# A phage T4 site-specific endonuclease, SegE, is responsible for a non-reciprocal genetic exchange between T-even-related phages

Farid A. Kadyrov*, Michael G. Shlyapnikov, Valentine M. Kryukov

*Laboratory of Genetic Enzymology, Institute of Biochemistry and Physiology of Microorganisms of Russian Academy of Sciences, Pushchino, Moscow region, 142292, Russia*

Received 8 August 1997

**Abstract The bacteriophage T4 *seg*E gene encoding site-specific endonuclease lies between the *hoc*.1 and *uvs*W genes. The similar region of T-even-related phage RB30 lacks the *seg*E gene. Here we demonstrate that the phage T4 *seg*E gene is inherited preferably by progeny of mixed infection with RB30. The preferred inheritance of the *seg*E gene depends on its own expression and is based on a non-reciprocal homologous recombination event providing the transfer of the gene from the *seg*E-containing to the *seg*E-lacking allele. The SegE endonuclease cleaves DNA in a site located at the 5′ end of the *uvs*W gene in the RB30 genome. The T4 DNA is also cleaved by the enzyme, but less efficiently. The cleavage at the RB30 site appears to initiate the observed conversion, which is stimulated by DNA homology and accompanied by co-conversion of flanking markers. Our findings provide a novel example of endonuclease-dependent generation of genetic variation in prokaryotes.**



*Key words:* Bacteriophage T4; Site-specific endonuclease; DNA break; Non-reciprocal genetic exchange; Genetic variation; Intron homing

## 1. Introduction

It is well documented that target DNA cleavage plays the key role in initiation of diverse genetic processes including intron homing, meiotic gene conversion, transposition and V(D)J recombination [1–3]. These data have resulted in the view that the target DNA cleavage and subsequent DNA repair are involved in molecular mechanisms leading to genetic variations. The study of homing of *Escherichia coli* phage T4 group I introns *td* and *sun*Y has demonstrated that endonuclease-dependent genetic variations appear to occur in mixed infections of related bacteriophages [4]. Endonucleases I-TevI [5] and I-TevII [6] encoded by ORFs of the *td* and *sun*Y introns, respectively, are required for homing of their own introns [4]. In each case, the enzyme initiates double-strand break repair, cleaving intronless DNA of the related T2 phage near the intron insertion site. Double-strand break repair results in the efficient unidirectional transfer of the intron to the intronless allele, converting the gene to an intron-containing allele. Furthermore, at least in the case of *td* intron homing the event is associated with co-conversion of flanking exon sequences [7,8]. Although the requirement for I-TevI and I-TevII for intron homing has been shown in plasmid-to-phage and phage-to-plasmid mobility experiments [4], it seems likely that these endonucleases can also promote the same processes during mixed infection between T4 and T2. Recently, a direct involvement of a site-specific endonuclease, I-HmuII, in the generation of genetic variation occurring in mixed infection of the closely related *Bacillus subtilis* phages SP01 and SP82 has been demonstrated [9]. The I-HmuII encoded by the ORF in the DNA polymerase intron of phage SP82 is responsible for the preferred inheritance of the intron and its flanking markers by progeny of mixed infection with SP01 harbouring a similar DNA polymerase intron [9]. Contrary to intron homing [1], this process is initiated by I-HmuII-mediated cleavage of SP01 DNA near the site of intron insertion and occurs via a conversion of the intron and flanking DNA sequence of phage SP01 to the corresponding region of the phage SP82 genome. The above results allow one to suggest that other phage site-specific endonucleases might be involved in similar processes of genetic exchange between related phages.

The *seg*E gene of phage T4 lies in the *hoc*.1-*uvs*W intergenic region and encodes a site-specific endonuclease [10–13]. SegE and related proteins (SegA–SegD) have a GIY-YIG motif, which is also present within I-TevI endonuclease [14]. We have found earlier that efficient transfer of the *seg*E gene from plasmids to the *hoc*.1-*uvs*W intergenic region of the T-even-related phage RB30 which naturally lacks this gene depends on *seg*E expression [15]. Here we demonstrate that the *seg*E gene is inherited preferentially by progeny of mixed infections between *seg*E-containing and *seg*E-lacking phages. This process is based on NHRE and depends on SegE endonuclease activity.

