Propionibacterium acnes is a commensal of human skin but is also implicated in the pathogenesis of acne vulgaris, in biofilm-associated infections of medical devices and endophthalmitis, and in infections of bone and dental root canals. Recent studies associate P. acnes with prostate cancer. As the species includes evolutionary lineages with distinct association with health and disease, there is a need for a high-resolution typing scheme. Recently, two multilocus sequence typing (MLST) schemes were reported, one based on nine and one based on seven housekeeping genes. In the present study, the two schemes were compared with reference to a phylogenetic tree based on 78 P. acnes genomes and their gene contents. Further support for a basically clonal population structure of P. acnes and a scenario of the global spread of epidemic clones of P. acnes was obtained. Compared to the Belfast scheme, the Aarhus MLST scheme (http://pacnes.mlst.net/), which is based on nine genes, offers significantly enhanced resolution and phylogenetic inferences more concordant with analyses based on a comprehensive sampling of the entire genomes, their gene contents, and their putative pathogenic potential.

**MATERIALS AND METHODS**

The comparison of the two MLST schemes for P. acnes, here referred to as the Aarhus scheme and the Belfast scheme, was based on sequences extracted from a total of 78 complete genomes. Seventy-two of these genomes are accessible in the NCBI database: NC_006085, KPA171202 (complete genome) (3); NC_014039, strain SK137 (complete genome); CP002409.1, strain 266 (complete genome) (3); NC_014039, strain SK137 (complete genome); CP002409.1, strain 266 (complete genome) (3); NC_014039, strain SK137 (complete genome); NC_014039, strain SK137 (complete genome); NC_014039, strain SK137 (complete genome); NC_014039, strain SK137 (complete genome). The alternative scheme (here referred to as the Belfast scheme) is based on partial sequences of seven housekeeping genes comprising a total of 4,287 nucleotides (nt) and is available at http://pacnes.mlst.net/. The alternative scheme (here referred to as the Belfast scheme) is based on partial sequences of seven housekeeping genes comprising a total of 3,135 nt (http://pacnes.mlst.net/). Here, we report a comparison of the schemes with regard to their discriminatory power and ability to identify and distinguish evolutionary lineages with distinct properties relevant for the disease association of subsets of P. acnes.

**T**he Gram-positive aerotolerant anaerobe Propionibacterium acnes is one of the predominant members of the commensal skin microbiota (12, 13, 17). It successfully colonizes healthy skin and becomes most prevalent around puberty on areas of skin with abundant sebaceous follicles, such as the face and the upper part of the back and chest (24). It is the only bacterium able to colonize the hostile environment of sebaceous follicles (2), where it often coexists with the fungus Malassezia and is also part of the nasal, oral, and gut microbiota.

The relevance of P. acnes in human medicine is its association with acne vulgaris and its isolation from a number of opportunistic infections. Although its role is intensely debated, there is increasing evidence that P. acnes is a powerful inducer of inflammation and that it plays a crucial role in the pathogenesis of acne in genetically disposed individuals (4, 8, 18, 30). The apparent contradiction with its role as a ubiquitous and predominant skin commensal may be explained by strain-dependent differences in pathogenic potential (11, 15, 21, 22, 25, 26, 33). In support of this explanation, we recently identified a distinct subpopulation of P. acnes, including an epidemic clone, which is associated with moderate to severe acne, whereas other parts of the P. acnes population are associated with healthy skin and with opportunistic infections (19). These findings were recently confirmed by an independent study (23).

Opportunistic infections from which P. acnes strains are frequently isolated include biofilm-associated infections of prosthetic shoulders, hips, heart valves, and other medical devices that may become contaminated with skin microorganisms, endophthalmitis following ocular surgery, bone infections, including orthopedic implants, and dental root canal infections (16, 27, 28, 29, 32, 34). Recently, P. acnes has been associated with prostate cancer due to its prevalence in affected prostate tissue, but its possible etiologic role has yet to be defined (1, 7, 9).

