Phylogenetic relationships of necrogenic *Erwinia* and *Brenneria* species as revealed by glyceraldehyde-3-phosphate dehydrogenase gene sequences

Eric W. Brown,^{1†} Renee M. Davis,¹ Chin Gouk² and Tom van der Zwet³

Author for correspondence: Eric W. Brown. Tel: +1 202 205 4172. Fax: +1 202 205 4408.
e-mail: ebrown1@cfsan.fda.gov

¹ Department of Biological Sciences, Loyola University of Chicago, Chicago, IL 60626, USA

² Ruakura Research Center, HortResearch, Hamilton, New Zealand

³ USDA-ARS, Appalachian Fruit Research Center, Kearneysville, WV 25430, USA

Recent examination of the relationships of the dry necrosis-inducing (necrogenic) erwinias using 16S rDNA sequences demonstrated that these bacteria comprise a polyphyletic group and, therefore, have been subdivided into three distinct genera, *Erwinia*, *Brenneria* and *Pectobacterium*, with the classical 'amylovora' group species now being distributed nearly evenly among the first two. To further assess the molecular evolutionary relationships between current necrogenic *Erwinia* and *Brenneria* species, as well as between these genera and the exclusively soft-rotting genus *Pectobacterium*, the glyceraldehyde-3-phosphate dehydrogenase (*gapDH*) genes from 57 *Erwinia* and *Brenneria* isolates along with *Pectobacterium* type strains were PCR-amplified, sequenced and subjected to phylogenetic analysis. Pairwise alignments of cloned *gapDH* genes revealed remarkably high interspecies genetic diversity among necrogenic isolates. Four evolutionary clades of necrogenic species were described that assorted more closely to known soft-rotting species than to each other. Interclade comparisons of *gapDH* nucleotide sequences revealed as much genetic divergence between these four necrogenic clades as existed between necrogenic and soft-rotting clades. An examination of the phylogenetic utility of the *gapDH* gene in light of current 16S rDNA clustering of these species revealed varying levels of taxonomic congruence between these genes for the structure of *Erwinia*, *Brenneria* and *Pectobacterium*. These analyses suggest that, while *gapDH* possesses sufficient genetic variation to fully differentiate *Erwinia* and *Brenneria* species, the gene may not accurately reflect interspecies taxonomic relatedness among all three phytopathogenic genera.

Keywords: *Erwinia*, *Brenneria*, *gapDH*, molecular evolution, differentiation, phylogeny

INTRODUCTION

Historically, the 'amylovora' group was one of two taxonomic phenons within the genus *Erwinia*. It was

[†] **Present address:** Molecular Biology Branch, DMBRE-OPA, Center for Food Safety & Applied Nutrition, US Food & Drug Administration, Mailstop HFS-237, 200 C Street SW, Washington, DC 20204, USA.

Abbreviations: CI, consistency index; *gapDH*, glyceraldehyde-3-phosphate dehydrogenase; RI, retention index.

The GenBank accession numbers for the *gapDH* sequences of 55 necrogenic isolates and seven soft-rotting strains, representing the genera *Erwinia*, *Brenneria* and *Pectobacterium*, are listed in Methods.

originally proposed to encompass the necrogenic, non-soft-rotting phytopathogens into a single congeneric cluster (Dye, 1968; van der Zwet & Keil, 1979). Necrogenic *Erwinia* species are recognized primarily as pathogens that induce dry necrosis and die back of their respective plant hosts (Vanneste, 1995; Elrod, 1941). The phenon was initially established with the type species *Erwinia amylovora* as its sole member, but was expanded to encompass other necrogenic species, including *Erwinia* (now *Brenneria*) *quercina*, *Erwinia psidii*, ' *Erwinia* (now *Brenneria*) *lupinicola*', *Erwinia mallotivora*, *Erwinia* (now *Brenneria*) *rubrifaciens*, *Erwinia* (now *Brenneria*) *salicis*, *Erwinia tracheiphila*,

Erwinia (now *Brenneria*) *nigrifluens* and *Erwinia* (now *Brenneria*) *alni* (Winslow *et al.*, 1920; Dye, 1968; Surico *et al.*, 1996; Neto *et al.*, 1987; Goto, 1976). The 'carotovora' group encompassed the other cluster of *Erwinia* species. This biochemically active phenon comprised species which employ pectate lyases and cellulases to induce soft-rotting of various herbaceous plant tissues (Dye, 1969).

Small-subunit-like (16S) ribosomal genes have been particularly helpful in discerning the relationships of many eubacterial lineages (Ludwig & Schleifer, 1999; Woese, 1987). Recently, 16S rDNA sequences have been obtained from *Erwinia* species in the 'amylovora' and 'carotovora' groups (Hauben *et al.*, 1998; Kwon *et al.*, 1997). Taxonomic inference from these sequences resulted in the partitioning of the *erwinias* into three distinct generic groupings: *Erwinia*, *Brenneria* and *Pectobacterium* (Hauben *et al.*, 1998). The genera *Erwinia* and *Brenneria* currently encompass all of the necrogenic phytopathogenic species in addition to *Erwinia rhapontici* and *Erwinia persicina*, while *Pectobacterium* is comprised solely of soft-rotting phytopathogenic species.

Due to the exceptionally slow substitution rates among ribosomal genes, the reliable separation of closely related bacterial species and strains remains difficult. While rDNA sequences are powerful in discerning the evolutionary relationships of more disparate prokaryotic lineages (Lawrence *et al.*, 1991; Woese, 1987), taxonomic inferences drawn between congeneric isolates are often based on only a few nucleotide substitutions (Lawrence *et al.*, 1991; Ochman & Wilson, 1988). To identify other regions of the bacterial chromosome that may also hold phylogenetic utility in the discrimination of bacterial species, several moderately evolving protein-coding loci have been evaluated (Lloyd & Sharp, 1993; Glassick *et al.*, 1996; Fukunaga & Koreki, 1996). The *gapDH* gene encodes glyceraldehyde-3-phosphate dehydrogenase (GapDH), an essential component of glycolysis (Prescott *et al.*, 1993). The *gapDH* locus has been used to describe the taxonomic positions of several taxa at the species level (Lawrence *et al.*, 1991; Schlaeper & Zuber, 1992; Liaud *et al.*, 1994). Phylogenetic utility of this gene has been attributed to the unusually slow third position substitution rate observed among *gapDH* codons (Lawrence *et al.*, 1991). Due to the presence of high codon bias, the *gapDH* gene is known to evolve more slowly at synonymous sites, making the molecule less prone to nucleotide reversals or ambiguous substitution events (Sharp & Li, 1987).

