Volume 17 Number 1 1989 Nucleic Acids Research

The inconsistent distribution of introns in the T-even phages indicates recent genetic exchanges

S.M.Quirk[1], D.Bell-Pedersen[1,2], J.Tomaschewski[3], W.Rüger[3] and M.Belfort[1,4]

[1]New York State Department of Health, Wadsworth Center for Laboratories and Research, PO Box 509, Empire State Plaza, Albany, NY 12201 and [2]Department of Biological Sciences, State University of New York, Albany, NY 12222, USA

Received August 19, 1988; Revised and Accepted October 19, 1988

ABSTRACT

Group I self-splicing introns are present in the td, nrdB and sunY genes of bacteriophage T4. We previously reported that whereas the td intron is present in T2, T4 and T6, the nrdB intron is present in T4 only. These studies, which argue in favor of introns as mobile genetic elements, have been extended by defining the distribution of all three T4 introns in a more comprehensive collection of T2, T4 and T6 isolates. The three major findings are as follows: First, all three introns are inconsistently distributed throughout the T-even phage family. Second, different T2 isolates have different intron complements, with T2H and T2L having no detectable introns. Third, the intron open reading frames are inherited or lost as a unit with their respective flanking intron core elements. Furthermore, exon sequences flanking sites where introns are inserted in the T4 td, sunY and nrdB genes were determined for all the different T-even isolates studied. Six of eighteen residues surrounding the junction sequences are identical. In contrast, a comprehensive comparison of exon sequences in intron plus and intron minus variants of the sunY gene indicate that sequence changes are concentrated around the site of intron occurrence. This apparent paradox may be resolved by hypothesizing that the recombination events responsible for intron acquisition or loss require a consensus sequence, while these same events result in sequence heterogeneity around the site.

INTRODUCTION

In bacteriophage T4, group I self-splicing introns are present in the td gene encoding thymidylate synthase (1, 2), the nrdB gene encoding nucleoside diphosphate reductase subunit B (3, 4) and in the sunY gene (split gene, unknown function, why?; refs. 5, 6). All three introns contain long open reading frames (ORFs) looped out of the core structures of the intron (5). We previously reported that the closely related T-even phages T2 and T6 also contained the td intron but lacked an intron in the nrdB gene (7). These results indicated that the nrdB intron is not essential for nrdB gene expression or phage viability. In addition, these findings posed interesting questions about how introns were lost or acquired from the T-even genomes during the course of evolution. The heterogeneous distribution of the nrdB intron in the T-even phages is particularly interesting because it is consistent with evidence obtained in eukaryotic systems for introns as mobile genetic elements (see Discussion). We were therefore prompted to undertake a more comprehensive study to characterize the distribution of group I introns in the T-even phages.

The T-even phages were originally identified and classified by their host ranges on a variety of resistant and sensitive mutants of Escherichia coli B. (8). Three T2 strains, T2H, T2L, and T2K, were isolated independently by Drs. Hershey, Luria and Kalmanson from the PC stock obtained from the Bronfenbrenner Laboratory (9, Introduction to ref. 10). T2H and T2L, the most widely used T2 strains, are very similar but are not isogenic. The T4 phages, T4B and T4D, which are also similar but not isogenic, were isolated independently by Drs. Benzer and Doerman from the Luria-Delbrück collection (9, Introduction to ref. 10). To the best of our knowledge there are no recognized variants of T6.

In the present study we have examined the distribution of the three T4 introns in several different isolates of T2, T4 and T6. We report the inconsistent distribution of each of the three introns throughout the T-even phage family. In addition, the study revealed the surprising finding that some T2 isolates appear to lack group I introns and therefore differ from another T2 isolate studied in this laboratory and from other members of the T-even phage family. These results argue strongly against a critical role for group I introns in T-even phage development. Further, the observations suggest that recent genetic exchanges account for the heterogeneous occurrence of introns in genes that have extremely closely related exon sequences.

MATERIALS AND METHODS

Strains and Media

Wild-type E. coli B was used as the host for infection with bacteriophages T2, T4 and T6. E. coli strains B/2 and B/6 (from Dr. B. Selick) and B/4 (from Dr. D. Shub) are indicator bacteria resistant to infection by wild-type phages T2, T6 and T4, respectively.

The T-even phages routinely used by us at the Wadsworth Center, Albany, NY are denoted T2W, T4W and T6W in this report. They were obtained from Dr. M. VanVunakis via Dr. G. Maley. The T2 and T6 routinely used at Ruhr-Universität Bochum are referred to here as T2R and T6R. They were obtained from Dr. S. Champe via Dr. E. Bautz. The precise lineage of the phages designated "W" or "R" is not known. Two different strains of T2 originally isolated by Dr. A. Hershey (T2H) and Dr. S. Luria (T2L) were provided to us by Dr. J. Wiberg. The T4 strain originally isolated by Dr. S. Benzer (T4B) was provided to us by Dr. D. Hall. RB69, a phage which is thought to be closely related to the T-even phages (obtained from Dr. W. McClain) and λcΔ(att-red)h80 (obtained from Dr. D. Wulff) were used as controls in some experiments.

E. coli strains were grown in TBYE containing 1% Bacto-tryptone (Difco Laboratories), 0.5% NaCl and 0.5% yeast extract. Phages were plated on E. coli B strains in a top agar overlay (0.5% NaCl, 1.0% tryptone, 0.7% Bacto-agar) on TB plates containing 0.5% NaCl, 1.0% tryptone and 1.1% Bacto-agar.