## 2. Materials and methods

### 2.1. Primers

Primers were designed to anneal to phage T4 DNA and were as follows: 1, 5′-TTTAGCTTTAGCTTCTTTG-3′; 2, 5′-CCCGACTAAGCCGAATG-3′; 3, 5′-CCATCAGCTTCGAAGCTAAAGAAATC-3′; 4, 5′-TTGGATCCTTATTTTAATGTTACG-3′; 5, 5′-CGAAAATAGCTTTACCATGCG-3′; 6, 5′-TTCATATGACTTTTACAGTTGATATAACTCC-3′; 7, 5′-TTAAGCTTGACCACCTGGATAGGTATAGATG-3′; 8, 5′-TTCATATGTACCACTTTGTATAT-3′; 9, 5′-ATGCCTTACTAAATTTGCCT-3′. Nucleotides mismatched with phage T4 sequences are underlined. The primer locations are shown in Fig. 1.

### 2.2. Plasmids and phages

Plasmid construction, PCR amplification and DNA sequencing were performed according to standard techniques [19]. The locations of cloned DNA fragments are shown in Fig. 1. The construction of plasmids pVT123, pMX3,4-2, pVTX and psEET-3a was described earlier [11,15,16]. Plasmids pEW19, pseT219 and pΔA19 contain fragments from the T4 *hoc*.1-*uvs*W region that were generated by PCR using primer pairs 1+5, 1+2, and 1+3, respectively, and inserted into

*Corresponding author. Fax: (7) (95) 923 3602.
E-mail: kadyrov@ibpm.serpukhov.su

*Abbreviations:* ORF, open reading frame; NHRE, non-reciprocal homologous recombination event; aa, amino acid


*PII* S0014-5793(97)01098-3

the *Sma*I site of pUC19. The sizes of DNA fragments cloned into the plasmids pEW19, pseT219 and pΔA19 are 1829 bp, 934 bp and 840 bp, respectively. pEWΔP19 was obtained by deletion of the 697 bp *Pst*I fragment from pEW19. In all plasmids, the *seg*E-containing insert is cloned in orientation opposite to the *lac* promoter. The structure of all the cloned fragments was verified by DNA sequencing. Plasmids pHW and pHΔEW19 contain fragments from the *hoc*.1-*uvs*W region of phage RB30 generated by PCR using primer pairs 1+5 and 4+5, respectively, and inserted into the *Sma*I site of pUC19. The sizes of DNA fragments cloned into the plasmids pHW and pHΔEW19 are 1212 bp and 1835 bp, respectively. These plasmids were used to determine a nucleotide sequence of the 1790 bp DNA fragment of the phage RB30 genome from the 3′ end of primer 4 to the 3′ end of primer 5 (the accession number in the EMBL Data Library is Y13611).

Phage RB30 was kindly provided by Prof. W.B. Wood (University of Colorado, USA) and Dr. E.S. Miller (North Carolina State University, USA). Phage T4 was from our laboratory's collection. Phage stocks were grown according to Steinberg and Edgar [17] using *E. coli* strain BL21 [18] as a host. To construct recombinant phages, crosses between RB30 and pVT123 and between RB30 and pVTX were performed. Negative colonies of phages RB30*seg*E and RB30*seg*EΔ*Xba*, which had acquired a copy of the native and the disrupted *seg*E gene, respectively, were identified using plaque hybridisation with a $^{33}$P-labelled *seg*E probe. Results of the sequencing of fragments amplified by PCR from the phage DNAs confirmed that DNA sequences from the 3′ end of primer 1 to the 3′ end of primer 2 surrounding both the native and the disrupted *seg*E are identical to those of phage T4.

*2.3. Mixed infections and crosses between phages and plasmids*

*E. coli* BL21 cells were grown at 37°C to an $OD_{600}$ of 0.6–0.7 ($5\times10^8$ cells/ml) in L broth supplemented with 2% casamino acids; just before infection, KCN was added to give a final concentration of 4 mM. Mixed infections were performed with a multiplicity of infection of 5 of each phage by mixing equal volumes of two phage dilutions (in L broth), each $5\times10^9$/ml, and then adding one volume of the mixture to an equal volume of bacteria. Adsorption was allowed to proceed for 10 min. The cultures were incubated after infection for 2 h at 37°C and then treated with chloroform to ensure complete lysis of bacterial cells. To perform crosses between phages and plasmids, phages were plated on *E. coli* BL21 transformants ($\sim2\times10^8$ cells) in top agar on LB plates and grown overnight at 37°C. Progeny phages from plates where confluent lysis occurred were extracted with TMN buffer (10 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 150 mM NaCl) for 2 h at 4°C. Dilutions of the resulting progeny of mixed infections or crosses between phages and plasmids were plated in top agar with *E. coli* BL21 and incubated overnight at 37°C.