To examine the molecular evolutionary relationships of *gapDH* alleles from necrogenic *Erwinia* and *Brenneria* species, as well as from soft-rotting species, the nucleotide sequences for *gapDH* from 55 necrogenic isolates and seven soft-rotting strains were determined. Resultant cladistic analysis of the *gapDH* nucleotides and computer-derived amino acid sequences has allowed for (i) a highly resolved mol-

ecular genetic differentiation of necrogenic *Erwinia* and *Brenneria* species, (ii) a description of *gapDH*-based phylogenetic relationships that exists between the necrogenic and soft-rotting enterobacterial phytopathogens, and (iii) an evaluation of the taxonomic utility of the *gapDH* gene in light of the currently available 16S rDNA partitions for this group of bacteria.

METHODS

Bacterial strains and culture conditions. A culture collection of 62 necrogenic and soft-rotting isolates, representing the genera *Erwinia*, *Brenneria* and *Pectobacterium*, was assembled for phylogenetic studies. Type cultures of all necrogenic *Erwinia* and *Brenneria* species were acquired from the American Type Culture Collection (ATCC, Manassas, VA, USA). The necrogenic (previously 'amylovora' group) isolates and the collections from which they originated are listed in Table 1. Type cultures of all soft-rotting (previously 'carotovora' group) species were also acquired from ATCC and used as reference isolates during phylogenetic differentiation of necrogenic species. These strains included *Pectobacterium carotovorum* subsp. *carotovorum* (15713^T), *Pectobacterium carotovorum* subsp. *beta-vasculorum* (43762^T), *Pectobacterium chrysanthemi* (11663^T), *Pectobacterium cacticida* (49481^T), *Pectobacterium cyripedii* (29267^T), *Erwinia rhapontici* (29283^T) and *Erwinia persicina* (33998^T). Strains of *Erwinia amylovora* and *Erwinia mallotivora* were grown on nutrient yeast dextrose agar; all other isolates were grown on tryptic soy agar (Difco).

Rapid preparation of bacterial DNA. Genomic DNA was isolated from bacterial strains using a commercially available extraction matrix (Bio-Rad). Briefly, cells were washed in saline, resuspended in Instagene DNA purification resin and incubated at 56 °C for 30 min. Cell preparations were then vortexed vigorously, incubated at 100 °C for 10 min and centrifuged at 13000 r.p.m. (9 g) for 6 min. The remaining supernatant, containing total genomic DNA, was decanted into a clean microtube.

PCR amplification of the *gapDH* gene. PCR (Ehrlich *et al.*, 1991) was performed with 20 µl DNA template, 10 × PCR buffer containing 1.5 mM MgCl₂ (Perkin-Elmer), 2.5 mM dNTPs (Pharmacia) and 1.5 U *Taq* DNA polymerase (Promega). Oligonucleotide primer pairs used to amplify the *gapDH* gene were added to a final concentration of 50 pmol and included gap1F (5'-TGAAATATGACTCCACTCACGG-3') and gap1R (5'-TAGAGGACGGGATGATGTTCTG-3'). Amplification was performed in a PTC-100 thermal cycler (MJ Research) under the following conditions: initial denaturation at 94 °C for 5 min, followed by 94 °C for 1 min, 52 °C for 1 min and 72 °C for 1.5 min (35 cycles), ending with incubation at 72 °C for 10 min. The segment of the *gapDH* gene amplified corresponded to base pair coordinates 134–625 of the enterobacterial *gapDH* locus (Lawrence *et al.*, 1991). The primers were designed from a *gapDH* gene alignment of several enterobacterial species and represent highly conserved regions of the *gapDH* molecule that are shared across a variety of plant- and animal-pathogenic bacteria.

Molecular cloning and sequence determination. Products from PCR amplification of bacterial DNA were column-purified using Centricon-100 spin columns (Amicon) and ligated to a 'sticky-T' cloning vector, previously synthesized

Table 1. Specific cultures of necrogenic *Erwinia* and *Brenneria* species analysed

Collection	Source	Strain codes*
LUC	Department of Biological Sciences, Loyola University of Chicago, IL, USA	Ea26eu, Ea4me, Ea7me, Ea4ca, Ea7ca
AFRS	USDA, ARS-Appalachian Fruit Research Center, Kearneysville, WV, USA	Ea581, Ea178, Ea179, Ea1829, Ea1804, Ea1450, Ea554, Ea601
Cornell	S.V. Beer, Department of Plant Sciences, Cornell University, Ithaca, NY, USA	Ea282, Br6d364, Bs7d32, Blw3l31, Bl346, Bl348
ATCC	American Type Culture Collection, Manassas, VA, USA	Ea15580 ^T , Et33245 ^T , Et11418, Em29573 ^T , Ep49406 ^T , Br29291 ^T , Bn13028 ^T , Bn29277, Bn29275, Bs15712 ^T , Bq29282, Bq29281 ^T , Bq27622
ICMP	HortResearch, Auckland, New Zealand	Et5845, Ep8427, Ep8428, Ep8429, Em8645, Em8646, Et1396, Br5946, Br5950, Br5951, Br4790, Bs1379, Bs9136, Bq1846, Bq1919, Bn1391, Bn1576
ERRC	J. Wells, USDA, Eastern Regional Research Center, Wyndmoor, PA, USA	Et9812, Et9815
NCPPB	Central Science Laboratory, Harpenden, UK	Bs2310, Bs2317, Ep3558, Bal12481 ^T

*Strain abbreviations are as follows: Ea, *Erwinia amylovora*; Ep, *Erwinia psidii*; Em, *Erwinia mallotivora*; Et, *Erwinia tracheiphila*; Bq, *Brenneria quercina*; Bl, '*Brenneria lupinicola*'; Br, *Brenneria rubrifaciens*; Bn, *Brenneria nigrifluens*; Bs, *Brenneria salicis*; Bal, *Brenneria alni*.

by attaching dTTP to the blunt end of a KS+ cloning phagemid (Stratagene). Ligations were incubated overnight at 14 °C and reactions were transformed into *Escherichia coli* DH5 α cloning cells by conventional heat-shock methods (Maniatis *et al.*, 1982). Overnight cultures were prepared from three clones for each PCR product and phagemids were then harvested through diatomaceous push-columns (Promega). Multiple phagemid clones were sequenced by the chain-termination method (Sanger *et al.*, 1977) on an ABI 377 automated sequencer (Perkin-Elmer; Applied Biosystems). M13 and T3/T7 primer-binding sites were used to initiate forward- and reverse-strand reactions. Sequence data were then transferred into the GCG (Genetics Computer Group, University of Wisconsin-Madison, USA) sequence handling program (Devereux *et al.*, 1984) and full-length *gapDH* fragments were assembled.

