JOURNAL OF BACTERIOLOGY, Aug. 1992, p. 5371–5381
0021-9193/92/165371-11$02.00/0

Vol. 174, No. 16

Multiple Repetitive Elements and Organization of the *lux* Operons of Luminescent Terrestrial Bacteria

EDWARD A. MEIGHEN* AND ROSE B. SZITTNER

Department of Biochemistry, McGill University, Montreal, Quebec, Canada H3G1Y6

Received 9 March 1992/Accepted 15 June 1992

The complete nucleotide sequences of the *luxA* to *luxE* genes, as well as the flanking regions, were determined for the *lux* operons of two *Xenorhabdus luminescens* strains isolated from insects and humans. The nucleotide sequences of the corresponding *lux* genes (*luxCDABE*) were 85 to 90% identical but completely diverged 350 bp upstream of the first *lux* gene (*luxC*) and immediately downstream of the last *lux* gene (*luxE*). These results show that the *luxG* gene found immediately downstream of *luxE* in luminescent marine bacteria is missing at this location in terrestrial bacteria and raise the possibility that the *lux* operons are at different positions in the genomes of the *X. luminescens* strains. Four enteric repetitive intergenic consensus (ERIC) or intergenic repetitive unit (IRU) sequences of 126 bp were identified in the 7.7-kbp DNA fragment from the *X. luminescens* strain isolated from humans, providing the first example of multiple ERIC structures in the same operon including two ERIC structures at the same site. Only a single ERIC structure between *luxB* and *luxE* is present in the 7-kbp *lux* DNA from insects. Analysis of the genomic DNAs from five *X. luminescens* strains or isolates by polymerase chain reaction has demonstrated that an ERIC structure is between *luxB* and *luxE* in all of the strains, whereas only the strains isolated from humans had an ERIC structure between *luxD* and *luxA*. The results indicate that there has been insertion and/or deletion of multiple 126-bp repetitive elements in the *lux* operons of *X. luminescens* during evolution.

The *lux* genes of luminescent bacteria have been isolated from marine bacteria of the *Vibrio* and *Photobacterium* genera (3, 5, 6, 14, 15, 18) and from terrestrial bacteria of the *Xenorhabdus* genus (9, 26, 29). *lux* genes that code for luciferase (*luxAB*) and the fatty acid reductase complex (*luxCDE*) involved in the synthesis of the fatty aldehyde substrate for the luminescence reaction have been detected (17) in all cases. In certain *Photobacterium* species, an additional gene (*luxF*), homologous in sequence to the luciferase genes, is also present (1, 12, 23). A new gene, *luxG*, closely linked to *luxE*, has been discovered in *Vibrio* and *Photobacterium* species (14, 24, 25). The order of the *lux* genes in these species is *luxCDAB(F)EG*, with the luciferase genes flanked by the *luxCD* and *luxE* genes.

Significant differences do, however, exist between the *lux* systems of the marine luminescent bacteria owing to the presence of other *lux* genes. Upstream of *luxC* in *Vibrio fischeri* species are two regulatory genes, *luxI* and *luxR* (6, 7), whereas in *Vibrio harveyi* and *Photobacterium* species, an AT-rich region of more than 500 bp with no extended open reading frames is present (19, 20). Downstream of *luxG* in *V. fischeri* is the *lux* operon termination site, while the *V. harveyi* *lux* operon contains another *lux* gene, *luxH*, before termination (24, 25). The roles of the *luxG* and *luxH* gene products in luminescence are not known. Genetic experiments using transposon insertion have not provided any evidence concerning their possible involvement in the emission of light (16).

Recently, *lux* genes have been cloned and transferred into *Escherichia coli* from three strains (Hb, Hm, and Hw) of the enteric terrestrial bacterium *Xenorhabdus luminescens* (9, 26, 29). Two of the strains, Hb and Hm, are very similar and can be isolated from the gut tracts of nematodes or from the hemocoel of insects (21). The nucleotide sequences of the *luxA* and *luxB* genes from these two strains (13, 26), as well as the Hm *luxD* gene (2), have been determined, and the sequences of the luciferase genes are almost identical. A transcribed sequence capable of forming a strong hairpin loop (−20.5 kcal [1 cal = 4.184 J]) was identified downstream of *luxB* in the insect Hb DNA (26). This sequence has recently been recognized as part of a 126-bp enterobacterial repetitive intergenic consensus (ERIC) sequence or intergenic repetitive unit found in multiple copies in the genomes of a number of enterobacteria, including *E. coli* and *Salmonella typhimurium* (11, 22, 28). However, such a sequence has not been found in multiple copies in one operon or at one site.

The nucleotide sequences of the *luxCDAB* genes, but not those of the *luxE* gene or flanking regions, have also been reported for the Hw strain (29). This strain is of particular interest, since it was isolated from a human wound rather than from a nematode and/or insect (4, 8). Although the nucleotide sequences of the intergenic regions flanking the luciferase genes were determined, repetitive structures were not recognized.

In this report, the complete nucleotide sequences of a 7-kbp DNA fragment of the Hb strain and a 7.7-kbp DNA fragment of the Hw strain that contains the *luxCDABE* genes, as well as the flanking regions, are described, including corrections to the previous Hw *lux* DNA sequence. As a result, four ERIC sequences have been discovered in the Hw DNA, including three ERIC sequences flanking the luciferase genes within the *lux* operon. Comparison of the nucleotide sequences of the Hw and Hb DNAs not only demonstrates that the organization of *lux* operons of terrestrial bacteria differs from that of marine luminescent bacteria but suggests that the *lux* operons are at different genomic locations in the *X. luminescens* strains.

MATERIALS AND METHODS

Materials. Restriction enzymes were from Bethesda Research Laboratories. [^{35}S]dATP (1,400 Ci/mol) was from

* Corresponding author.

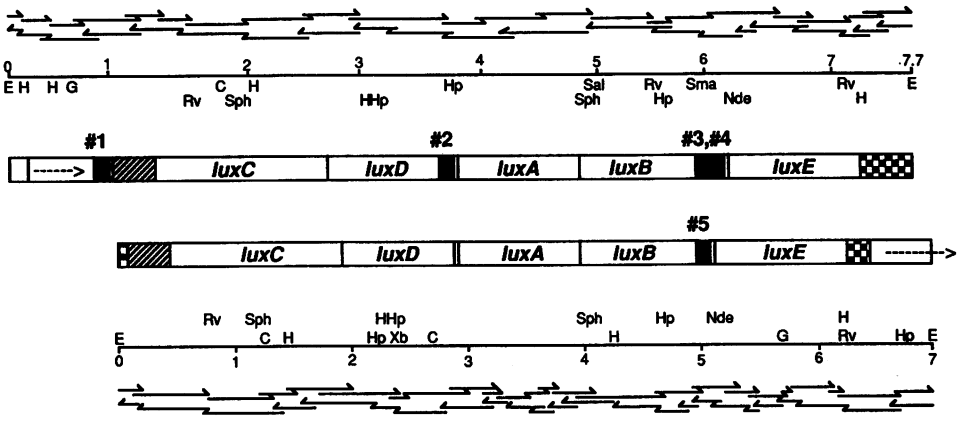


FIG. 1. Organization, restriction maps, and sequencing strategy for *Eco*RI DNA fragments containing the *lux* structural genes from *X. luminescens* Hw (upper) and Hb (lower). Open reading frames corresponding to the *lux* genes are labelled, and the direction of transcription is from left to right. Two additional open reading frames are indicated by the boxes with arrows. The numbered black areas correspond to the five repetitive ERIC sequences. The checkered areas upstream and downstream of the *lux* genes are noncoding regions that cannot be aligned, whereas the areas labelled with diagonal lines are noncoding regions with a high degree of sequence homology between the two *X. luminescens* DNAs. Common restriction sites between the two DNAs are given closest to the center of the diagram, whereas restriction sites unique to the Hw and Hb DNAs are given immediately adjacent to the respective linear scales of 0 to 7.7 kbp and 0 to 7 kbp for the respective DNAs. All restriction sites for *Bgl*II (G), *Cla*I (C), *Eco*RI (E), *Eco*Rv (Rv), *Hin*dIII (H), *Hpa*I (Hp), *Nde*I (Nde), *Sal*I (Sal), *Sma*I (Sma), *Sph*I (Sph), and *Xba*I (Xb) are given. There are no sites in either DNA fragment for *Bam*HI, *Nco*I, *Pst*I, *Sac*I, *Sna*B, *Stu*I, or *Xho*I. A single *Pvu*I site (not shown) is located between the *Eco*RI and *Hin*dIII sites at the start of the Hw DNA.