*2.4. Hybridisation of progeny plaque DNAs to $^{33}$P-labelled segE probe*

Primers 8 and 9, homologous to the *seg*E ORF and the template DNA of psEET-3a, were used for $^{33}$P-labelling of DNA by PCR according to Mertz and Rashtchian [20]. Hybridisation of progeny plaque DNAs to the $^{33}$P-labelled *seg*E probe was performed as described by Sambrook et al. [19]. After hybridisation filters were washed four times for 10 min in 0.5×SSC at 60°C.

*2.5. The endonuclease SegE preparation*

The endonuclease SegE was expressed in *E. coli* and purified as described earlier [11].

Table 1
Inheritance of the *seg*E gene by the progeny of mixed infections between the *seg*E-containing and *seg*E-lacking phages

| Cross | *seg*E-containing progeny (%)[a] |
|---|---|
| 1. T4+RB30 | 99.5±0.4 |
| 2. RB30*seg*EΔ*Xba*+RB30 | 52.4±4.2 |
| 3. RB30*seg*E+RB30 | 99.5±0.4 |

[a]Percentages with standard deviations determined from four separate experiments were calculated from the number of progeny plaques giving a positive signal with the $^{33}$P-labelled probe of the *seg*E gene, divided by the total number of progeny, times 100. The percentage of RB30*seg*EΔ*Xba* progeny phage for cross 3 is presented.

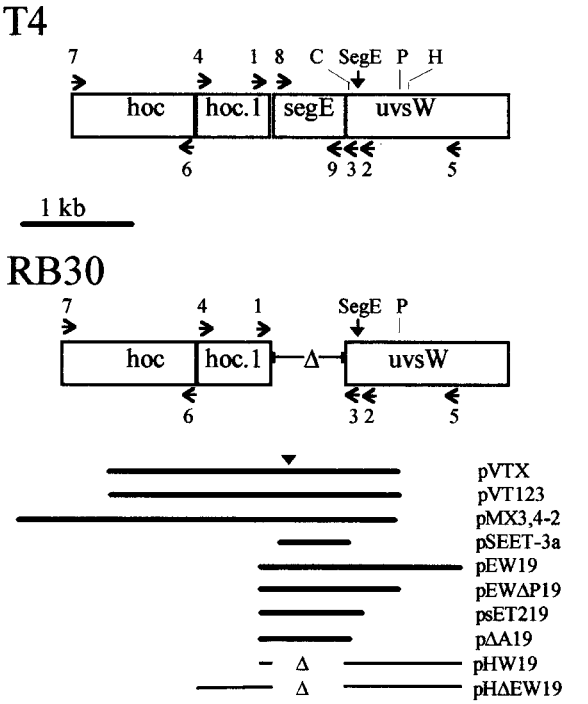


Fig. 1. Physical maps of the T4 and RB30 genome regions, containing the *hoc* gene through to the *uvs*W gene and summary of plasmids containing inserts from the corresponding region of these phages. Locations of genes *hoc*, *hoc*.1, *seg*E and *uvs*W are indicated by boxes. Cleavage sites of the SegE endonuclease are shown by vertical arrows. Maps also indicate locations of selected restriction sites (C, *Cla*I; P, *Pst*I; H, *Hin*dIII) and primers 1–9 (horizontal arrows) used in the study. Bold and thin lines designate the cloned T4 and RB30 sequences, respectively. Arrowhead shows the location of TAG triplet of *seg*E in pVTX.

## 3. Results

To determine whether the phage T4 *seg*E gene is inherited preferentially by progeny of mixed infection, *E. coli* BL21 was coinfected with phages T4 and RB30. The progeny were analysed for the presence of the *seg*E gene by hybridisation with the $^{33}$P-labelled *seg*E probe. The results presented in Table 1 (cross 1) indicate that 99.5% of the progeny inherit the *seg*E gene. In order to explore the requirement of *seg*E gene expression for its own preferential inheritance, the recombinant phages RB30*seg*E and RB30*seg*EΔ*Xba* were constructed. Unlike RB30*seg*E, RB30*seg*EΔ*Xba* contains the disrupted *seg*E gene in which translation of SegE is truncated after 51 aa. The phages RB30*seg*E, RB30*seg*EΔ*Xba* and RB30 in single infections had comparable burst sizes indicating that the *seg*E DNA does not influence phage propagation. When coinfection of RB30 with RB30*seg*EΔ*Xba* occurred, uniform distribution of the disrupted *seg*E gene in the progeny was observed (Table 1, cross 2). In contrast, the overwhelming majority of progeny from mixed infection of RB30 with RB30*seg*E inherited the native *seg*E gene (Table 1, cross 3). These data indicate that the preferred inheritance of the *seg*E gene absolutely depends on *seg*E gene expression.

