Propionibacterium acnes: disease-causing agent or common contaminant? Detection in diverse patient samples by next generation sequencing

Mollerup, Sarah; Friis-Nielsen, Jens; Vinner, Lasse; Richter, Stine Raith; Fridholm, Helena; Herrera, Jose Romero; Lund, Ole; Brunak, Søren; Gonzalez-Izarzugaza, Jose Maria; Mourier, Tobias

Total number of authors:
12

Published in:
Journal of Clinical Microbiology

Link to article, DOI:
10.1128/jcm.02723-15

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Propionibacterium acnes: Disease-Causing Agent or Common Contaminant? Detection in Diverse Patient Samples by Next-Generation Sequencing

Sarah Mollerup, Jens Friis-Nielsen, Lasse Vinner, Thomas Arn Hansen, Stine Raith Richter, Helena Fridholm, Jose Alejandro Romero Herrera, Ole Lund, Søren Brunak, Jose M. G. Izarzugaza, Tobias Mourier, Lars Peter Nielsen, Anders Johannes Hansen

Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark; Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Kongens Lyngby, Denmark; Disease Systems Biology Program, Panum Institute, University of Copenhagen, Copenhagen, Denmark; Department of Autoimmunology and Biomarkers, Statens Serum Institut, Copenhagen S, and Aalborg University, Health Sciences, Aalborg, Denmark

Propionibacterium acnes is the most abundant bacterium on human skin, particularly in sebaceous areas. P. acnes is suggested to be an opportunistic pathogen involved in the development of diverse medical conditions but is also a proven contaminant of human clinical samples and surgical wounds. Its significance as a pathogen is consequently a matter of debate. In the present study, we investigated the presence of P. acnes DNA in 250 next-generation sequencing data sets generated from 180 samples of 20 different sample types, mostly of cancerous origin. The samples were subjected to either microbial enrichment, involving nuclease treatment to reduce the amount of host nucleic acids, or shotgun sequencing. We detected high proportions of P. acnes DNA in enriched samples, particularly skin tissue-derived and other tissue samples, with the levels being higher in enriched samples than in shotgun-sequenced samples. P. acnes reads were detected in most samples analyzed, though the proportions in most shotgun-sequenced samples were low. Our results show that P. acnes can be detected in practically all sample types when molecular methods, such as next-generation sequencing, are employed. The possibility of contamination from the patient or other sources, including laboratory reagents or environment, should therefore always be considered carefully when P. acnes is detected in clinical samples. We advocate that detection of P. acnes always be accompanied by experiments validating the association between this bacterium and any clinical condition.

Propionibacterium acnes is a facultative anaerobic Gram-positive bacterium present on human skin as part of the normal flora, as well as in the oral cavity, large intestine, conjunctiva, and external ear canal (1). P. acnes is the most prevalent bacterium in sebaceous areas of the skin (2, 3) but is also abundant in dry areas (3). P. acnes predominates in the pilosebaceous follicles of the skin (4).

P. acnes is proposed to play a role in the development of acne vulgaris (5) and is considered an opportunistic pathogen causing postoperative infections (6). It has been isolated from patients with endocarditis (7), synovitis, acne, pustulosis, hyperostosis, and osteitis (SAPHO) syndrome (8), and sarcoidosis (9), among other syndromes. Some studies also identify P. acnes to be a contaminant of blood products, tissue cultures, and surgical wounds (10–13), and the significance of this widely abundant skin commensal is consequently debated.

The postulated role of P. acnes in various conditions is often highly speculative and based on the mere detection of the bacterium (8, 9, 14–19). To expand the knowledge of the abundance of P. acnes, we investigated the proportions of P. acnes sequences in samples of a wide variety of cancer types, as well as blood samples. Next-generation sequencing was performed on DNA extracted either after enrichment for intact microbes or directly from the samples. This study is the first large-scale analysis investigating the abundance of P. acnes DNA in such diverse tissue types by next-generation sequencing, and we show that P. acnes can be readily detected in literally all sample types investigated, particularly when samples are subjected to microbial enrichment.