Sequence alignment and phylogenetic reconstruction. The nucleotide and resultant computer-translated amino acid sequences were first subjected to alignment using CLUSTAL W v. 1.5 (Thompson *et al.*, 1992). CLUSTAL W minimized gaps by invoking a predetermined penalty for every insertion or deletion in the alignment. Gaps in the analysis were treated as missing data points, while transitions and transversions in the alignment were assigned equal character weights. Genetic distances between *gapDH* nucleotide sequences were calculated using the method of Jukes & Cantor (1969).

The *gapDH* nucleotide and amino acid sequence matrices were subjected to phylogenetic analysis by using a total of 447 bp and 149 aa residues, respectively. The phylogenetic method employed utilized the principle of maximum-parsimony (Farris, 1983) and was available in PAUP (phylogenetic analysis using parsimony) v. 3.11 (Swofford, 1991; Swofford & Olsen, 1990). Most parsimonious trees were sought using heuristic search methods combined with tree bisection-reconnection (TBR) branch-swapping and random addition order of taxa. When necessary, a strict

consensus method was applied to reduce the number of equally parsimonious cladograms into a single tree such that every relationship present in the consensus tree was found in each of the original trees (Forey *et al.*, 1992). All tree branches with a step length of zero were collapsed into polytomies. Cladograms were rooted using *gapDH* sequences from two *Bacillus* species, *Bacillus subtilis* and *Bacillus megaterium* (GenBank accession numbers X13011 and X54520, respectively). Character support for internal tree nodes was determined by 100 iterations of bootstrapping (Felsenstein, 1985) and was available in PAUP v. 3.11 (Swofford, 1991). Relative levels of homoplasy were measured among *gapDH* nucleotide and amino acid sequences using two separate tree indices, the consistency index (CI; Forey *et al.*, 1992) and the retention index (RI; Farris, 1989). Tree structure and skewedness (out of 10000 randomly generated trees) was measured using the *G*-statistic (*G*₁; Hillis & Huelsenbeck, 1992), available in the RANDOM TREES program of PAUP v. 3.11.

Nucleotide sequence accession numbers. The nucleotide sequence data reported in this paper have been deposited in the GenBank sequence database under the following accession numbers (strain codes are defined in Table 1). Necrogenic isolates: AF164969 (Ea7ca), AF164970 (Ea4ca), AF164971 (Ea26eu), AF164973 (Ea4me), AF164974 (Ea7me), AF164976 (Ea178), AF164978 (Ea601), AF164972 (Ea1450), AF164975 (Ea15580^T), AF164977 (Ea179), AF164979 (Ea1804), AF164980 (Ea1829), AF164981 (Ea282), AF164982 (Ea581), AF164983 (Ea554), AF164984 (Em29573^T), AF164985 (Em8645), AF164986 (Em8646), AF165002 (Ep49406^T), AF165003 (Ep8427), AF165004 (Ep8428), AF165006 (Ep3558), AF165005 (Ep8429), AF165007 (Et1396), AF165009 (Et33245^T), AF165010 (Et11418), AF165011 (Et5845), AF165012 (Et9815), AF165013 (Et9812), AF164988 (Bq1846), AF164989 (Bq1919), AF164990 (Bq27622), AF164992 (Bq29282), AF164991 (Bq29281^T), AF164993 (Blw3l31), AF164994

Fig. 1. Maximum-parsimony phylogenetic tree of *Erwinia*, *Brenneria* and *Pectobacterium* strains as revealed by *gapDH* nucleotide sequences. The topology shown depicts only those relationships present in each and every one of the resulting

(B1346), AF164995 (B1348), AF164996 (Br29291^T), AF164998 (Br5946), AF164999 (Br5950), AF164997 (Br4790), AF165001 (Br6d364), AF165000 (Br5951), AF165019 (Bs1379), AF165022 (Bs9136), AF165023 (Bs2310), AF165024 (Bs2317), AF165021 (Bs7d32), AF165020 (Bs15712^T), AF165025 (Ba112481^T), AF165014 (Bn13028^T), AF165015 (Bn1391), AF165017 (Bn29275), AF165018 (Bn29277), AF165016 (Bn1576). Soft-rotting isolates: AF165032 (*Pectobacterium carotovorum* subsp. *carotovorum* 15713^T), AF165031 (*Pectobacterium carotovorum* subsp. *betavasculorum* 43762^T), AF165030 (*Pectobacterium chrysanthemi* 11663^T), AF165026 (*Pectobacterium cacticida* 49481^T), AF165027 (*Pectobacterium cypripedii* 29267^T), AF165029 (*Erwinia rhapontici* 29283^T) and AF165028 (*Erwinia persicina* 33998^T).

RESULTS

Phylogenetic analysis of *gapDH* nucleotide sequences from *Erwinia*, *Brenneria* and *Pectobacterium* species

The nucleotide sequence of a 447 bp segment of the *gapDH* locus from 55 necrogenic strains, representing 10 *Erwinia* and *Brenneria* species, was combined with sequences from five soft-rotting *Pectobacterium* species and the remaining two soft-rotting *Erwinia* species and subjected to maximum-parsimony phylogenetic analysis.

