of this repair synthesis the DNA polymerase (or an enzyme complex containing the polymerase) were to encounter a simple sequence repeat it would have the potential to

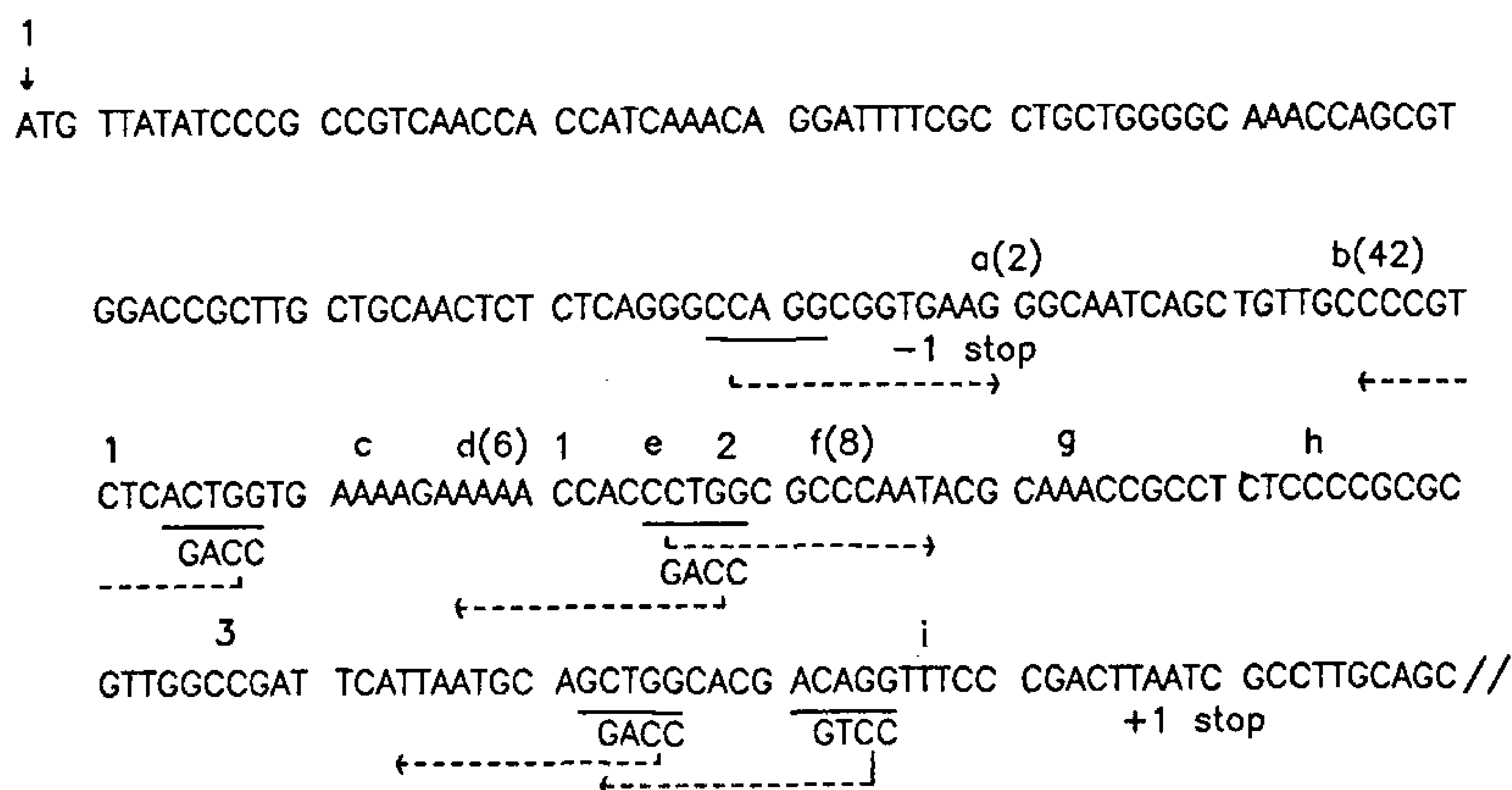


Figure 1. Sequence of 244 nucleotides of a region of the *lacI–lacZ* fusion carried by the F′ conjugative plasmid in the strain FC40 which has a deletion in the chromosomal *lac* operon. A +1 frameshift due to an insertion mutation CCC to CCCC at bp 119–121 creates a stop codon at bp 229–231 thus making the cell Lac$^-$. Lac$^+$ reversion can occur by any one of a number of possible mutations within the −1 and +1 stop signals that restore the *lacZ* reading frame without creating a stop codon. These mutations may be 1 bp deletions or more complex DNA rearrangements. The early and late arising Lac$^+$ alleles were amplified using PCR primers outside this stretch and sequenced by Rosenberg *et al.*[15] and Foster and Trimachi[21]. A total of 66 late arising Lac$^+$ mutants were sequenced. All had 1 bp deletions. The sites of these deletions and the numbers recovered at each site are shown. Letters (**a**) to (**i**) demark mononucleotide stretches of length ≥ 3 nt. The most striking feature here is that 56/65 (86%) of all the mutants were confined to only three of these tracts **b**, **d** and **f** at positions 119–122 (CCCC), 140–144 (AAAAA) and 156–158 (CCC) while no mutants were produced at any of the five similar repeats, **c**, **e**, **g**, **h** and **i**; only two mutations mapped at **a**. The putative VSP repair sites are underlined. The 'C' residues where 5 methyl-C to T mutation (or C to U) would generate a putative target site[49–51] for the Vsr endonuclease are immediately above the start of the dashed lines which indicate the range and the direction of the repair tract predicted by the model. In four cases (positions 131, 152, 210 and 219), these C residues lie on the complementary strand for which the sequence in the relevant regions is shown.

produce short deletions by enzyme slippage. Alternatively, ligation of the newly synthesized strand and the old strand before the synthesis of the last excised base(s) may also generate a deletion. If the VSPMR is indeed the source of the mutation produced in stationary state cells, one prediction would be that the mononucleotide repeats which serve as mutational hotspots in *lacI33–lacZ* should have potential Vsr target sites in their vicinity on the 5′ side while other mutationally cold repeats should not. Examination of the 135 bp sequence between −1 and +1 stop signals in Figure 1 showed it to have at the desired places (i.e. near the deletion hot spots) sequence contexts where C → T transition would generate potential Vsr sites. These (C → T) transitions can be generated by methylation of the 'C' residues by the enzyme Dcm (DNA cystosine methylase) and deamination of the resulting 5-methyl cytosine[27].

The model