Dupont. *X. luminescens* Hb (ATCC 29999) and three other strains (ATCC 29304, ATCC 48949, and ATCC 48951) were obtained from the American Type Culture Collection, while strain Hw (ATCC 48950) was the kind gift of P. Colepicolo and J. W. Hastings (Harvard University). The *lux* genes of the Hw DNA were cloned as previously described for strain Hb (26). The genomic DNA was restricted with *Eco*RI, ligated into plasmid pT7-5, transformed into *E. coli* HB101, and screened for light.

DNA sequence analysis. Subfragments of the *lux* DNAs created by restriction digestion with four- and six-base cutters were inserted into M13mp18 or M13mp19 and analyzed by sequencing with dideoxynucleotides. In a few instances, synthetic oligonucleotides (McGill Biotechnology Centre) were used to complete missing regions. In this manner, the sequences of the entire lengths of both *Eco*RI fragments from strains Hb and Hw were determined in both directions with overlaps at each junction. Analysis of nucleotide and amino acid sequence data was performed by using the DNASIS and PROSIS programs provided by Hitachi Software Engineering Co. Ltd.

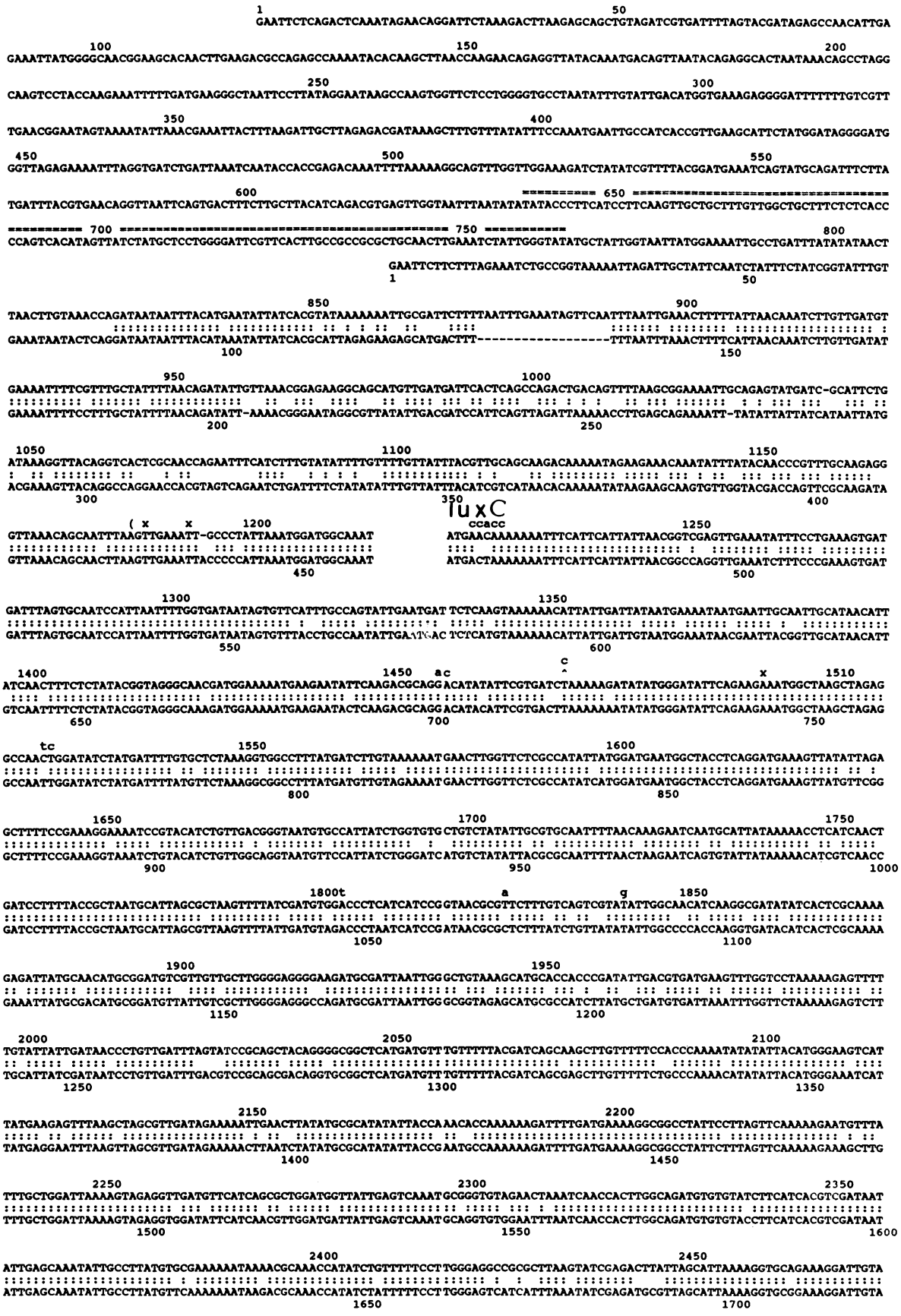
The sequence data for the Hw and Hb *lux* DNAs of the *X. luminescens* strains can be obtained under GenBank accession numbers M90092 and M90093, respectively.

Polymerase chain reaction (PCR). The PCR mixture (50 μl) contained 50 pmol of each primer, 300 ng of genomic DNA, a 1.25 mM concentration of each of the four deoxynucleoside triphosphates, and 0.25 U of Vent DNA polymerase (New England BioLabs) in the standard buffer supplied with the polymerase. PCR was performed in an automated thermal cycler with cycles of 95°C for 1 min for denaturation, 55°C for 1 min for annealing, and 72°C for 1 min for extension. After 35 cycles, an aliquot (10 μl) of the PCR mixture was resolved by electrophoresis on 3% agarose (Nu-Sieve) gels in Tris-borate-EDTA buffer containing 0.5 μg of ethidium bromide per ml.

Oligonucleotides for PCR. Synthetic oligonucleotides were synthesized by the McGill Biotechnology Centre. The ERIC primer was designed from the most highly conserved ERIC sequences within the Hw and Hb *lux* DNAs so that it would direct transcription in the coding direction of the *lux* operon. Three places of degeneracy in the sequence [GTGAC(T/C)GGGGTG(A/C)GTGAA(A/C)GCAGCCAAC] were introduced so that it would be identical to the ERIC structures in both the Hw and Hb *lux* DNAs immediately in front of *luxE* and *luxA*. The reverse *luxA* (CCAAACTCCGTGAAAT GATGCTC) and *luxE* (TGTGCCTGACAGTAGTGACG) primers were selected so that they were completely complementary to a region in close proximity to the 5′ terminus of the coding strands for these genes in both the Hw and Hb *lux* DNAs.