The resultant group of equally parsimonious trees yielded a CI of 0.59, an RI of 0.85 and a G_1 of -0.51, indicating a substantial level of skewedness in the *gapDH* nucleotide data. The derived *gapDH* consensus tree generated several interesting findings (Fig. 1). First, the *gapDH* nucleotide data completely resolved necrogenic *erwinias* and *brennerias* at the species level, demonstrating that the *gapDH* locus possesses sufficient genetic variation to fully differentiate the members of these two genera. Second, all of the strains exhibited species monophyly; that is, every strain had as its nearest neighbour another strain from the same species. Third, four disparate phylogenetic lineages or clades of necrogenic *Erwinia* and *Brenneria* species, designated clades I–IV, were resolved that appear to have diverged from at least three deep radiations in the tree. Interestingly, these four clades were dispersed among three separate clusters of soft-rotting species (A–C) such that every clade of necrogenic species was more closely related to a soft-rotting clade than to any of the other necrogenic clades (Fig. 1).

Pairwise comparisons of *gapDH* sequences revealed remarkable diversity between necrogenic *Erwinia* and *Brenneria* species. Interspecies distances for members of the two genera and *Pectobacterium* species are reported in Table 2. Surprisingly, interspecies *gapDH*

Table 2. Genetic distances between *Erwinia*, *Brenneria* and *Pectobacterium* species based on *gapDH* sequence divergence

Species*	Genetic distance (%)†															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1 <i>Erwinia amylovora</i>																
2 <i>Erwinia mallotivora</i>	15.6															
3 <i>Erwinia tracheiphila</i>	22.8	21.0														
4 <i>Erwinia psidii</i>	18.3	12.5	12.0													
5 <i>Brenneria quercina</i>	15.9	18.6	25.4	18.6												
6 <i>Brenneria lupinicola</i>	15.3	20.7	26.7	20.0	7.2											
7 <i>Brenneria rubrifaciens</i>	22.2	24.0	29.5	24.1	22.2	23.6										
8 <i>Brenneria salicis</i>	24.1	25.1	32.8	26.2	22.5	24.1	21.8									
9 <i>Brenneria alni</i> ‡	17.4	17.9	24.7	18.8	13.9	15.1	21.4	12.0								
10 <i>Brenneria nigrifluens</i>	17.3	18.4	24.0	17.7	15.1	16.9	21.1	20.6	10.3							
11 <i>Pectobacterium carotovorum</i> subsp. <i>carotovorum</i>	17.7	16.5	23.6	16.9	16.5	16.9	22.0	20.4	12.1	14.5						
12 <i>Pectobacterium carotovorum</i> subsp. <i>betavasculorum</i>	16.1	16.6	23.3	16.7	15.8	15.8	20.7	20.6	11.4	14.0	7.8					
13 <i>Pectobacterium cacticida</i>	17.6	18.3	23.4	17.6	17.2	17.4	22.1	24.2	18.3	19.3	19.9	19.2				
14 <i>Pectobacterium chrysanthemi</i>	14.3	15.4	21.2	14.7	14.5	14.5	20.6	22.0	22.4	16.3	17.0	16.3	3.6			
15 <i>Erwinia rhapontici</i>	12.4	14.3	22.6	16.8	15.8	16.9	24.3	24.3	14.5	17.6	14.8	14.1	16.8	14.2		
16 <i>Pectobacterium cypripedii</i>	13.6	15.4	24.5	18.9	18.0	18.9	25.2	25.2	17.4	19.0	17.2	16.3	18.8	15.7	5.8	
17 <i>Erwinia persicina</i>	11.8	13.2	22.8	16.5	15.8	16.4	22.8	24.5	15.0	16.7	14.3	13.9	16.3	13.2	3.4	3.4

* Underlining denotes a necrogenic species.

† Percentages are reported as mean genetic distances (Jukes & Cantor, 1969) and were derived from pairwise sequence comparisons of *gapDH* nucleotides.

‡ Includes type strain only.

478 equally most parsimonious trees, all having a nucleotide-based tree length of 973 steps. Bootstrap values are reported below each node in italics. The brackets to the right of each clade define the locations of the four major necrogenic clades, designated I–IV, as well as the positions of the three soft-rotting clades, A–C. Roman numerals, positioned within the tree, denote the specific node from which each of the four necrogenic clades diverged. Individual branch lengths are presented above each respective branch, while the adjacent values in parentheses denote the number of unambiguous substitutions that mapped to the tree only once.

Fig. 2. Minimum and maximum percentage nucleotide divergence of *gapDH* sequences between the four major necrogenic *Erwinia/Brenneria* clades as well as between necrogenic clades and soft-rotting clades. The Roman numeral designations of the four necrogenic clades are indicated within the larger circles (I–IV). Soft-rotting *Pectobacterium* and *Erwinia* clades are represented by the smaller squares and designated A, B and C. Individual *Erwinia* and *Brenneria* species comprising each clade are as follows: I, *Erwinia amylovora*; II, *Erwinia mallotivora*, *Erwinia psidii* and *Erwinia tracheiphila*; III, *Brenneria rubrifaciens*, *Brenneria salicis*, *Brenneria nigrifluens* and *Brenneria alni*; IV, *Brenneria quercina* and '*Brenneria lupinicola*'; A, *Pectobacterium cypripedii*, *Erwinia rhapontici* and *Erwinia persicina*; B, *Pectobacterium chrysanthemi* and *Pectobacterium cacticida*; C, *Pectobacterium carotovorum* subsp. *carotovorum* and *Pectobacterium carotovorum* subsp. *betavascularum*.

sequence diversity was greatest between two necrogenic species, rather than between a necrogenic and a soft-rotting species. *Brenneria rubrifaciens*, the pathogen of Persian walnut, and *Erwinia tracheiphila*, the cucumber-wilt pathogen, varied by nearly 30% at the nucleotide level. Taken together, these results have demonstrated an extraordinary amount of genetic diversity present among necrogenic species within and between the *Erwinia* and *Brenneria*.

The genetic distances observed between the four distinct necrogenic species clades further reiterated the phylogenetic disparity noted among necrogenic species (Fig. 2). The greatest levels of *gapDH* nucleotide divergence were observed between necrogenic clades (clades II and III are 15.9–33.3% divergent) rather than between soft-rotting and necrogenic clades (clades III and A are 14.5–27.1% divergent). In nearly every case, each of the four necrogenic clades was more distant from the other three clades of necrogenic species than from the three soft-rotting clades A, B or C. The only exception to this trend was clade III which was slightly more distant from A (maximum diver-

gence of 27.1%) than from clades I and IV (26.6 and 25.5% maximum sequence divergence, respectively). These data indicate that *gapDH*-based molecular evolutionary relationships between necrogenic species are far more disparate than previously thought and further support the rDNA sequence-based partitioning of these species into distinct taxonomic subgroupings.