These facts and a careful examination of other reported characteristics of the late arising Lac⁺ revertants suggested the following model which rests on three postulates: (i) Certain enzymatic processes continue to function in stationary state cells though, probably, at a low level. These include, methylation of appropriately located cytosine residues in DNA, deamination of the 5 meCs to give T : G mismatches, and repair of these mismatches. (ii) Of the two competing repair pathways available for the repair of T : G mismatches, namely, the adenine methylation dependent (general) mismatch repair pathway (MMR), which is also called the long patch mismatch repair system (LPMR), and the very short patch mismatch repair (VSPMR)[23] system specific to the T : G mismatches, the former accounts for most of the repair in growing (WT) cells but in stationary cells (or even in growing cells lacking a functional LPMR) the relative contribution of the latter increases. This may be due to down-regulation of the LPMR genes (especially *mutS*) which would make the target (T : G mismatch) more easily accessible to the Vsr endonuclease which initiate the VSPMR, or reduced ability of the LPMR to effect the repair due to non-availability of hemi-methylated GATC sequence in the absence of regular chromosomal replication, or poor supply of dNTPs required for excision and resynthesis of long tracts of DNA associated with LPMR. (iii) The DNA replication/recombination associated with VSPMR, is error prone and can produce new mutations in the short repair tract (of less than 10–12 nt) 5′ to the excised T. These mutations may include deletions in tracts of repeated nucleotides. VSPMR is thus proposed to be both anti-mutagenic (in reverting the C → T transitions) and mutagenic (in producing new mutations in the neighbourhood of the T : G mismatch). In contrast, LPMR is largely anti-mutagenic as the repair replication associated with it is largely error-free. It should be mentioned here that the mutagenic activity of VSPMR proposed by us is different from another mutagenic activity of this system present in growing cells, where the correction of a T : G mismatch produced by an A → G transition can fix the mutation[28]. (iv) If the mutation so produced is on the transcribed strand and adaptive (i.e. with the potential of promoting growth immediately), the transcription of the mutant gene and translation of the mRNA leads to production of functional protein (β-galactosidase in FC40) which promotes growth and DNA replication resulting in fixation of the mutation. If the mutation is on the non-transcribed (coding) strand then it may not lead immediately to the production of a functional protein and its own fixation by replication, but the mismatch produced may be corrected by LPMR operating on the complementary strand, initiated presumably at a nearby nick or an unmethylated GATC. This mismatch correction could lead to transfer of the mutation to the transcribed strand followed by its fixation via transcription, cell growth and chromosomal replication.