Preparation of Phage RNA and DNA

RNA was extracted from T2- T4- or T6-infected E. coli B 9 minutes after infection using the lysozyme freeze-thaw method, followed by phenol extraction and ethanol precipitation as described previously (11).

DNA was prepared after centrifuging high titer lysates of the T-even phages (90 min at 27,000 g) to pellet phages. Following extraction with phenol (3 times) and chloroform/isoamyl alcohol (24:1), DNA was dialysed extensively against 10 mM Tris-HCl (pH 8.0), 0.1 mM EDTA.

Labeling of T-even phage RNA with [^{32}P]GTP

RNA from phage-infected E. coli B was incubated with [α-^{32}P]GTP (3000 Ci/mmol; 1 μCi per μg of RNA) at 42°C for 1 hr under self-splicing conditions (12) and the reaction products were separated on a 5% acrylamide-8M urea gel.

RNA and DNA blot hybridization

RNA from T-even phage-infected E. coli B was applied to Hybond-N filters (Amersham) under vacuum with a Minifold II Slot Blot apparatus (Schleicher and Schuell, Keene, N.H.; 1 μg RNA/slot). Filters were UV irradiated for 5 minutes to fix the RNA. To prepare DNA blots, dilutions of phage lysates (in 10 mM Tris-HCl, 1 mM $MgCl_2$) that would induce confluent lysis were spotted on a lawn of E. coli B and grown overnight at 37°C. Plaques were lifted onto Hybond-N filters which were treated successively with solutions of 1) 0.5 M NaOH-1.5 M NaCl, 2) 1 M Tris-HCl (pH 7.5)-1.5 M NaCl and 3) 2 X SSC, and UV irradiated for 5 minutes to fix DNA. Hybridization of RNA and DNA blots with [5'-^{32}P]-labeled synthetic oligonucleotides defining specific exon and intron regions of the nrdB, td and sunY genes of T4 (Table 1) was carried out as described previously (13). Filters were washed three times for 30 min each in 6 X SSC at 37°C or 46°C before autoradiography.

Southern hybridization analysis

Phage DNA (1 μg) was digested with NdeI and fractionated on 1% agarose gels (40 mM Tris-acetate [pH 8.0], 2 mM EDTA), which were stained with 1 μg/ml ethidium bromide and photographed under long-wave UV light. Transfer of DNA to Hybond-N filters and subsequent hybridization with [^{32}P]-labeled probes were performed as described previously (14). Following hybridization, filters were washed at room temperature. Filters were stripped of [^{32}P]-labeled probes by incubating sequentially in 0.4 M NaOH and 0.1 X SSC, 0.1% SDS, 0.2 M Tris-HCl (pH 7.5) for 30 min each at 45°C. Efficacy of probe removal was assessed by autoradiography of filters before rehybridization with additional probes.

RNA and DNA sequencing by primer extension

Dideoxy sequencing of RNA was carried out as described previously (11) with the modifications noted. Annealing of 40 μg total phage-infected E. coli B RNA with

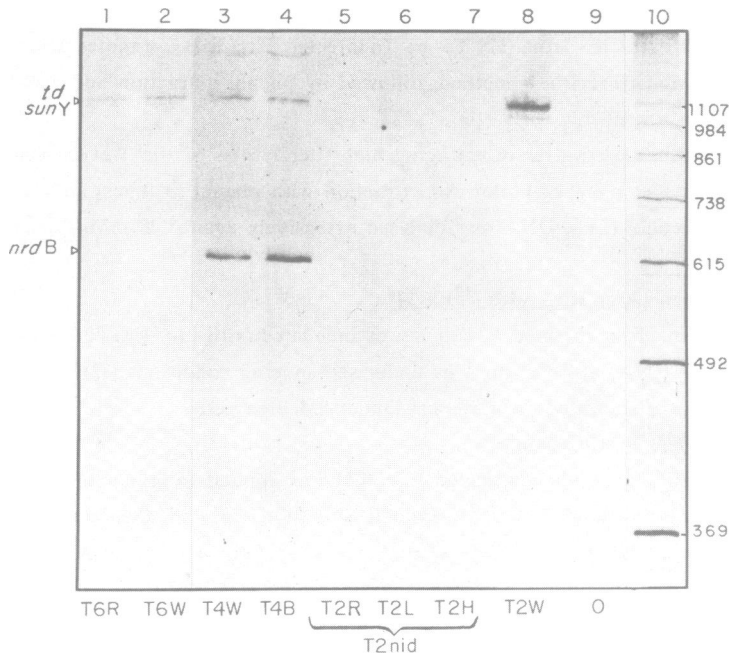

Figure 1. Labeling of T-even phage RNA with [α-^{32}P]GTP. RNA was extracted from E. coli B 9 min after infection with one of the T-even phages (lanes 1-8) or from uninfected E. coli B (lane 9). After incubating the deproteinized RNA with [α-^{32}P]GTP at 42°C for 1 h under self-splicing conditions, samples were separated on a 5% acrylamide 8M urea gel. A 123-bp DNA ladder was used as molecular size marker (lane 10). Positions of the nrdB intron (599 nt) and of the comigrating td and sunY introns (1017 and 1034 nt, respectively) are indicated. The unmarked low mol wt band in lanes 1 and 2 has not been identified and may represent another group I intron. The uppermost band in lanes 3 and 4 has not been unambiguously identified but preliminary evidence indicates that it is a splicing intermediate.