Typing by various means is an important tool for the identification of subsets of bacterial species with particular pathogenic potential and for epidemiological analysis. A major advance in typing methodology was the introduction of multilocus sequence typing (MLST), which is based on sequences of fragments of usually six to seven housekeeping genes that can be stored in internet-based databases for easy comparison and storage of new data, thus enabling the generation of global epidemiological records (20).

Recently, MLST schemes for P. acnes were reported by us (19) and McDowell et al. (23). The scheme reported by us (here referred to as the Aarhus scheme) is based on partial sequences of nine housekeeping genes comprising a total of 4,287 nucleotides (nt) and is available at http://pacnes.mlst.net/. The alternative scheme (here referred to as the Belfast scheme) is based on partial sequences of seven housekeeping genes comprising a total of 3,135 nucleotides (nt) and is available at http://pacnes.mlst.net/.
HL005PA4; PRJNA49271, strain HL007PA1; PRJNA49169, strain HL013PA1; PRJNA49161, strain HL020PA1; PRJNA49211, strain HL025PA1; PRJNA49213, strain HL025PA2; PRJNA49257, strain HL027PA1; PRJNA49243, strain HL027PA2; PRJNA49241, strain HL030PA1; PRJNA49243, strain HL030PA2; PRJNA49247, strain HL036PA1; PRJNA49249, strain HL036PA2; PRJNA49251, strain HL036PA3; PRJNA49279, strain HL037PA1; PRJNA49281, strain HL037PA2; PRJNA49283, strain HL037PA3; PRJNA49203, strain HL038PA1; PRJNA49175, strain HL043PA1; PRJNA49177, strain HL043PA2; PRJNA49253, strain HL044PA1; PRJNA49167, strain HL045PA1; PRJNA49221, strain HL046PA1; PRJNA49223, strain HL046PA2; PRJNA49233, strain HL050PA1; PRJNA49237, strain HL050PA2; PRJNA49239, strain HL050PA3; PRJNA49163, strain HL053PA1; PRJNA49165, strain HL053PA2; PRJNA49273, strain HL056PA1; PRJNA49275, strain HL056PA2; PRJNA49277, strain HL059PA1; PRJNA49279, strain HL059PA2; PRJNA49201, strain HL060PA1; PRJNA49261, strain HL063PA1; PRJNA49263, strain HL063PA2; PRJNA49255, strain HL067PA1; PRJNA49179, strain HL072PA1; PRJNA49181, strain HL072PA2; PRJNA49183, strain HL074PA1; PRJNA49173, strain HL078PA1; PRJNA49275, strain HL082PA1; PRJNA49277, strain HL082PA2; PRJNA49207, strain HL083PA1; PRJNA49209, strain HL083PA2; PRJNA49219, strain HL086PA1; PRJNA49195, strain HL087PA1; PRJNA49197, strain HL087PA2; PRJNA49199, strain HL087PA3; PRJNA49205, strain HL092PA1; PRJNA40727, strain HL096PA2; PRJNA40729, strain HL096PA3; PRJNA40715, strain HL097PA1; PRJNA40717, strain HL099PA1; PRJNA40721, strain HL103PA1; PRJNA49187, strain HL110PA1; PRJNA49189, strain HL110PA2; PRJNA49191, strain HL110PA3; and PRJNA49193, strain HL110PA4 (all whole-genome shotgun sequencing projects that are part of the human microbiome project).

The remaining six are genomes in the annotation phase in our laboratory and represent the following Aarhus sequence types (STs): ST3 (strain 15.1.R1), ST18 (strain 12.1.L1), ST20 (strain 12.1.R1), ST27 (strain 30.2.L1), ST29 (strain 15.2.L1), and ST36 (strain 21.1.L1).