Expression of *lux* DNA in the pT7 plasmids. Plasmids pT7 (ampicillin resistance) and pGP-1-2 (kanamycin resistance) were the generous gifts of S. Tabor and C. Richardson (27). *E. coli* K38 cells were transformed with plasmid pGP-1-2, which codes for T7 RNA polymerase and plasmid pT7,

FIG. 2. Nucleotide sequences of the Hw and Hb DNAs containing the *lux* genes. The upper sequence, corresponding to the Hw DNA, extends from nucleotides 1 to 7669; the lower sequence, corresponding to the Hb DNA, extends from nucleotides 1 to 6960. Spaces have been left before and after the initiation and termination codons of the *lux* genes. The sequences have been aligned starting at nucleotide 82 and extending to nucleotide 6258 of the Hb DNA, and identical nucleotides are denoted by colons. A few gaps have been introduced for maximum alignment and are indicated by spaces for long stretches and hyphens for short stretches. Sequence homologies were not detected outside these regions. Differences previously reported for the Hw DNA between nucleotides 1182 and 6508 are indicated above the line in lowercase with additional nucleotides indicated by a cicumflex and missing nucleotides indicated by crosses. Frameshifts in coding regions in the Hw DNA occur between nucleotides 1471 and 1498 in *luxC* and from residue 4840 to the end of *luxA*. The sequences corresponding to the repetitive ERIC structures are indicated above and below the Hw and Hb DNAs, respectively, by double or single dashes.


