Journal of Clinical Microbiology

Multilocus Sequence Typing and Phylogenetic Analysis of Propionibacterium acnes

Mogens Kilian, Christian F. P. Scholz and Hans B. Lomholt
J. Clin. Microbiol. **2012, 50(4):1158. DOI: 10.1128/JCM.r06129-11.**
Published Ahead of Print 18 January 2012.

JCM
Journals.ASM.org

Multilocus Sequence Typing and Phylogenetic Analysis of *Propionibacterium acnes*

Mogens Kilian, Christian F. P. Scholz, and Hans B. Lomholt

Department of Biomedicine, Faculty of Health Sciences, Aarhus University, DK-8000 Aarhus C, Denmark

***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.**

The 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 nt (http://pubmlst.org/pacnes/). 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*.

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); CP002815.1, strain 6609 (14); CP002409.1, strain 266 (complete genome) (5); NZ_ADJL00000000, strain J165 (whole-genome shotgun sequencing project); NZ_ADJM00000000, strain SK187 (whole-genome shotgun sequencing project); NZ_ADFS00000000, strain J139 (whole-genome shotgun sequencing project); PRJNA49245, strain HL001PA1; PRJNA49265, strain HL002PA1; PRJNA49267, strain HL002PA2; PRJNA49269, strain HL002PA3; PRJNA49225, strain HL005PA1; PRJNA49227, strain HL005PA2; PRJNA49229, strain HL005PA3; PRJNA49231, strain

Received 25 October 2011 Returned for modification 5 December 2011
Accepted 15 December 2011
Published ahead of print 28 December 2011
Address correspondence to Mogens Kilian, kilian@microbiology.au.dk.