Sequences extracted from the genomes were those used in the two MLST schemes, which represent fragments of the genes cel, coa, fba, gms, lac, oxc, pak, recA, and zno (Aarhus MLST) and aroE, atpD, gmk, guaA, lepA, recA, and sodA (Belfast MLST). The locations of the respective genes in the genome of strain KPA171202 and all other closed genomes of P. acnes are illustrated in Fig. 1.

For the construction of a robust reference phylogenetic tree, a more comprehensive sampling of the genomes was performed. Full sequences of 76 annotated housekeeping genes, which are evenly distributed throughout the genomes, are present in all strains, and are not located adjacent to genes annotated as encoding surface-exposed proteins, were retrieved (coordinates are according to the KPA171202 genome): DNA polymerase III subunit beta (2219 to 3379), D-ribose pyranase (22214 to 22597), formate-tetrahydrofolate ligase (47426 to 49135), sugar transport permease BglB (56040 to 56870), aldehyde dehydrogenase (85493 to 86950), catalase (112227 to 113678), thiamine-phosphate pyrophosphorylase (132955 to 134664), M16 family peptidase (146394 to 147767), reductase ferredoxin (167045 to 168424), dihydroorotate dehydrogenase (191680 to 192687), lysyl-tRNA synthetase (220963 to 222495), putative heat shock protein (251890 to 252255), putative endonuclease III (277186 to 278112), putative helicase (287834 to 290098), putative flavin-containing amine oxidase (304329 to 305678), GTP cyclohydrolase I (315622 to 316197), dihydropteroate synthase (321592 to 322437), allophanate hydrolase subunit 1 (332920 to 333540), aspartate-semialdehyde dehydrogenase (369582 to 370895), putative c-type cytochrome biogenesis protein (376261 to 377772), putative phosphoenolpyruvate-protein kinase (402218 to 403936), superfamily II DNA-RNA helicase (413845 to 414840), cysteinyl-tRNA synthetase (431282 to 432754), S-ribosylhomocysteinase (495789 to 496241), superfamily II DNA-RNA helicase (521933 to 52401), molybdenum cofactor biosynthesis protein A (550932 to 551972), flavin adenine dinucleotide (FAD)-dependent oxidoreductase (573023 to 574270), alcohol dehydrogenase (590844 to 592403).
annotated genomes of strains KPA171202 and 266 as the reference. Using the genes identified in the two complete genomes, a phylogenetic analysis was undertaken including the eBURST algorithm, which was developed for the species based on significantly fewer gene sequences (19, 23). The mean genetic distance within the population shown in Fig. 2 was 0.008 ± 0.0001. Three strains (HL037PA2, HL0PA3, and HL044PA1) excluded from the figure were highly divergent, with a mean distance to the remaining population of 0.146. It is likely that these strains represent a separate species. The alignment of 16S rRNA gene sequences from the three strains confirmed the distance to \( P. acnes \), but a search against the RNA database (release 10) at \( http://rdp.cme.msu.edu/ \) showed homology no closer to any recognized species than that to \( P. acnes \) (not shown). Thus, the strains warrant recognition as a new species. During the preparation of the manuscript, the genome sequence of one highly similar strain, tentatively designated "\( P. pacnes \)" was announced (6). However, the novel species has yet to be effectively and validly described.

The separate examination of each of the 76 genes included in the concatenated complete sequences of 76 housekeeping genes constituting a total of 92,577 bp revealed a tree consisting of seven distinct clades with multiple strains (corresponding to CC3, CC18, CC31, ST27, CC28, CC36, and CC60 plus ST62 and ST72 according to the Aarhus scheme) and four strains that formed separate lineages (HL078PA1, HL086PA1, SK187, and HL097PA1) (Fig. 2). The overall topology of the tree is similar to those previously detected for the species based on significantly fewer gene sequences (19, 23). The mean genetic distance within the population shown in Fig. 2 was 0.008 ± 0.0001. Three strains (HL037PA2, HL0PA3, and HL044PA1) excluded from the figure were highly divergent, with a mean distance to the remaining population of 0.146. It is likely that these strains represent a separate species. The alignment of 16S rRNA gene sequences from the three strains confirmed the distance to \( P. acnes \), but a search against the RNA database (release 10) at \( http://rdp.cme.msu.edu/ \) showed homology no closer to any recognized species than that to \( P. acnes \) (not shown). Thus, the strains warrant recognition as a new species. During the preparation of the manuscript, the genome sequence of one highly similar strain, tentatively designated "\( P. pacnes \)" was announced (6). However, the novel species has yet to be effectively and validly described.