luxC

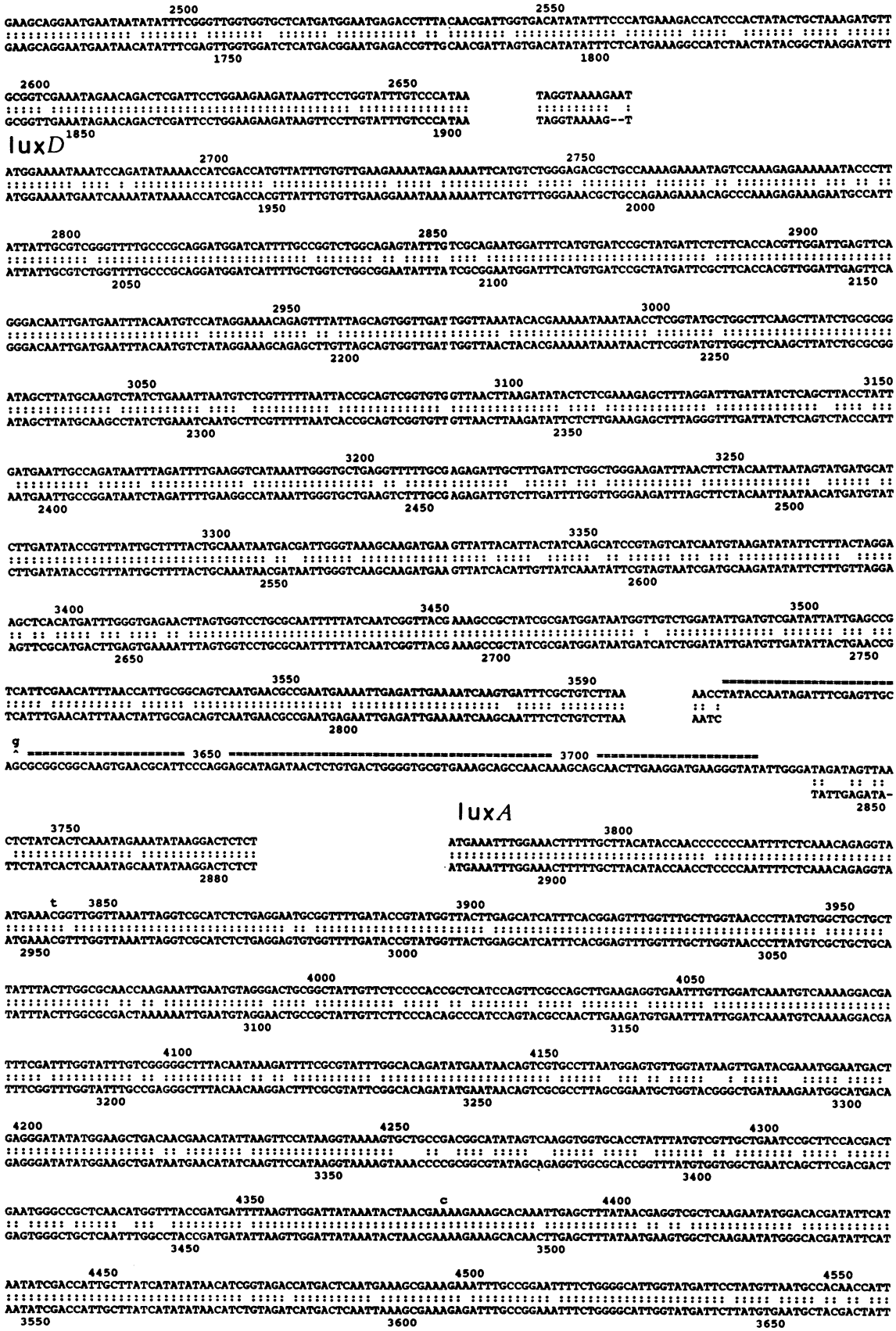


FIG. 2—*Continued.*

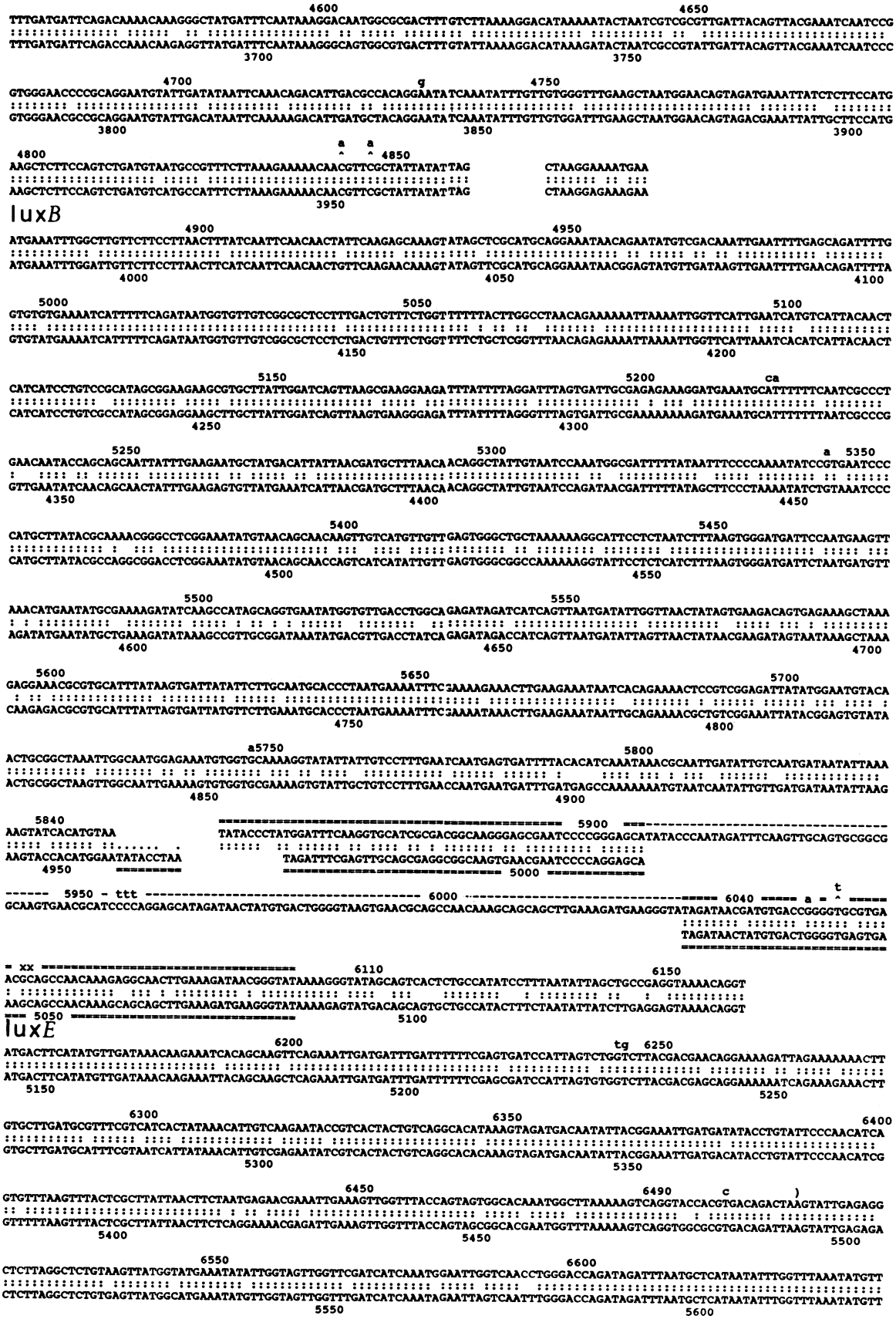


FIG. 2—*Continued.*

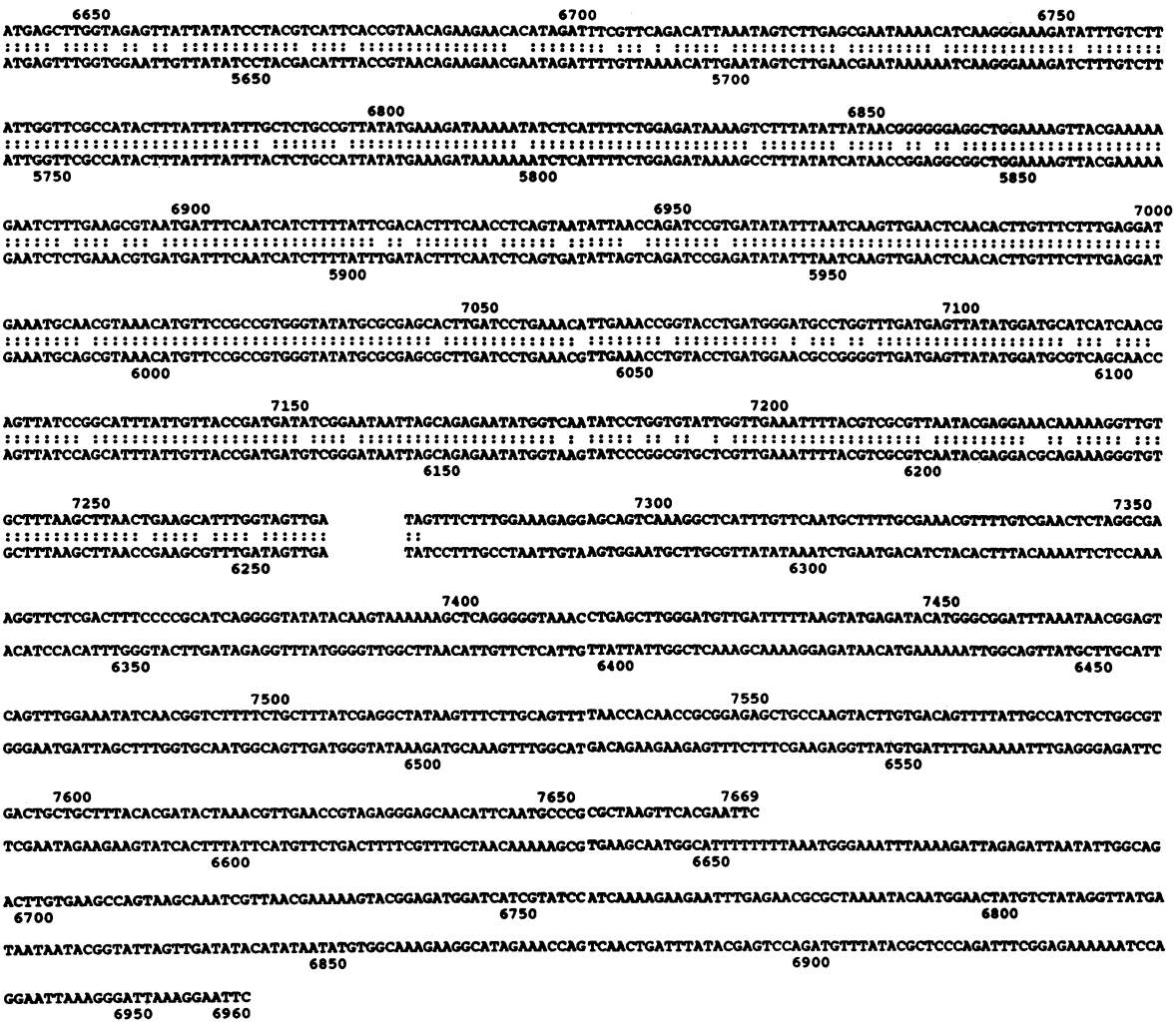


FIG. 2—*Continued.*

which contains the *X. luminescens* DNA inserted at the *Eco*RI site. Cells were grown in Luria-Bertani broth containing 50 μg of ampicillin per ml and 50 μg of kanamycin per ml at 37°C. Under these conditions, there is sufficient synthesis of T7 RNA polymerase that the *lux* DNA under control of a T7 promoter is fully expressed. In vivo luminescence (in light units per milliliter of culture) was determined as a function of cell density (optical density at 660 nm). One light unit corresponds to 2×10^{10} quanta/s on the basis of the light standard of Hastings and Weber (10).

TABLE 1. Comparison of the locations and nucleotide sequences of the *lux* genes of two *X. luminescens* strains

Gene	Initiation in strain:		Termination in strain:		No. of codons[a]	Identity (%)
	Hw	Hb	Hw	Hb		
luxC	1215	459	2655	1899	480	89.1
luxD	2671	1913	3592	2834	307	90.5
luxA	3776	2885	4856	3965	360	90.2
luxB	4873	3982	5845	4963	324 (327)	86.7
luxE	6160	5144	7270	6254	370	91.1

[a] *lux* genes in *X. luminescens* Hw and Hb have the same number of codons, except for *luxB*. Hb *luxB* has three additional codons compared with Hw *luxB*.

RESULTS

Figure 1 gives the sequencing strategy, restriction maps, and *lux* gene organization of the *Eco*RI restriction fragments containing the *luxCDABE* genes from strains Hb and Hw. Both DNA fragments were sequenced completely in both directions. A comparison of the nucleotide sequences is given in Fig. 2, with the Hw DNA extending about 800 bp farther upstream from the *lux* genes than the Hb DNA. A high degree of sequence homology between the two DNAs exists, starting at nucleotides 82 of the Hb DNA and 821 of the Hw DNA. Computer analysis revealed no sequence homology between the DNA extending upstream of these points. This homology extends downstream over 350 nucleotides to the start of *luxC* and continues through the *lux* genes to just after the termination codon of *luxE*. The extent of sequence identity in this upstream region is 72%, requiring only four single-nucleotide gaps and one 18-bp gap.

The positions of translation initiation and termination of the *lux* genes are summarized in Table 1. The five *lux* genes