1.2 pmole of [5'-^{32}P]-labeled oligonucleotide primer was accomplished by heating to 60°C for 3 min followed by quick freezing. Primer extension with dNTPs, ddNTPs and AMV reverse transcriptase (Life Sciences; 0.4 U/reaction tube) was carried out at 48°C for 30 min. Primer extension products were then heated to 90°C and analyzed on an 8% acrylamide-8M urea sequencing gel. A similar procedure was used to sequence phage DNA except that annealing of 40 μg DraI-digested DNA to the oligonucleotide primer was at 90°C for 3 min and primer extension required more reverse transcriptase (1.5 U/reaction tube) and was carried out at 52°C for 30 min.

RESULTS

Distribution of self-splicing introns in the T-even phages

The initial event in splicing of group I introns involves covalent addition of a guanosine residue to the 5' end of the intron by a transesterification reaction. Group I

Table 1. Description of Oligonucleotides

Gene	Designation[1]	nt[2]	Specificity[3]	Complement of[4]
nrdB	a	22	Ex II	1864-1885
	d	20	In(ORF)	1439-1458
	g	23	In(S)	1273-1295
td	b	21	Ex I	369-389
	e	19	In(ORF)	1174-1189
	h	20	In(S)	1713-1732
	j	22	In(S)	817-838
	k	20	Ex II	1839-1858
	l	19	Ex II	1872-1890
sunY	c	21	Ex II	3959-3979
	f	21	In(ORF)	3067-3087
	i	22	In(S)	2834-2855
	m	22	Ex II	3865-3886

1 Letters correspond to the designation of probes in Figures 2, 3, and 4.
2 nt = number of nucleotides in oligomer.
3 In(S) = intron secondary structure; In(ORF) = intron open reading frame; Ex I = exon I; Ex II = exon II.
4 Numbering of residues in the nrdB, td and sunY genes that are complementary to oligomers (numbering according to ref. 1, 2, 3 and 6). The 5' and 3' splice sites are upstream of nucleotides 1201 and 1799 in nrdB, 769 and 1785 in td, and 2773 and 3806 in sunY.

introns have been detected in bacteriophage by labeling RNAs with [α-^{32}P]GTP under self-splicing conditions (4, 7, 15). This technique was used in the present study to characterize the distribution of self-splicing introns in eight T-even phage isolates (Fig. 1). Consistent with earlier observations (7), a labeled band corresponding in size to the nrdB intron (599 nt; refs. 3, 4) was present in RNAs from the T4 isolates (T4W and T4B) and was absent in RNAs from the T2 (T2R, T2L, T2H, T2W) and T6 (T6R, T6W) isolates. (Intron lengths correspond to the coded sequence plus one nucleotide, to account for the non-coded guanosine). Labeled bands corresponding in size to comigrating td and sunY introns (1017 and 1034 nt, respectively; refs. 1, 2, 6) were present in RNA from T2W and all of the T4 and T6 isolates but were absent from T2R, T2L and T2H. These results indicate that T2R, T2L and T2H differ from T2W in that they lack both the td and sunY introns. In addition, unlike their close relatives T2W, T4 and T6, these other T2 phages are devoid of detectable group I self-splicing introns. We therefore refer to T2R, T2L and T2H collectively as "T2nid" (no introns detectable).

The distribution of introns in the T-even phages was further defined by hybridization with [5'-^{32}P]-labeled oligonucleotide probes (described in Table 1) specific for exon and intron sequences of the T4 nrdB, td and sunY genes. Exon-specific probes gave strong hybridization signals to RNA and DNA blots of all T2, T4 and T6 isolates studied (Fig. 2). In contrast, probes specific for intron sequences showed a variable pattern of

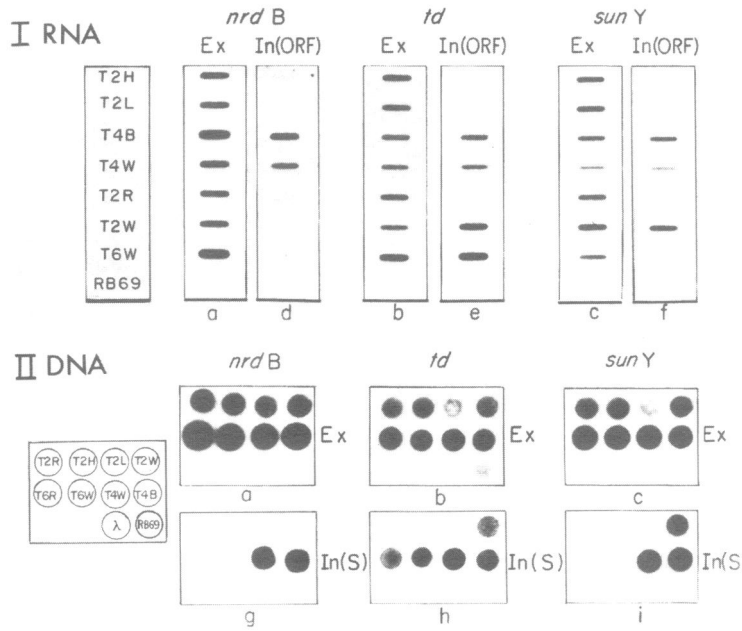

Figure 2. Hybridization analysis of T-even phage RNA (I) and DNA (II). I. Total cellular RNA extracted from E. coli B 9 min after infection with one of the T-even phages was deproteinized and applied to Hybond-N filters (1 μg/slot). II. Plaques or clearings were lifted onto Hybond-N filters, which were then processed to fix DNA. The arrays of phage RNA and DNA on the filters are indicated on the templates on the left of the figure. DNA and RNA blots were probed with [^{32}P]-labeled oligonucleotides defining specific exon and intron regions of the nrdB, td and sunY genes [Ex = exon, In(ORF) = intron open reading frame, In(S) = intron core structure; probes a-i are described in Table 1].