The separate examination of each of the 76 genes included in the concatenated complete sequences of 76 housekeeping genes constituting a total of 92,577 bp revealed a tree consisting of seven distinct clades with multiple strains (corresponding to CC3, CC18, CC31, ST27, CC28, CC36, and CC60 plus ST62 and ST72 according to the Aarhus scheme) and four strains that formed separate lineages (HL078PA1, HL086PA1, SK187, and HL097PA1) (Fig. 2). The overall topology of the tree is similar to those previously detected for the species based on significantly fewer gene sequences (19, 23). The mean genetic distance within the population shown in Fig. 2 was 0.008 ± 0.0001. Three strains (HL037PA2, HL0PA3, and HL044PA1) excluded from the figure were highly divergent, with a mean distance to the remaining population of 0.146. It is likely that these strains represent a separate species. The alignment of 16S rRNA gene sequences from the three strains confirmed the distance to \( P. acnes \), but a search against the RNA database (release 10) at \( http://rdp.cme.msu.edu/ \) showed homology no closer to any recognized species than that to \( P. acnes \) (not shown). Thus, the strains warrant recognition as a new species. During the preparation of the manuscript, the genome sequence of one highly similar strain, tentatively designated "\( P. pacnes \)" was announced (6). However, the novel species has yet to be effectively and validly described.

The separate examination of each of the 76 genes included in the concatenated complete sequences of 76 housekeeping genes constituting a total of 92,577 bp revealed a tree consisting of seven distinct clades with multiple strains (corresponding to CC3, CC18, CC31, ST27, CC28, CC36, and CC60 plus ST62 and ST72 according to the Aarhus scheme) and four strains that formed separate lineages (HL078PA1, HL086PA1, SK187, and HL097PA1) (Fig. 2). The overall topology of the tree is similar to those previously detected for the species based on significantly fewer gene sequences (19, 23). The mean genetic distance within the population shown in Fig. 2 was 0.008 ± 0.0001. Three strains (HL037PA2, HL0PA3, and HL044PA1) excluded from the figure were highly divergent, with a mean distance to the remaining population of 0.146. It is likely that these strains represent a separate species. The alignment of 16S rRNA gene sequences from the three strains confirmed the distance to \( P. acnes \), but a search against the RNA database (release 10) at \( http://rdp.cme.msu.edu/ \) showed homology no closer to any recognized species than that to \( P. acnes \) (not shown). Thus, the strains warrant recognition as a new species. During the preparation of the manuscript, the genome sequence of one highly similar strain, tentatively designated "\( P. pacnes \)" was announced (6). However, the novel species has yet to be effectively and validly described.

The separate examination of each of the 76 genes included in the concatenated complete sequences of 76 housekeeping genes constituting a total of 92,577 bp revealed a tree consisting of seven distinct clades with multiple strains (corresponding to CC3, CC18, CC31, ST27, CC28, CC36, and CC60 plus ST62 and ST72 according to the Aarhus scheme) and four strains that formed separate lineages (HL078PA1, HL086PA1, SK187, and HL097PA1) (Fig. 2). The overall topology of the tree is similar to those previously detected for the species based on significantly fewer gene sequences (19, 23). The mean genetic distance within the population shown in Fig. 2 was 0.008 ± 0.0001. Three strains (HL037PA2, HL0PA3, and HL044PA1) excluded from the figure were highly divergent, with a mean distance to the remaining population of 0.146. It is likely that these strains represent a separate species. The alignment of 16S rRNA gene sequences from the three strains confirmed the distance to \( P. acnes \), but a search against the RNA database (release 10) at \( http://rdp.cme.msu.edu/ \) showed homology no closer to any recognized species than that to \( P. acnes \) (not shown). Thus, the strains warrant recognition as a new species. During the preparation of the manuscript, the genome sequence of one highly similar strain, tentatively designated "\( P. pacnes \)" was announced (6). However, the novel species has yet to be effectively and validly described.