Evidence that unmethylated GATC stretches are present in stationary state cells is provided by the observation that overproduction of the DNA adenine methylase (Dam) which would reduce the LPMR activity, considerably enhances the recovery of the late arising Lac⁺ revertants in FC40, an effect also shown by *mutL* : : Tn*10* mutations in this background[17]. Apparently, LPMR is also operative in stationary state cells[29] though not as efficiently as in growing cells where it seems to repair most of the T : G mismatches and makes them inaccessible to the error-prone VSPMR. However, mutational inactivation of *mutS* or *mutL* may make T : G mismatches accessible to VSPMR even in growing cells which then gives the same spectrum of Lac⁺ mutations as seen in (WT) stationary state cells[18]. Beletskii and Bhagwat[30] have argued that processes like repair, recombination and transcription which make ds DNA transiently single stranded enhance deamination and shown that transcription of a gene enhances the frequency of C → T transition at the Dcm sites in the non-transcribed strand. This implies that in spite of the additional repair step required to transfer a mutation from a non-transcribed strand to a transcribed strand, the contribution of an appropriately located Dcm site on the former (non-transcribed) strand to the apparently adaptive Lac⁺ mutations may be quite high. This may explain why comparable numbers of late arising Lac⁺ are recovered at sites **d** and **f**.

The explanation of the late arising Lac⁺ mutant spectrum in FC40 based on the above model is as follows. There are a total of nine mononucleotide repeats of length 3 nt or more in the sequence between −1 and +1 stop codons (Figure 1). Of these, three (**b**, **d** and

f) account for 86% (56/65) of the post plating Lac$^+$ revertants sequenced by the groups of Rosenberg and Foster[15,21], a fourth **a** for some 3% and the other five for none. It is remarkable that each one of **a**, **b**, **d** and **f** lies within 12 bp of a methylatable cytosine residue (at the start of the arrows in Figure 1) in which, a C → T mutation would generate a putative Vsr target sequence, while the neighbourhoods of the other four repeats (**e**, **g**, **h** and **i**) have no such cytosines. In spite of the fact that the repeat **e** (CCC), which overlaps with a canonical Dcm site is not mutated in any of the 65 sequenced Lac$^+$ revertants, this correlation is much too strong to be coincidental and strongly suggests a role for methylatable cytosine in the production of the observed late-arising Lac$^+$ mutants.

Of the five C residues, three at positions 92, 150 and 152 lie in canonical Dcm methylation sequences (5′CCWGG3′) and can be converted to T by methylation and deamination. The Vsr targets thus generated at these sites could account for the mutations seen in the tracts **a**, **d** and **f**. The C residues at positions 110 and 132 do not lie in canonical Dcm methylation sequences but their conversion to T would generate the sequences 5′TTAGC3′ and 5′CTAGT3′ respectively, both of which are known targets of Vsr (ref. 26). We propose that these 'C' residues, and probably all 'C' residues whose conversion to T is likely to generate a Vsr target site, can serve as substrates of the Dcm methyl transferase. Several lines of evidence support the above proposal. First, the gene (*vsr*) encoding the Vsr endonuclease is coordinately regulated with *dcm* from a common promoter. It is reasonable to expect the two enzymes to act in a common pathway, with matching sequence specificities. Second, the canonical sequence context for the C methylation by Dcm, namely 5′CCWGG3′ (W = A or T), was proposed by May and Hattman[31] based on the structure of pyrimidine fragments generated by depurination of radioactively labelled DNA. A careful perusal of the data of these workers[31] shows that though they had emphasized methylation only at the canonical sequences, a low frequency methylation of cytosine in non-canonical contexts is not ruled out. Third, Bhagwat and his coworkers[32] have shown that *M. Eco*RII, which is a methylase sharing with Dcm a common canonical methylation sequence and other characteristics including active site structure, can methylate cytosine in the non-canonical context 5′GCAGG3′ at about 10% of the efficiency of the canonical context. They have proposed that *M. Eco*RII may methylate any sequence matching the canonical sequence at four out of the five bases. It is reasonable to believe that Dcm would have similarly relaxed sequence specificity. Indeed the results of Glasner *et al.*[26], who investigated the efficiency of various pentanucleotides as Vsr targets, suggest that many sequences differing from the canonical target sequence at two residues are also active. This may be so even for Dcm.

