Psychothertolerant spore-forming bacteria represent a major challenge to the goal of extending the shelf life of pasteurized dairy products. The objective of this study was to identify prominent phylogenetic groups of dairy-associated aerobic sporeformers and to characterize representative isolates for phenotypes relevant to growth in milk. Analysis of sequence data for a 632-nucleotide fragment of rpoB showed that 1,288 dairy-associated isolates (obtained from raw and pasteurized milk and from dairy farm environments) clustered into two major divisions representing (i) the genus *Paenibacillus* (737 isolates, including the species *Paenibacillus odorifer*, *Paenibacillus graminis*, and *Paenibacillus amylyticus* sensu lato) and (ii) *Bacillus* (*n* = 467) (e.g., *Bacillus licheniformis* sensu lato, *Bacillus pumilus*, *Bacillus weihenstephanensis*) and genera formerly classified as *Bacillus* (*n* = 84) (e.g., *Viridibacillus* spp.). When isolates representing the most common rpoB allelic types (ATs) were tested for growth in skim milk broth at 6°C, 6/9 *Paenibacillus* isolates, but only 2/8 isolates representing *Bacillus* subtypes, grew ≥5 log CFU/ml over 21 days. In addition, 38/40 *Paenibacillus* isolates but only 3/47 *Bacillus* isolates tested were positive for β-galactosidase activity (including some isolates representing *Bacillus licheniformis* sensu lato, a common dairy-associated clade). Our study confirms that *Paenibacillus* spp. are the predominant psychrotolerant sporeformers in fluid milk and provides 16S rRNA gene and rpoB subtype data and phenotypic characteristics facilitating the identification of aerobic spore-forming spoilage organisms of concern. These data will be critical for the development of detection methods and control strategies that will reduce the introduction of psychrotolerant sporeformers and extend the shelf life of dairy products.

Members of the genus *Paenibacillus*, once considered group 3 bacilli (8), appear to occupy diverse ecological niches and have been isolated from various sources, including soil (60, 67, 99), rhizosphere (63, 96), honeybee larvae (5, 31), compost (2, 93), humans (76), and cow feces (95). *Paenibacillus* spp. have also been isolated from dairy products, including raw milk (18, 78), various pasteurized foodstuffs (25, 33, 39), and even commercial ultrahigh-temperature (UHT)-treated milk (79), suggesting that at least some *Paenibacillus* isolates can survive short-time heat treatments over 100°C. Although *Paenibacillus* persistence on processing equipment (e.g., fillers) has not been established, certain *Paenibacillus* spp. have been shown to produce exopolysaccharide (2) or to form biofilms (89), which, if present in appropriate locations, may lead to postpasteurization contamination of fluid milk. Consistent with this, at least one study has reported evidence of *Paenibacillus* contamination of fluid milk originating from in-plant sources (41). Overall, the presence of *Paenibacillus* in farm and processing environments suggests a number of different potential sources of fluid milk contamination with these organisms (40). While some studies have provided information on dairy-associated *Paenibacillus* species and subtypes (18, 72, 78), a general lack of information on the ecology and diversity of dairy-associated *Paenibacillus* spp., including the lack of specific detection methods for common psychrotolerant *Paenibacillus* spp., has
limited the ability to develop control strategies, in both milk production and processing, for this increasingly important group of spoilage organisms (72).

The goal of this study was to identify and characterize prominent psychrotolerant sporeformers in dairy processing systems. To this end, we used DNA sequence-based approaches (i.e., maximum-likelihood [ML] phylogenetic analysis of partial rpoB and 16S rRNA gene sequence data) to systematically identify and classify a large set of isolates (most of which have been described previously) representing dairy-associated Gram-positive sporeformers. Isolates representing specific clades and rpoB allelic types (ATs) commonly associated with pasteurized milk spoilage were targeted as relevant phenotypes (i.e., growth in milk at refrigeration temperatures and β-galactosidase activity). A comprehensive maximum-likelihood phylogenetic analysis of this large set of dairy-associated sporeformer isolates, which until recently was computationally prohibitive, will provide a better understanding of fluid milk spoilage due to Gram-positive sporeformers and will provide new insights into sporeformer diversity and ecology in dairy systems. The results of this study will facilitate the development of strategies to reduce food spoilage by sporeforming bacteria in different food systems, including the development of specific DNA-based detection systems.

MATERIALS AND METHODS

Isolate collection and selection. Of the 1,288 isolates used for the study reported here (see Table S2 in the supplemental material), 1,279 have been described previously (25, 39–41, 71, 73). As detailed in these previous studies, isolates were obtained from raw milk, environmental samples collected on dairy farms (e.g., feed, bedding materials, manure, soil, and milking parlor wash water), and pasteurized milk tested over its shelf life by using standard methods for the examination of dairy products (24), including (i) spore counts (i.e., heat treatment at 80°C for 12 min, followed by isolation on standard plate count (SPC) agar plates incubated at 32°C) on raw and pasteurized milk and (ii) in lactoperoxidase counts. Typically, colonies representing each visually distinct morphology (ranging from 1 to 10 colonies per sample) were selected, streaked for purity on brain heart infusion (BHI) agar (BD, Franklin Lakes, NJ), characterized for the Gram reaction by using a 3-step Gram stain kit (Becton, Dickinson and Co., Sparks, MD), and subsequently frozen at −80°C in 15% glycerol. Only isolates representing Gram-positive sporeformers were included in the study reported here. In addition to the isolates reported previously, eight farm isolates and one pasteurized milk isolate not previously reported were included in the study reported here because they represented unique, previously unreported rpoB ATs. Overall, the 1,288 isolates included here were obtained from raw milk (n = 201), dairy farm environments (n = 85), and HTST pasteurized milk (n = 1,002), which included in-line (n = 213) and packaged (n = 789) products. All isolates were obtained from samples representing the U.S. dairy system, with the majority of isolates (73.8%) obtained from milk that was produced or processed in New York State.