In order to determine whether the preferential inheritance of the *seg*E gene is associated with the exclusion of the RB30 region lacking this sequence, genomic DNAs extracted from 48 progeny plaques of the mixed infection between RB30 and RB30*seg*E were amplified by PCR with primers 4 and 5. In all

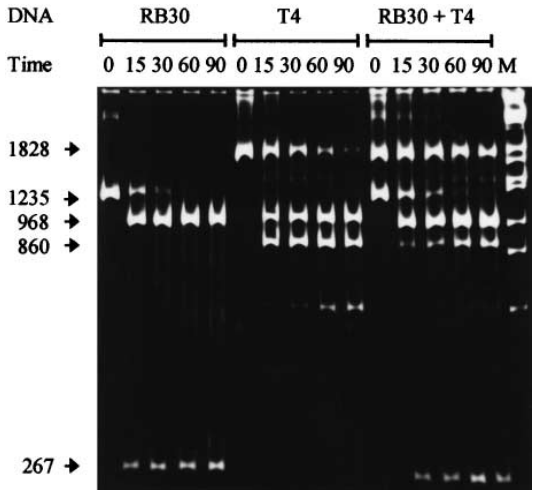


Fig. 2. The endonuclease SegE prefers to cleave RB30 DNA. 1.5 μg of individual DNA fragments that were amplified from T4 or RB30 DNA using primers 1 and 5 and purified by Wizard PCR Preps (Promega), or 3 μg of an equal mixture of these fragments (1:1) was incubated at 30°C with 5 ng of the SegE in a 75 μl reaction mixture containing 20 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$ and 50 μg/ml bovine serum albumin. Aliquots of 15 μl were removed at 0, 15, 30, 60 and 90 min. Reactions were stopped by adding 10 mM EDTA and transfer to 0°C. The reaction products separated on 6% polyacrylamide gel. Sizes of reaction products (bp) are indicated on the left. M, lambda DNA digested with *Eco*RI and *Hin*dIII.

48 cases, the amplification of a single DNA fragment containing the *seg*E ORF was observed (data not shown). This might account for either SegE-dependent restriction of DNA lacking *seg*E or SegE-dependent transfer of the *seg*E gene from the RB30*seg*E genome to the *hoc*.1-*uvs*W intergenic region of RB30 via NHRE. The first suggestion was eliminated, because propagation of RB30 in transformants harbouring a plasmid containing the native *seg*E gene did not change its burst size.

To further confirm the non-reciprocal character of *seg*E inheritance, the frequencies of loss and acquisition of the *seg*E gene by progeny of crosses between phages and *E. coli* transformants harbouring plasmids were measured. The plasmids which were used in these and subsequent experiments are shown in Fig. 1. When crosses between *seg*E-containing phage (T4 or RB30*seg*E) and transformants harbouring the cloned RB30 DNA fragment (pHW19 or pHΔEW19) were performed, the frequency of loss of the *seg*E gene by progeny was less than 0.2% (the expected level of a reciprocal recombination only). On the other hand, 16–71% of progeny acquired the gene from crosses between RB30 and transformants harbouring the T4 cloned DNAs (Table 2, crosses 1–3). Thus, the frequency of gene acquisition is approximately two orders of magnitude greater than the level of gene loss. These results indicate that preferential inheritance occurs by NHRE ensuring transfer of the *seg*E gene from a donor allele to a recipient one.

The dependence of *seg*E gene transfer on its own expression strongly suggested that the endonuclease SegE might be capable of initiating the transfer. We have previously demonstrated that a site at the 5′ end of the *uvs*W gene of both T4 and RB30 (Fig. 1) is cleaved by SegE endonuclease [12]. Furthermore, primer extension analysis confirmed that cleavage at the site of both *seg*E-lacking and *seg*E-containing phages occurred when cells were infected by RB30 and RB30*seg*E but never took place during co-propagation of RB30 and RB30*seg*EΔ*Xba* (Kadyrov et al., unpublished results). These results were provocative because the same sensitivity of both DNAs to SegE would not result in efficient *seg*E gene transfer. This suggests that a probable step defining the transfer is conditioned by a preferential cleavage of the RB30 site. To verify this hypothesis we studied SegE-mediated cleavage of PCR fragments of phages RB30 and T4 both separately and in a mixture. As shown in Fig. 2, cleavage at the expected sites occurred. In addition, weak cleavage was observed at a secondary site of T4 DNA located within *seg*E [11,12]. When individual DNA was incubated with SegE, a preferential cutting of RB30 DNA was found. This effect was more pronounced when both substrates were incubated with the enzyme simultaneously. The presence of T4 DNA did not influence the efficiency of RB30 DNA cleavage but the degree of T4 DNA cleavage decreased sharply. Therefore the SegE endonuclease prefers cleaving phage RB30 DNA indicating that preferential cleavage is responsible for the efficient *seg*E gene transfer.