MATERIALS AND METHODS

Ethics statement. Human sample collection, handling, and analysis were performed under ethical protocol H-2-2012-FSP2 (Regional Committee on Health Research Ethics, Capital Region of Denmark) and case no. 1304226 (National Committee on Health Research Ethics, Denmark). In accordance with Danish national legislation (Sundhedsloven), all human samples were processed anonymously.
were wearing a lab coat and gloves, as well as sleeves during extraction and amplification. All laboratory work was conducted by individuals who directed work flow. Separate laboratories were used for reagent preparation, whereas pelleted cells were used for shotgun sequencing. The supernatants were subsequently filtered through 5-pore-size centrifuge filters (Millipore, Darmstadt, Germany). To remove unprotected nucleic acids while retaining the nucleic acids within bacterial cells or viral particles, the filtrates were nuclease digested using 14 μl Turbo DNase (2 U/μl; Ambion, Thermo Fisher Scientific, Waltham, MA, USA), 12 μl Baseline-Zero DNase (1 U/μl; Epicentre, Madison, WI, USA), 16 μl RNase cocktail enzyme mix (Ambion), and 40 μl 10× Turbo DNase buffer in a total volume of 400 μl. Nucleic acids from the enriched samples were extracted using the High Pure viral RNA kit (Roche, Basel, Switzerland) according to the manufacturer’s instructions with the addition of 10 μg linear acrylamide carrier (Applied Biosystems, Foster City, CA, USA).

DNA libraries were built using the Nextera or Nextera XT DNA sample preparation kit (Illumina, San Diego, CA, USA), according to the manufacturer’s instructions. Nextera libraries were amplified by 20 cycles of PCR, and Nextera XT libraries were amplified by 12 cycles of PCR. The amplified and indexed libraries were purified using the

**Patient samples.** (i) **Samples collected in Denmark.** Malignant melanoma biopsy specimens and cells from patients with acute myeloid leukemia (AML), B-cell chronic lymphocytic leukemia (B-CLL), and chronic myelogenous leukemia (CML) sorted by fluorescence-activated cell sorting (FACS) were obtained from Aarhus University Hospital. Samples from patients with T-lineage acute lymphoblastic leukemia (T-ALL) were obtained from either Aarhus University Hospital (cells sorted by FACS) or Rigshospitalet (Copenhagen University Hospital; bone marrow samples; n = 9). Samples (bone marrow) from patients with B-cell precursor acute lymphoblastic leukemia (BCP-ALL) and biopsy specimens from patients with oropharyngeal or oral cavity head and neck cancer and testicular cancer were obtained from Rigshospitalet. Biopsy specimens from patients with basal cell carcinoma and mycosis fungoides (cutaneous T-cell lymphoma) were obtained from Bispebjerg Hospital. Bladder, breast, and colon cancer biopsy specimens, blood samples from colon cancer patients, and ascitic fluid samples from colon, ovarian, and pancreatic cancer patients were obtained from or collected in collaboration with the Danish Cancer Biobank, Herlev Hospital. Ascitic fluid samples were subjected to low-speed centrifugation upon receipt to pellet intact host cells and cellular debris. The supernatant was used for microbial enrichment, whereas pelletted cells were used for shotgun sequencing. Cryopreserved, fully transformed B-cell lymphoma cell lines (OCI Ly3_M, OCI-Ly7_M, OCI-Ly8, SU-DHL-4, SU-DHL-5, and U698M) and multiple myeloma cell lines (KMS-12-BM, KMS-12-PE, MOLP-2, MOLP-8, RPMI-8226, and U266) were obtained from Aalborg University Hospital.

(ii) **Samples collected outside Denmark.** Vulva cancer biopsy specimens were obtained from the National Institute of Oncology, Budapest, Hungary. All samples are listed in Table 1.

**Laboratory facilities.** All sample processing was performed under conditions designed to diminish contamination through the use of a unidirectional work flow. Separate laboratories were used for reagent preparation, sample extraction and sequencing library preparation, and library amplification. All laboratory work was conducted by individuals who were wearing a lab coat and gloves, as well as sleeves during extraction and library preparation. All sample processing was carried out in laminar-flow cabinets equipped with UV lamps. All benchtops were cleaned with 5% bleach and 70% ethanol before and after any lab work. Only sterile, disposable lab utensils were used. Filter tips were used for all pipetting.