Lysate preparation. Lysates for PCR were prepared, from overnight cultures grown in BH1 at 32°C, as described by Furrer et al. (29) with slight modifications. Briefly, 250 μl of overnight culture was centrifuged at 13,000 rpm for 10 min, and pellets were resuspended in 95 μl of 1× PCR buffer (Promega, Madison, WI). Lysozyme was added to achieve a final concentration of 2.0 to 2.5 mg/ml. After 15 min of incubation at room temperature, 1 μl of a proteinase K solution (20 mg/ml) was added, and the mixture was incubated at 58°C for 1 h. Enzymes were subsequently inactivated by heating at 95°C for 8 min.

rpoB sequencing. Molecular typing of all isolates was performed based on the DNA sequence data for a 632-nucleotide (nt) fragment of rpoB, which encodes the beta subunit of RNA polymerase, as described previously (41). Briefly, the rpoB fragment was amplified using previously described PCR primers (23) and PCR conditions (25). rpoB PCR products were purified using the QIAquick PCR purification kit (Qiagen Inc., Valencia, CA) and were quantified with a NanoDrop ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, DE). Bidirectional sequencing with PCR primers was performed at Cornell University’s Life Sciences Core Laboratory Center (Ithaca, NY) using the ABI 3730 DNA analyzer (Applied Biosystems, Foster City, CA). DNA sequences were assembled and proofread in SeqMan (Lasergene; DNASTar, Madison, WI), and high-quality, double-stranded sequence data were used for further analyses. rpoB sequence data for 1,279 isolates had been reported in previous publications by our group (25, 28, 39–41, 72, 73).

Sequences were aligned in MegAlign (Lasergene), and 632-nt rpoB fragments (25), corresponding to nt 2455 to 3086 of the 3,534-nt rpoB open reading frame of Bacillus cereus ATCC 10987 (GenBank accession number AE017194; locus tag BCE_0102), were used for subsequent analyses. Partial rpoB sequencing was used, because its discriminatory power allows for the differentiation of isolates beyond the species level (25) and because this approach is more economical than most banding pattern-based methods, such as ribotyping or pulsed-field gel electrophoresis.

AT assignment. rpoB allelic types (ATs) were assigned essentially as described by Huck et al. (41), using BioEdit Sequence Alignment Editor, version 7.0.9.0 (34). A unique rpoB AT was assigned to every gene sequence that differed from any previously obtained sequence by one or more nucleotides. The first isolate of each new rpoB AT was designated the reference strain for that AT; partial 16S rRNA gene sequencing was performed for each AT reference strain, as described below, to facilitate species identification.

Sequencing of 16S rRNA genes. A 700-nt segment of the 16S rRNA gene was amplified as described previously (25, 28) using primers PE7 (75) and DG74 (28). Subsequent DNA sequencing of PCR products was performed as described previously (41) using primers PE7 and P3SH (70). 16S rRNA gene sequences for 274 isolates representing different rpoB ATs have been reported previously (25, 28, 39–41, 72, 73); 16S rRNA gene sequences for isolates representing the other 9 rpoB ATs were determined as part of the study reported here. The isolates representing these previously unreported rpoB ATs were from farm samples (8 isolates; ATs 280 to 287) and from pasteurized milk (1 isolate; AT288). If forward and reverse sequences indicated the presence of two nucleotides at a given position, indicating chromosomal rRNA operons with different sequences within a given isolate (53), 16S rRNA gene sequences were reported with appropriate nucleotide ambiguity codes as described by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology. 16S rRNA gene sequence alignments were performed using MegAlign (Lasergene), and sequences for each isolate were trimmed to correspond to a 616-nt fragment (nt 823 to 1438) of the 1,508-nt 16S rRNA gene in B. cereus ATCC 10987 (GenBank accession number AE017194; locus tag BCE_5738) (41).

Alignment, tree construction, and species identification. An rpoB maximum-likelihood (ML) phylogenetic tree was constructed using the rapid maximum-likelihood algorithm RAxML (84) with rapid bootstrapping (100 bootstrap replicates). Because of the absence of an appropriate outgroup, the rpoB tree was midpoint rooted. rpoB ATs were grouped according to their phylogenetic positions; only clades with bootstrap support (BS) values of >70 were considered well supported.

For species identification, partial 16S rRNA gene sequences for isolates representing each unique rpoB AT were queried against type strain 16S rRNA gene sequences using the "Seqmatch" function in the Ribosomal Database Project (RDP) database (17). To confirm species identifications, we also constructed, using RAXML, a maximum-likelihood phylogenetic tree containing partial 16S rRNA gene sequences for (i) each unique 16S rRNA gene AT identified among the isolates representing the 283 unique rpoB ATs and (ii) relevant type strains obtained from the RDP. Partial 16S rRNA gene sequences for three different Staphylococcus species (i.e., Staphylococcus simiae, Staphylococcus aureus, and Staphylococcus lunata).
were included as an outgroup. Both RDP similarity scores (percentage of sequence identity over all pairwise comparable positions [17]) and the 16S rRNA gene phylogeny were used to assign species identifications (IDs) to all 283 rpoB ATs. An isolate with a similarity score of ≥99% against a type strain was assigned the species ID of that type strain; for isolates that had similarity scores of ≥99% against more than one type strain and grouped with more than one type strain in the 16S rRNA gene tree, the “sensu lato” notation was used to indicate that the 16S rRNA gene sequence showed a high level of similarity with multiple closely related species. For example, the partial 16S rRNA gene sequence for the isolate representing rpoB AT212 matched both Bacillus subtilis (99%) and Bacillus valivistirmortis (99%) and hence was assigned the species ID Bacillus subtilis sensu lato. For isolates that showed <98% sequence similarity but grouped with one or more type strains in the 16S rRNA gene tree, the “confer” (cf.) notation was used to denote taxonomic uncertainty. For isolates that showed identity scores of <98% and that did not group with any type sequences in the 16S rRNA gene tree, the AT was assigned a genus but no species (e.g., Paenibacillus sp. clade 1), indicating that these isolates could not be assigned to a species; as multiple clades with such isolates were identified, these clades were also given numerical identifiers (e.g., clade 1, clade 2).