doi:10.1128/JCM.06129-11

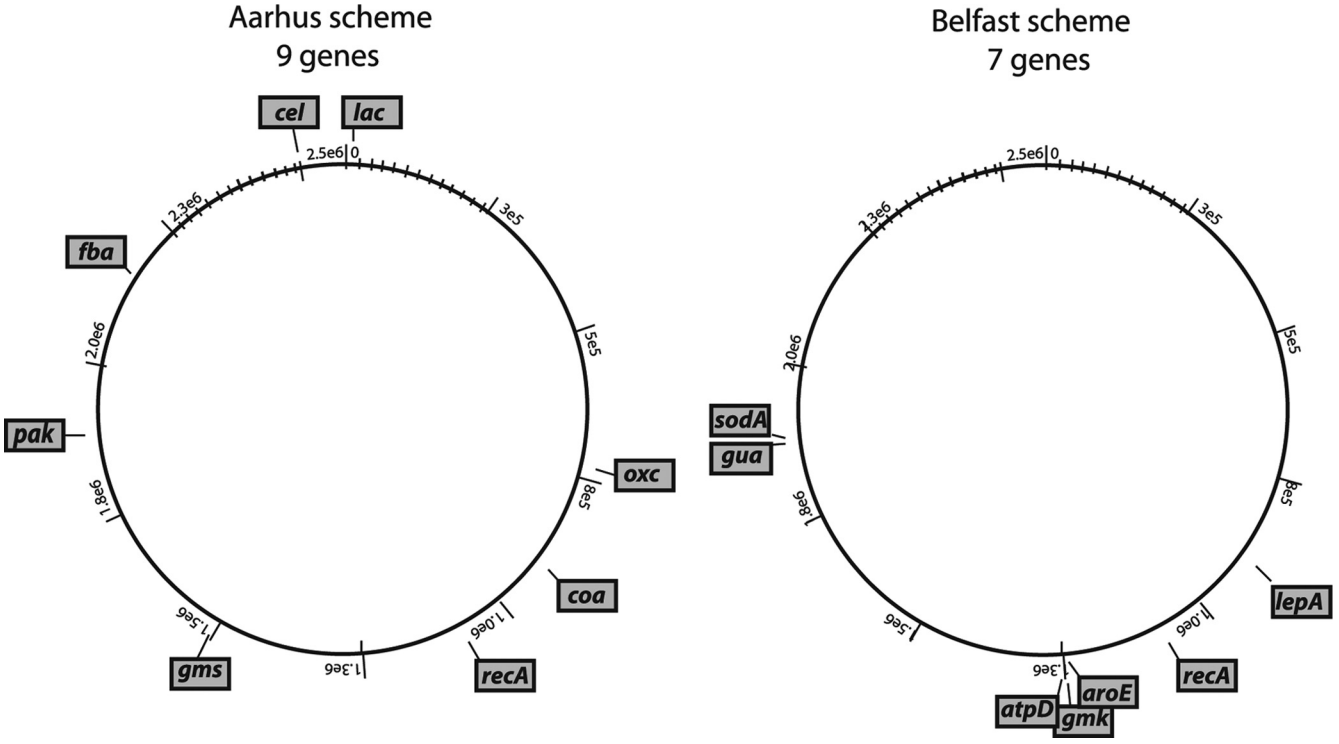

FIG 1 Map of the *P. acnes* genome with the location of the nine genes used in the Aarhus MLST scheme and seven genes used in the Belfast MLST scheme.

HL005PA4; PRJNA49271, strain HL007PA1; PRJNA49169, strain HL013PA1; PRJNA49171, strain HL013PA2; PRJNA49161, strain HL020PA1; PRJNA49211, strain HL025PA1; PRJNA49213, strain HL025PA2; PRJNA49257, strain HL027PA1; PRJNA49259, 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; PRJNA49215, strain HL059PA1; PRJNA49217, 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 are 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), ferrochelatase (HemH) (359218 to 361269), aspartate-semialdehyde dehydrogenase (369582 to 370895), putative *c*-type cytochrome biogenesis protein (376261 to 377772), putative phosphoenolpyruvate-protein kinase (402218 to 403936), L-asparaginase I (413845 to 414840), cysteinyl-tRNA synthetase (431282 to 432754), L-serine dehydratase (440439 to 441875), zinc-binding dehydrogenase (455425 to 456441), cobalt transport protein CbiQ (477344 to 478189), 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

591899), uracil-DNA glycosylase (618243 to 618959), serine protease (662715 to 664187), zinc-binding dehydrogenase (686124 to 687173), Xaa-Pro aminopeptidase I (714480 to 715976), lipoyl synthase (758856 to 759848), putative acyltransferase (795741 to 796496), UDP-N-acetylmuramyl pentapeptide synthase (828645 to 830102), oxidoreductase and putative D-lactate dehydrogenase (850574 to 853003), Zn-dependent hydrolase (884716 to 885372), GTP-binding protein LepA (981470 to 983305), putative metal-dependent hydrolase (1020344 to 1021834), acyl carrier protein (1065584 to 1065829), LexA repressor (1108732 to 1109406), tryptophanase (1163084 to 1164460), putative alpha-amylase (1211284 to 1213410), DNA polymerase III subunit alpha (1249423 to 1253085), bifunctional phosphopantothenoylcysteine (1297294 to 1298526), uracil-DNA glycosylase (1342891 to 1343727), sensory transduction histidine kinase (1371210 to 1372685), metal-dependent hydrolase (1418651 to 1419319), D-alanine-D-alanine ligase (1483346 to 1484443), inorganic polyphosphate ATP-NAD kinase (1513522 to 1514478), amido hydrolase family protein (1559524 to 1560687), arabinose operon protein AraM (1604590 to 1605963), enoyl (acyl carrier protein) reductase (1653180 to 1653953), trigger factor (1702280 to 1703872), putative long-chain fatty acid-coenzyme A (CoA) ligase (AMP-binding enzyme) (1773902 to 1775809), glutamate racemase (1828307 to 1829200), putative fatty acid-CoA ligase (AMP-binding enzyme) (1878129 to 1879802), cardiolipin synthetase (1923523 to 1924791), DNA-directed RNA polymerase subunit beta (2045208 to 2048687), M13 family metallopeptidase (2070668 to 2072650), sugar phosphatase-hydrolase (2115351 to 2116169), adenylosuccinate synthetase (2166495 to 2167781), chaperone protein DnaJ (2211718 to 2212860), TetR family transcriptional regulator (2254400 to 2255035), putative two-component sensor kinase (2286813 to 2288243), putative intracellular protease (general stress protein) (2335002 to 2335670), glutamine amidotransferase/GMP synthase (2369851 to 2370552), PTS system fructose-mannitol specific IIA subunit (2414767 to 2415228), D-serine-D-alanine-glycine transporter (2456880 to 2458379), OsmC-like protein (2513545 to 2513982), and FemAB family protein (2556010 to 2557101). The individual genes were concatenated in the mentioned order into a single sequence of 92,577 bp. Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 5 (31). Phylogenetic trees were constructed using the minimum evolution algorithm, and bootstrap analysis was carried out using 500 replications.