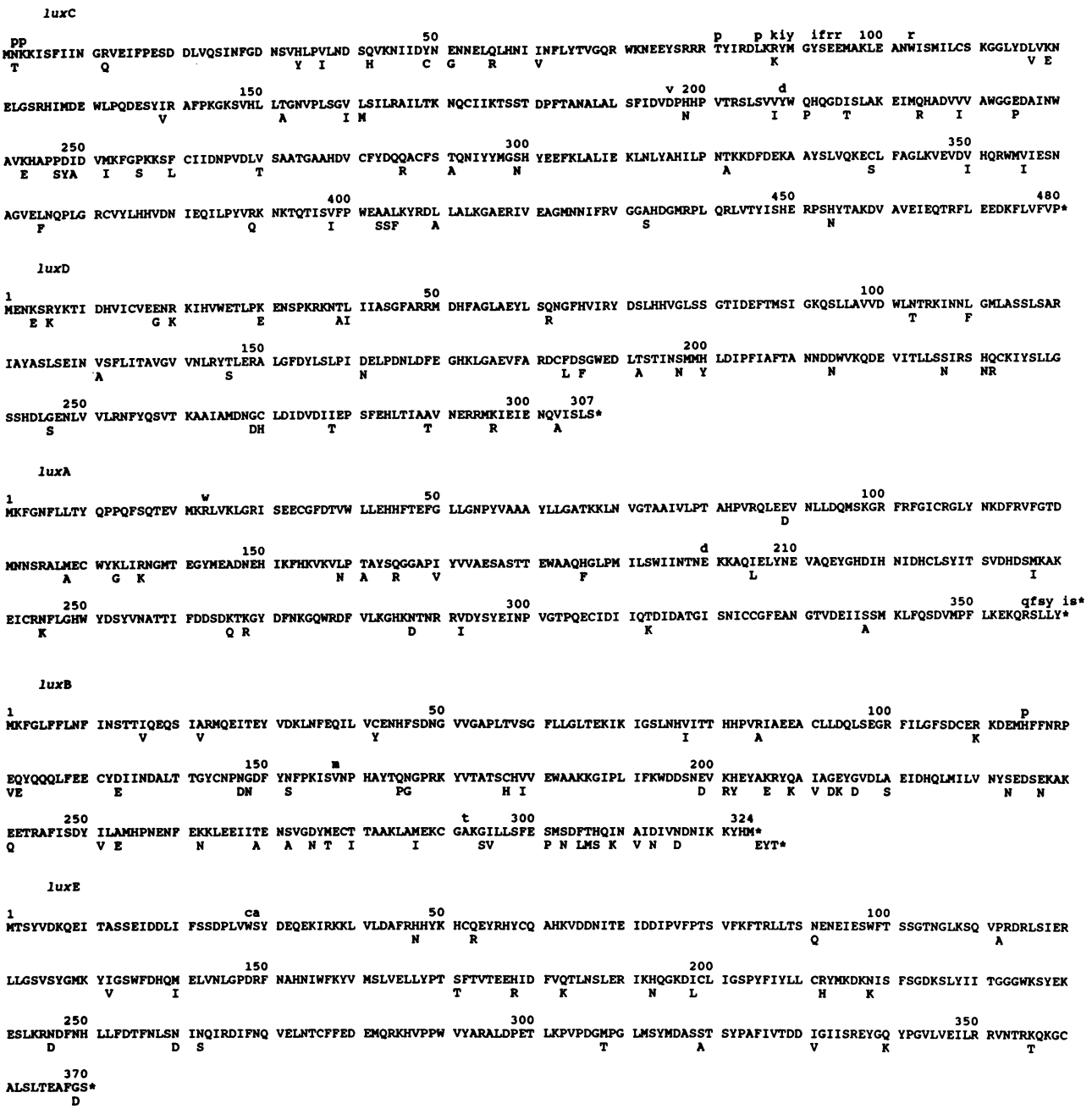


FIG. 3. Amino acid sequences encoded by the *luxCDABE* genes of the *X. luminescens* Hw and Hb DNAs. The amino acid sequences are given for the Hw *lux* genes, and any differences for the Hb *lux* genes are recorded below the line. Indicated in lowercase above the sequences are the amino acids previously reported for the proteins encoded by the Hw *lux* genes.

can be readily aligned in the order *luxCDABE*, providing that gaps of 126 bp are introduced into the Hb intergenic *luxDA* and *luxBE* regions (Fig. 2). In contrast, the DNA in the intergenic *luxCD* and *luxAB* regions can be aligned with only an additional 2-bp gap between *luxC* and *luxD* in the Hb DNA. The sequence identity for the *lux* genes is between 87 and 91% on comparison of the respective coding regions for the two strains (Table 1).

Indicated also in Fig. 2, between nucleotides 1182 and 6508 in the Hw DNA sequence, are the different nucleotides previously reported by other workers (29). Thirty-six differences were detected over this range in the present report, including 10 additions or deletions of nucleotides. Corrections in the nucleotide sequence of the Hw DNA are localized to the ends of the DNA with 15 changes in the 5′ end of *luxC* and the remaining corrections after the *luxD* sequence. Four of these corrections were in coding regions resulting in frameshifts. In the *luxC* gene, deletion of one nucleotide and addition of another nucleotide further downstream caused a localized frameshift extending over 9 codons, whereas two deletions near the end of the *luxA* gene caused a shift in the sequence at the carboxyl terminal of the

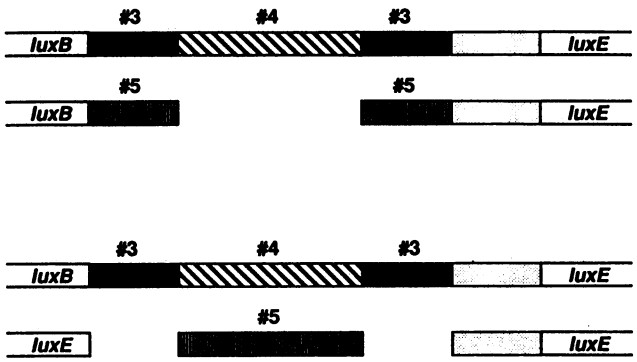


FIG. 4. Alternate alignments of the ERIC sequences between *luxB* and *luxE* in the *lux* operons of *X. luminescens* Hw and Hb. The upper gene organization corresponds to that for the Hw strain in both sets of alignments. The ERIC sequences (3 to 5) are numbered on the basis of Fig. 1, with the direction of transcription from left to right. The unlabelled region in front of *luxE* corresponds to an intergenic region of approximately 60 bp with a high degree of homology between the two strains.

α subunit of luciferase. Both corrections result in a significant improvement in alignment and homology of the amino acid sequences coded for by the Hw *luxC* and *luxA* genes with the corresponding sequences of the Hb strain (Fig. 3). The amino acid sequences of the α and β subunits of Hw luciferase differ by 5 and 15%, respectively, from the sequences of the same luciferase subunit encoded by the Hb DNA. The reductase (*luxC*), transferase (*luxD*), and synthetase (*luxE*) polypeptides of the fatty acid reductase differ by 10, 10, and 5%, respectively, on comparison of the amino acid sequences of the respective subunits of the two strains (Fig. 3).

Intergenic regions and ERIC sequences. One of the most striking features of the *X. luminescens* sequences containing the *lux* genes is the presence of multiple copies of a 126-bp repetitive element (ERIC or intergenic repetitive unit). A total of five ERIC sequences, labelled 1 to 5 (Fig. 1), were detected in the *X. luminescens* DNA; four of these repetitive elements are in the Hw DNA, and one is in the Hb DNA. Only one of these ERIC structures (no. 1) is not present in the *lux* intergenic region but is located about 350 nucleotides upstream of the Hw *luxC* gene (Fig. 2). This structure occurs immediately after the end of an open reading frame (Fig. 1).

```
  1                                                                  63
c TATACCCaAAATAATTCGaGTTGCaGcAAGGC GGCAAGtGAGTGAATCCCCAGGAGCTTACAT
1 -------TTC--CC---A------T--TTT-TT---TGC-TTC-CTCA------TCA-A--GT-
2 ------ --T-G-T-------------GC--- ---------AC-C--T---------A--G--
3 -------T-TGG-T---A--G----T-GC-A- ------G---C--------G-----A--G--
4 ---------T-G-T---A--------TGC--- ---------AC-C------------A--G--
5 ------T--T-G-T-------------G---- ---------AC-------------A--G--

                                                                  126
c AAGTAAGTGACTGGGGTGAGCGAACGcAGCCAACGCAgctGCAgCtTGAAatATGAcGgGTAT
1 -TC--T-CTC------ATTCGTTCACTT---GC--- ------A--------C-ATT------
2 --C-CT------------C-T---A---------AA---A---A------GG----A------
3 --CG-T-----C------C-T-------------AA--AG---A-------G--A--------
4 --C--T-----------A--T-------------AA---A----------G----A-----
5 --C--T-------------T---A---------AA---A----------G----A------
```

FIG. 5. Alignment of the *X. luminescens* ERIC sequences with the consensus ERIC sequence for enterobacteria. The consensus sequence (c) comprises 111 sites where the base is statistically significantly different (uppercase) and 15 sites where the base is found in the majority of cases (lowercase) in *E. coli* and *S. typhimurium* (23). Horizontal bars in ERIC structures 1 to 5 indicate that the same nucleotide is present as in the consensus sequence. A gap in the consensus sequence has been left so that ERIC 1 can be aligned, as well as single-base gaps in ERICs 1 and 2.

TABLE 2. Differences in *X. luminescens* ERIC sequences[a]

Sequence	No. of changes in comparison with ERIC:				
	1	2	3	4	5
c	54	25	29	21	18
ERICs					
1		58	62	56	56
2			21	10	8
3				19	20
4					7

[a] Includes deletions and additions as well as substitutions. c, consensus sequence for enterobacteria. Numbers refer to *X. luminescens* ERIC sequences as given in Fig. 1 and 4.