A comparison of the RB30 DNA sequence from the 3′ end of primer 4 to the 3′ end of primer 5 with that of phage T4 reveals that with the exception of the *seg*E ORF, these sequences share 95.6% identity. It is known that after cleavage and nucleotide degradation, the 3′ ends of the recipient allele are annealed with complementary sequences of the donor allele generating a donor-recipient duplex, an essential step of NHRE initiated by double-strand breakage [1]. In the case of *seg*E gene transfer, recipient sequences both upstream of the site lacking the *seg*E gene and downstream of the SegE cleavage site might be involved in the formation of duplexes with homologous donor sequences. In order to determine whether the recipient phage requires the presence of the homologous donor sequences to efficiently acquire the gene, transformants containing the native *seg*E gene and different portions of the flanking sequences were infected with RB30. Efficient acquis-

Table 2
Influence of the length of the cloned sequences flanking the *seg*E ORF on gene acquisition by the progeny of crosses with RB30

| Cross | 5′ length[a] (bp) | 3′ length[b] (bp) | *seg*E-acquiring progeny (%)[c] |
|---|---|---|---|
| 1. pMX3,4-2 | 2614 | 400 (272) | 16.4 ± 1.2 |
| 2. pEW19 | 115 | 1096 (968) | 71 ± 2.5 |
| 3. pEWΔP19 | 115 | 422 (294) | 19.2 ± 1.5 |
| 4. psET219 | 115 | 201 (73) | 0.8 ± 0.5 |
| 5. pΔA19 | 115 | 107 | < 0.01 |

[a]Length of cloned T4 DNA upstream of the *seg*E ORF.
[b]Length of cloned T4 DNA downstream of the *seg*E ORF. The downstream DNA homology region between a donor plasmid and the cleaved recipient phage genome (3′ length of the cloned T4 DNA minus 128 bp) is indicated in parentheses.
[c]Percentages are presented with standard deviations for exception of cross 5, where *seg*E-acquiring progeny was not observed, and were calculated as in Table 1.

ition of the *seg*E gene by RB30 occurred when the flanking sequences contained at least 115 bp at the 5′ end and 272 bp at the 3′ end (Table 2, crosses 1–3) homologous to the cleaved recipient DNA upstream of the site lacking the *seg*E gene and downstream of the SegE cleavage site, respectively. The frequency of *seg*E gene inheritance by progeny declined approximately threefold when the downstream DNA homology was decreased from 968 bp to 294 bp. Further decreasing the downstream DNA homology to 73 bp or its absence did not result in efficient acquisition of the *seg*E gene by RB30 (Table 2, crosses 4 and 5). Thus, efficient *seg*E gene transfer is stimulated by extensive DNA homology between the donor and recipient alleles since the downstream DNA homology required for the transfer is within a range of 73–272 bp.

To study whether the conversion to *seg*E is accompanied by co-conversion of flanking sequences, the frequency of appearance of T4 markers surrounding the gene in *seg*E-acquiring progeny was analysed using a PCR-based approach. Phage T4 *hoc* gene located 717 bp upstream of the *seg*E was used as a 5′ DNA marker (Fig. 1), since a fragment amplified with the *hoc*-specific primers 6 and 7 from RB30 DNA is about 200 bp longer than the analogous PCR product of T4 (Fig. 3). The observed difference in size appears to reflect the different molecular weights of the T4 Hoc (40 kDa) and RB30 Hoc (50 kDa) proteins [21]. PCR analyses of 20 *seg*E-acquiring phage plaque DNAs for the presence of the T4 *hoc* gene were performed and some of the results are shown in Fig. 3. The amplification products of five out of 20 DNAs contained the fragment corresponding to the T4 *hoc* gene, indicating that this gene is transferred together with *seg*E at a frequency of 25%. In order to test co-conversion of 3′ flanking sequences, *Cla*I and *Hin*dIII sites, which are located 48 bp and 440 bp downstream of the 3′ end of the *seg*E gene, respectively (Fig. 1), in T4 but not RB30 DNA, were used as markers. Thirteen out of 14 patterns generated by PCR followed by digestion showed fragments similar to the T4 DNA pattern. In only one case was the loss of the *Hin*dIII site observed (Fig. 4, lane 3). These data indicate that the conversion of the *seg*E$^-$ to the *seg*E$^+$ allele is accompanied by highly efficient co-conversion of the 3′ flanking markers.