hybridization. In all cases, hybridization profiles with probes directed against the intron open reading frames (ORF) and the core secondary structure were similar. The nrdB intron probes hybridized to T4 RNA and DNA blots only, whereas the td intron probes hybridized to RNA and DNA blots of T2W and all of the T4 and T6 isolates studied but failed to hybridize to blots of the T2nid phages. The sunY intron probes hybridized to blots of T2W and to the T4 isolates but failed to hybridize to T6 isolates and the T2nid phages. The results of these hybridization studies are in accord with results of autocatalytic GTP-labeling experiments (above) and further indicate that the [^{32}P]GTP-labeled band at 1 kb (Fig. 1) corresponds to the td intron in the T6 isolates and to comigrating td and sunY introns in the T2W and T4 isolates. These data, which demonstrate the variable distribution of introns throughout the T-even phage family and among different isolates of T2, are summarized in Table 2. Notably, the T2nid phages are without detectable group I self-splicing introns.

Table 2. Distribution of introns in the T-even bacteriophages[a,b]

gene	T2W	T2(H,L,R) (= T2nid)	T4(W,B)	T6(W,R)
td	+	-	+	+
nrdB	-	-	+	-
sunY	+	-	+	-

[a] Nomenclature of phage strains is described in Materials and Methods.
[b] "+" = intron present, "-" = intron absent

The hybridization data were further corroborated by sequence analysis (see below). Sequencing not only confirmed the absence of introns from specific sites, but also demonstrated that when any particular intron is present in the genome, it always occurs at the same site in its respective gene.

Characterization of T-even phage isolates

In light of differences in the distribution of introns within the T-even phages and among different isolates of T2 we wished to assess the separate identities of the phages and to address the possibility that T2W is a simple hybrid between one of the T2nid phages and T4. As a first step, we tested the phages' ability to infect indicator bacterial strains B/2, B/4 and B/6, which are resistant to infection by T2, T4 and T6, respectively (Table 3). The two T4 isolates (T4W, T4B) and the two T6 isolates (T6W, T6R) failed to infect B/4 and B/6, respectively. Furthermore, the three T2nid isolates failed to infect B/2, while T2W plated with 10-fold reduced efficiency compared to its infection of wild type E. coli B. Although T2W differs from the T2nid phages in that it is not completely excluded by B/2, T2W resembles the T2nid phages and differs from T4 and T6 strains in its ability to infect B/4, B/6 and E. coli B with equal efficiency.

As another means of assessing differences between phage isolates, NdeI restriction of phage DNAs was carried out followed by Southern hybridization analysis. The NdeI restriction patterns of T2H and T2L (and T2R, not shown) were identical and differed from that of T2W by only a few fragments (notably fragments of about 4 and 9 kb; Fig. 3A). However, the restriction patterns of the T2nid phages and T2W differed markedly from that of T4B.

Restriction fragments were hybridized to [^{32}P]-labeled probes specific for exon and intron sequences of the T4 td and sunY genes. Representative data with exon I probes are shown in Fig. 3B. Results with other probes and restriction enzymes are summarized in the text. T4B DNA fragments hybridizing to td and sunY exon I probes (3.5 and 9.5 kb, respectively; Fig. 3) agree with those expected based on the NdeI restriction map of T4 (Fig. 3C; ref. 16). Whereas DNA fragments hybridizing to the td and sunY exon I

Table 3. Efficiency of plating of phages on indicator bacteria[a]

Bacterial host	T2(W)	T2(H,L,R) (=T2nid)	T4(W,B)	T6(W,R)
B/2	10^{-1}	$< 10^{-5}$	1	1
B/4	1	1	$< 10^{-5}$	1
B/6	1	1	1	$< 10^{-5}$

[a] Compared to an efficiency of plating on wild type E. coli B = 1.

probes were the same size in T2nid and T2W (4.4 and 7.4 kb, respectively; Fig. 3B), they were clearly different from the fragments obtained with T4B. These results indicate differences from T4 in the location of NdeI restriction sites in all of the T2 isolates, irrespective of the presence or absence of introns. As expected, fragments hybridizing to td and sunY intron probes were present in digests of T2W and T4B but not in T2H or T2L (data not shown). These results, and the results of similar hybridization analyses with DraI- and TaqI-digested phage DNA (data not shown), indicate that although T2W resembles T4 in the presence of introns in the td and sunY genes, it is closely related to the T2nid phages (and not to T4) in restriction enzyme sites in exon sequences both 5' and 3' to the two genes (see Discussion).