A comparable ERIC sequence cannot be recognized at this site in the Hb DNA.

A second ERIC structure (no. 2) is present between *luxD* and *luxA* in the Hw DNA, whereas this sequence is absent in the intergenic region in the Hb DNA. Indeed, the difference in the *luxDA* intergenic regions is almost completely accounted for by the presence of the second (no. 2) ERIC sequence in the Hw DNA.

The remaining three ERIC structures are located between the *luxB* and *luxE* genes. In the Hb strain of *X. luminescens*, one ERIC sequence of 126 bp is present (no. 5), whereas in the Hw strain, sequences corresponding to two ERIC sequences (no. 3 and 4) can be recognized (Fig. 1 and 2). In the latter case, it appears that an ERIC sequence (no. 4) has been inserted almost exactly at the middle of a second ERIC structure (no. 3) (Fig. 2).

Two alignments of the intergenic regions between *luxB* and *luxE* for the Hw and Hb strains are possible (Fig. 4). In the first alignment, ERIC sequence no. 5 of the Hb DNA can be split so that it aligns with split ERIC sequence no. 3 of the Hw DNA with a 125-bp gap corresponding to the extra ERIC sequence (no. 4) in the Hw DNA. In the second alignment, the Hb ERIC sequence (no. 5) is aligned with ERIC sequence no. 4 of the Hw DNA. In this case, two gaps must be introduced in the Hb DNA corresponding to the flanking ERIC sequence (no. 3) in the Hw DNA.

Comparison of the nucleotide sequences of ERIC structures 1 to 5 to the consensus sequence for enterobacteria shows a very high degree of conservation, particularly for ERIC structures 2 to 5 inside the *X. luminescens lux* operons (Fig. 5). The sequences can be aligned with the consensus sequence (23) over the entire 126-bp region with only one additional nucleotide in ERIC structure 1, a single-nucleotide gap in each of ERIC structures 1 and 2, and a missing nucleotide at the end of ERIC structure 4.

The numbers of nucleotide differences between the con-

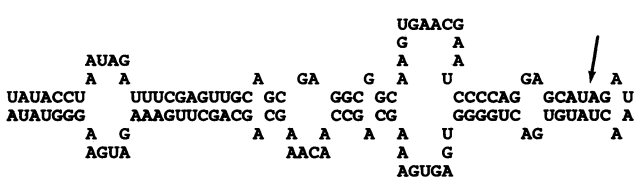


FIG. 6. Potential mRNA stem-loop structure formed from the ERIC sequence (no. 5) between *luxB* and *luxE* in *X. luminescens* Hb. The structure with maximum stability (−48.6 kcal/mol) in this region was calculated by using the DNASIS programs provided by Hitachi Software Engineering Co. Ltd.

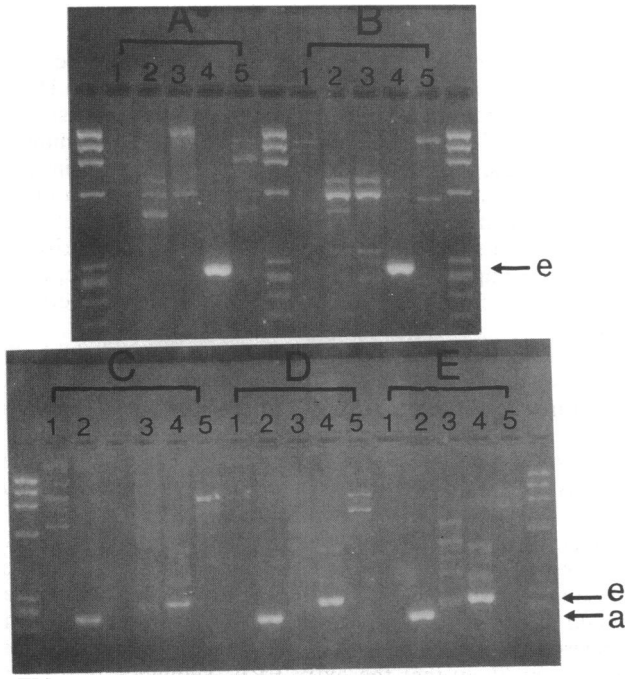


FIG. 7. Agarose gel electrophoresis of PCR products obtained from genomic DNAs from *X. luminescens* strains amplified with ERIC and *lux* gene primers. The genomic DNAs from five isolates or strains (A, ATCC 29999, Hb; B, ATCC 29304; C, ATCC 43950, Hw; D, ATCC 43949; E, ATCC 43951) were subjected to PCR. The sequences of the oligonucleotides, referred to as the *luxA*, ERIC, and *luxE* primers, are given in Materials and Methods. The primers used were *luxA* (lanes 1), *luxA* and ERIC (lanes 2), ERIC (lane 3), *luxE* and ERIC (lanes 4), and *luxE* (lanes 5). The arrows indicate the positions of the PCR products for an ERIC structure upstream of *luxA* (a) and *luxE* (e). Both products were observed in DNAs from *X. luminescens* strains isolated from humans (C, D, and E), whereas only the latter product was present in DNAs from *X. luminescens* strains isolated from insects (A and B). Unlabelled lanes contained molecular weight standards (1.4 and 1.1 kbp and 872, 603, 310, 281, 271, 234, and 194 bp), except for a blank lane between lanes 2 and 3 in sample C.

sensus sequence and the *X. luminescens* ERIC structures are given in Table 2. The differences from the consensus sequence for the ERIC structures in the *lux* genes are between 18 and 29 nucleotides out of 126 bp (14 to 23% difference), a relationship as close to the consensus as ERIC sequences from *E. coli* and *S. typhimurium* (11). The ERIC sequence upstream of Hw *luxC* (no. 1) has diverged the furthest, with 54 differences from the consensus sequence. Moreover, all four of the ERIC structures (no. 2 to 5) in the *lux* DNA have even greater differences from ERIC structure 1 (56 to 62 changes) than from the consensus sequence.

The data in Table 2 also indicate that three of the ERIC sequences (no. 2, 4, and 5) are much more closely related to each other (7 to 10 differences) than to the split ERIC structure (no. 3) in the Hw DNA (19 to 21 changes). Although these data indicate that the alignment of ERIC structures 4 and 5 in the *luxBE* intergenic region may be the preferred alignment (Fig. 4), it should be noted that this alignment requires two insertions or deletions. In contrast, alignment of ERIC structures 3 and 5 could be created by a simple insertion (or deletion) of an ERIC structure. Moreover, the remaining intergenic DNA between the *luxB* and *luxE* genes of the Hb and Hw strains can be aligned without any additions or deletions. It is possible that once a second ERIC structure (no. 4) was inserted into ERIC structure 3, a much faster divergence of the nucleotide sequence of the first ERIC occurred.

As all ERICs contain a central core with an inverted repeat (23), it is possible to generate a stem-loop structure of very high stability from the transcribed mRNA. Figure 6 shows a potential stem-loop structure for ERIC no. 5 generated by the Hitachi DNASIS program with a stabilization energy of −48.6 kcal/mol. The arrow at the start of the loop indicates the corresponding position in ERIC sequence no. 3 of the Hw DNA where insertion of a second ERIC sequence (no. 4) would have given rise to the double ERIC sequences in the Hw DNA.

To determine whether ERIC structures are present in front of *luxA* and/or *luxE* in other *X. luminescens* strains, the genomic DNAs from two other isolates from humans (ATCC 48949 and ATCC 48951) and one from insects (ATCC 29304), as well as the Hb and Hw DNAs, were analyzed by PCR. An oligonucleotide (28-mer) starting at residue 70 in the *lux* ERIC sequences (Fig. 5) with three places of degeneracy so that it would be identical to the ERIC structures immediately in front of *luxA* and/or *luxE* (no. 2, 4, and 5) was used in conjunction with the *luxA* or *luxE* primer in the reverse direction. The predicted PCR products would be 258 bp if an ERIC structure were located in front of the *luxA* genes, as in the Hw DNA, and 303 bp if an ERIC structure were located in front of the *luxE* gene, as in the Hb and Hw DNAs.