The identification of clonal complexes (CCs) and their founders based on allele profiles was achieved by eBURST analysis at http://eburst.mlst.net/ using the eBURSTST version 2 clustering algorithm, which was developed and is hosted by Imperial College London and is based on principles originally described by Feil et al. (10).

The comparison of gene contents in the draft genomes and in the clonal complexes identified in the study was performed by Blastn at http://blast.ncbi.nlm.nih.gov/ using the genes identified in the two complete annotated genomes of strains KPA171202 and 266 as the reference.

RESULTS

MLST genes and typing. All genes relevant to MLST were successfully retrieved from all of the draft and complete genomes. Each gene fragment was assigned to an allele type in the two MLST schemes by introducing them into the respective databases as a query sequence and by tree construction using reference sequences for all recognized alleles. The respective allelic profiles of loci were used for the assignment of the 78 genome-sequenced strains to sequence types (STs) according to each of the two MLST schemes. Among the 78 strains, three (HL037PA2, HL0PA3, and HL044PA1) showed highly divergent sequences and were excluded from further analysis (see below). According to the Aarhus scheme, a total of 32 distinct STs were identified (here referred to as ST1, ST2, etc.) among the remaining 75 strains. A total of 19 STs were distinguished according to the Belfast scheme (here referred to as ST-1, ST-2, etc.). According to the Aarhus scheme, 48 of the 75 strains were assigned to one of the 59 STs previously demonstrated among a total of 210 predominantly northern European isolates (19). The remaining 27 strains were assigned to 19 new STs (ST60 to ST78) added to the Aarhus MLST database at http://pacnes.mlst.net/. According to the Belfast scheme, all but one (21.1.L1) of the 75 strains could be assigned to one of the 58 STs recognized as of August 2011 in the database at http://pubmlst.org/pacnes/.

Reference phylogenetic tree. The phylogenetic analysis of 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 rRNA 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 "*Propionibacterium humerusii*," 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 concatemer identified 11 housekeeping genes that were almost completely conserved in the entire collection of genomes and an additional 14 genes that separated CC60 (see below) strains but failed to distinguish between the remaining clusters.

eBURST analysis. The BURST algorithm identifies mutually exclusive groups of related genotypes in a bacterial population and attempts to identify the founding genotype (ST) of each group. The analysis of each of the two MLST data sets based on the 75 genome-sequenced strains combined with all strains in the two respective MLST databases identified eight CCs according to the Aarhus scheme and seven clonal complexes according to the Belfast scheme. A few strains were singletons (Fig. 3). The assignment of each of the 75 genome-sequenced strains to the individual CCs defined by the two data sets is shown in Fig. 2 as they relate to the phylogenetic tree based on a 92,577-bp concatemer of 76 housekeeping genes evenly distributed over the entire genome. As shown in the figure, clonal clusters identified by the eBURST analysis based on data generated by the MLST scheme correlated, almost without exception, with clades identified by the phylogenetic analysis.