With the above relaxation in the Dcm sites, the mutagenic role of the C residues at positions 92, 110, 132, 150 and 152 in generating the hot spots **a**, **b**, **c**, **d** and **f** is obvious. It may be noted that 5′ATT GG3′ is not a Vsr substrate[26], therefore, if C → T transition were to take place at position 129, it would not initiate VSPMR which would explain why 4As at **c** do not form a mutational hot spot. However, some quantitative anomalies need further explaining. Why are the frequencies of mutations at different hot spots so different? **d** and **f**, which are near a canonical Dcm site yielded only 6 and 8 mutations respectively, while **b** which has only non-canonical Dcm sites in its neighbourhood, gave 42 mutants. Further, though the VSPMR mediated mutation at **d** would lie on the transcribed strand those at **f** would be on the non-transcribed strand and therefore not immediately adaptive; however, these two repeats contribute to nearly the same extent to the adaptive Lac$^+$ colonies. How? Again **a**, which is also near a canonical Dcm site, gave only two mutations and **e** gave none though it overlaps a canonical site. The answers to these questions would obviously have to await elucidation of the details of the molecular mechanisms involved in production of the T : G mismatches and in their (error-prone) repair. However, the available facts permit some speculations.

The length of the repair tract associated with VSPMR is known to be very small[25]. If it is postulated that this length falls rapidly after 10 nt then the low mutation frequency at **a**, which is 12 nt away from the methylatable cytosine at position 92, will be expected. If it is further postulated that up to 10 nt the probability of recovering a mutation (at least a deletion) increases with the tract length then both the non-recovery of any mutation at **e** (whose first base is only two bases from the putative repair initiation site) and the high recovery at **b** (where the first base is 10 nt away from the repair initiation site on right and 9 nt from that on left (Figure 1), can be understood. Several models with such properties and which explain the differential contribution of different hot spots consistent with the known genetic requirements of the phenomenon are conceivable (to be discussed elsewhere). However, a schematic diagram depicting the essential features of the model as proposed here is shown (Figure 2).

Why are the late arising Lac$^+$ mutants not recovered (formed?) during starvation on plates lacking lactose or another essential nutrient? We propose that such mutants are actually formed but in the absence of immediate cell growth most of them lose their viability due to production in the same cells of additional (deleterious) mutations or are repaired by LPMR during prolonged incubation in the absence of selection or both. In our

hypothesis, during starvation all Dcm sites act as potential sources of new mutations in their vicinity. A cell which has experienced a *lac*⁺ reversion (i.e. a –1 deletion in one of the appropriate tracts) can be recovered as a Lac⁺ colony only if transcription of the mutant allele and the consequent cell growth and replication of the mutant chromosomal strand occurs before the latter experiences a deleterious mutation near another methylatable cytosine. If the revertant allele cannot be transcribed due to absence of an appropriate inducer or if the cell cannot grow due to the lack of an appropriate nutrient then the cell is likely to lose viability due to formation of a lethal mutation elsewhere on the chromosome.

Why does the number of adaptive Lac⁺ revertants increase almost linearly with time? Several scenarios consistent with our hypothesis are possible. One which we favour is suggested by the observations that during starvation on lactose plates there is hardly any loss of viability of the FC40 cells but its *mutS* derivative loses viability rapidly. MutS is also the protein which recognizes T : G (and other mismatches) and initiates their repair by the LPMR. The absence (or low supply) of this protein makes the T : G mismatches accessible to Vsr leading to VSPMR which is responsible for the production of both the adaptive and the lethal mutations.