Phylogenetic mapping of pathological phenotypes

The polyphyletic distribution of necrogenic and soft-rotting clades, as revealed by the *gapDH* cladogram in Fig. 1, suggested that one or both of the pathological phenotypes acquired by *Erwinia*, *Brenneria* and *Pectobacterium* species may have evolved multiple times during the radiation of the genus. To investigate the phylogenetic history of these pathological traits, necrosis and soft-rot were converted into phylogenetic characters and mapped to the *gapDH* tree, again using maximum-parsimony (Fig. 3). Various optimization attempts were made to determine the least number of times that the two traits could have arisen. Several

Fig. 3. Phylogenetic mapping of the two major pathological states elicited by *Erwinia*, *Brenneria* and *Pectobacterium* species. The topology shown is a simplified version of the *gapDH* nucleotide tree presented in Fig. 1 and reveals each of the major necrogenic (N) and soft-rotting (SR) lineages. The four necrogenic clades are denoted I–IV while soft-rotting clades are designated A, B and C. Three different scenarios for the evolution of necrosis and soft-rot are depicted on the *gapDH* tree and are denoted by open, filled or striped rectangles (necrosis) and ovals (soft-rot).

different parsimony solutions were found that mapped the characters to the *gapDH* phylogeny. Two of these scenarios required four evolutionary transformations (striped and filled tree symbols), while one required only three transformations (open tree symbols) making it the most parsimonious mapping solution for the evolution of these two pathological characters. In this least-assumptive scenario, where necrosis was inferred to be the ancestral state, soft-rot had to evolve three separate times along the three different lineages of soft-rotting bacteria to give rise to the current polyphyletic distribution of the pectolytic isolates. Assuming that an evolutionary hierarchy in the *gapDH* molecule exists for these species, these findings would indicate independent phylogenetic origins for the soft-rotting phenotype during the radiation of these genera within the family *Enterobacteriaceae*.

Phylogenetic analysis of GapDH amino acid sequences

Phylogenetic reconstruction of the GapDH amino acid sequences from necrogenic and soft-rotting species was derived with PAUP v. 3.11 (Swofford, 1991). The most parsimonious tree (CI = 0.86, RI = 0.79, $G_1 = -1.50$) derived from the combined amino acid sequence alignment is presented in Fig. 4. The GapDH amino acid tree reiterated the relationships that were previously described for the *gapDH* nucleotide analysis

and revealed the four major necrogenic clades (I–IV) which, again, assorted more closely to the same three soft-rotting clades presented in Fig. 1. The only topological incongruence between the amino acid and nucleotide phylogenies was seen in the placement of clade I which diverged earliest in the protein tree but last in the DNA tree. Also, the nodes separating the four major groups of necrogenic bacteria did not exhibit the depth of clade divergence in the GapDH amino acid tree that was observed in the nucleic acid tree. This is likely a reflection of the structural conservation of the GapDH molecule resulting from its critical role in the bacterial glycolytic pathway (Prescott *et al.*, 1993).

Evaluation of the aligned GapDH computer-translated amino acid sequences from necrogenic *Erwinia* and *Brenneria* strains revealed the existence of shared residues that are unique to specific necrogenic clades or species. These signature synapomorphic residues, their position in the CLUSTAL W alignment and the groups that they distinguish are listed in Table 3. Synapomorphic residues at positions 35, 36, 59, 86 and 91 are unique to one of the four major necrogenic clades, while amino acids at positions 3, 102 and 147 distinguish pairs of necrogenic *Erwinia* and *Brenneria* species within major clades. Finally, there are 26 synapomorphic sites that are unique to a single necrogenic *Erwinia* or *Brenneria* species. These residues may be used in conjunction with PCR amplification and automated DNA sequencing techniques to identify novel necrogenic *Erwinia* and *Brenneria* species as they are isolated from nature.

Phylogenetic comparisons of *gapDH* and 16S rDNA sequence evolution

The *gapDH* nucleotides from the type strains of *Erwinia*, *Brenneria* and *Pectobacterium* species were combined with *gapDH* sequences from reference isolates of 16 other enterobacterial species/subgroups for which *gapDH* sequences have been reported in GenBank. The sequences were aligned and analysed using the cladistic methods described above (Swofford, 1991). The single most parsimonious tree (CI = 0.62, RI = 0.61, $G_1 = -0.58$) to result from this analysis was topologically congruent with rDNA taxonomy in several instances (Fig. 5). First, the *erwinias* formed a monophyletic clade that reiterated 16S rDNA cluster I (Hauben *et al.*, 1998). The only exception was *Pectobacterium cypripedii* which assorted with the other two soft-rotting *erwinias* (*Erwinia rhapontici* and *Erwinia persicina*) in the *gapDH* analysis. Second, *gapDH* sequence divergence further subdivided the *erwinias* into two subclades (an *Erwinia amylovora* subgroup and an *Erwinia mallotivora* subgroup) which were also supported by 16S rDNA relationships. Finally, *Pectobacterium* and *Brenneria* species formed distinct clusters disparate from the *Erwinia* species clade. Interestingly, the *gapDH* phylogeny differed from known 16S rDNA groupings in one significant aspect in that it further partitioned *Brenneria* and *Pecto-*

Fig. 4. Maximum-parsimony phylogenetic tree of *Erwinia*, *Brenneria* and *Pectobacterium* species as revealed by *GapDH* amino acid sequences. The topology shown depicts the single most parsimonious tree and was drawn to scale with an overall tree length of 217 aa substitutions. Conical branches at the tips of the tree denote the monophyletic pattern of strain variation observed when every necrogenic strain was included in the same analysis. Individual branch lengths are presented above each respective branch and the values following in parentheses denote the number of unambiguous substitutions. Bootstrap values are reported below each node in italics. The brackets to the right of each clade define the locations of the four necrogenic clades (I–IV) as well as the positions of the three soft-rotting clades (A–C).

bacterium each into two distinct clusters (Fig. 5). Furthermore, these clusters were distributed in different locations on the tree and failed to form monophyletic lineages for each of these two genera. This observation would indicate that while the *gapDH* molecule does accurately reflect the taxonomic partitions of *Erwinia* species, it does not fully reconstruct the monophyletic lineages of *Brenneria* or *Pectobacterium* as defined by the 16S rRNA locus.