Resolution achieved by the two MLST schemes. A comparison of the typing results achieved by the two MLST schemes reveals notable differences (Fig. 2). Thirty-five strains (47%) were assigned to a single ST, ST-6, according to the Belfast MLST scheme. Among these same 35 strains, the Aarhus scheme distinguished a total of 11 STs: ST3 ($n = 10$), ST8 ($n = 2$), ST12 ($n = 1$), ST18 ($n = 7$), ST27 ($n = 2$), ST29 ($n = 1$), ST31 ($n = 5$), ST58 ($n = 2$), ST59 ($n = 3$), ST63 ($n = 1$), and ST77 ($n = 1$). The

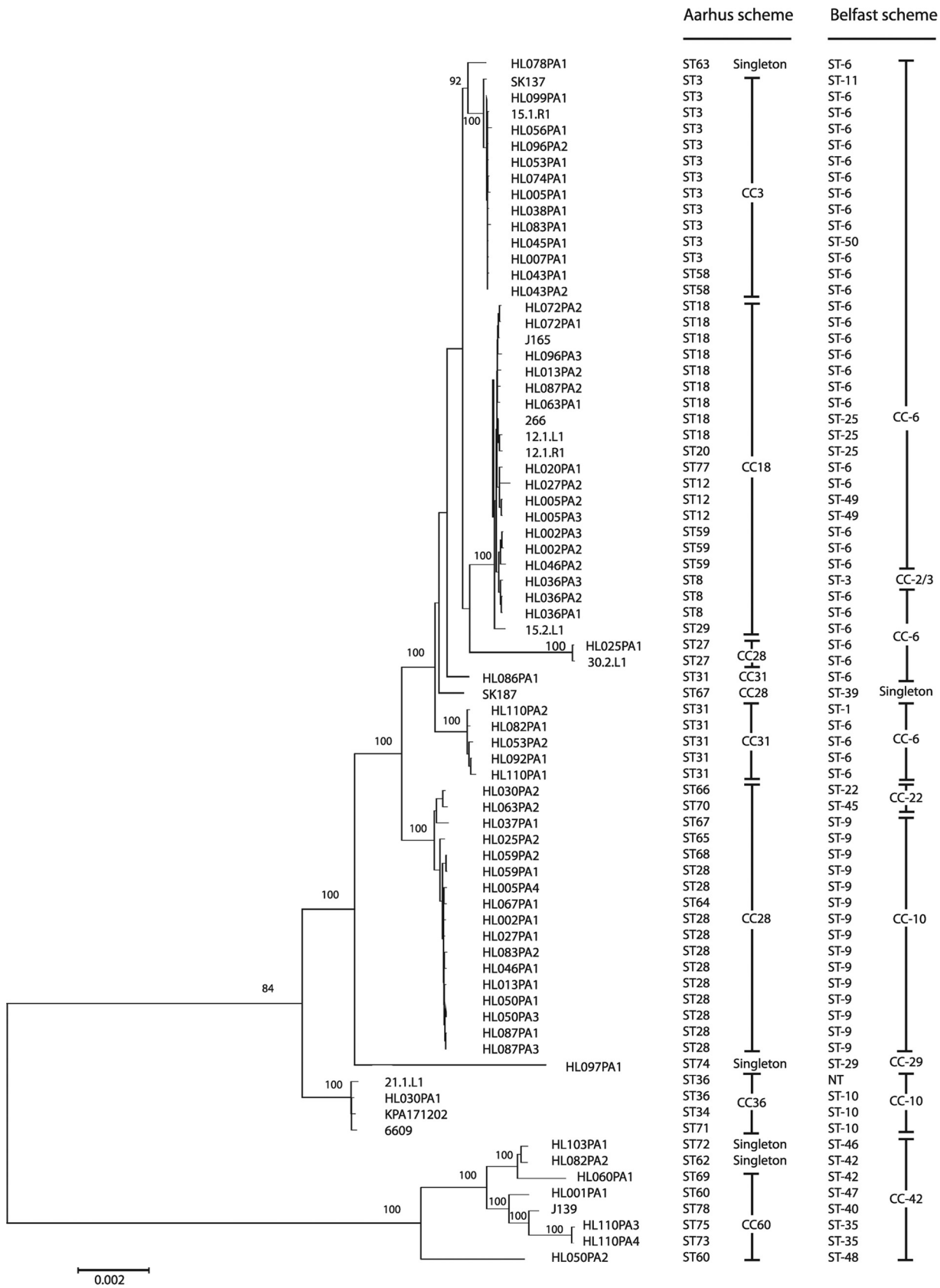

FIG 2 Phylogenetic tree of 75 strains of *P. acnes* constructed by the minimum evolution algorithm in MEGA version 5 and based on a 92,577-bp concatemer of complete sequences of 76 housekeeping genes evenly distributed across the entire genome. For each strain the ET and CC determined according to the 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.