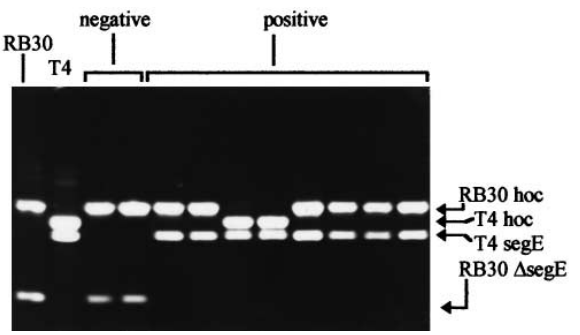


Fig. 3. The conversion to *seg*E is accompanied by co-conversion of the 5′ flanking marker. DNA of single plaques of progeny from the cross between RB30 and pMX3,4-2 (Table 2, cross 1), which gave negative or positive signals with the $^{33}$P-labelled probe of *seg*E DNA, were amplified using the primer pairs 1+2 and 6+7. Phage RB30 and T4 DNAs were amplified with the same primers and used as controls. PCR products were separated on 1.2% agarose gels. Arrows on the right indicate the positions of the RB30 *hoc*-, T4 *hoc*-, T4 *seg*E- and RB30 Δ*seg*E-containing fragments.

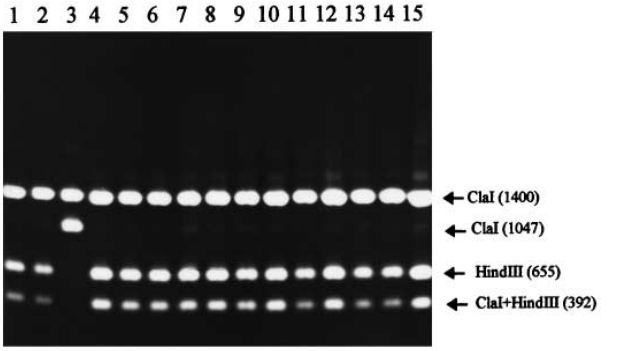


Fig. 4. The conversion to *seg*E is accompanied by co-conversion of the 3′ flanking markers. DNA of single plaques of progeny from the cross between RB30 and pEW19 (Table 1, cross 2), which acquired *seg*E, were amplified using primer pair 4+5 and cleaved with *Hin*dIII and *Cla*I (lanes 1–14). The same pattern obtained from phage T4 DNA was used as a control (lane 15). The cleavage products were separated on 1.2% agarose gels. Arrows on the right indicate the positions of the cleavage products of the 2447 bp PCR fragment at the *Cla*I, *Cla*I+*Hin*dIII and *Hin*dIII sites. Their sizes (bp) are indicated in parentheses.

## 4. Discussion

In this study we have demonstrated that the phage T4 optional gene *seg*E encoding the site-specific endonuclease is inherited preferentially by progeny of mixed infection between the *seg*E-containing and *seg*E-lacking phages. This phenomenon depends on SegE endonuclease activity and occurs via NHRE resulting in the efficient transfer of the *seg*E$^+$ allele with its flanking sequences to the *seg*E$^-$ allele. It is interesting to note that SegE, as well as Endo.SceI [22] and I-HmuII [9], are able to cleave both their own and the heterologous DNA, although the enzymes demonstrate an obvious preference in cutting the latter, which appears to define the direction of transfer of genetic information. In the case of SegE this capability is conditioned by differences in nucleotide sequences of the sites recognised by the enzyme [12]. In contrast, other site-specific endonucleases involved in initiation of NHRE are able to cleave only DNA of recipient alleles [1]. There are many local genetic determinants of phage T4 that are preferentially inherited by progeny of mixed infection with a related phage [23]. Genetic studies suggested that different optional genes are required for the preferential inheritance of various T4 genetic determinants [23,24]. Recently, the involvement of the SegA endonuclease in preferential inheritance of genetic markers located at both sides of its gene by progeny of mixed infection with phage T2 has been reported (A. Belle and D.A. Shub, personal communication). There are some data indicating that the phage T4 site-specific endonucleases SegB and SegD can also initiate similar processes of non-reciprocal genetic exchange in mixed infections with a close relative (Kadyrov et al., unpublished data). The ability of intron-encoded endonucleases and related enzymes, HO, PI-SceI and Endo.-SceI, to initiate NHRE has been described in other organisms as well [1,9,22,25,26].