**Microbial enrichment, nucleic acid extraction, and library preparation.** Samples used for enrichment were fresh frozen after collection, and no nucleic acid preservatives were added. Each batch of processed samples included a nontemplate control containing only phosphate-buffered saline (PBS), which was run in parallel with the patient samples (n = 19). Samples were kept on ice unless otherwise stated.

Biopsy specimens were thawed and transferred to 400 μl cold PBS. For the samples from leukemia patients, 150 to 400 μl of cell suspension was mixed with cold PBS to a total volume of 400 μl. Two stainless steel beads of 2 to 3 mm in diameter were added to the samples, which were subsequently homogenized using a TissueLyser II tissue disrupter (Qiagen, Hilden, Germany) for 6 min at 30 Hz. All tissue homogenates, samples from leukemia patients, and ascitic fluid specimens (1 ml) devoid of pelleted cells were centrifuged for 2 min at 800 × g to remove tissue debris, and the supernatants were subsequently filtered through 5-μm pore-size centrifuge filters (Millipore, Darmstadt, Germany). To remove unprotected nucleic acids while retaining the nucleic acids within bacterial cells or viral particles, the filtrates were nuclease digested using 14 μl Turbo DNase (2 U/μl; Ambion, Thermo Fisher Scientific, Waltham, MA, USA), 12 μl Baseline-Zero DNase (1 U/μl; Epicentre, Madison, WI, USA), 16 μl RNase cocktail enzyme mix (Ambion), and 40 μl 10× Turbo DNase buffer in a total volume of 400 μl and incubated at 37°C for 2 h. Nucleic acids from the enriched samples were extracted using the High Pure viral RNA kit (Roche, Basel, Switzerland) according to the manufacturer’s instructions with the addition of 10 μg linear acrylamide carrier (Applied Biosystems, Foster City, CA, USA).
Agencourt AMPure XP PCR purification system (Beckman Coulter, Indianapolis, IN, USA). In case of insufficient amplification, libraries were reamplified using AccuPrime Pfx DNA polymerase (Life Technologies, Carlsbad, CA, USA) and P5 and P7 sequence primers.

**Total DNA extraction and library preparation.** Total DNA was extracted from samples using the QIAamp DNA minikit (Qiagen) following the manufacturer’s instructions. DNA libraries were prepared from 1 µg of DNA using either the TruSeq DNA protocol (catalog number PE-940-2001; Illumina, San Diego, CA, USA) or an in-house protocol (20) using NEBnext reagents (protocol E6070; New England BioLabs, Ipswich, MA, USA).

**Sequencing.** Paired-end sequencing (2 reads of 100 bp) was performed on the Illumina HiSeq 2000 platform at BGI-Europe (Copenhagen N, Denmark).

**Analysis of sequence data.** (i) Preprocessing and depletion of human sequence reads. Paired-end sequencing reads were trimmed of adapter sequences, and overlapping read pairs were merged using AdapterRemoval software (version 1.5.3) (21). Reads shorter than 30 nucleotides after trimming were excluded from further analysis. For digital subtraction of human sequence reads, the remaining reads were mapped to the human genome (hg38) using the mem algorithm implemented in Burrows-Wheeler Aligner (BWA; version 0.7.7) (22). Reads of a pair were evaluated independently. Reads containing 25 bp or more of low-complexity regions were filtered out using the DustMasker algorithm (version 1.0.0) (23).

(ii) **Analysis of reads mapping to** *P. acnes*. Genomic sequences from 97 *P. acnes* strains representing complete genomes, scaffolds, or contigs were downloaded from the NCBI genome resource (http://www.ncbi.nlm.nih.gov/genome/genomes/1140) (see Table S1 in the supplemental material). Reads filtered of human and low-complexity sequences were mapped to the *P. acnes* genome using the aln algorithm implemented in BWA (version 0.7.7) (22). Reads of a pair were evaluated independently. Reads containing 25 bp or more of low-complexity regions were filtered out using the DustMasker algorithm (version 1.0.0) (23).