Cold growth. For selected rpoB clades that included a considerable number of dairy-associated isolates, an isolate representing the most common AT in the clade was chosen for cold growth analysis. These isolates were plated on BHI agar and were incubated at 32°C overnight. A single colony was then inoculated into 5 ml of BHI broth. After aerobic incubation (agitation at 200 rpm) at 32°C for 18 to 24 h, 1 ml of this culture was pelleted at 13,000 rpm for 10 min, followed by resuspension of the cell pellet in 1 ml of phosphate buffer. A 1-ml volume of an appropriate serial dilution of this culture was used to inoculate 9 ml of sterile skim milk broth (SMB) for a final inoculum level of 102 CFU/ml. SMB samples were plated on BHI agar and were incubated at 32°C overnight. A single colony was then inoculated into 5 ml of BHI broth. After aerobic incubation, as well as after 6, 10, 13, 17, 20, and 24 days of incubation at 6°C, the cultures were inoculated, as well as after 6, 10, 13, 17, 20, and 24 days of incubation at 6°C. One without an overlay of 100 µl solution of bromo-chloro-indolyl-galactopyranoside (X-Gal), followed by incubation at 32°C for 24 h. Blue colonies on the plates containing X-Gal were indicative of β-galactosidase activity. For evaluation of β-galactosidase activity, bacterial cultures were streaked onto two BHI agar plates, one with and one without an overlay of 100 µl of a 40-µg/ml solution of bromo-chloro-indolyl-galactopyranoside (X-Gal), followed by incubation at 32°C for 24 h. Blue colonies on the plates containing X-Gal were indicative of β-galactosidase activity. A phylogenetic clade was considered β-galactosidase positive if all representative isolates tested from that clade were positive. A clad was considered “β-galactosidase variable” if some isolates from the clad were positive and others were negative. Isolates that showed weak β-galactosidase activity were designated weakly positive.

RESULTS
Dairy-associated sporeformers represent two major phylogenetic divisions, one representing the genus Paenibacillus and the other including the genus Bacillus and related genera. An overall analysis of rpoB sequence data for 1,288 dairy-associated aerobic sporeformer isolates from pasteurized and raw milk (25, 39–41, 71, 73) and from dairy farm environments (40) identified 283 unique rpoB allelic types (ATs), including 274 that had been reported previously (25, 39–41, 71, 73). The nine new ATs identified here represent Psychrobacillus spp. (AT280 and AT283 to AT286), Bacillus subtilis sensu lato (AT282), Bacillus clausii (AT287), Bacillus psycrosacharolyticus (AT281), and a Bacillus sp. closely related to Bacillus circulans (AT288). The isolates representing these new ATs were isolated from packaged pasteurized milk (AT288) and from farm samples such as manure (AT281, AT283, AT284), soil (AT285 to AT287), and water (AT282).

To further probe the diversity and relatedness of all isolates, we constructed a maximum-likelihood (ML) phylogenetic tree based on an alignment of sequences representing all 283 unique rpoB ATs. The overall alignment revealed a total of 330 polymorphic sites among the 632 nt aligned. Analysis of the rpoB alignment with DNAsp (59) showed an overall per site nucleotide diversity (π) of 0.213 and an average number of nucleotide differences (κ) of 134.44. Analysis for horizontal gene transfer, performed by calculating the dxy statistic (15), revealed no evidence for lateral gene transfer among these sequences (P = 0.168).

The ML tree of the 283 unique partial rpoB sequences revealed a primary division into two major phylogenetic groups. One of these divisions (Fig. 1) represents Bacillus and closely related genera (such as Solibacillus, Luzinibacillus, and Psychrobacillus), while the other division represents isolates that cluster with the genus Paenibacillus (Fig. 2). Within each of these two divisions, we identified monophyletic clades that represent major phylogenetic groups (i.e., groups I to IV in the Bacillus division and groups V to XI in the Paenibacillus division). Overall, while both the 16S rRNA gene and rpoB trees supported the same well-supported clades, differences between the phylogenies were generally found where bootstrap support was low or lacking.

Isolates in the Bacillus division represent Bacillus spp., as well as one clad representing non-Paenibacillus genera that were formerly classified in the genus Bacillus. The rpoB phylogeny (Fig. 1) shows that sequences in the Bacillus division can be further separated into two major subdivisions: (i) a well-supported (bootstrap support [BS], 90) subdivision consisting of species that were formerly classified in the genus Bacillus (Fig. 1, group IV) but are now considered to belong to different genera (e.g., Viridibacillus) and (ii) a second subdivision consisting of Bacillus spp. including ATs classified in the genus Bacillus (Fig. 1, groups I, II, and III) and other Bacillus spp. that do not represent clear clades, including two ATs classified as Oceanobacillus. Overall, of the 150 rpoB ATs in the Bacillus division, 132 ATs are in the four main groups (i.e., groups I to IV) and 18 ATs were not assigned to groups.

Group I is a phylogenetically well supported group (BS, 84) representing B. subtilis and related species and composed of 324 isolates representing 67 unique rpoB ATs (Fig. 1). Isolates within this group were identified as Bacillus safensis, Bacillus pumilus, and Bacillus aerophilus sensu lato (B. aerophilus sensu lato includes B. aerophilus, Bacillus stratosphericus, and Bacillus alitidinis), as well as members of the “Bacillus subtilis species complex,” which includes B. subtilis, Bacillus mojavensis, Bacillus valivistirmortis, and Bacillus licheniformis (74). The Bacillus licheniformis sensu lato clad was the second most frequently isolated clad, containing 8 ATs that represent 188 (14.6%) of the 1,288 dairy-associated isolates characterized here (Fig. 1). 16S rRNA gene sequences for the 8 B. licheniformis sensu lato ATs showed >99% 16S rRNA gene similarity to B. licheniformis, Bacillus aeriis, and Bacillus sonorenis; 16S rRNA gene phylogeny further confirmed the similarity of these 8 ATs to these closely related species (see Fig. S1 in the supplemental material). The B. licheniformis sensu lato clad could be further divided into three well-supported subgroups that were designated “B. licheniformis sensu lato subgroups 1, 2, and 3.” B. licheniformis sensu lato subgroup 1 represents four different rpoB ATs (Fig. 1), including AT001, which was the second most frequently isolated rpoB AT.