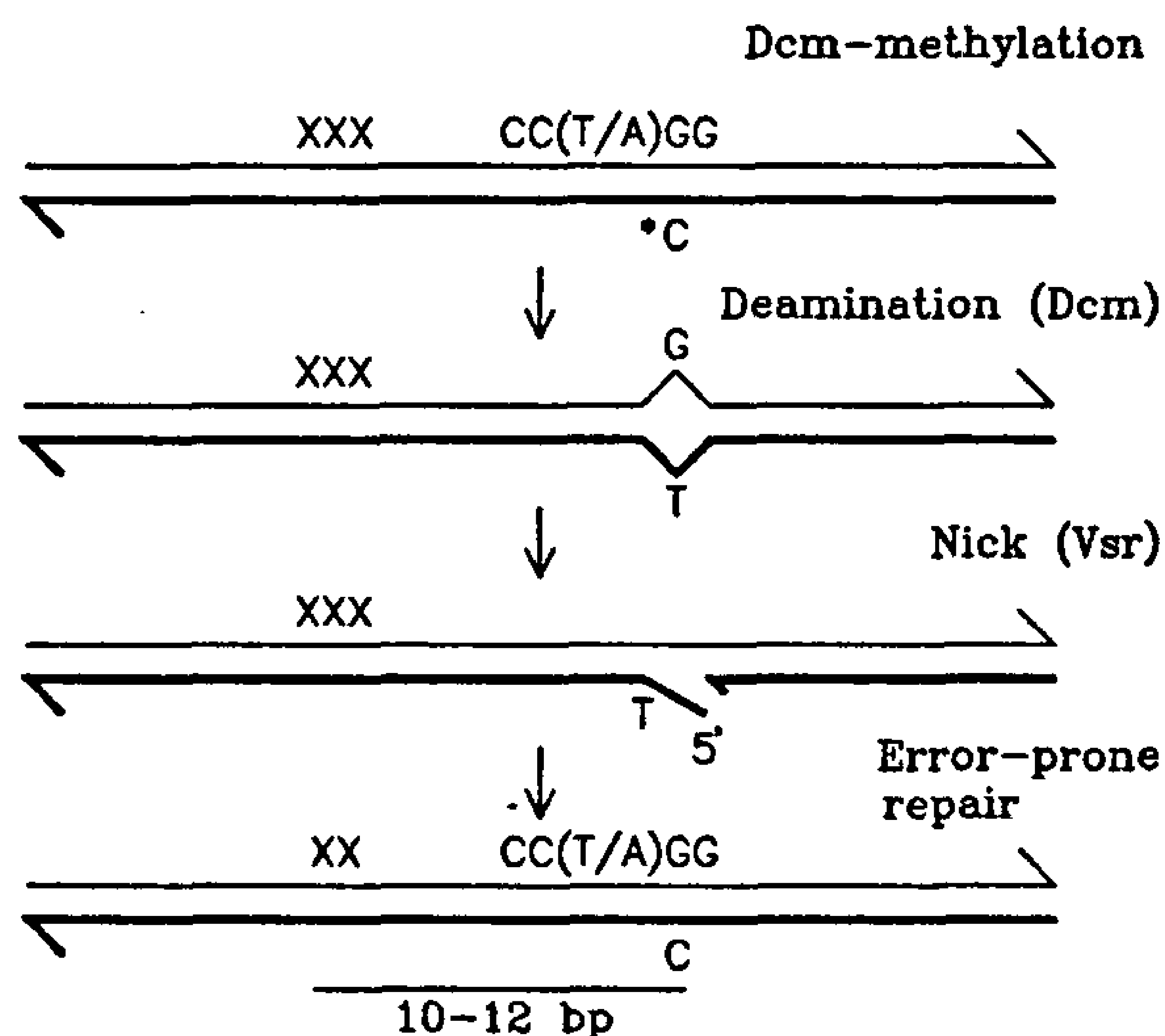


Figure 2. Schematic view of the proposed model for explaining the occurrence of adaptive Lac⁺ revertants in FC40. *C is a cytosine residue which is vulnerable to methylation due to its presence in the sequence context recognized by Dcm. Only the canonical Dcm site is shown though several degenerate sites with relaxed specificity of its recognition can also be targets of Dcm and of Vsr (after C → T mutation) (see text for details). 'XXX' represents a mononucleotide tract with 3 or more residues. In FC40 all selected adaptive mutations are single base deletions from such tracts. We propose that the error-prone VSPMR may also lead to point mutations within its range in other systems (see the text for details).

If it is assumed that in starving FC40 cells on lactose plates, at any given time the MutS concentration falls below a critical level (necessary for the accessibility of the T : G mismatches to Vsr and activation of VSPMR) only in a small proportion of cells which can now undergo hypermutagenesis in the neighbourhood of Dcm sites leading to either adaptive Lac⁺ or lethal mutations, then the constant rate of the formation of Lac⁺ colonies would be predicted. Stochastic processes which can produce widely different levels of the products of a specific gene in different cellular subpopulations have recently been proposed[33].

The model seems also to be consistent with the mutational spectra of adaptive His⁺ revertants obtained by Prival and Cebula[34] with two isogenic strains of *S. typhimurium* harbouring two different mutations *hisG46* and *hisG428* at widely separate locations in the same coding sequence. No putative VSPMR initiation site exists in the neighbourhood of the ochre mutation in *hisG46*. However, the strain *hisG428* gave time-dependent mutations. The ochre codon of *hisG428* lies entirely within 10 base pairs of the site, 3′-GCTGA-5′ which contains the core methylatable CTG of the Dcm recognition site and is not one of the known VSPMR non-vulnerable sites[26]. Presumably, methylation occurs despite substitution of the second G by A.