The separate examination of each of the 76 genes included in the concatenated complete sequences of 76 housekeeping genes constituting a total of 92,577 bp revealed a tree consisting of seven distinct clades with multiple strains (corresponding to CC3, CC18, CC31, ST27, CC28, CC36, and CC60 plus ST62 and ST72 according to the Aarhus scheme) and four strains that formed separate lineages (HL078PA1, HL086PA1, SK187, and HL097PA1) (Fig. 2). The overall topology of the tree is similar to those previously detected for the species based on significantly fewer gene sequences (19, 23). The mean genetic distance within the population shown in Fig. 2 was 0.008 ± 0.0001. Three strains (HL037PA2, HL0PA3, and HL044PA1) excluded from the figure were highly divergent, with a mean distance to the remaining population of 0.146. It is likely that these strains represent a separate species. The alignment of 16S rRNA gene sequences from the three strains confirmed the distance to \( P. acnes \), but a search against the RNA database (release 10) at \( http://rdp.cme.msu.edu/ \) showed homology no closer to any recognized species than that to \( P. acnes \) (not shown). Thus, the strains warrant recognition as a new species. During the preparation of the manuscript, the genome sequence of one highly similar strain, tentatively designated "\( P. pacnes \)" was announced (6). However, the novel species has yet to be effectively and validly described.

The separate examination of each of the 76 genes included in the concatenated complete sequences of 76 housekeeping genes constituting a total of 92,577 bp revealed a tree consisting of seven distinct clades with multiple strains (corresponding to CC3, CC18, CC31, ST27, CC28, CC36, and CC60 plus ST62 and ST72 according to the Aarhus scheme) and four strains that formed separate lineages (HL078PA1, HL086PA1, SK187, and HL097PA1) (Fig. 2). The overall topology of the tree is similar to those previously detected for the species based on significantly fewer gene sequences (19, 23). The mean genetic distance within the population shown in Fig. 2 was 0.008 ± 0.0001. Three strains (HL037PA2, HL0PA3, and HL044PA1) excluded from the figure were highly divergent, with a mean distance to the remaining population of 0.146. It is likely that these strains represent a separate species. The alignment of 16S rRNA gene sequences from the three strains confirmed the distance to \( P. acnes \), but a search against the RNA database (release 10) at \( http://rdp.cme.msu.edu/ \) showed homology no closer to any recognized species than that to \( P. acnes \) (not shown). Thus, the strains warrant recognition as a new species. During the preparation of the manuscript, the genome sequence of one highly similar strain, tentatively designated "\( P. pacnes \)" was announced (6). However, the novel species has yet to be effectively and validly described.
FIG 2 Phylogenetic tree of 75 strains of \textit{P. acnes} constructed by the minimum evolution algorithm in MEGA version 5 and based on a 92,577-bp concatemer of complete sequences of 76 housekeeping genes evenly distributed across the entire genome. For each strain the ET and CC determined according to the Aarhus MLST scheme and the Belfast scheme are indicated. Bootstrap values exceeding 50 are shown. The phylogenetic analysis identified seven clades that corresponded to CCs identified by eBURST analysis based on data generated according to the Aarhus scheme.
A total of 15 strains assigned to ST-9 according to the Belfast scheme were distributed across five STs according to the Aarhus scheme: ST28 (n = 11), ST64 (n = 1), ST65 (n = 1), ST67 (n = 1), and ST68 (n = 1). However, in contrast to ST-6 strains, all 15 strains belonged to a tight phylogenetic cluster.