Group II is a well-supported group (BS, 99) composed of 81 isolates representing 22 rpoB ATs. Based on 16S rRNA gene sequence data, isolates in this group were identified as species belonging to the B. cereus group, which includes B. cereus, Bacillus thuringiensis, Bacillus weihenstephanensis, Bacillus antarctic, Bacillus pseudomyxoides, and Bacillus mycoides (68). All isolates in this
FIG 1 Midpoint-rooted maximum-likelihood (ML) phylogenetic tree of partial rpoB sequences from Bacillus spp. and related species isolated from pasteurized milk (red), raw milk (blue), and dairy farm environments (green). The scale represents the estimated number of nucleotide substitutions per site. Source 1856 aem.asm.org Applied and Environmental Microbiology
group showed ≥98% 16S rRNA gene sequence similarity to the *B. cereus* type strain. The *rpoB* phylogeny clearly separated group II isolates into one clade that represented *B. weihenstephanensis* and *B. mycoides*. Both 16S rRNA gene (see Fig. S1 in the supplemental material) and *rpoB* (Fig. 1) phylogenies further separated these sequences into a *B. weihenstephanensis* and a *B. mycoides* clade; the RDP type strains for these two species clustered into the appropriate 16S rRNA gene clades. Another well-supported clade in group II included isolates with 16S rRNA gene sequences that had >99% 16S rRNA gene sequence similarity to both *B. cereus* and *B. thuringiensis* and were thus designated *Bacillus cereus* sensu lato. Although *Bacillus cereus* sensu lato isolates represented a wide diversity of sources, all *B. cereus* sensu lato isolates with AT158 came from a single processing plant, and AT158 has been determined previously to be a plant-specific contaminant (73); therefore, only one AT158 isolate was included in the isolate count in this study. While 94% of bacterial isolates from our study came from pasteurized dairy products, the *B. weihenstephanensis* clade includes fewer pasteurized milk isolates (*n* = 17) than raw milk isolates (*n* = 31; obtained from silos (*n* = 28), farm tanks (*n* = 2), and a milk-hauling truck (*n* = 1).

Group III is composed of 18 isolates representing 12 *rpoB* ATs. While this group received very low bootstrap support in our analyses, we kept this group for convenience and because it is supported by other studies (87). Isolates in this group were identified, based on 16S rRNA gene data, as *Bacillus cf. firmus* (4 isolates) and *Bacillus farraginis* (1 isolate). Additional *rpoB* ATs in group III (i.e., *Bacillus* sp. clades 1, 2, and 3) represent species that did not closely match any of the type strains in the RDP database and were distinct from all the 16S rRNA gene type sequences (see Fig. S1 in the supplemental material). The 16S rRNA gene sequence of *Bacillus* sp. clade 1 showed the highest similarity (93%) to the type strain of *Bacillus niacini*, while the 16S rRNA gene sequence of *Bacillus* sp. clade 3 closely matched (95%) the type strain of *Bacillus pocheonensis*. 16S rRNA gene sequence data could not be obtained for *Bacillus* sp. clade 2, and therefore, this isolate could not be assigned to any specific species.

Several small clades representing a diversity of *rpoB* ATs fell outside major groups (i.e., groups I to IV) and showed ambiguous phylogenetic relationships to other groups. These clades were identified as containing *Bacillus gibbonii* (1 isolate), *B. clausii* (4 isolates), *Bacillus barbaricus* (1 isolate), *Bacillus psychrosaccharolyticus* (2 isolates), *Brevibacterium frigoritolerans* (2 isolates), *Bacillus realsonii* (1 isolate), and *Oceanobacillus chironomi* (1 isolate), a distinct genus in the family *Bacillaceae* (62, 88). *Brevibacterium frigoritolerans* was described as a *Brevibacterium* species; however, 16S data clearly show this to be a species that should be placed in the *Bacillaceae*, consistent with previous reports (30). Clades identified as *Bacillus cf. megaterium* (4 isolates) and *Bacillus cf. horikoshii* (1 isolate) also fell outside well-supported major groups. Overall, *Bacillus* isolates that could not be phylogenetically assigned to groups I, II, III, and IV represented 1.5% of all isolates in this study.

Group IV is composed of 84 isolates representing 31 *rpoB* ATs; isolates in this group largely represent recently described genera that were formerly classified as group 2 *Bacillus* spp. (7). High bootstrap support (BS, 90) was observed for group IV (Fig. 1), confirming that these genera, which included *Vitribacillus* spp. (4) (46 isolates), *Lysinibacillus* spp. (3) (9 isolates), *Solibacillus* spp. (55) (5 isolates), *Psychrobacillus* spp. (56) (12 isolates), and a *Paenibacillus* spp. (54) (1 isolate), are distinct from *Bacillus* spp. Although the majority (94%) of bacterial isolates in our study came from raw or pasteurized milk, all *Psychrobacillus* sp. isolates (14 isolates), which represented 10 *rpoB* ATs (Fig. 1), were isolated from animal bedding, soil, and manure samples collected on a single dairy farm.

**Isolates in the division that represents the genus *Paenibacillus* represent 7 major groups, including a number of clades that cannot be assigned a species identification.** The part of the ML tree that represents the *rpoB* sequences for the 737 isolates grouped into the genus *Paenibacillus* showed that these isolates represent seven major groups (groups V to XI; described in more detail below). A number of specific clades consisted of a single species ID based on 16S rRNA gene data (i.e., *Paenibacillus odorifer* clades 1 to 3, *Paenibacillus graminis*, *Paenibacillus cf. peoriae*, and *Paenibacillus amyloyticus* sensu lato), allowing for clear species identification of 677 *Paenibacillus* isolates (i.e., 92% of all *Paenibacillus* isolates). On the other hand, most of these seven major groups also included clades (though typically with 4 or fewer ATs) that could not be assigned a species; these clades were designated *Paenibacillus* sp. clades 1 to 11 (Fig. 2). Overall, of the 133 *rpoB* ATs in the *Paenibacillus* division, 126 ATs are in the seven main groups (i.e., groups V to XI), while 7 *rpoB* ATs were not assigned to groups.

*Paenibacillus* group V is well supported (BS, 98) and includes 506 isolates representing 45 *rpoB* ATs; this group consists of three distinct and well-supported clades that were identified as *P. odorifer* and were designated *P. odorifer* clades 1, 2, and 3. *P. odorifer* was the most frequently isolated species of *Paenibacillus*, representing 68.7% of all *Paenibacillus* isolates, with *P. odorifer* clade 1 containing the most isolates (*n* = 463) (Fig. 2).

*Paenibacillus* group VI consists of one well-supported (BS, 100) clade composed of 8 isolates representing 4 *rpoB* ATs (*Paenibacillus* clade 1) (Fig. 2) that could not be identified to the species level. 16S rRNA gene sequences for the 4 ATs in this group did not show a >99% match to any type strain but showed >97% 16S rRNA gene sequence similarity to both *P. odorifer* and *Paenibacillus borealis*. 16S rRNA gene phylogenetic analysis (see Fig. S2 in the supplemental material) also did not allow for species identification of the isolates in this clade. Thus, this clade appears to represent a taxonomically uncharacterized species.