Exon sequences in td, nrdB and sunY genes of phages with and without introns

A comparison of exon sequences flanking introns in the td, nrdB and sunY genes and the corresponding residues in phages that lack introns indicates substantial similarity in the regions surrounding the junctions (Fig. 4). Out of 18 nucleotides immediately surrounding the junctions, 6 are invariant (shading) among different isolates of the T-even phages in the three different genes examined. The possibility that these invariant nucleotides may have a role in the genetic exchanges that have resulted in the acquisition or loss of introns at these specific sites in the T-even genomes is being considered.

Group I introns are characterized by a nucleotide sequence at their 5' ends referred to as the internal guide sequence (IGS), which is thought to pair with an external guide sequence (EGS) consisting of nucleotides at the 3' end of exon I (see Fig. 4). Phages which lack introns in the td, nrdB and sunY genes have nucleotide substitutions in regions corresponding to the EGS of exon I relative to phages which have introns in these genes (see consensus at bottom of Fig. 4). A guanosine residue located 2 nucleotides upstream of the splice junction in phages with introns in the nrdB and sunY genes is replaced by adenine or thymine in phages lacking introns (Fig. 4). Similarly, in phages lacking introns in the td gene, 3 out of 7 nucleotides corresponding to the EGS of exon I have been substituted. Speculations on the possible significance of these differences in sequence between phages with and without introns are entertained in the Discussion.

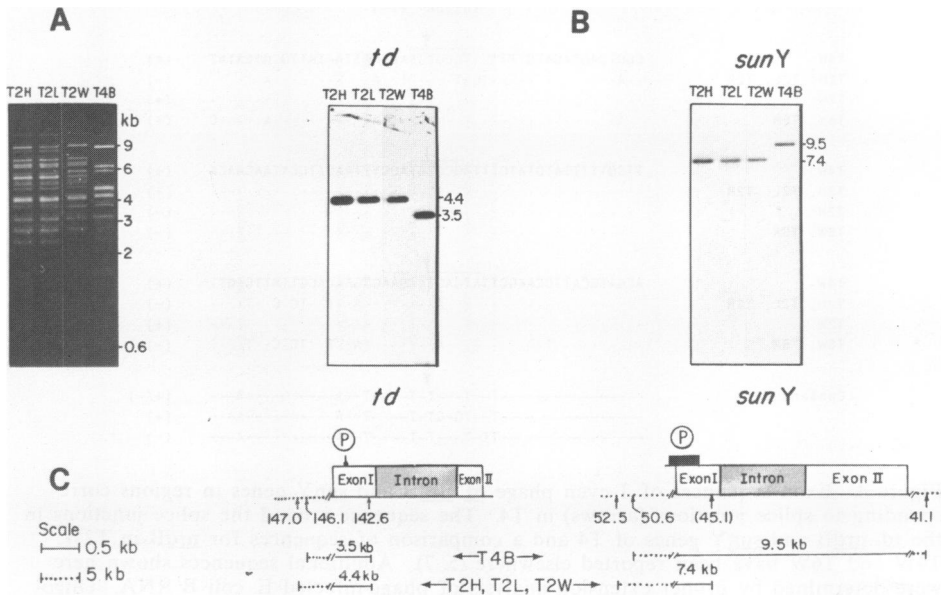

Figure 3. Southern hybridization analysis of NdeI-digested DNA from T-even phages. (A) NdeI-digested DNA from T-even phages was separated on a 1% agarose gel alongside molecular size standards and stained with 1 μg/ml ethidium bromide. (B) Fragments were transferred to a Hybond-N membrane, which was sequentially probed, stripped and re-probed with various [^{32}P]-labeled probes directed to td and sunY sequences of T4. (C) The T4 td and sunY genes are drawn to scale with positions of NdeI restriction sites corresponding to T4 map coordinates indicated below the maps (arrows; ref. 16) and hybridization probes (P) indicated above the maps. The td probe corresponds to b in Table 1, whereas the sunY probe is a random primer-labeled 605 bp SphI-HindIII exon I fragment. Experimentally determined fragments hybridizing to the probes are correlated with the restriction maps. An NdeI site that is present in the sunY gene of T2R but is absent from T4 is shown in parentheses at map position 45.1 (see Table 4, change at nucleotide position 2663). Restriction data indicate that the NdeI site at 45.1 is also present in T2H, T2L and T2W. NdeI sites in T4 at map positions 146.1 (td) and 50.6 (sunY) are apparently absent in T2W and all the T2nid phages.

The sequence of the sunY gene of T2R, one of the T2nid isolates, was determined and compared to the previously reported sequence of the sunY gene of T4 (ref. 6; Table 4). Within 22 nucleotides in the immediate vicinity of the junction there are 7 substitutions in T2R relative to T4 (32%) (Fig. 4, Table 4). In another 1001 exon nucleotides of the sunY gene in T2R that were sequenced there were only 24 nucleotide substitutions relative to T4 (2.4%) (Table 4). The concentration of nucleotide differences between T2R and T4 in the region immediately surrounding the junction suggests that nucleotide substitutions might occur more frequently at sites that either house or are capable of housing introns than in other regions of the exons.