Figure 7 shows the electrophoresis of the PCR products from the five strains when either a single oligonucleotide (*luxA*, ERIC, or *luxE* primer; lanes 1, 3, and 5, respectively) was added to the genomic DNA or the ERIC primer was added with the *luxA* (lane 2) or *luxE* (lane 4) primer. For *X. luminescens* strains isolated from insects (samples A and B), one major band corresponding to the predicted 303-bp product (labelled e) for an ERIC structure in front of *luxE* was detected. For *X. luminescens* strains isolated from humans (samples C, D, and E), two major bands (labelled a and e) corresponding to the predicted 258- and 303-bp products for ERIC structures in front of *luxA* and *luxE*, respectively, were detected. These same PCR products were observed when the cloned *lux* DNA from the Hw strain was used instead of the genomic DNA. Similarly, only the 303-bp product was observed when the cloned *lux* DNA from the Hb strain was used in the PCR (data not shown). Although a number of other bands can be detected in the different PCRs with the genomic DNA, these products are significantly smaller in amount and arise primarily when a single oligonucleotide is added to the reaction mixture.

DISCUSSION

The present results have demonstrated that differences exist between the *lux* operons of an *X. luminescens* strain isolated from an insect compared with those of a strain isolated from a human wound. Although a 10 to 15% variation in the nucleotide sequence of the *lux* genes within the same species might not seem that surprising, it was unexpected to find the Hw and Hb sequences diverged upstream of *luxC* and immediately downstream of *luxE*. These results suggest that the *lux* operons are at different positions in the genomes of these two *X. luminescens* strains and demonstrate that the *luxG* gene located immediately downstream of *luxE* in marine luminescent bacteria is absent, at least at this location, in terrestrial bacteria.

The DNA extending upstream of the *luxC* gene in the Hw

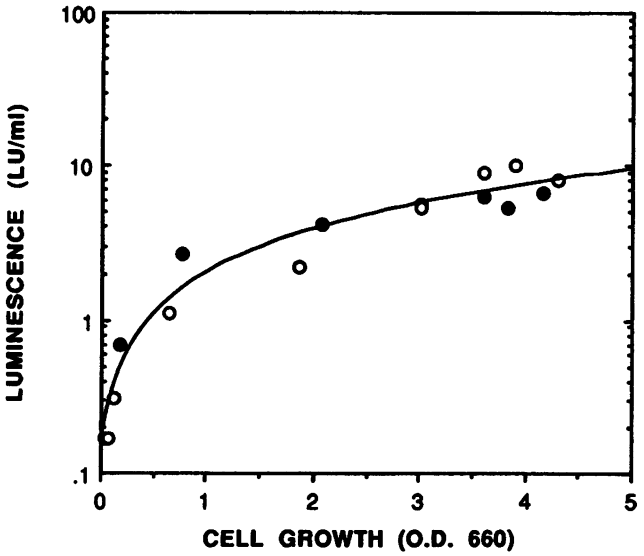


FIG. 8. Expression of the *Eco*RI *lux* DNA fragments from *X. luminescens* Hb and Hw in *E. coli*. *lux* DNA was inserted into plasmid pT7 and expressed in *E. coli* K38 with the assistance of plasmid pGp1, which codes for T7 RNA polymerase. Additional details are given in Materials and Methods. Luminescence (in light units per milliliter) is plotted against optical density at 660 nm for *E. coli* containing the Hw (●) and Hb (○) DNAs.

and Hb strains could be aligned with 72% nucleotide sequence identity for over 350 bp before the sequences diverged. This conserved DNA would be expected to contain the promoter and operator regions for the *X. luminescens lux* operon. Within this noncoding region, specific areas were highly conserved, including a 73-bp region with 94% nucleotide sequence identity over 250 bp from the start of *luxC*. These highly conserved sequences may represent DNA-binding sites for proteins that control the expression of the *lux* system. No sequence homology exists between the upstream *X. luminescens* DNA and that from other luminescent bacteria, showing that factors capable of recognizing the regulatory regions in marine and terrestrial bacteria may be different.

The unique feature of the *lux* DNA from *X. luminescens* was the presence of multiple ERIC structures. Although a single ERIC sequence is present in the Hb *lux* DNA (no. 5), it is the presence of multiple ERIC sequences in the Hw *lux* DNA that makes analyses of the ERIC structures of particular relevance. These repetitive structures have been recognized in *E. coli*, *S. typhimurium*, and a number of other gram-negative bacteria (23–25). In the *Xenorhabdus lux* systems, four of the ERIC structures are closely related to the consensus sequence with between 76 and 85% identity. A fifth ERIC sequence, upstream of *luxC* by about 400 bp in the Hw strain, is quite different, having only 60% identity with the consensus sequence.

The ERIC structures in *X. luminescens* differ in several properties from those reported for *E. coli* and *S. typhimurium* (23). In particular, ERIC sequences in other bacteria have been found to occur only singly at different genomic locations, whereas there are multiple ERIC sequences within the Hw *lux* operon with two ERIC sequences at the same site between the *luxB* and *luxE* genes. Moreover, the coding region for the last two amino acids and the termination codon of the β subunit of Hb luciferase are part of the single ERIC structure between *luxB* and *luxE*. Although this result does show that ERIC sequences can occur in coding regions, it should be noted that an ERIC structure is at exactly the same nucleotide position in the Hw DNA. However, since the Hw *luxB* gene is three codons shorter than the Hb *luxB* gene, the ERIC structure starts immediately after the termination codon of *luxB* in this strain. It is likely that a mutation in the *luxB* termination codon in the Hb DNA occurred after introduction of the ERIC sequence, resulting in a slightly larger *luxB* gene. This result provides the first example of translation of an ERIC sequence, as all other ERICs have been found exclusively in the intergenic regions.

The differences in the number of ERIC structures in the *X. luminescens* strains plus the presence of multiple ERIC structures at one site raise some interesting questions about whether the differences in the number of ERIC sequences in the two *X. luminescens* strains arose by insertion or deletion. As the start of the ERIC structures at the end of *luxB* are at identical nucleotide positions in the Hw and Hb DNAs, the first ERIC sequence at this position was likely inserted prior to divergence of the luminescence systems of these two strains. Subsequently, the other ERIC structures would have been added to the Hw DNA or deleted from the Hb DNA to give the present distribution of ERIC structures. It appears likely that the extra ERIC sequence in the Hw DNA after *luxB* would have been inserted at a later date rather than deletion from the Hb DNA, as the latter mechanism would require exact removal of one ERIC structure.

The PCR experiments also indicate that multiple ERIC structures are present in the *lux* DNAs from three different isolates of *X. luminescens* obtained from humans. Two of these strains (ATCC 43950 and ATCC 43951) were isolated from patients in San Antonio, Tex., in 1986 and 1987; the third strain (ATCC 43949) was isolated from a patient in Maryland in 1977 (8). These strains appear to be very similar, particularly if one compares the PCR products by using only the ERIC primer (Fig. 7). Hybridization of the labelled ERIC probe to Southern blots of restricted genomic DNA (data not shown) also suggests that these strains are similar, particularly with respect to the number and distribution of ERIC structures in the *X. luminescens* strains isolated in San Antonio. By the same criteria (A and B in Fig. 7), the two strains from insects appear to be different in terms of the number and distribution of ERIC structures in the genomic DNA, although both strains lack the ERIC structure in front of *luxA* and have an ERIC structure in front of *luxE*. Although these results indicate a clear distinction between *X. luminescens* strains isolated from humans and insects in terms of the distribution of ERIC structures in the *lux* operon, additional *X. luminescens* strains must be analyzed to reach a more definitive conclusion.