Group VII comprises 52 isolates representing 23 *rpoB* ATs.
FIG 2 Midpoint-rooted maximum-likelihood phylogenetic tree of partial rpoB sequences from Paenibacillus isolated from pasteurized milk (red), raw milk (blue), and dairy farm environments (green). The scale represents the estimated number of nucleotide substitutions per site. Source information is shown for.
This group includes two clades identified as *P. graminis* (46 isolates), as well as two other clades (i.e., *Paenibacillus* clades 2 and 3) that could not be identified at the specie level. Isolates representative of *rpoB* ATs clustered into *Paenibacillus* clades 2 and 3 showed 16S rRNA gene sequence similarities between 96 and 97% to *P. borealis*, *P. graminis*, and *P. odorifer* type strains.

Group VIII contains 15 isolates representing 7 *rpoB* ATs. Isolates in this group were identified as *Paenibacillus lautus* (4 isolates), *Paenibacillus lactis* (3 isolates), *Paenibacillus rhizosphaericae* (1 isolate), *Paenibacillus glucanolyticus* (6 isolates), and *Paenibacillus cookii* (1 isolate) (Fig. 2; see also Fig. S2 in the supplemental material).

Group IX is comprised of 24 isolates representing 9 *rpoB* ATs. All isolates in this group were designated *Paenibacillus cf. peoriae*. Isolates representing *rpoB* ATs in this group showed ≥97% 16S rRNA gene sequence similarity to *Paenibacillus peoriae*. *Paenibacillus jumilae*, *Paenibacillus kribbensis*, and *Paenibacillus polymyxa*, although 16S rRNA gene phylogenetic analysis showed evidence (BS, >70) that the 16S ATs within group IX may be distinct from any of the type strains (see Fig. S2 in the supplemental material).

Group X comprises 116 isolates representing 33 *rpoB* ATs. For one clade with 101 isolates, the 16S rRNA gene sequences for most ATs showed ≥98% 16S rRNA gene sequence similarity to the closely related species *P. amyloyticus*, *Paenibacillus xylanexedens*, and *Paenibacillus tundrae*, but 16S rRNA gene phylogeny did not allow for discrimination among type sequences or ATs within this clade. Therefore, this clade was identified as *Paenibacillus amylolyticus sensu lato* (Fig. 2). Also within group X is a well-supported clade (BS, 85) consisting of 12 isolates representing 4 *rpoB* ATs. Isolates within this clade showed >98% 16S rRNA gene sequence similarity to *Paenibacillus xylanilyticus*, although representatives of all 4 ATs within this clade also showed >98% sequence similarity to *Paenibacillus paubli* and *Paenibacillus taichungensis*, and 16S rRNA gene phylogeny did not allow for clear species identification (see Fig. S2 in the supplemental material). Therefore, this clade was referred to as *Paenibacillus cf. xylanilyticus* (Fig. 2). Also included in group X was AT192, which showed 96.7% similarity to *P. xylanexedens*, and AT005, which showed 98.1% 16S rRNA gene sequence similarity to *P. xylanilyticus* and *P. taichungensis*. Based on 16S rRNA gene phylogeny, AT005 grouped with *Paenibacillus cf. xylanilyticus* isolates, and AT192 did not group with any of the type strain sequences and was therefore identified as *Paenibacillus sp* clade 9 (see Fig. S2).

Group XI is composed of two well-supported (BS, ≥97) clades with unknown phylogenetic relationships to each other and includes 9 isolates representing 5 *rpoB* ATs. One clade within group XI contains 7 isolates; the isolates representing the three *rpoB* ATs within this clade showed ≥95% 16S rRNA gene sequence similarity to *Paenibacillus sepulcri*, although 16S rRNA gene phylogeny suggested that they represent a distinct, yet uncharacterized species (see Fig. S2 in the supplemental material). Group XI also includes AT156 and AT149; isolates representing these ATs showed 97.9% and 96.9% similarity to *Paenibacillus castaneae*. 16S rRNA gene phylogeny supported the identification of AT156 as *Paenibacillus cf. castaneae*, since the isolate representing AT156 grouped with the *P. castaneae* type strain (see Fig. S2). However, 16S rRNA gene phylogenic analysis showed that the representative AT149 isolate did not group with any of the type strains, indicating that it may represent a distinct uncharacterized species; this isolate was therefore designated *Paenibacillus sp* clade 11.

Two *Paenibacillus* clades and one *Paenibacillus rpoB* AT fell outside major monophyletic groups. One clade consisted of isolates identified as *Paenibacillus mucraens* (AT238), *Paenibacillus sp* clade 6 (AT187), and *Paenibacillus sp* clade 7 (AT266). *Paenibacillus sp* clade 6 showed 97.5% sequence similarity to *Paenibacillus barengoltzii*, and the closest 16S rRNA gene sequence matches to *Paenibacillus sp* clade 7 were *Paenibacillus motobuenensis* (95.7%) and *Paenibacillus alkaliterreae* (95.7%), although 16S rRNA gene phylogeny did not allow for species identification, indicating that these isolates may represent uncharacterized *Paenibacillus* species (see Fig. S2). Finally, *Paenibacillus sp* clade 8, consisting of a single *rpoB* AT (AT057), fell outside major monophyletic groups. The AT057 isolate characterised showed 97.9% 16S rRNA gene sequence similarity to *Paenibacillus provencensis*, consistent with the 16S rRNA gene phylogeny, which also grouped this isolate with *P. provencensis* (see Fig. S1 in the supplemental material).

Representatives from major *Paenibacillus* clades grow in milk at refrigeration temperatures, whereas, with the exception of *B. weihenstephanensis*, representatives from major *Bacillus* clades do not. To evaluate their potential to grow in milk under refrigeration, isolates representing common clades in both the *Bacillus* division and the *Paenibacillus* division were tested for growth in skim milk broth (SMB) over 21 days at 6°C. The *Bacillus* isolates that were tested represented AT001 (*Bacilluslicheniformis sensu lato* clade 1; 2 isolates), AT003 (*B. weihenstephanensis*), AT017 (*Viridibacillus sp*), AT020 (*B. pumilus* clade 1), AT135 (*Bacillus aerophilus sensu lato*), AT141 (*B. safensis*), and AT158 (*Bacillus cereus sensu lato clade 1*). Only two of these eight isolates (i.e., *B. weihenstephanensis* [AT003] and the *Viridibacillus* sp. [AT017]) showed evidence of growth under these conditions; both of these isolates showed >6.0 log CFU/ml growth between day 0 and day 21 (Fig. 3A). The clades to which these two isolates belonged included 51 (*B. weihenstephanensis*) and 46 (*Viridibacillus*) isolates.