The delayed Trp⁺ and Tyr⁺ revertants obtained from two ochre mutants *trpE6* and *tyrA* analysed by Bridges[35,36] differed from that of ochre mutant *hisG428*. In both the cases, revertants were neither altered at the ochre site nor were suppressor mutations detected when probed by ochre suppressor T4 test phages. We analysed the sequences of wild type alleles of both and found an abundance of putative VSPMR initiation sequences (unpublished results). Although other possibilities cannot be ruled out, it is more likely that the revertants harbour some intragenic frameshift mutation in both cases and that the poor growth was probably a consequence of partially defective products in the revertants.

The two mutator tRNA genes identified by Slupska *et al.*[37] have been suggested to cause mutator phenotype by substitution of C to T in the anticodon in one of the four wild type copies of the gene present in the cell. We note that the anticodon overlaps partially with the noncanonical sequence (underlined), 5′C<u>CAAGG</u>3′. All the three independent mutations occurred at the adjacent C at the 5′ end which is the 4th base from the putative methylatable C in the complementary strand and is likely to have arisen by VSPMR process as hypothesized in our model. The methylatable C is presumably conserved by VSPMR[27,38] which gives a selective advantage for enhancing variability within a short stretch in the neighbourhood without losing the site for initiation of the repair on the chromosome. Such mutations may, therefore, occur more frequently under stationary state

and give rise to hypermutable subpopulations. Hypermutable subpopulalions have long been proposed as an explanation[13,39,40] for adaptive mutagenesis though the mechanism of their emergence was not known. Recent reports based on computer simulation[41] as well as experimental studies[42] show that such subpopulations indeed arise though no specific molecular mechanisms have been proposed.

The model proposed here makes several testable predictions. For instance, the relative contribution of any of the three hot tracts (namely **b, d** and **f**) in Figure 1 should be reduced by altering the neighbouring putative methylatable Cs and their sequence contexts so as to make them nonamenable to Dcm methylation or VSPM repair. Conversely, creation of an appropriate Dcm recognition site close to (10 to 12 nt) any of the cold tracts in *lacI33–lacZ* should increase the mutation frequency at these tracts. Similar rules should also apply to production of mutations in other genes, i.e. mutations located within 10 nt of an appropriately oriented Dcm methylatable C should give high frequencies of apparently adaptive while others not so located should not. One interesting case in this context would be the production of Val^R mutations studied by Jayaraman[43]. The spectra of Val^R mutations in growing and stationary state cell would be predicted to have Dcm sites near the latter. Another set of predictions of the model would be that mutational inactivation of any of the genes involved in the creation of $C \rightarrow T$ mutations (i.e. *dcm*) or in the repair of the T : G mismatches thus created (e.g. *vsr* or *polA*) should decrease the frequency of late arising Lac^+ mutants. Jayaraman[43] also showed that the adaptive Val^R mutations in *mutS* background were *recA*$^+$ dependent while those in *mutL* background are not. Since the late arising Lac^+ in FC40 are also produced by a *recA*$^+$-dependent mechanism, one would be tempted to predict that the adaptive Val^R mutants in *mutS* cells are produced by a Dcm directed VSPMR mutations while those in *mutL* background are produced by another mechanism. The mutational spectra obtained in these two backgrounds should answer this question. Additionally, since our model proposes a role for VSPM repair in production of the adaptive mutations, the genes which on inactivation block the production of Lac^+ revertants in FC40, would be predicted to be involved in VSPMR mechanism. Therefore, if our hypothesis is valid in general, VSPMR should be inactive in *recA*, *recB* or *recC* hosts in the classical lambda genetic crosses performed by Lieb[25] where the phenomenon was first observed and later suggested to explain gene conversion.