(iii) **Analysis of the** *P. acnes* **strains present.** The presence of the different *P. acnes* strains was assessed on the basis of the mapping to the *P. acnes* genomic sequences (above). Only reads mapping exclusively to a single strain were considered for this analysis. Information regarding the phylogenetic type of the different strains was obtained from the *P. acnes* multilocus sequence typing (MLST) database (http://pubmlst.org/pacnes/).

(iv) **Analysis of the coverage of the** *P. acnes* **genome.** To visualize the genomic coverage of *P. acnes*, all nonhuman reads were mapped to the 12 full *P. acnes* genomes currently present in GenBank using the mem algorithm implemented in BWA (version 0.7.7) (22). The resulting alignments are represented as Circos plots (25).

**Nucleotide sequence accession numbers.** Sequencing reads mapping to *P. acnes* have been uploaded to the European Nucleotide Archive (ENA) and are available under accession number PRJEB12636.

**RESULTS**

*P. acnes* in microbe-enriched and shotgun-sequenced samples. The relative number of *P. acnes* sequencing reads in 180 samples of 20 different sample types was analyzed (Table 1). Ascitic fluid samples and lymphoma/myeloma cell lines were analyzed as each being one common sample category. One hundred forty-three samples were processed to enrich for intact microbes before extraction of DNA and preparation of sequencing libraries. Sequencing reads from which human reads were removed were mapped to a custom database containing the 97 *P. acnes* genomes currently present in GenBank as full genomes or genome assemblies. *P. acnes* levels were assessed as the number (parts per million) of unique *P. acnes* reads relative to the total number of demultiplexed reads (see Materials and Methods). Using this approach, we detected relative levels of *P. acnes* reads exceeding 300 ppm in samples from patients with basal cell carcinoma (median, 576 ppm), mycosis fungoides (median, 480 ppm), malignant melanoma (median, 446 ppm), and breast cancer (median, 373 ppm) (Fig. 1, black bars; see also Table S2 in the supplemental material).
The proportions of *P. acnes* reads detected in the remaining enriched samples were below 100 ppm (median, 0.8 to 89 ppm). *P. acnes* reads were also detected in nontemplate control samples containing PBS run in parallel with the enriched samples (*n*/H1100519). In these samples, the relative proportion of *P. acnes* reads was high (median, 10,377 ppm) (see Discussion).

Total DNA extracted from 107 samples was subjected to shotgun sequencing (Table 1). Only samples from patients with basal cell carcinoma (median, 3.9 ppm), malignant melanoma (median, 1.2 ppm), and breast cancer (median, 1.1 ppm) contained relative *P. acnes* levels above 1 ppm (Fig. 1, gray bars; see also Table S2 in the supplemental material). The fold increase in the median number of *P. acnes* reads (parts per million) in enriched samples relative to the number in shotgun-sequenced samples ranged from 35 in colon cancer biopsy specimens (12 versus 0.3 ppm) to >1,000 in samples from patients with mycosis fungoides (480 versus 0.5 ppm). The blood samples contained the lowest relative *P. acnes* levels of all samples included in this study (median, 0.05 ppm).

**Distribution of *P. acnes* strains.** To assess the *P. acnes* strains and the phylogenetic groups present in our samples, we analyzed the distribution of reads mapping to the 97 *P. acnes* genomes used in our analysis. Only reads mapping unambiguously to one *P. acnes* strain were included in the analysis. Using this criterion, an average of 8.6% of the *P. acnes* reads were included in the analysis (see Table S3 in the supplemental material); however, with this approach we ensured a high stringency when assigning the reads. Figure 2 shows the relative abundance of strains for each sample. The phylogenetic type is indicated next to the individual strain names. For 20 of the strains included, the phylogenetic type was not determined, and these are designated unknown (Unkn.). Overall, two strains predominated across samples for both enriched and shotgun-sequenced samples (Table 2; see also Table S4).