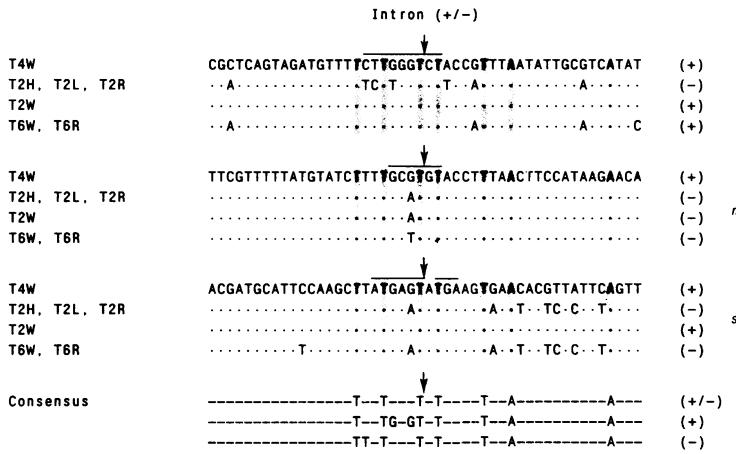

Figure 4. Exon sequences of T-even phage td, nrdB and sunY genes in regions corresponding to splice junctions (arrows) in T4. The sequence around the splice junctions in the td, nrdB and sunY genes of T4 and a comparison of sequences for nrdB in T2W, T4W and T6W have been reported elsewhere (5, 7). Additional sequences shown here were determined by primer extension analysis of phage-infected E. coli B RNA. Oligonucleotide probes used to prime cDNA synthesis are described in Table 1 (td - j, k and l; nrdB - a; sunY - i and m). In several instances where RNA secondary structure caused ambiguities in sequence, data were corroborated by sequencing phage DNA (td gene of T2H and the sunY gene of T2W). Dots indicate identity with the respective T4 sequence. Sequence identities among the various phage in the three different genes are shaded. The bottom of the figure shows the consensus sequence for all of the phage in the three different genes (intron +/-) and further defines the consensus sequence for phages with (+) and without (-) introns in the 3 different genes. Exon sequences thought to interact with the intron and align the splice sites are overscored. The alignment sequence in the 5' exon is referred to as the EGS (see text).

DISCUSSION

The present study demonstrates the variable occurrence of group I self-splicing introns among members of the T-even phage family. The results confirm our previous report that the nrdB gene of T4W is interrupted by an intron that is absent from T2W and T6W (7), and that an intron is uniformly present in the td gene of those three phages (7, 15). However, the td intron is absent from the T2nid phages, which are T2 isolates that appear to have no detectable group I introns. The previously identified intron in the sunY gene of T4 (5, 6) was shown to be present in T2W and absent in T6 phage isolates and the T2nid phages. The T2nid phages, therefore, differ from T2W and the other T-even phages studied in their apparent absence of introns.

The T-even phages are very closely related. They are morphologically indistinguishable, recombine in mixed infections (17) and are serologically related (18, 19). Heteroduplex analysis indicates 88% homology (20, 21) and sequence analysis of selected homologous genes have indicated up to 97% homology (22, 23). Nevertheless, the occurrence of

Table 4. Nucleotide changes in exon I and exon II of the sunY gene of T2R relative to T4[a]

Exon I			Exon II					
nucleotide position	T4	T2R	nucleotide position	T4	T2R	nucleotide position	T4	T2R
2264	A	G	3813	G	A	4039	C	T
2426	C	T	3816	C	T	4047	C	T
2427	G	T*	3819	G	T	4128	G	A
2429	G	C*	3820	T	C	4150	A	G‡
2525	G	A	3822	A	C	4194	C	T
2552	A	C	3825	C	T	4284	G	A
2585	G	A	3840	G	A	4285	C	T
2588	A	G	3843	T	G	4287	A	G
2624	G	C	3921	T	C	4290	T	C
2663	C	A	3936	T	C			
2771	G	A	4029	A	G			

[a] Numbering is according to the reported sequence of the T4 genome between map positions 48.266 kb and 39.16 kb, the region which includes the sunY gene (6). The nucleotide sequence of T2R corresponding to the last 537 exon I nucleotides (positions 2235 to 2772) and the first 512 exon II nucleotides of T4 (positions 3806 to 4318) was determined. The 7 nt changes that cluster around the junction are underscored (see also Fig. 4). Nucleotide changes in T2R relative to T4 resulted in two amino acid changes, from val to phe (*) and from ser to gly (‡).

introns varies dramatically among the T-even phages and among different isolates of T2. These differences are in contrast to the highly conserved nature of the phage genomes and of the coding sequences of the homologous genes that either do or do not harbor introns. These results suggest that the acquisition or loss of introns has occurred relatively recently since the divergence of these phages from a common ancestor.

Do these introns play a role in the phage life cycle? The fact that the occurrence of introns varies among the T-even phages and that some T2 isolates lack detectable group I introns indicates that introns are not essential for phage development. However, the possibility that the intron imparts a subtle selective advantage or disadvantage to the phage remains to be tested. One view holds that introns provide the potential for interesting regulatory scenarios. One consequence of introns in the T-even phages is an increase in the variety of proteins coded by the individual genes. For example, the presence of an intron in the td gene of T4 results in the expression of thymidylate synthase (TS) coded by the spliced mRNA as well as of the N-terminal polypeptide NH_2-TS (of unknown function) coded by exon I (24). Another consequence of the td intron is that ex-

pression of TS is delayed relative to that of an intronless gene by the time required for splicing (25). In addition, experiments using lacZ fusions to the ORFs in the td, nrdB and sunY introns have indicated that ORF proteins are expressed late in T4 infection (26). It is possible that these features of the split genes alter the developmental potential of a phage, particularly under limiting growth conditions.