The nine *Paenibacillus* isolates tested for growth in SMB at 6°C represented AT015 (*P. odorifer* clade 1), AT023 and AT111 (*Paenibacillus amylolyticus sensu lato*), AT039 (*P. graminis* clade 2), AT045 (*P. graminis* clade 1), AT100 (*Paenibacillus cf. xylanilyticus*), AT157 (*Paenibacillus cf. peoriae*), AT159 (*P. lautus*), and...
AT260 (P. odorifer clade 3). While six of these isolates showed more than 5.0 log CFU/ml growth between 0 and 21 days, one isolate (representing AT159) showed no growth. Two isolates (with AT100 and AT039) showed limited growth (1.98 and 3.28 log CFU between days 0 and 21, respectively [Fig. 3B]).

Most Paenibacillus isolates were positive for β-galactosidase activity, whereas most Bacillus isolates were not. β-Galactosidase catalyzes the hydrolysis of β-galactosidic bonds and thus facilitates growth in milk by catalyzing the breakdown of lactose to glucose and galactose. A total of 87 isolates representing common clades in both the Bacillus (47 isolates representing 39 ATs) and Paenibacillus (40 isolates representing 39 ATs) divisions were tested for β-galactosidase activity. While the isolates selected typically included one isolate representing a common clade, multiple isolates representing a given clade or AT were tested in a few instances to confirm unusual phenotypes. Among the 47 Bacillus isolates tested, only 3 were positive for β-galactosidase activity (i.e., 1 B. nealsonii isolate and 2 of the 5 Bacillus licheniformis sensu lato isolates tested), with another 3 isolates (i.e., 1 B. megaterium isolate, 1 Oceanobacillus chironomi isolate, and 1 of the 5 Bacillus licheniformis sensu lato isolates tested) showing weak β-galactosidase activity (Table 1; see also Table S1 in the supplemental material).

**TABLE 1** Frequency of isolation and β-Gal activity of select rpoB clades isolated more than 10 times

<table>
<thead>
<tr>
<th>Clade ID</th>
<th>Group</th>
<th>No. of isolates in clade</th>
<th>Representative AT</th>
<th>β-Gal activity b (no. of isolates tested)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacillus aerophilus sensu lato</td>
<td>I</td>
<td>24</td>
<td>135</td>
<td>(1)</td>
</tr>
<tr>
<td>Bacillus pumilus clade 1</td>
<td>I</td>
<td>52</td>
<td>072</td>
<td>(1)</td>
</tr>
<tr>
<td>Bacillus safensis</td>
<td>I</td>
<td>30</td>
<td>141</td>
<td>(1)</td>
</tr>
<tr>
<td>Bacillus licheniformis sensu lato clade 1</td>
<td>I</td>
<td>181</td>
<td>001</td>
<td>+ (2); − (2); wp (1)</td>
</tr>
<tr>
<td>Bacillus subtilis sensu lato clade 1</td>
<td>I</td>
<td>17</td>
<td>065</td>
<td>(1)</td>
</tr>
<tr>
<td>Bacillus cereus sensu lato</td>
<td>II</td>
<td>48</td>
<td>059</td>
<td>(1)</td>
</tr>
<tr>
<td>Bacillus weihenstephanensis</td>
<td>IV</td>
<td>51</td>
<td>003</td>
<td>(1)</td>
</tr>
<tr>
<td>Viridibacillus spp.</td>
<td>IV</td>
<td>46</td>
<td>017</td>
<td>(1)</td>
</tr>
<tr>
<td>Paenibacillus odorifer clade 1</td>
<td>V</td>
<td>463</td>
<td>015</td>
<td>(1)</td>
</tr>
<tr>
<td>Paenibacillus odorifer clade 3</td>
<td>V</td>
<td>36</td>
<td>260</td>
<td>(1)</td>
</tr>
<tr>
<td>Paenibacillus graminis clade 1</td>
<td>VII</td>
<td>23</td>
<td>045</td>
<td>(1)</td>
</tr>
<tr>
<td>Paenibacillus graminis clade 2</td>
<td>VII</td>
<td>23</td>
<td>039</td>
<td>(1)</td>
</tr>
<tr>
<td>Paenibacillus cf. peoriae</td>
<td>IX</td>
<td>24</td>
<td>157</td>
<td>wp (1)</td>
</tr>
<tr>
<td>Paenibacillus amylolyticus sensu lato</td>
<td>X</td>
<td>101</td>
<td>023</td>
<td>(2)</td>
</tr>
<tr>
<td>Paenibacillus cf. xylanilyticus</td>
<td>X</td>
<td>13</td>
<td>100</td>
<td>(1)</td>
</tr>
</tbody>
</table>

a The complete list of all 87 isolates tested for β-galactosidase (β-Gal) activity is presented in Table S1 in the supplemental material. This table also includes unique Cornell Food Safety Lab (FSL) isolate identifiers (e.g., FSL H8-493), which can be used to access additional isolate information at www.pathogentracker.net.

b Phylogenetic group number; see Fig. 1 and 2.

c rpoB allelic type of the representative isolate(s) that was characterized for β-Gal activity.

d The representative isolates tested were classified as positive (+), negative (−), or weakly positive (wp) for β-Gal activity. Details on all isolates tested are available in Table S1 in the supplemental material.
Table S1 in the supplemental material). Except for 1 representative isolate from *P. graminis* clade 2 that was β-galactosidase negative and 1 *Paenibacillus cf. peoriae* isolate that was weakly β-galactosidase positive, all 40 *Paenibacillus* isolates tested were positive for β-galactosidase activity (see Table S1).