Thus, we believe that a mutagenic VSP-mismatch repair process conserves the site of mismatch and is consistent with the occurrence and nonoccurrence of adaptive mutations in most of the cases though mutational events in some have not been fully characterized. We also attribute a significant biological role, as a tool of evolvability, to the deaminase activity of Dcm and propose that it is true for 5-methylcytosine methylases in general. The coregulation of *dcm* and *vsr* as part of the same operon[44] makes sense since both are necessary for efficient creation of adaptive mutations under stress. The suggestion that the degenerate sites of the Vsr-endonuclease action could be involved in genome evolution[26] is provided a testable mechanistic basis by our hypothesis involving VSPMR as a mutagenic process. Chi site (5′-GCTGGTGG-3′) stimulation of recombination[45] may be a consequence of initiation of VSPMR process since it includes an asymmetric putative Dcm recognition sequence that could be responsible for its polarity. The repetitive sequence, REP, in *E. coli* also has one such putative methylatable tetranucleotide on both arms of the conserved inverted repeat each being nearly equal to the maximum VSP mismatch repair tract length[46]. Such sequences and trinucleotide repeats containing methylatable cytosine such as C(T/A)G have been suggested to play a role in evolution of the genome and in diseases like cancer in as yet unknown ways[46–48].

1. Brock, T. D., in *The Emergence of Bacterial Genetics*, 1990, Cold Spring Harbor Laboratory.
2. Delbruck, M., in *Cold Spring Harbor Symp. Quant. Biol.*, 1947, **11**, 139–155 (while discussing Andre Lwoff's results).
3. Ryan, F. J., Okada, T. and Nakata, T., *J. Gen. Microbiol.*, 1963, **30**, 193–199.
4. Cairns, J., Overbaugh, J. and Miller, S., *Nature*, 1988, **335**, 142–145.
5. Foster, P. L., *Annu. Rev. Microbiol.*, 1993, **47**, 467–507.
6. Lenski, R. E. and Mittler, J. E., *Science*, 1993, **259**, 188–194.
7. Sneigowski, P. D., *J. Mol. Evol.*, 1995, **40**, 94–101.
8. Davis, B. D., *Proc. Natl. Acad. Sci.*, 1989, **86**, 5005–5009.
9. Stahl, F. W., *Nature*, 1988, **335**, 112–113.
10. Cairns, J. and Foster, P. L., *Genetics*, 1991, **128**, 695–701.
11. Boe, L., *Mol. Microbiol*, 1990, **4**, 597–601.
12. Stahl, F. W., *Genetics*, 1992, **132**, 865–867.
13. Hall, B. G., *Genetics*, 1990, **126**, 5–16.
14. Foster, P. and Cairns, J., *Genetics*, 1992, **131**, 783–789.
15. Harris, R. S., Longerich, J. and Rosenberg, S. M., *Science*, 1994, **264**, 258–260.
16. Rosenberg, S. M., Harries, R. S., Longerich, S. and Galloway, A. M., *Mutat. Res.*, 1996, **350**, 69–76.
17. Foster, P. L., Gudmundsson, G., Trimachi, J. M., Cai, H. and Goodman, M. F., *Proc. Natl. Acad. Sci.*, 1996, **92**, 7951–7955.
18. Longerich, S., Galloway, A. M., Harris, R. S., Wong, C. and Rosenberg, S. M., *Proc. Natl. Acad. Sci.*, 1995, **92**, 12017–12020.
19. Rosenberg, S. M., Harris, R. S., Longerich, S. and Galloway, A. M., *Mutat. Res.*, 1996, **350**, 69–76.
20. Foster, P. L. and Trimachi, J. M., *Proc. Natl. Acad. Sci.*, 1995, **92**, 7951–7955.
21. Foster, P. L. and Trimachi, J. M., *Science*, 1994, **265**, 407–410.
22. Foster, P. L., *J. Bacteriol.*, 1997, **179**, 1550–1554.
23. Modrich, P., *Annu. Rev. Genet.*, 1991, **25**, 229–253.
24. Hennecke, F., Kolmar, H., Brundl, K. and Fritz, H-J., *Nature*, 1991, **353**, 776–778.
25. Lieb, M., Allen, E. and Read, D., *Genetics*, 1986, **114**, 1041–1060.
26. Glasner, W., Merkl, R., Schellenberger, V. and Fritz, H-J., *J. Mol. Biol.*, 1995, **245**, 1–7.