DISCUSSION

This study has demonstrated the phylogenetic divergence of necrogenic members of *Erwinia* and *Brenneria* as discerned by *gapDH* nucleotide sequence evolution. Cladistic analysis of these necrogenic phyto-

pathogens gave rise to four robust and distinct clades which, when examined in conjunction with other species, were found to be taxonomically disrupted by several clades of soft-rotting *Pectobacterium* and *Erwinia* strains. In addition, *inter se* genetic distances for necrogenic clades were as great as genetic distances between necrogenic and soft-rotting clades. These evolutionary findings, which attempt to explain the radiation of necrogenic *Erwinia* and *Brenneria* species, subvert classical dogma for the taxonomy of the genus which has long supported a dichotomous split between those erwinias that induce necrosis and those that induce soft-rot (Dye, 1981; Lelliott & Dickey, 1985). These historic divisions remained central to the classification of members of the genus for nearly three decades (Starr & Chatterjee, 1972; Brenner & Falkow,

Table 3. GapDH residues unique to *Erwinia* and *Brenneria* species and clades

Species/clades defined	Clade no.	Position no.*	Residue
1. <i>Erwinia amylovora</i>	I	35†	A
2. <i>Erwinia mallotivora</i> / <i>psidii/tracheiphila</i>	II	36†	I
		91†	A
3. <i>Brenneria alni</i> / <i>nigrifluens/rubrifaciens</i> / <i>salicis</i>	IV	36†	A
		86†	D
4. ' <i>Brenneria lupinicola</i> '/ <i>Brenneria quercina</i>	III	59†	T
5. <i>Erwinia psidii</i> / <i>tracheiphila</i>	II	3	E
6. <i>Brenneria rubrifaciens</i> / <i>salicis</i>	IV	102	V
		147	R
7. <i>Erwinia tracheiphila</i>	II	2	S
		4	R
		15	A
		32	N
		41	F
		42	F
		63	N
		76	A
		95	D
		98	W
		108	Y
		128	N
8. <i>Brenneria nigrifluens</i>	IV	88	E
9. <i>Brenneria rubrifaciens</i>	IV	14	I
		117	A
10. <i>Brenneria salicis</i>	IV	11	S
		12	Y
		24	E
		71	P
		85	A
		107	T
		125	N
		127	I
		139	Y

*Determined in the CLUSTAL W multiple sequence alignment of 64 computer-translated GapDH amino acid sequences.

†Signature synapomorphic residues that define an entire necrogenic clade.

1971). Rather, these results support the more recent 16S rDNA partitioning of the genus *Erwinia* which clearly demonstrates that the original species dichotomy has little evolutionary validity (Hauben *et al.*, 1998). DNA sequence analysis of small-subunit ribosomal genes have indicated that necrogenic species are distributed into at least two separate lineages within the *Enterobacteriaceae* (Hauben *et al.*, 1998; Kwon *et al.*, 1997). The phylogenetic conclusions drawn from *gapDH* sequence analysis are consistent with these

findings and have served to underscore the phylogenetic inaccuracy of the 'amylovora' group as a valid operational taxonomic unit.

While all of the *gapDH* parsimony trees definitively supported a polyphyletic distribution for necrogenic *Erwinia* and *Brenneria* species, it was noted that statistical support for these lineages was not always high. Although bootstrap values were instrumental in supporting more terminal relationships in the trees, the tests generally failed to support structure among the deep nodes with any measurable levels of confidence. While this observation does not refute the *gapDH* phylogenetic hypothesis, it does suggest that this deep node resolution is supported by a limited number of phylogenetic characters, and furthermore, would account for the loss of deep node hierarchy during bootstrap replicates. The most likely explanation for this outcome was the omission of these characters in subsequent bootstrap iterations during the test, which is designed to generate and evaluate a series of random submatrices to quantify nodal support (Felsenstein, 1985). Regardless of this observation, however, the single most parsimonious phylogenetic hypothesis generated from these data remains one that supports deep node resolution and necrogen polyphyly.

A phylogenetic mapping approach allowed examination of the evolutionary history of the two opposing pathological states found among *Erwinia*, *Brenneria* and *Pectobacterium* species in the context of an independent phylogenetic measure, as derived from *gapDH* sequences. While three different optimization solutions were explored, the scenario which assumed necrosis as the ancestral state with soft-rot arising concomitantly along three separate lineages of pectolytic bacteria was recognized as the most parsimonious mapping solution. This conclusion is further supported by two separate lines of evidence. First, when the soft-rotting phenotype is mapped to the 16S rDNA tree originally reported by Hauben *et al.* (1998), it also demonstrates a polyphyletic distribution indicative of multiple evolutionary events giving rise to a similar pathogenicity trait. Second, this observation is supported by examining the prevalence of specific pathogenicity systems employed among other enterobacterial taxa. Necrogenic *Erwinia* and *Brenneria* species rely on type III secretion systems to export effector proteins onto the plant surface, thereby initiating host-induced hypersensitivity on non-host and pathology on host organisms (Barinaga, 1996). This conserved mechanism, as well as the genes which comprise it, has been adopted by a diverse group of enterobacterial pathogens that vary widely in their host affinities to include a variety of plant and animal species (Galan & Collmer, 1999; Brown *et al.*, 1998). From an evolutionary perspective, the observation that this pathogenicity system is functionally conserved and widespread among a diverse group of enterobacterial species is entirely consistent with its position as the ancestral pathological trait in the *gapDH*-based molecular phylogeny.

Fig. 5. Phylogenetic relationships among *Erwinia*, *Brenneria* and *Pectobacterium* species in light of recent 16S rDNA clustering of these genera. The topology shown was the most parsimonious tree that resulted from analysis of *gapDH* nucleotide sequences generated from this study along with the remaining enterobacterial *gapDH* loci currently available in GenBank. The tree was drawn to scale with an overall length of 1083 nt substitutions. The brackets to the right indicate the clusters of phytopathogenic species as revealed by 16S rDNA nucleotide sequences (Hauben *et al.*, 1997). *GapDH* sequences obtained from this study are depicted in bold type. The GenBank sequence accession numbers are given in brackets to the right of each enteric reference species.