**FIG 2 Relative distribution of *P. acnes* strains in the individual samples.** The y axis shows the percentage accounted for by each strain, and each vertical line represents a sample. The bottom horizontal lines indicate enriched or shotgun-sequenced samples, and the bottom vertical ticks delimit the sample categories. The phylogenetic type of the *P. acnes* strains is indicated next to the strain name in the key on the right. The strains mentioned in the main text are highlighted in bold. NTC, nontemplate control; Unkn., unknown.

**TABLE 2 Prevalence of the 20 most frequently detected *P. acnes* strains in samples investigated by microbial enrichment, shotgun sequencing, and all samples combined.**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Prevalence (%) in:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All samples</td>
</tr>
<tr>
<td>I19_PAVI</td>
<td>82</td>
</tr>
<tr>
<td>SK182B-JCVI</td>
<td>78</td>
</tr>
<tr>
<td>HL025PA1</td>
<td>56</td>
</tr>
<tr>
<td>JCM 18920</td>
<td>45</td>
</tr>
<tr>
<td>HL050PA2</td>
<td>44</td>
</tr>
<tr>
<td>HL001PA1</td>
<td>34</td>
</tr>
<tr>
<td>HL411PA1</td>
<td>32</td>
</tr>
<tr>
<td>C1</td>
<td>31</td>
</tr>
<tr>
<td>HL110PA4</td>
<td>29</td>
</tr>
<tr>
<td>ATCC 6919</td>
<td>27</td>
</tr>
<tr>
<td>JCM 18909</td>
<td>26</td>
</tr>
<tr>
<td>HL082PA2</td>
<td>24</td>
</tr>
<tr>
<td>J165</td>
<td>24</td>
</tr>
<tr>
<td>HL201PA1</td>
<td>20</td>
</tr>
<tr>
<td>HL110PA1</td>
<td>19</td>
</tr>
<tr>
<td>HL030PA1</td>
<td>16</td>
</tr>
<tr>
<td>SK187</td>
<td>15</td>
</tr>
<tr>
<td>HL099PA1</td>
<td>15</td>
</tr>
<tr>
<td>HL078PA1</td>
<td>12</td>
</tr>
<tr>
<td>J139</td>
<td>11</td>
</tr>
<tr>
<td>HL087PA2</td>
<td>15</td>
</tr>
<tr>
<td>KPA171202</td>
<td>14</td>
</tr>
<tr>
<td>HL056PA1</td>
<td>7</td>
</tr>
<tr>
<td>HL110PA3</td>
<td>6</td>
</tr>
<tr>
<td>HL025PA2</td>
<td>4</td>
</tr>
<tr>
<td>HL046PA2</td>
<td>3</td>
</tr>
<tr>
<td>HL074PA1</td>
<td>3</td>
</tr>
</tbody>
</table>
in the supplemental material), namely, strains 119_PAVI and SK182B-JCVI (both of unknown phylogenetic type). Strains HL025PA1 and HL050PA2 (type IA, and type II, respectively) were also among the 10 most abundant strains detected by both methods. Certain strains, such as JCM 18920 (unknown type), HL001PA1 (type II), and C1 (type IA2), were mainly detected in enriched samples, while other strains, such as HL110PA4 (type II) and HL411PA1 (unknown type), were mainly detected in shotgun-sequenced samples.

Generally, greater strain diversity was seen in enriched tissue samples (Fig. 2, left, indicated by the striated pattern seen for those samples; see also Fig. S1 in the supplemental material) than in shotgun-sequenced samples, ranging from an average of 10 different strains in samples from patients with testicular cancer to an average of 29 different strains in samples from patients with mycosis fungoides. The remaining samples contained an average of 2 to 9 different strains.