This report has not addressed the role of the ERIC structures. Expression of the two different fragments under control of a T7 RNA polymerase promoter demonstrates that the luminescence is the same for both *X. luminescens lux* DNAs at all stages of growth (Fig. 8). Consequently, it appears that the presence of multiple ERIC structures in the Hw DNA or translation of the initial part of the single ERIC structure in the Hb DNA as part of *luxB* has caused no significant difference in expression at the transcriptional or translation level under these conditions. These results are consistent with the proposal that the repetitive elements function as "selfish DNA" without any apparent role (11, 22). However, the strong conservation of the ERIC sequences in bacteria from different genera and their high

frequency in the *lux* operon of the *X. luminescens* strain isolated from humans is so remarkable that it is possible that they reflect a function essential for the bacteria. Among the potential functions that have been suggested are structural signals involved in DNA metabolism, organization, or rearrangement (11, 22, 28). If the frequency and function of such signals are related to the environmental, nutritional, and/or physical location of the bacteria, the presence of multiple repetitive elements in the *lux* operons of *X. luminescens* strains isolated from humans will be of particular interest.

ACKNOWLEDGMENTS

We thank Pio Colepicolo and J. W. Hastings (Harvard University) for the *X. luminescens* Hw strain and S. Tabor and C. Richardson (Harvard University) for the pT7 and pGP-1-2 plasmids.

This research was supported by grant MT7672 from the Medical Research Council of Canada.

REFERENCES

1. **Baldwin, T. O., J. H. Devine, R. C. Heckel, J.-W. Lin, and G. S. Shadel.** 1989. The complete nucleotide sequence of the *lux* regulon of Vibrio fischeri and the *lux*ABN region of Photobacterium leiognathi and the mechanism of control of bacterial bioluminescence. J. Biolumin. Chemilumin. **4**:326–341.
2. **Cochrum, L., K. S. Hruska, E. B. Rucker, and T. C. Johnston.** 1990. The nucleotide sequence of the *lux*D gene of *Xenorhabdus luminescens* Hm. Nucleic Acids Res. **18**:5570.
3. **Cohn, D. H., R. C. Ogden, J. N. Abelson, T. O. Baldwin, K. H. Nealson, M. J. Simon, and A. J. Milehan.** 1983. Cloning of the Vibrio harveyi luciferase genes: use of a synthetic oligonucleotide probe. Proc. Natl. Acad. Sci. USA **80**:120–123.
4. **Colepicolo, P., K.-W. Cho, G. O. Poinar, and J. Hastings.** 1989. Growth and luminescence of the bacterium *Xenorhabdus luminescens* from a human wound. Appl. Environ. Microbiol. **55**: 2601–2606.
5. **Delong, E. F., D. Steinhauer, A. Israel, and K. H. Nealson.** 1987. Isolation of the *lux* genes from Photobacterium leiognathi and expression in Escherichia coli. Gene **54**:203–210.
6. **Engebrecht, J., K. H. Nealson, and M. Silverman.** 1983. Bacterial bioluminescence: isolation and genetic analysis of functions from Vibrio fischeri. Cell **32**:773–781.
7. **Engebrecht, J., and M. Silverman.** 1984. Identification of genes and gene products necessary for bacterial bioluminescence. Proc. Natl. Acad. Sci. USA **81**:4154–4158.
8. **Farmer, J. J., III, J. H. Jorgensen, P. A. D. Grimont, R. J. Akhurst, G. O. Poinar, Jr., E. Ageron, G. V. Pierce, J. A. Smith, G. P. Carter, K. L. Wilson, and F. W. Hickman-Brenner.** 1989. *Xenorhabdus luminescens* (DNA hybridization group 5) from human clinical specimens. J. Clin. Microbiol. **27**:1594–1600.
9. **Frackman, S., M. Anhalt, and K. H. Nealson.** 1990. Cloning, organization, and expression of the bioluminescence genes of *Xenorhabdus luminescens*. J. Bacteriol. **172**:5767–5773.
10. **Hastings, J. W., and G. Weber.** 1963. Total quantum flux of isotropic sources. J. Opt. Soc. Am. **53**:1140–1415.
11. **Hulton, C. S. J., C. F. Higgins, and P. M. Sharp.** 1991. ERIC sequences: a novel family of repetitive elements in the genomes of Escherichia coli, Salmonella typhimurium, and other enterobacteria. Mol. Microbiol. **5**:825–834.
12. **Illarionov, B. A., V. M. Blinov, A. P. Donchenko, M. V. Protopopova, V. A. Karginov, N. P. Mertvetsov, and J. I. Gitelson.** 1990. Isolation of bioluminescent functions from Photobacterium leiognathi: analysis of *lux*A, *lux*B, *lux*G and neighboring genes. Gene **86**:89–94.
13. **Johnston, T. C., E. B. Rucker, L. Cochrum, K. S. Hruska, and V. Vandegrift.** 1990. The nucleotide sequence of the *lux*A and *lux*B genes of *Xenorhabdus luminescens* HM and a comparison of the amino acid sequences of luciferases from four species of bioluminescent bacteria. Biochem. Biophys. Res. Commun. **170**:407–415.
14. **Lee, C. Y., R. Szittner, and E. A. Meighen.** 1991. The *lux* genes of the luminous bacterial symbiont, *Photobacterium leiognathi*, of the ponyfish. Eur. J. Biochem. **201**:161–167.
15. **Mancini, J. A., M. Boylan, R. R. Soly, A. F. Graham, and E. A. Meighen.** 1988. Cloning and expression of the *Photobacterium phosphoreum* luminescence system demonstrates a unique *lux* gene organization. J. Biol. Chem. **263**:14308–14314.
16. **Martin, M., R. Showalter, and M. Silverman.** 1989. Identification of a locus controlling expression of luminescence genes in *Vibrio harveyi*. J. Bacteriol. **171**:2406–2414.
17. **Meighen, E. A.** 1988. Enzymes and genes from the *lux* operons of bioluminescent bacteria. Annu. Rev. Microbiol. **42**:151–176.
18. **Meighen, E. A.** 1991. Molecular biology of bacterial bioluminescence. Microbiol. Rev. **55**:123–142.
19. **Miyamoto, C. M., A. F. Graham, M. Boylan, J. F. Evans, K. W. Hasel, E. A. Meighen, and A. F. Graham.** 1985. Polycistronic mRNAs code for polypeptides of the *Vibrio harveyi* luminescence systems. J. Bacteriol. **161**:995–1001.
20. **Miyamoto, C. M., A. F. Graham, and E. A. Meighen.** 1988. Nucleotide sequence of the *lux*C gene and the upstream DNA from the bioluminescent system of *Vibrio harveyi*. Nucleic Acids Res. **16**:1551–1562.
21. **Schmidt, T. M., K. Kopecky, and K. H. Nealson.** 1989. Bioluminescence of the insect pathogen *Xenorhabdus luminescens*. Appl. Environ. Microbiol. **55**:2607–2612.
22. **Sharples, G. J., and R. G. Lloyd.** 1990. A novel repeated sequence located in the intergenic regions of bacterial chromosomes. Nucleic Acids Res. **18**:6503–6508.
23. **Soly, R. R., J. A. Mancini, S. R. Ferri, M. Boylan, and E. A. Meighen.** 1988. A new *lux* gene in bioluminescent bacteria codes for a protein homologous to the bacterial luciferase subunits. Biochem. Biophys. Res. Commun. **155**:351–358.
24. **Swartzman, E., S. Kapoor, A. Graham, and E. Meighen.** 1990. A new *Vibrio fischeri lux* gene precedes a bidirectional termination site for the *lux* operon. J. Bacteriol. **172**:6797–6802.
25. **Swartzman, E., C. Miyamoto, A. Graham, and E. Meighen.** 1990. Delineation of the transcriptional boundaries of the *lux* operon of *Vibrio harveyi* demonstrates the presence of two new *lux* genes. J. Biol. Chem. **265**:3513–3517.
26. **Szittner, R., and E. Meighen.** 1990. Nucleotide sequence, expression and properties of luciferase coded by *lux* genes from a terrestrial bacterium. J. Biol. Chem. **265**:16581–16587.
27. **Tabor, S., and C. C. Richardson.** 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82**:1074–1078.
28. **Versailovic, J., T. Koeuth, and J. R. Lupski.** 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. Nucleic Acids Res. **18**: 6503–6508.
29. **Xi, L., K.-W. Cho, and S. C. Tu.** 1991. Cloning and nucleotide sequences of *lux* genes and characterization of luciferase of *Xenorhabdus luminescens* from a human wound. J. Bacteriol. **173**:1399–1405.