The other view holds that the introns are selfish DNA elements that merely parasitize the phage. This view is based on the dispensibility of introns in the T-even phages isolated from the wild and T4 laboratory constructs containing precise deletions of the naturally occurring introns (unpublished results). In this case the introns would simply be exploiting the T-even phage genome as a vehicle for their perpetuation.

Hybridization analyses indicate that the ORFs contained within the core structure of the three T4 introns consistently occur as units with the intron core, implying coinheritance of the ORF and the core structure. This may, however, be an oversimplification, given the extensive homology shared between the core structure of the three T4 introns contrasted with the dissimilarity of the three ORFs (5). These observations suggest that the ORFs are indeed capable of genetic exchanges independent of the core structures. This situation would be analogous to that in Neurospora in which otherwise homologous introns in the mitochondrial ND1 gene of two closely related species contain completely dissimilar ORFs located at different positions within the intron core structural elements (27).

The fact that T2W differs from other T2 isolates in having introns in the nrdB and sunY genes suggested the possibility that T2W is a hybrid phage produced by homologous recombination between an intronless T2 phage and T4. Although we have not formally eliminated this possibility, the likelihood that T2W is a T2-T4 hybrid is low for the following reasons: First, the host range of T2W is T2-like and differs from that of T4 and T6. Second, the restriction patterns of T2W and T2nid phage DNA are similar (but not identical) and strikingly different from T4. Third, Southern hybridization data (Fig. 3) indicated that T2W and T2nid phages are similar and differ from T4 close to the td and sunY intron boundaries. Restriction site polymorphisms between the T2 phages and T4 occurred at 100 nt from the 5' end and 2.5 kb from the 3' end of the sunY intron and at 3.6 kb from the 5' end and 3 kb from the 3' end of the td intron (Fig. 3 and data not shown). Fourth, sequencing regions of the td and sunY genes indicated that T2W and T2nid phages were similar and differed from T4 at sites as close as 34 and 37 nt from the 3' end of the sunY intron and at 28 nt from the 3' end of the td intron (unpublished data). Therefore, it is unlikely that T2W acquired introns by simple homologous recombination with T4, since recombination would have had to occur within extremely narrow limits of the td and sunY introns and without major genetic recombination at other sites in the 160 kb genome.

One possible explanation for the seemingly haphazard distribution of introns in the

T-even phages is that introns may be mobile genetic elements, with the potential for insertion into or precise excision from the genome. This would also explain the interspecies variability of the Tetrahymena group I rRNA intron (28), as well as intraspecies non-uniformity of intron occurrence in fungal mitochondrial genes (29, 30). The best evidence for introns as mobile elements is the site specific transposition of a group I intron in the large rRNA of Saccharomyces cerevisiae mitochondria from omega$^+$ strains to omega$^-$ strains (29, 31, 32). Furthermore, a close relative of this intron is present in the ATPase subunit 9 gene of Kluyveromyces fragilis, suggesting that transposition of the intron to a different gene may have occurred (33). Of interest is also the resemblance between the T4 td intron ORF and the ORF of three group I introns in the mitochondria of filamentous fungi, suggesting the possibility of horizontal gene transfer between close ancestors of these fungi and the T-even phages (34).

Comparison of exon sequences surrounding sites where introns are inserted in the td, sunY and nrdB genes of the various phages revealed a high degree of sequence similarity, allowing a consensus sequence to be drawn for the three different genes (Fig. 4). These sequence similarities may be merely fortuitous, or they may be explained by the high A-T content of the phage genomes. Alternatively, this sequence conservation may reflect a preferred site where introns are incorporated into or excised from the T-even genomes.

A comparison of exon sequences in the sunY gene of T4 and one of the intronless T2 phages (T2R) indicated a clustering of nucleotide differences in the region surrounding the splice junction in T4 relative to exon sequences more distal to the junction. Similarly, in two different strains of S. cerevisiae which differ in the occurrence of an intron in the cox1 gene, sequence differences were localized to the splice junction (35). Likewise, in different strains of Schizosaccharomyces pombe variations in exon sequences between mosaic and unsplit forms of the cob gene were localized to the region of the splice junction (30). Therefore, in several evolutionarily distant organisms, variations in sequence between related strains that differ in their complement of introns occur more frequently around sites where introns are inserted into the genomes than in other regions of the exons. As suggested by Hensgens et al. (35), clustered nucleotide changes may result from intron insertion or loss at these sites, with possible mechanisms including gene conversion, non-homologous recombination or error-prone repair. Indeed, transposition of the r1 intron from omega$^+$ to omega$^-$ strains results in co-conversion of the flanking exon sequences (29, 36).

The variable occurrence of the T-even phage introns in three different highly conserved protein coding sequences provides strong circumstantial evidence for the innate mobility of these elements. Lateral movement of introns across species lines might then account for the extreme structural and functional relatedness of group I introns across such evolutionarily divergent genetic systems as protist nuclei, fungal mitochondria, plant chloroplasts and eubacterial viruses.

ACKNOWLEDGEMENTS

We thank members of our laboratory at the Wadsworth Center for their help during the course of this work. We also appreciate the help with random primer labeling which we received in Dr. Lorraine Flaherty's laboratory. We thank all those who provided us with phage and bacterial strains. This work was supported by National Institutes of Health Grant GM39422 and National Science Foundation Grant DMB8502961 to M.B.