In T-even-related phages, NHRE leading to T4 *td* intron homing has been well characterised using varios model systems [4,7,8,27–30]. These studies have demonstrated that the I-TevI endonuclease initiates the process primarily via introduction of a double-strand break into the intronless allele. After the cleavage occurs, *td* intron homing proceeds through

multiple pathways and depends on the expression of T4 genes involved in phage recombination-dependent DNA replication. The general feature of the *td* intron homing pathways consists of using DNA of the intron-containing allele as the template for repairing cleaved and processed 3′ DNA ends of the intronless allele. There are a number of similarities between *seg*E gene transfer and *td* intron homing. First, the efficient inheritance of both the *td* intron and the *seg*E gene by the corresponding recipient allele absolutely depends on the expression of their own ORFs. When the ORFs were disrupted, the frequencies of transfer of both the *td* intron and the *seg*E gene from plasmids to phage genomes decreased by one to two orders of magnitude [4,15]. Second, in both cases, NHRE ensures transfer of flanking DNA sequences. The study of transfer frequencies of DNA markers closely linked to the *seg*E gene demonstrated that the efficiency of the process drops with increasing distance from the *seg*E gene. This is in agreement with the observed gradient in co-conversion efficiency of sequences of the phage T2 *td* gene accompanying *td* intron homing [7,8]. Third, the mobility of both the *td* intron and the *seg*E gene is stimulated by DNA homology between the donor and recipient alleles. However, it appears that for an efficient transfer the *td* intron and the *seg*E gene require different lengths of the DNA homology regions. The *td* intron homing occurred efficiently even when the length of DNA homology was decreased by up to 35–50 bp [30], while *seg*E gene mobility did not take place when the downstream DNA homology region was reduced to 73 bp. It is also known that expression of both the I-TevI ORF and the *seg*E gene is controlled by T4 late promoters [4,13]. Thus, although the detailed molecular mechanism by which *seg*E gene transfer occurs remains to be determined, the process promoted by the SegE endonuclease appears to be associated with the phage recombination-dependent DNA replication occurring at the late phase of the T4 life cycle [31]. All the observations summarised above indicate that homing of the *td* intron and the transfer of *seg*E gene are generally similar phenomena. This allows us to designate the *seg*E gene as a 'mobile endonuclease gene'. We speculate that the mobile *td* intron and the mobile *seg*E gene encoding endonucleases of the GIY-YIG family [14] arose from a common progenitor endonuclease-encoding DNA which was inserted into the *td* intron and the *hoc*.1-*uvs*W intergenic regions, respectively, during divergent evolution of T-even-related phages. This is consistent with the idea that ORFs of site-specific endonucleases rather than group I introns were the original mobile elements and the introns simply were optimal locations for the ORFs where their insertion has a minimal impact on the host genome [5,32,33].

It has been postulated that phage evolution proceeds by generation and subsequent selection of genetic variations [34,35]. Shuffling of genetic material by homologous recombination between closely related phages during their propagation in the same host cell is one of the potential sources of genetic variation [34,35]. In any single event, some genetic variation occurring among recombinant progeny could result in selective advantages and subsequent spread in nature. In accordance with this scenario, 'mobile endonucleases' of T-even-related phages appear to be responsible for an extensive non-reciprocal genetic exchange in mixed infections of close relatives leading to efficient generation of genetic variations. If it were generally true, then the distinctive ability of the T4 intron-encoded and Seg proteins to cleave highly polymorphic sequences would expand the potential number of initiating sites for non-reciprocal homologous recombination [12,36, 37]. This would contribute efficient generation of genetic variations in mixed infection of T4 with a related phage. In this context it will be interesting to determine whether non-reciprocal genetic exchange between T4 and related phages is initiated by SegE-mediated DNA cleavage at any other site.

*Acknowledgements:* We thank W.B. Wood and E.S. Miller for kindly providing the collection of RB phages, D.V. Mavrodi, V.I. Tanyashin, I.E. Granovsky, M.B. Evgen'ev, A.A. Zimin, A.V. Kaliman and T.V. Ivashina for their interest in this work and valuable discussions. This work was supported by a grant from the Russian Foundation for Basic Research (Grant 94-04-12669a).