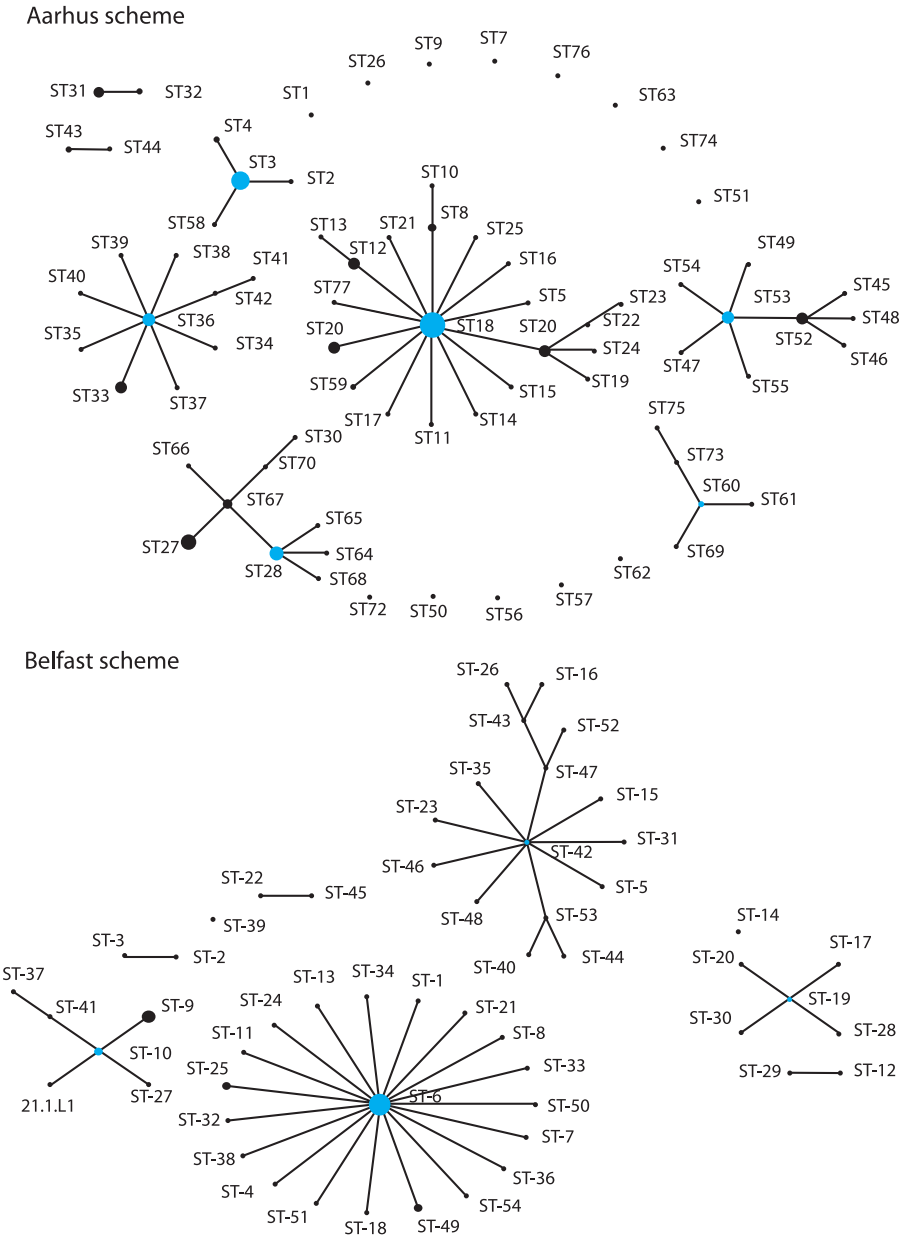

FIG 3 Population snapshots of *P. acnes* generated by eBURST analysis of MLST allele profiles determined according to the Aarhus and Belfast MLST schemes. The analysis was based on data sets that include all STs in the two MLST databases supplemented with data retrieved from 75 genome-sequenced strains. Clonal clusters (CCs) are named after the assumed founding ST, which are indicated by blue dots. Note that ST numbers in the two schemes do not correspond to each other.

significant phylogenetic diversity of these strains is confirmed by the tree in Fig. 2 and supported by significant bootstrap values (100%).

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 *tly* (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), protein associated with putative adhesion protein (332676057), hypothetical protein PAZ_c13420 (332675693), hypothetical protein PAZ_c13390 (332675690), linear gramicidin synthase 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_13380-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 concatemer 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 concatemer, 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 concatemer of complete sequences of 76 housekeeping genes constituting more than 92,000 nt (genetic distance, 0.8% $\pm$ 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

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 phylogenic 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-like genomic region were found in all CC18 strains but were absent from all CC3 strains. Accordingly, our previous study demonstrated that strains in CC3 and CC28 differ from CC18 strains by biotype and by distinctive alleles of the two putative virulence genes *camp5* and *tly*.

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. Aanensen 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

1. **Alexeyev OA, et al.** 2007. Direct visualization of *Propionibacterium acnes* in prostate tissue by multicolor fluorescent in situ hybridization assay. J. Clin. Microbiol. **45**:3721–3728.
2. **Bek-Thomsen M, Lomholt H, Kilian M.** 2008. Acne is not associated with yet uncultured bacteria. J. Clin. Microbiol. **46**:3355–3360.