Comparison of the clonal complexes detected by the eBURST analysis of the data sets generated by the two MLST schemes showed a similar pattern of discrepancies. Most notably, CC-6 (Belfast scheme) comprised three clonal complexes distinguished by the Aarhus scheme (CC3, CC18, and CC31) in addition to two strains of CC28 (HL025PA1 and 30.2.L1), which show striking divergence from other CC28 strains in the reference phylogenetic tree (Fig. 2). A major part of CC-10 (Belfast scheme) was congruent with CC28 (Aarhus scheme), with two notable exceptions. CC-10 included strains assigned to CC36 according to the Aarhus scheme which, according to the phylogenetic tree, are highly divergent from the remaining strains of CC-10. In contrast, two other strains of CC28 (HL030PA2 and HL063PA2), which were detected as distinct STs by both schemes, were part of CC28 according to the Aarhus scheme, whereas they were allocated to a clonal complex (CC-22) distinct from CC-10 according to the
Belfast scheme. Their separation as being distinct from other strains in CC28 is supported by the phylogenetic tree in Fig. 2 and by a distinct 16S rRNA allele (not shown).

**Gene content, phenotype, and differences in resolution.** To examine the potential relevance of the discrepancies in resolution achieved by the two MLST schemes, we compared the related phenotypic differences and gene contents. According to our previous study, both CC3 and CC31, which were distinguished from CC18 by the Aarhus scheme but not by the Belfast scheme (all CC-6), are distinct because they lack hemolytic activity, biotype, and alleles of the two putative virulence factor genes camp5 and fly (19). A comparison of genes that are present in or absent from the 69 annotated genome sequences showed a remarkable consistency within the clonal complexes identified by MLST data generated by the Aarhus scheme. The few exceptions are discussed below.

In support of the distinction between CC3 and CC18 (Aarhus scheme), the following genes were present in CC18 but absent from CC3: cysteine synthase-ornithine cyclodeaminase (332675701), hypothetical protein PAZ_c13380 (332675689), hypothetical protein PAZ_c13400 (332675691), putative thioesterase (332675698), tyrocidine synthase 3 (332675699), hypothetical protein PAZ_c13540 (332675705), putative siderophore biosynthesis protein SbnA (332675702), putative protein associated with putative adhesion protein (332676057), hypothetical protein PAZ_c13420 (332675693), hypothetical protein PAZ_c13390 (332675690), linear gramicidin synthesize subunit B (332675700), hypothetical protein PAZ_c13460 (332675697), hypothetical protein PAZ_c13440 (332675695), hypothetical protein PAZ_c13450 (332675696), biosurfactant production protein (332675703), hypothetical protein PAZ_c07360 (332675104), hypothetical protein PAZ_c13640 (332675715), ATP-dependent RNA helicase HrpB (332674556), hypothetical protein PAZ_c13410 (332675692), hypothetical protein PAZ_c13530 (332675704), and hypothetical protein PAZ_c13430 (332675694). These genes are clustered in an island-like region (Island 2; PAZ_13580-PAZ_13460) identified in the KPA171202 and 266 genomes and includes a cluster of genes (PPA1284 to PPA1292) encoding proteins with homology to nonribosomal peptide synthetases (NRPS). Conceivably it was acquired via plasmid insertion (4, 5). The absence of these genes was unique to CC3 strains (and to the aberrant strain HL097PA1). One notable exception was strain HL099PA1, which in spite of its clear relationship to other CC3 strains possessed all of the genes mentioned above.

The comparison of strains of CC36 and CC28, which are all part of CC-10 according to the Belfast scheme, showed 86 genes (details not shown) that were unique to CC36 strains, including several genes involved in carbohydrate metabolism and genes encoding a bacteriophage, as previously observed (3, 5, 19). As no closed genome of strains belonging to CC31 was available, we were not able to identify genes that were uniquely absent from or present in this clonal complex.