[3]Ruhr-Universität Bochum, Arbeitsgruppe Molekulare Genetik, Postfach 102148, 4630 Bochum 1, FRG and [4]Department of Biomedical Sciences, School of Public Health, State University of New York and New York State Department of Health, Albany, NY 12201, USA

REFERENCES

1. Chu, F.K., Maley, G.F., Maley, F. and Belfort, M. (1984) Proc. Natl. Acad. Sci. USA **81**, 3049-3053.
2. Chu, F.K., Maley, G.F., West, D.K., Belfort, M. and Maley, F. (1986) Cell **45**, 157-166.
3. Sjöberg, B.-M., Hahne, S., Mathews, C.Z., Mathews, C.K., Rand, K.N. and Gait, M.J. (1986) EMBO J. **5**, 2031-2036.
4. Gott, J.M., Shub, D.A. and Belfort, M. (1986) Cell **47**, 81-87.
5. Shub, D.A., Gott, J.M., Xu, M.Q., Lang, B.F., Michel, F., Tomaschewski, J., Pedersen-Lane, J. and Belfort, M. (1988) Proc. Natl. Acad. Sci. USA **85**, 1151-1155.
6. Tomaschewski, J. and Rüger, W. (1987) Nucleic Acids Res. **15**, 3632-3633.
7. Pedersen-Lane, J. and Belfort, M. (1987) Science **237**, 182-184.
8. Demerec, M. and Fano, U. (1945) Genet. **30**, 119-136.
9. Doermann, A.H. (1983) In Mathews, C.K., Kutter, E.M., Mosig, G. and Berget, P.B. (eds), Bacteriophage T4, American Society for Microbiology, Washington, D.C., pp. 1-7.
10. Baylor, M.B. (1977) Virology **83**, 380-389.
11. Belfort, M., Pedersen-Lane, J., West, D., Ehrenman, K., Maley, G., Chu, F., and Maley, F. (1985) Cell **41**, 375-382.
12. Ehrenman, K., Pedersen-Lane, J., West, D., Herman, R., Maley, F. and Belfort, M. (1986) Proc. Natl. Acad. Sci. USA **83**, 5875-5879.
13. Hall, D.H., Povinelli, C.M., Ehrenman, K., Pedersen-Lane, J., Chu, F. and Belfort, M. (1987) Cell **48**, 63-71.
14. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, pp. 382-389, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
15. Chu, F.K., Maley, F., Martinez, J., and Maley, G.F. (1987) J. Bacteriol. **169**, 4368-4375.
16. Kutter, E., Guttman, B., Ruger, W., Tomaschewski, J. and Mosig, G. (1986) T4 News **1**, 3-22.
17. Delbrück, M. and Bailey, W.T. (1946) Cold Spring Harbor Symp. Quant. Biol. **11**, 33-37.
18. Delbrück, M. (1946) Biol. Rev. **21**, 30-40.
19. Lanni, F. and Lanni, Y.T. (1953) Cold Spring Harbor Symp. Quant. Biol. **18**, 159-168.
20. Kim, J. and Davidson, N. (1974) Virology **57**, 93-111.
21. Cowie, D.B., Avery, R.J. and Champe, S.P. (1971) Virology **45**, 30-37.
22. Paddock, G.V. and Abelson, J. (1975) J. Biol. Chem. **250**, 4207-4219.
23. Gram, H. and Rüger, W. (1986) Mol. Gen. Genet. **202**, 467-470.
24. Belfort, M., Pedersen-Lane, J., Ehrenman, K., Chu, F.K., Maley, G.F., Maley, F., McPheeters, D.S. and Gold, L. (1986) Gene **41**, 93-102.
25. West, D.K., Belfort, M., Maley, G.F. and Maley, F. (1986) J. Biol. Chem. **264**, 13446-13450.
26. Gott, J.M., Zeeh, A., Bell-Pedersen, D., Ehrenman, K., Belfort, M. and Shub, D. (1988) Genes and Devel. In press.
27. Mota, E.M. and Collins, R.A. (1988) Nature **332**, 654-656.

28. Sogin, M.L., Ingold, A., Karlok, M., Nielsen, H. and Engberg, J. (1986) EMBO J. **5**, 3625-3630.
29. Jacquier, A. and Dujon, B. (1985) Cell **41**, 383-394.
30. Zimmer, M., Welser, F., Oraler, G. and Wolf, K. (1987) Curr. Genet. **12**, 329-336.
31. Colleaux, L., d'Auriol, L., Betermier, M., Cottarel, G. Jacquier, A., Galibert, F. and Dujon, B. (1986) Cell **44**, 521-533.
32. Macreadie, I.G., Scott, R.M., Zinn, A.R. and Butow, R.A. (1985) Cell **41**, 395-402.
33. Dujon, B., Colleaux, L., Jacquier, A., Michel, F. and Monteilhet, C. (1986) In Wickner, R.B., Hinnebusch, A., Lambowitz, A.M., Gunsalus, I.C. and Hollaender, A. (eds), Extrachromosomal elements in lower eukaryotes, Plenum Press, New York, pp. 5-27.
34. Michel, F. and Dujon, B. (1986) Cell **46**, 323.
35. Hensgens, L.A.M., Bonen, L., de Haan, M., van der Horst, G. and Grivell, L.A. (1983) Cell **32**, 379-389.
36. Zinn, A.R. and Butow, R.A. (1985) Cell **40**, 887-895.