The taxonomic utility of the *gapDH* gene has been documented in several other enterobacterial genera (Lawrence *et al.*, 1991). Levels of taxonomic congruence exhibited between the *gapDH* and 16S rDNA phylogenies varied between the three phytopathogenic genera examined here. While our data largely reiterated the proposed constitution of *Erwinia*, 16S rDNA species compositions of the two novel genera were not completely supported in the context of *gapDH* sequence evolution. *Pectobacterium* was precluded by the fact that *Pectobacterium chrysanthemi* was removed from the *Pectobacterium carotovorum* clade, forming a separate soft-rotting lineage further up the tree along with *Pectobacterium cacticida*. Small-subunit RNA sequences grouped *Brenneria quercina* with a clade containing *Brenneria rubrifaciens*, while

the *gapDH* phylogeny placed *Brenneria quercina* in a distinct necrogenic clade (*gapDH* clade IV) along with its sister taxon '*Brenneria lupinicola*' (see Figs 1, 4 and 5). These subtle taxonomic discrepancies observed between *gapDH* and rDNA molecules may be interpreted two different ways. First, the *gapDH* gene may be subject to some level of mutational saturation among these taxa. This conclusion is buttressed by the observation that homoplasy is slightly elevated in the *gapDH* nucleotide tree (CI = 0.59). However, it should be noted that not all homoplasy indices supported this explanation. The retention index (RI = 0.91), which corrects for uninformative characters in the form of autapomorphies, suggested that homoplasy did not significantly contribute to internal tree structure for this gene. The alternative explanation implicates the

small-subunit ribosomal genes as rDNA molecules may simply lack sufficient genetic variation to resolve these closely related phytopathogenic species. This observation might be expected for bacteria that have diverged only recently in eubacterial evolution (Ahern, 1994). Whatever the final explanation, supplemental analyses, including the phylogenetic examination of additional protein-coding loci, should aid in clarifying the extent to which the *gapDH* molecule may be useful as a molecular taxonomic marker for studying enterobacterial evolution (Brown *et al.*, 1996; Brown & van der Zwet, 1998).

Other agriculturally significant proteobacterial genera, such as *Xanthomonas* and *Pantoea*, exist that maintain well defined taxonomies as a result of small-subunit rRNA analyses (Hauben *et al.*, 1997, 1998; Mergaert *et al.*, 1993). It remains to be seen to what extent *gapDH* diversity can reiterate these relationships. In this example, *gapDH* nucleotide sequences uncovered remarkable levels of genetic diversity in several phytopathogenic genera in the family *Enterobacteriaceae*. This provided an opportunity to investigate the phylogenetic utility of a moderately evolving protein-coding gene for the discrimination of a commercially and agriculturally significant group of enterobacterial taxa.

ACKNOWLEDGEMENTS

We are very grateful to Drs Jose Lopez, Domenic Castignetti, Marc Allard and John Wells for their useful discourse; to Drs S. V. Beer and J. Wells for kindly contributing several isolates; and to Loyola University Information Technologies and The Frederick Biomedical Supercomputing Center (NCI-FCRDC) for allotting required computer time used for sequence analyses. This work was supported in part by a Loyola University intramural research award and by the USDA, Agricultural Research Service's Cooperative Agreement Program.

REFERENCES

- Ahern, H. (1994). Does a molecular clock tick for bacteria? *Am Soc Microbiol News* **60**, 600–603.
- Barinaga, M. (1996). A shared strategy for virulence, *Science* **272**, 1261–1263.
- Brenner, D. J. & Falkow, F. (1971). Molecular relationships among members of the *Enterobacteriaceae*, *Adv Genet* **16**, 81–118.
- Brown, E. W. & van der Zwet, T. (1998). Phylogenetic analysis of the tribe *Erwiniae* reveals evolutionary polyphyly among isolates which induce necrosis, *Phytopathology* **88**, S128–S129.
- Brown, E. W., Janisiewicz, W. & van der Zwet, T. (1996). Nucleotide sequence analysis of the *Erwinia amylovora omp3A* gene: molecular genetic differentiation of the fire blight bacterium and other necrogenic *Erwinia* species, *Phytopathology* **86**, 89.
- Brown, E. W., Allard, M. W. & van der Zwet, T. (1998). Phylogenetic characterization of the eubacterial *lcrD* gene family: molecular evolutionary aspects of pathogen-induced hypersensitivity in plants, *Cladistics* **14**, 49–62.
- Devereux, J., Haeberli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX, *Nucleic Acids Res* **12**, 387–395.
- Dye, D. W. (1968). A taxonomic study of the genus *Erwinia* I: the 'amylovora' group, *N Z J Sci* **11**, 590–607.
- Dye, D. W. (1969). A taxonomic study of the genus *Erwinia* II: the 'carotovora' group, *N Z J Sci* **12**, 81–97.
- Dye, D. W. (1981). A numerical taxonomic study of the genus *Erwinia*, *N Z J Agric Res* **24**, 223–231.
- Ehrlich, H. A., Gelford, D. & Sninsky, J. J. (1991). Recent advances in the polymerase chain reaction, *Science* **252**, 1643.
- Elrod, R. P. (1941). Biochemical and serological studies of the *Erwiniae*, *Bot Gaz* **103**, 123–131.
- Farris, J. S. (1983). The logical basis of phylogenetic analysis. In *Proceedings of the 2nd Meeting of the Willi Hennig Society. Advances in Cladistics* 2, pp. 7–36. Edited by N Platnick & V. Funk. New York: Columbia University Press.
- Farris, J. S. (1989). The retention index and the rescaled consistency index, *Cladistics* **5**, 417–419.
- Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap, *Evolution* **39**, 783–791.
- Forey, P. L., Humphries, C. J., Kitching, I. L., Scotland, R. W., Siebert, D. J. & Williams, D. M. (1992). *Cladistics: a Practical Course in Systematics*. Oxford: Clarendon Press.
- Fukunaga, M. & Koreki, Y. (1996). A phylogenetic analysis of *Borrelia burgdorferi sensu lato* isolates associated with Lyme disease in Japan by flagellin gene sequence determination, *Int J Syst Bacteriol* **46**, 416–421.
- Galan, J. E. & Collmer, A. (1999). Type III secretion machines: bacterial devices for protein delivery into host cells, *Science* **284**, 1322–1328.
- Glassick, T., Giffard, P. & Timms, P. (1996). Outer membrane protein 2 gene sequences indicate that *Chlamydia pecorum* and *Chlamydia pneumoniae* cause infections in koalas, *Syst Appl Microbiol* **19**, 457–464.
- Goto, M. (1976). *Erwinia mallotivora*, the causal organism of bacterial leaf spot of *Mallotus japonicus*, *Int J Syst Bacteriol* **26**, 467–473.
- Hauben, L., Vauterin, L., Swings, J. & Moore, E. R. (1997). Comparison of 16S ribosomal DNA sequences of all *Xanthomonas* species, *Int J Syst Bacteriol* **47**, 328–335.
- Hauben, L., Moore, E. R., Vauterin, L., Steenackers, M., Mergaert, J., Verdonck, L. & Swings, J. (1998). Phylogenetic position of phytopathogens within the *Enterobacteriaceae*, *Syst Appl Microbiol* **21**, 384–397.
- Hillis, D. M. & Huelsenbeck, J. P. (1992). Signal, noise, and reliability in molecular phylogenetic analysis, *J Hered* **83**, 189–195.
- Jukes, T. H. & Cantor, C. R. (1969). Evolution of protein molecules. In *Mammalian Protein Metabolism*, pp. 21–132. Edited by H. N. Munro. New York: Academic Press.
- Kwon, S. W., Go, S. J., Kang, H. W., Ryu, J. C. & Jo, J. K. (1997). Phylogenetic analysis of *Erwinia* species based on 16S rRNA gene sequences, *Int J Syst Bacteriol* **47**, 1061–1067.
- Lawrence, J. G., Ochman, H. & Hartl, D. L. (1991). Molecular and evolutionary relationships among enteric bacteria, *J Gen Microbiol* **137**, 1911–1921.
- Lelliott, R. A. & Dickey, R. S. (1985). Genus VII *Erwinia*. In *Bergey's Manual of Systematic Bacteriology*, pp. 469–476. Edited by N. R. Krieg & J. G. Holt. Baltimore: Williams & Wilkins.
- Liaud, M. F., Valentin, C., Martin, W., Bouget, F. Y., Kloareg, B. & Cerff, R. (1994). The evolutionary origin of red algae as deduced from the nuclear genes encoding cytosolic and chloroplast