**DISCUSSION**

This study provides a comprehensive analysis of the diversity of aerobic bacterial sporeformers that are associated with fluid milk production systems in the United States, with specific emphasis on isolates obtained from pasteurized milk. While the majority of isolates and DNA sequence data analyzed here have been reported previously (25, 39–41, 72, 73), meta-analysis and phylogenetic characterization of rpoB and 16S sequence data for >1,200 aerobic Gram-positive sporeformer isolates from different segments of the dairy production continuum allowed for identification of key spore-forming spoilage organisms of concern and provided phenotype data on isolates representative of the diversity that was identified and characterized through this comprehensive study. Our data specifically show that a few *Bacillus, Viridibacillus, and Paenibacillus* species and clades represent the majority of dairy-associated aerobic sporeformers. Among the isolates representing these clades, *Paenibacillus* spp. could generally be distinguished from *Bacillus* spp. by their ability to grow in milk at 6°C and their ability to display β-galactosidase activity.

A few *Bacillus* and *Paenibacillus* species and clades represent the majority of dairy-associated aerobic sporeformers. Our analysis of 1,288 aerobic sporeformer isolates representing 283 unique *rpoB* sequences found that a relatively small number of species and clades represent the majority of dairy-associated sporeformers. A few *Bacillus* spp. (i.e., *B. pumilus, Bacillus licheniformis* sensu lato, *Bacillus cereus* sensu lato, and *B. weihenstephanensis*) and *Paenibacillus* spp. (i.e., *P. odorifer, Paenibacillus amylyticus* sensu lato, and *P. graminis*) accounted for more than 80% of the dairy-associated sporeformer isolates characterized (with most isolates obtained from pasteurized milk). While a number of these *Bacillus* species have been isolated previously from raw and processed milk as well as from dairy-associated environments (19, 21), only a few studies (18, 19), in addition to those that detailed the isolates characterized here (25, 39–41, 72, 73), have reported the identification and characterization of *Paenibacillus* species from dairy products and dairy-associated environments. Interestingly, a number of the predominant dairy-associated species identified here have also been isolated previously from non-dairy-associated environments (e.g., secluded Antarctic experimental stations [88] and clean rooms [32, 57, 77, 82]). Additionally, a number of studies have reported the identification of spoilage *Bacillus* spp. identified here (e.g., *B. cereus, B. licheniformis, B. subtilis,* and *B. weihenstephanensis*) in nondairy foods, including bread, liquid eggs, seafood, and sous vide products, further illustrating the importance of spore-forming bacilli in our food system (16, 20, 44, 83).

*B. pumilus, Bacillus licheniformis* sensu lato, *Bacillus cereus* sensu lato, and *B. weihenstephanensis* represented 26.3% of all isolates in our study. These species have been isolated previously from raw milk (18) and farm environments, including dairy cattle feed (40, 91) and feces (98). For example, in a study of Belgian dairy farms, Coorevits et al. (18) reported that, of 40 identified species of Gram-positive sporeformers, *B. licheniformis* and *B. pumilus* accounted for 55% of all raw milk isolates. Therefore, our results, along with the results of others, indicate that these *Bacillus* species, and *B. licheniformis* in particular, are commonly found in dairy environments across geographical regions. Several of the species that clustered in group 1 (i.e., *B. safensis, Bacillus aerophilus* sensu lato, and *B. pumilus* clades), which included 22% of non-*Paenibacillus* isolates in our study, have been isolated previously from spacecraft and the environment of spacecraft assembly facilities (57, 77, 82). *B. pumilus* in particular has shown high resistance to spacecraft clean room decontamination methods, such as UV light or rigorous cleaning measures (32, 61). The presence of these extremely resistant organisms in raw milk and dairy-associated environments may thus present a particular challenge for the dairy and food industries.

In our study, *P. odorifer, Paenibacillus amylyticus* sensu lato, and *P. graminis* accounted for more than 80% of *Paenibacillus* dairy-associated isolates. These and other *Paenibacillus* species have been isolated from the milk storage compartments of milk trucks and raw milk silos (39) and from processing lines (41), as well as from packaged pasteurized milk (39, 40). Interestingly, *P. odorifer* and *P. graminis* were originally isolated from plant roots as well as from pasteurized pureed vegetables (13), suggesting that these organisms are also a potential spoilage concern in nondairy foods. In general, *Paenibacillus* species have been isolated from a number of environments, such as soil (37, 60, 67, 99), rhizospheres (63, 96), aquatic environments (9, 10, 66, 86), and compost (94). *Paenibacillus* has only recently been recognized as a genus separate from *Bacillus* (8), and as many new species of *Paenibacillus* continue to be identified (9, 10, 12, 45, 46, 48–52, 64, 66, 86, 90, 92, 94, 97), it is becoming evident that members of this genus occupy diverse environmental niches. The presence of *Paenibacillus* spp. in a wide range of environments, including dairy farms, presents a challenge for efforts to prevent these organisms from entering raw milk supplies.

The fact that we have identified 11 previously uncharacterized *Paenibacillus* clades not only indicates that a number of species within the genus *Paenibacillus* remain to be characterized and described but also shows that we still lack a complete understanding of the bacterial diversity associated with dairy products. The isolates reported here represent an important starting point for efforts to characterize and describe additional new dairy-associated *Paenibacillus* species. Further characterization of different *Paenibacillus* spp., including an improved understanding of their ecology and physiology, will be critical for the development of novel detection systems, as well as for improved control strategies for these spoilage organisms.

*Paenibacillus* spp. can generally be distinguished from *Bacillus* spp. by their ability to grow in milk at 6°C and by their β-galactosidase activity. Except for one *B. weihenstephanensis* isolate, isolates representing common *Bacillus* clades (including one *Bacillus cereus* sensu lato isolate) were unable to grow in SMB at 6°C. While *B. weihenstephanensis* was initially identified as a psychrotolerant species within the *Bacillus cereus* sensu lato clade (58, 68), several studies have demonstrated the abilities of different species within the *Bacillus cereus* sensu lato clade, such as *B. cereus* (18), *B. thuringiensis* (11), and *B. weihenstephanensis* (27, 85), to grow at temperatures of ≤7°C; since these species all share high 16S rRNA gene similarity (6), it is possible that *B. cereus* or *B. thuringiensis* was misidentified in at least some of these studies. Furthermore, in most of these studies, growth was determined in media, such as tryptc soy agar (18) or plate count medium (27),
that contain glucose, whereas in our study, growth studies were conducted in rehydrated skim milk, in which lactose is the primary carbohydrate source. Interestingly, although isolates representing *B. weihenstephanensis* and the *Viridibacillus* clade showed growth in SMB at 6°C, all of the *B. weihenstephanensis* and *Viridibacillus* isolates tested here were negative for β-galactosidase activity at a higher temperature (i.e., 32°C). While further experiments are needed to determine whether these species hydrolyze lactose in milk at refrigeration temperatures, these findings indicate that *B. weihenstephanensis* and *Viridibacillus* spp. may have a β-galactosidase enzyme that is specifically expressed or active at low temperatures, like a thermolabile β-galactosidase that has been characterized in *Planococcus* sp. strain L4 (38). Since a number of *Bacillus* spp. have been isolated from dairy products and fluid milk (including the isolation of *B. weihenstephanensis* and other *Bacillus cereus* sensu lato species from raw and heat-treated milk [11, 80]), it should be noted that even *Bacillus* spp. that cannot grow in milk at refrigeration temperatures may negatively affect shelf life or safety, for example, if products are not kept at proper refrigeration temperatures throughout distribution and storage.