**Aberrant strains.** The separate examination of each of the 76 genes included in the concatenator offered explanations for the few strains that showed unexpected eBURST clusterings relative to the phylogenetic tree. In general, the overall topologies of trees based on single genes were never in disagreement with the tree based on the 92,577-bp concatenator, indicating limited recombination between the subclusters of *P. acnes*. Notable exceptions were strains HL025PA1, HL086PA1, SK187, and HL097PA1, which were either singletons or strains that unexpectedly were assigned to clonal clusters to which they were not closely related according to the phylogenetic tree.

A detailed analysis of each of the 76 genes by tree construction revealed a mosaic pattern in the strains HL025PA1, SK187, and HL086PA1. HL025PA1, which was assigned to CC28 according to the Aarhus scheme and CC-6 according to the Belfast scheme (although it was clearly distinct from other strains of these clusters [Fig. 2]), showed unique alleles in 16 of the 76 genes, whereas other genes were closely related or were identical to alleles of CC3, CC18, CC31, or CC60. Nevertheless, neither the nine nor the seven genes used in the two MLST schemes were able to disclose its aberrant genetic nature. HL086PA1, which unexpectedly was assigned to CC31 according to the Aarhus scheme and CC-6 according to the Belfast scheme, showed a mosaic of a majority of gene alleles that were in agreement with its assignment to CC31 but combined with alleles in seven genes that clustered with CC3 strains. Likewise, the majority of gene alleles in SK187 were in agreement with its assignment to CC28 according to the Aarhus scheme. However, in seven genes the alleles clustered with CC3, CC18, or CC31. While the unique genetic constitution of SK187 was disclosed by the Belfast MLST scheme, the Aarhus scheme assigned it to ST67/CC28, although its genome was distinct in parts from the genome of strain HL037PA1, which was also assigned to ST67/CC28. Finally, the very distinct position of singleton strain HL097PA1 in the tree is explained by the fact that alleles of 28 out of the 75 genes analyzed were unique to this strain.

**DISCUSSION**

The measure of success of any bacterial typing scheme is its discriminatory power and its ability to identify and distinguish clones or subpopulations with distinct pathogenic potential or ecological and epidemiological characteristics. As expected, the Aarhus MLST scheme for *P. acnes*, which is based on nine loci constituting 4,286 nt, showed significantly better resolution than the Belfast scheme, which is based on seven loci that constitute a total of 3,135 nt. While the Aarhus scheme distinguished 32 distinct STs among 75 strains, the number of STs distinguished by the Belfast scheme was 19. Most MLST schemes that have been developed so far use six or seven loci according to the initially published scheme for *Neisseria meningitidis* (20). We chose nine loci for our scheme, because the comparison of sequences generated during the development of the scheme revealed an unexpected degree of sequence conservation among members of the species. This is in agreement with the fact that 25 out of 76 housekeeping genes analyzed in this study showed no or very limited discrimination between strains. The relative conservation of the *P. acnes* populations is further reflected in the limited genetic distance between strains calculated on the basis of a concatenator of complete sequences of 76 housekeeping genes constituting more than 92,000 nt (genetic distance, 0.8% ± 0.01%). In comparison, species such as *Haemophilus influenzae*, *Streptococcus mitis*, *Streptococcus oralis*, and *Neisseria meningitidis* show intraspecies genetic distances of 2, 5, 9.7, and 4.6%, respectively (distances calculated in MEGA version 5 from concatenated sequences of MLST genes from the respective MLST databases).

Separate and comparative analyses of each of 76 housekeeping genes confirmed our conclusion that the population structure of *P. acnes* is basically clonal but with occasional examples of recombination. Among the 69 annotated genomes examined in detail, three showed evidence of extensive recombination, with mosaics

April 2012 Volume 50 Number 4 jcm.asm.org
of all major parts of the population. Accordingly, eBURST analysis based on the MLST genes showed relationships to clonal clusters that were not supported by the phylogenetic tree based on a 92,577-bp concatemer (Fig. 2). As expected, this was most pronounced with the Aarhus scheme, which samples more evenly distributed parts of the genome (Fig. 1).