27. Wyszynski, M., Gabbara, S. and Bhagwat, A. S., *Proc. Natl. Acad. Sci.*, 1994, **91**, 1574–1578.
28. Merkl, R., Kroger, M., Rice, P. and Fritz, H.-J., *Nucleic Acids Res.*, 1992, **20**, 1657–1662.
29. Reddy, M. and Gowrishankar, J., *Genetics*, 1997, **147**, 981–1001.
30. Beletskii, A. and Bhagwat, A. S., *Proc. Natl. Acad. Sci.*, 1996, **93**, 13919–13924.
31. May, M. S. and Hattman, S., *J. Bacteriol.*, 1975, **122**, 129–138.
32. Bandaru, B., Gopal, J. and Bhagwat, A. S., *J. Biol. Chem.*, 1996, **271**, 7851–7859.
33. McAdams, H. H. and Arkin, A., *Proc. Natl. Acad. Sci.*, 1997, **94**, 814–819.
34. Prival, M. J. and Cebula, T. A., *Genetics*, 1992, **132**, 303–310.
35. Bridges, B. A., *Mutat. Res.*, 1993, **302**, 173–176.
36. Bridges, B. A., *Mutat. Res.*, 1994, **307**, 149–156.
37. Slupska, M. M., Baikalov, C., Lloyed, R. and Miller, J. H., *Proc. Natl. Acad. Sci.*, 1996, **93**, 4380–4385.
38. Lieb, M. and Rehmat, S., *Proc. Natl. Acad. Sci.*, 1997, **94**, 940–945
39. Bridges, B. A., *Nature*, 1997, **387**, 557–558.
40. Moxon, E. R. and Thaler, D. S., *Nature*, 1997, **387**, 659–662.
41. Taddel, F., Radman, M., Maynard-Smith, J., Toupance, B., Gouyon, P. H. and Godelle, B., *Nature*, 1997, **387**, 700–702.
42. Sniegowski, P. D., Gerrish, P. J. and Lenski, R. E., *Nature*, 1997, **387**, 703–705.
43. Jayaraman, R., *J. Genet.*, 1992, **71**, 33–41.
44. Sohail, A., Leib, M., Dar, M. and Bhagwat, A. S., *J. Bacteriol.*, 1990, **172**, 4214–4221.
45. Myers, F. S. and Stahl, F., *Annu. Rev. Genet.*, 1994, **28**, 49–70.
46. Higgins, C. F., McLaran, R. S. and Newbury, S. F., *Gene*, 1988, **72**, 3–14.
47. Darlow, J. M. and Leach, D. R. F., *Genetics*, 1995, **141**, 825–832.
48. Smith, S. S. and Baker, D. J., *Biochem. Biophys. Res. Commun.*, (in press).
49. Gabbara, S., Wyszynski, M. and Bhagwat, A. S., *Mol. Gen. Genet.*, 1994, **243**, 244–248.
50. Yebra, M. J. and Bhagwat, A. S., *Biochemistry*, 1995, **34**, 14752–14757.
51. Doiron, K. M. J., Viau, S., Koutroumanis, M. and Cupples, C. G., *J. Bacteriol.*, 1996, **178**, 4294–4296.

ACKNOWLEDGEMENTS. All the sequences analysed were from the GeneBank Database. The program SQM, kindly provided by N. D. Shirke of Computer Division, BARC, was used for sequence analysis. We thank Manjula Mathur, Jyoti Ramachandani, K. G. Khot and Inderjeet Singh for their help.

Received 4 December 1997; revised accepted 7 March 1998

MEETINGS/SYMPOSIA/SEMINARS

National Seminar on Polymer Research in Academy, Industry and R&D Organization

Date: 26–27 June 1998
Place: Calcutta

Contact:
Prof. S. K. Dolui
Convener
Department of Chemical Sciences
Tezpur University
Tezpur 784 025
Phone: (03712) 27004-6 (3 lines)
Fax: (03712) 21539
e-mail: tul@tezpuru.ren.nic.in

Dr Basudam Adhikari
Joint Convener
Materials Science Centre
Indian Institute of Technology, Kharagpur 721 302
Phone: (03222) 55221-4 (4 lines) Extn. 4782 (Off.), 78041 (Res.)
Fax: (03222) 55303
e-mail: ba@matsc.iitkgp.ernet.in

Current Concepts in Immunology

Date: 30 Oct. to 1 Nov. 1998
Place: Jawaharlal Nehru University, New Delhi

A two and a half-day education workshop in immunology will be held in JNU from 30 October to 1 November 1998. This workshop will immediately precede the 10th International Congress of Immunology to be held in New Delhi from 1 November to 6 November 1998. The basic aim of the workshop is to provide an update on current concepts of immunology.

Contact: Prof. Rajiv. K. Saxena
School of Life Sciences
Jawaharlal Nehru University
New Delhi 110 067
Fax: 91-11-6187338
e-mail: rksaxena@jnuniv.ernet.in
Website: www.afl.co.in/indtravels/confer.htm