Interestingly, we also identified a number of isolates representing genera formerly classified as group 2 (7) *Bacillus* species (i.e., *Viridibacillus, Lysinibacillus*, and *Psychrobacillus*), indicating that these organisms occupy dairy environments. Our observation that an isolate representing *Viridibacillus* was also able to grow in SMB at 6°C indicates that *Viridibacillus* in particular represents a dairy-associated psychrotolerant spoilage organism. *Viridibacillus* has only recently been recognized as a genus distinct from *Bacillus* (4), and representatives of this species (*Viridibacillus arenosi*, *Viridibacillus arvi*, and *Viridibacillus neidei*) were originally described as soil bacteria belonging to the genus *Bacillus* (35, 65).

Our results provide the first direct experimental evidence that a number of *Paenibacillus* sp. isolates are able to grow in milk at refrigeration temperatures, supporting an emerging body of evidence demonstrating that this genus includes important dairy and food spoilage organisms. Previous studies have shown that, while both *Paenibacillus* and *Bacillus* spp. are commonly isolated directly after pasteurizing, *Paenibacillus* spp. are more frequently isolated late in the shelf life of refrigerated HTST pasteurized fluid milk (28, 71, 72). In addition, a previous study found that storage of pasteurized vegetable purées at 4°C favored the predominance of *Paenibacillus*, whereas *Bacillus* spp. predominated in purées stored at 20 to 25°C (33). Taken together, these results indicate that, in general, storage of food at refrigeration temperatures (e.g., 4 to 6°C) selects for *Paenibacillus* sp., supporting a potentially broad importance for *Paenibacillus* spp. as spoilage organisms in foods, where postprocessing contamination with spoilage organisms that grow more rapidly at refrigeration temperatures and outcompete *Paenibacillus* (e.g., *Pseudomonas* spp.) has been controlled.

Interestingly, a cold-active β-galactosidase has been identified in *Paenibacillus* strain C7 (81). While this enzyme may contribute to the ability of *Paenibacillus* to utilize lactose at low temperatures, hence facilitating growth in milk under refrigeration temperatures, it is not known whether the C7 cold-active β-galactosidase is conserved across *Paenibacillus* spp. Overall, our understanding of cold tolerance among *Paenibacillus* as well as *Bacillus* spp. is limited, even though a number of studies have explored mechanisms used by *B. subtilis* to adapt to temperatures around 15°C (14, 36). Further studies on mechanisms of cold growth in *Paenibacillus* spp. will thus be needed, including the identification of potential target genes that could be used for molecular detection of these spoilage organisms.

Our finding that the majority of dairy-associated *Paenibacillus* subtypes characterized in this study produce β-galactosidase activity at 32°C, while most of the non-*Paenibacillus* subtypes were β-galactosidase negative, suggests that β-galactosidase indicator plates may allow for rapid and easy discrimination of Gram-positive sporeformers into putative *Paenibacillus* and non-*Paenibacillus* sp. isolates. While this is important, since Bergey’s *Manual of Systematic Bacteriology* currently lists no distinguishing *Paenibacillus* phenotype (69), isolates representative of *Bacillus licheniformis* sensu lato were positive for β-galactosidase activity, and some *Paenibacillus* isolates were negative for β-galactosidase. Therefore, one cannot rely solely on testing for β-galactosidase activity to distinguish *Bacillus* spp. from *Paenibacillus* spp., and as shown here, such testing may not detect all *Paenibacillus* spp. Screening for β-galactosidase activity does appear to have some potential for use as an initial screening method and may, in particular, be useful for detecting *Paenibacillus* spp. in raw milk. Further characterization of *Paenibacillus* isolates from nondairy sources is needed, though, in order to determine whether β-galactosidase activity is common among all *Paenibacillus* isolates. Ultimately, identification of *Paenibacillus*-specific gene targets and the subsequent design of rapid, DNA-based systems to detect and confirm *Paenibacillus* spp. will be needed to facilitate specific detection of these spoilage organisms.

**Conclusion.** Psychrotolerant sporeformers represent a particular concern, since these organisms can both survive heat treatments commonly used in food processing and also grow in foods that are held under refrigeration temperatures after processing. Our data reported here identify the genus *Paenibacillus*, which has recently been recognized as a separate genus (8), as a diverse group of organisms that appear to be predominantly psychrotolerant, with an ability to grow in milk and possibly other foods at temperatures as low as 6°C. Improved control of these organisms along the dairy production chain and other food chains will be critical for reducing the spoilage of various heat-treated food products. To that end, our study not only has identified β-galactosidase activity as a potential screening tool that will facilitate the detection of *Paenibacillus* spp. but also provides a comprehensive characterization of *Paenibacillus* diversity that will facilitate further research on the taxonomy, diversity, ecology, and evolution of this genus. Future efforts in this area should also lead to novel approaches that will contribute to the control of these spoilage organisms in the food supply.

**ACKNOWLEDGMENTS**

We acknowledge the contributions of the staff of the Milk Quality Improvement Program (MQIP) at Cornell University to this project. The research at the MQIP, including this work, is supported by the New York State Milk Promotion Advisory Board (through the New York State Department of Agriculture), representing New York State dairy farmers committed to producing high-quality milk.

**REFERENCES**


Kim BC, et al. 2009. Paenibacillus pinium sp. nov., a cellulolytic bacte-