## References

[1] Lambowitz, A.M. and Belfort, M. (1993) Annu. Rev. Biochem. 62, 587–622.
[2] Shinohara, T. and Ogawa, T. (1995) Trends Biochem. Sci. 20, 387–391.
[3] Jackson, S.P. and Jeggo, P.A. (1995) Trends Biochem. Sci. 20, 412–415.
[4] Quirk, S.M., Bell-Pedersen, D. and Belfort, M. (1989) Cell 56, 455–465.
[5] Bell-Pedersen, D., Quirk, S.M., Clyman, J. and Belfort, M. (1990) Nucleic Acids Res. 18, 3763–3770.
[6] Chu, F.K., Maley, G., Pedersen-Lane, J., Wang, A.-M. and Maley, F. (1990) Proc. Natl. Acad. Sci. USA 87, 3574–3578.
[7] Bell-Pedersen, D., Quirk, S.M., Aubrey, M. and Belfort, M. (1989) Gene 82, 119–126.
[8] Mueller, J.E., Smith, D. and Belfort, M. (1996) Genes Dev. 10, 2158–2166.
[9] Goodrich-Blair, H. and Shub, D.A. (1996) Cell 84, 211–221.
[10] Kadyrov, F.A., Kryukov, V.M., Shlyapnikov, M.G. and Bayev, A.A. (1994) Dokl. Acad. Nauk. (Moscow) 339, 404–406.
[11] Kadyrov, F.A., Kaliman, A.V. and Kryukov, V.M. (1996) Mol. Biol. (Moscow) 30, 1096–1106.
[12] Kryukov, V.M., Shlyapnikov, M.G. and Kadyrov, F.A. (1996) Mol. Biol. (Moscow) 30, 1307–1315.
[13] Kaliman, A.V., Khasanova, M.A., Kryukov, V.M., Tanyashin, V.I. and Bayev, A.A. (1990) Nucleic Acids Res. 18, 4277.
[14] Sharma, M., Ellis, R.L. and Hinton, D.M. (1992) Proc. Natl. Acad. Sci. USA 89, 6658–6662.
[15] Kadyrov, F.A. and Kryukov, V.M. (1996) Dokl. Acad. Nauk. (Moscow) 346, 700–704.
[16] Kaliman, A.V., Khasanova, M.A., Kryukov, V.M., Tanyashin, V.I. and Bayev, A.A. (1988) Dokl. Akad. Nauk. USSR (Moscow) 303, 1486–1489.
[17] Steinberg, C.M. and Edgar, R.S. (1962) Genetics 47, 187–208.
[18] Studier, F.W. and Moffatt, B.A. (1996) J. Mol. Biol. 189, 113–130.
[19] Sambrook, J., Fritch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd. edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[20] Mertz, L.M. and Rashtchian, A. (1994) Anal. Biochem. 221, 160–165.
[21] Yamagushi, Y. and Yanagida, M. (1980) J. Mol. Biol. 141, 175–193.
[22] Nakagawa, K., Morishima, N. and Shibata, T. (1992) EMBO J. 11, 2707–2715.
[23] Russel, R.L. and Huskey, R.J. (1974) Genetics 78, 989–1014.
[24] Pees, E. (1970) Mutat. Res. 9, 345–348.
[25] Kostriken, R., Strathern, J.N., Klar, A.J.S., Hicks, J.B. and Heffron, F. (1983) Cell 35, 167–174.
[26] Gimble, F.S. and Thorner, J. (1992) Nature 357, 301–306.
[27] Clyman, J. and Belfort, M. (1992) Genes Dev. 6, 1269–1279.
[28] Mueller, J.E., Clyman, J., Hyang, Yi.-J., Parker, M.M. and Belfort, M. (1996) Genes Dev. 10, 351–364.
[29] George, J.W. and Kreuzer, K.N. (1996) Genetics 143, 1507–1520.
[30] Parker, M.M., Court, D.A., Preiter, K. and Belfort, M. (1996) Genetics 143, 1057–1068.
[31] Mosig, G. (1983) in: Bacteriophage T4 (Mathews, C.K., Kutter,

E.M., Mosig, G. and Berget, P.B., Eds.), pp. 120–130, American Society for Microbiology, Washington, DC.
[32] Perlman, P.S. and Butow, R.A. (1989) Science 246, 1106–1109.
[33] Clyman, J., Quirk, S. and Belfort, M. (1994) in: Molecular Biology of Bacteriophage T4 (Karam, J.D., Ed.), pp. 83–88, American Society for Microbiology, Washington, DC.
[34] Campbell, A. and Botstein, D. (1983) in: Lambda II (Hendrix, R.W., Roberts, J.W., Stahl, F.W. and Weisberg, R.A., Eds.), pp. 365–380, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[35] Botstein, D. and Herskowitz, I. (1974) Nature 251, 584–589.
[36] Bryk, M., Quirk, S.M., Mueller, J.M., Loizos, N., Lawerence, C. and Belfort, M. (1993) EMBO J. 12, 2141–2149.
[37] Sharma, M. and Hinton, D.M. (1994) J. Bacteriol. 176, 6439–6448.