- glyceraldehyde-3-phosphate dehydrogenases from *Chondrus crispus*, *J Mol Evol* **38**, 319–327.
- Lloyd, A. T. & Sharp, P. M. (1993). Evolution of the *recA* gene and the molecular phylogeny of bacteria, *J Mol Evol* **37**, 399–407.
- Ludwig, W. & Schleifer, K. H. (1999). Phylogeny of bacteria beyond the 16S rRNA standard, *ASM News* **65**, 752–757.
- Maniatis, T., Fritsch, E. & Sambrook, J. (1982). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Mergaert, J., Verdonck, L. & Kersters, K. (1993). Transfer of *Erwinia ananas* and *Erwinia stewartii* to the genus *Pantoea* and description of *Pantoea stewartii* ssp. *indologenes*, *Int J Syst Bacteriol* **43**, 162–173.
- Neto, J. R., Robbs, C. F. & Yomashiro, T. (1987). A bacterial disease of guava caused by *Erwinia psidii* sp. nov., *Fitopatol Bras* **12**, 345–350.
- Ochman, H. & Wilson, A. C. (1988). Evolution in bacteria: evidence for a universal substitution rate in cellular genomes. *J Mol Evol* **26**, 74–86.
- Prescott, L. M., Harley, J. P. & Klein, D. A. (1993). *Microbiology*, 2nd edn. Dubuque, IA: William C. Brown Publishers.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors, *Proc Natl Acad Sci USA* **74**, 5463–5467.
- Schlaeper, B. S. & Zuber, H. (1992). Cloning and sequencing of the genes encoding glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate kinase, and triose phosphate isomerase (*gap* operon) from mesophilic *Bacillus megaterium*: comparison with corresponding sequences from thermophilic *Bacillus* species, *Gene* **122**, 53–63.
- Sharp, P. M. & Li, W. H. (1987). Rates of synonymous substitution in enterobacterial genes is inversely related to codon usage bias, *Mol Biol Evol* **4**, 222–230.
- Starr, M. P. & Chatterjee, A. K. (1972). The genus *Erwinia*: enterobacteria pathogenic to plants and animals, *Annu Rev Microbiol* **26**, 389–426.
- Surico, G., Mugnai, L., Pastorelli, R., Giovannetti, L. & Stead, D. E. (1996). *Erwinia alni*, a new species causing bark cankers of Alder (*Alnus Miller*) species, *Int J Syst Bacteriol* **46**, 720–726.
- Swofford, D. L. (1991). Phylogenetic analysis using parsimony (PAUP), version 3.11. Washington, DC: The Smithsonian Institution.
- Swofford, D. L. & Olsen, G. J. (1990). Phylogeny reconstruction. In *Molecular Systematics*, pp. 411–501. Edited by D. M. Hillis & C. Moritz. Sunderland, MA: Sinaur & Associates.
- Thompson, J. D., Higgins, D. & Gibson, T. J. (1992). CLUSTAL W: a novel multiple sequence alignment program, *Nucleic Acids Res* **22**, 4673–4680.
- Vanneste, J. L. (1995). *Erwinia amylovora*. In *Pathogenesis and Host Specificity in Plant Diseases: Histopathological, Biochemical, Genetic and Molecular Bases. Vol. 1. Prokaryotes*, pp. 21–41. Edited by U. S. Singh, R. P. Singh & K. Kohmoto. Oxford: Pergamon.
- Winslow, C. E., Broadhurst, J. & Buchanan, R. E. (1920). The families and genera of the bacteria *Erwiniae*, *J Bacteriol* **5**, 209.
- Woese, C. R. (1987). Bacterial evolution, *Microbiol Rev* **51**, 221–271.
- van der Zwet, T. & Keil, H. L. (1979). *Fire Blight, a Bacterial Disease of Rosaceous Plants*. United States Department of Agriculture Handbook no. 510.