3. **Brüggemann H, et al.** 2004. The complete genome sequence *of Propionibacterium acnes*, a commensal of human skin. Science **305**:671–673.
4. **Brüggemann H.** 2005. Insights in the pathogenic potential of *Propionibacterium acnes* from its complete genome. Semin. Cutan. Med. Surg. **24**:67–72.
5. **Brzuszkiewicz E, et al.** 2011. Comparative genomics and transcriptomics of *Propionibacterium acnes.* PLoS One **6**:e21581.
6. **Butler-Wu SM, Sengupta DJ, Kittichotirat W, Matsen FA, III, Bumgarner RE.** 2011. Genome sequence of a novel species, *Propionibacterium humerusii.* J. Bacteriol. **193**:3678.
7. **Cohen RJ, Shannon BA, McNeal JE, Shannon T, Garrett KL.** 2005. *Propionibacterium acnes* associated with inflammation in radical prostatectomy specimens: a possible link to cancer evolution? J. Urol. **173**:1969–1974.
8. **Dessinioti C, Katsambas AD.** 2010. The role of *Propionibacterium acnes* in acne pathogenesis: facts and controversies. Clin. Dermatol. **28**:2–7.
9. **Fassi Fehri L, et al.** 2011. Prevalence of *Propionibacterium acnes* in diseased prostates and its inflammatory and transforming activity on prostate epithelial cells. Int. J. Med. Microbiol. **301**:69–78.
10. **Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG.** 2004. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J. Bacteriol. **186**:1518–1530.
11. **Grange PA, et al.** 2009. Production of superoxide anions by keratinocytes initiates P. acnes-induced inflammation of the skin. PLoS Pathog. **5**:e1000527 1–14.
12. **Grice EA, et al.** 2009. Topographical and temporal diversity of the human skin microbiome. Science **324**:1190–1192.
13. **Grice EA, Segre JA.** 2011. The skin microbiome. Nat. Rev. Microbiol. **9**:244–253.
14. **Hunyadkürti J, et al.** 2011. Complete genome sequence of *Propionibacterium acnes* type IB strain 6609. J. Bacteriol. **193**:4561–4562.
15. **Jappe U, et al.** 2004. Evidence for diversity within *Propionibacterium acnes*: a comparison of the T-cell stimulatory activity of isolates from inflammatory acne, endocarditis and the laboratory. J. Eur. Acad. Dermatol. Venereol. **18**:450–454.