The difference in discriminatory power between the two systems was most strikingly demonstrated by ST-6 in the Belfast scheme, which in the study reported by McDowell et al. (23) included 50% of 123 strains and the majority of isolates from acne. According to the Aarhus scheme, strains assigned to ST-6 encompassed 13 distinct STs distributed across four phylogenetic clades supported by significant bootstrap values (Fig. 2) and characterized by distinct differences in gene content. The comparison of the phylogenetic tree with the typing results obtained by use of the Aarhus MLST scheme (Fig. 2) shows that this scheme detects virtually all genetic diversity revealed by the analysis of more than 92,577 bp of housekeeping gene sequences. The more limited resolution achieved by the Belfast scheme is also demonstrated by the 15 strains assigned to ST-9 (Belfast scheme), which were distributed across six STs according to the Aarhus scheme. In addition to the smaller sample of the genome probed by the Belfast scheme, the more limited coverage of all parts of the genome (Fig. 1) are likely to contribute to the lower resolution.

Analysis of the gene contents of the genomes analyzed in this study demonstrated a remarkable correlation with the clonal clusters identified by the Aarhus scheme, whereas the lower resolution achieved by the Belfast scheme resulted in less correlation. For example, 21 genes partly located in an island-lower resolution achieved by the Belfast scheme resulted in less clonal clusters identified by the Aarhus scheme, whereas the more limited coverage of all parts of the genome (Fig. 1) are likely to contribute to the lower resolution.

One of the striking results of our previously reported population genetic analysis of P. acnes (19) was the detection of an epidemic clone, ST18, which in a conserved form has been circulating for at least 85 years and constituted 25 out of 143 nonredundant isolates from Scandinavia and the United Kingdom. As shown in Fig. 2, this study identified seven ST18 strains from the United States and one from Germany and further confirmed the almost complete conservation of this clone, even based on the concatenated sequence of more than 92,000 bp (genetic distance among nine strains, 0.02%). This successfully disseminating clone is particularly interesting because of its strong association with moderate to severe acne (19).

Differential association with health or disease is one of the most interesting results of the application of population genetic analysis to P. acnes (19, 23). Apart from ST18, our previous study suggested an association of CC3, CC18 (which contains ST18), and CC31 with acne. In contrast, CC36, CC60, and ST27 seemed to be associated with healthy skin but also with opportunistic infections associated with medical implant devices (19). Similar patterns were observed by McDowell and coworkers (23). However, it is not known if all or only part of the clusters CC3, CC18, CC28, and CC31 are associated with acne. Therefore, further association studies and the identification of virulence-associated properties are essential, emphasizing the need for methods that can distinguish between virulent and nonvirulent clones of P. acnes.

Previous reports introduced various designations for the major P. acnes clusters based on sero-, bio-, or genotype. As MLST offers much-improved resolution based on portable data that are directly comparable between laboratories, we recommend using STs and CCs as well-defined units in future studies and communications.

In conclusion, the comparison of the two MLST schemes for P. acnes that have been recently reported showed that the Aarhus MLST scheme (http://pacnes.mlst.net/) offers significantly increased discriminatory power and phylogenetic inferences concordantly with analysis results based on a comprehensive sampling of the entire genomes, gene contents, and putative pathogenic potential. While the Belfast scheme, based on seven genes, is less costly and less time-consuming, it loses resolution power and fails to distinguish between subpopulations of P. acnes that differ in gene contents, and according to preliminary data they seem to differ in their association with acne and general disease potential.

ACKNOWLEDGMENTS

We are grateful to David M. Ansen for setting up the MLST database within the framework of databases hosted by the laboratory of Brian Spratt at Imperial College London and to Holger Brüggemann, Aarhus University, for the critical reading of the manuscript.

The study was supported by a grant from the Karen Elise Jensen Foundation.

REFERENCES