16. **Javey G, Albini TA, Jr, Flynn HW.** 2010. Resolution of pigmented keratic precipitates following treatment of pseudophakic endophthalmitis caused by *Propionibacterium acnes.* Ophthalmic Surg. Laser Imaging **9**:1–3.
17. **Kong HH.** 2011. Skin microbiome: genomics-based insights into the diversity and role of skin microbes. Trends Mol. Med. **17**:320–328.
18. **Kurokawa I, et al.** 2009. New developments in our understanding of acne pathogenisis and treatment. Exp. Dermatol. **18**:821–832.
19. **Lomholt H, Kilian M.** 2010. Population genetic analysis of *Propionibacterium acnes* identifies a subpopulation and epidemic clones associated with acne. PLoS One **5**:e12277.
20. **Maiden MC, et al.** 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc. Natl. Acad. Sci. U. S. A. **95**:3140–3145.
21. **McDowell A, et al.** 2005. *Propionibacterium acnes* types I and II represent phylogenetically distinct groups. J. Clin. Microbiol. **43**:326–334.
22. **McDowell A, Perry AL, Lambert PA, Patrick S.** 2008. A new phylogenetic group of *Propionibacterium acnes.* J. Med. Microbiol. **57**:218–224.
23. **McDowell A, et al.** 2011. A novel multilocus sequence typing scheme for the opportunistic pathogen *Propionibacterium acnes* and characterization of type I cell surface-associated antigens. Microbiology **157**:1990–2003.
24. **Mourelatos K, Eady EA, Cunliffe WJ, Clark SM, Cove JH.** 2007. Temporal changes in sebum excretion and propionibacterial colonization in preadolescent children with and without acne. Br. J. Dermatol. **156**:22–31.
25. **Nagy I, et al.** 2005. Distinct strains of *Propionibacterium acnes* induce selective human beta-defensin-2 and interleukin-8 expression in human keratinocytes through toll-like receptors. J. Investig. Dermatol. **124**:931–938.
26. **Nagy I, et al.** 2006. *Propionibacterium acnes* and lipopolysaccharide induce the expression of antimicrobial peptides and proinflammatory cytokines/chemokines in human sebocytes. Microbes Infect. **8**:2195–2205.
27. **Niazi SA, et al.** 2010. *Propionibacterium acnes* and *Staphylococcus epidermidis* isolated from refractory endodontic lesions are opportunistic pathogens. J. Clin. Microbiol. **48**:3859–3869.
28. **Sampedro MF, et al.** 2009. Species of *Propionibacterium* and *Propionibacterium acnes* phylotypes associated with orthopedic implants. Diagn. Microbiol. Infect. Dis. **64**:138–145.
29. **Schaeverbeke T, et al.** 1998. *Propionibacterium acnes* isolated from synovial tissue and fluid in a patient with oligoarthritis associated with acne and pustulosis. Arthritis Rheum. **41**:1889–1893.
30. **Szabó K, Kemény L.** 2011. Studying the genetic predisposing factors in the pathogenesis of acne vulgaris. Hum. Immunol. **72**:766–773.
31. **Tamura K, et al.** 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28**:2731–2739.
32. **Uçkay I, et al.** 2010. Spondylodiscitis due to *Propionibacterium acnes*: report of twenty-nine cases and a review of the literature. Clin. Microbiol. Infect. **16**:353–358.
33. **Valanne S, et al.** 2005. CAMP factor homologues in *Propionibacterium acnes*: a new protein family differentially expressed by types I and II. Microbiology **151**:1369–1379.
34. **Zeller V, et al.** 2007. *Propionibacterium acnes*: an agent of prosthetic joint infection and colonization. J. Infect. **55**:119–124.