**Coverage of the *P. acnes* genome.** To evaluate the genomic coverage of *P. acnes*, nonhuman sequencing reads mapping to the 12 full genomes were presented using the Circos program (see Fig. S2 at http://www.cbs.dtu.dk/~txema/Supplementary/propioni_web/index.html). The plots show evenly distributed sequencing reads leading to almost full coverage of the *P. acnes* genome in some samples, whereas more sporadic coverage was seen in samples containing lower *P. acnes* levels. A disproportionately high sequencing depth was seen in two or three regions, depending on which reference genome was used for mapping, corresponding to the regions containing the genes encoding the rRNAs. The high depth of sequencing could be a result of reads originating from bacteria other than *P. acnes* mapping to the rRNA regions as a result of high sequence similarity between different bacterial rRNAs, as the algorithm used for generating the plots allows some mismatches, unlike the more stringent mapping algorithm used for quantification. To investigate if an excess of reads in the rRNA regions impacted the quantitative proportions of *P. acnes*, we reanalyzed these data by excluding reads mapping to the rRNA regions. For the majority of the samples, the reanalyzed values of the median relative levels of unique *P. acnes* reads were <1% lower than the values presented above, except for some sample types containing very low numbers of *P. acnes* reads (see Table S2 in the supplemental material). These results indicate that an excess of reads mapping to rRNA regions generally was not found in our quantification analysis.

**DISCUSSION**

During metagenomic analysis of next-generation sequencing data from patient samples, we identified considerable amounts of genetic material from the bacterium *P. acnes*. *P. acnes* is known to contaminate human samples, and this initial discovery therefore prompted us to investigate the abundance of *P. acnes* in metagenomic data from a range of different human clinical samples. In the present study, we investigated the proportions of *P. acnes* DNA in data from samples subjected to enrichment for intact microbes and from samples subjected to shotgun sequencing of the total pool of DNA. We detected the highest relative levels of *P. acnes* in samples subjected to enrichment, with relative *P. acnes* levels being the highest in skin tissue-derived samples and breast cancer biopsy specimens (Fig. 1). Basal cell carcinoma samples contained the highest *P. acnes* levels among the shotgun-sequenced samples, while most other shotgun-sequenced samples as well as enriched samples from patients with BCP-ALL contained *P. acnes* at levels below 1 ppm (see Table S2 in the supplemental material).

Prior to surgical procedures or penetration of the skin for sampling of blood, the skin is routinely sterilized using antiseptics. However, antiseptic preparation of the skin does not kill all bacteria or affect bacteria present in the deeper layers of the skin (26, 27), and bacteria from surrounding areas of the skin may rapidly colonize the antiseptically treated areas (28, 29). Moreover, antiseptics do not remove extracellular bacterial DNA present on the skin. Therefore, detection of *P. acnes* in skin tissue samples is not unexpected. As the breast tissue lies just beneath the subcutaneous adipose tissue in an area of the skin rich in *P. acnes*, contamination from the skin during surgery is conceivable.

In our study, microbial enrichment was achieved by reducing the amount of host endogenous DNA by centrifugation and filtration followed by nuclease treatment of fluid samples or tissue homogenates. This procedure is widely used for enrichment of intact viral particles in virus discovery studies (30–33) and results in an increase in the amount of intact viral nucleic acids relative to the amount of host nucleic acids (34). We did not culture bacteria from the samples, and therefore, we do not know if the detected bacterial DNA originated from live bacteria capable of forming colonies. As the DNA of live bacteria is protected from degradation by DNases (35, 36), the up to >1,000-fold increase in the median proportion of *P. acnes* reads achieved upon enrichment could suggest that the *P. acnes* DNA detected in our samples at least partially originated from live bacteria.

Over the years, *P. acnes* has been detected in samples from patients with diverse diseases, which has led to its proposed association with several conditions. Among these, *P. acnes* has been implicated in the pathogenesis of acne vulgaris, although its role in this condition is still debated (37). Other diseases for which the bacterium is proposed to play a role are sarcoidosis (9), diverse spine conditions (15–17), SAPHO syndrome (8), aortic aneurysms (18), and prostate cancer (19). Common to these studies is that detection of *P. acnes* is based on culture, PCR, or a combination of the two. In most of the studies, roughly half of the samples are positive, and the identified bacterium is *P. acnes* in about half of the cases. Besides the proposed role in certain diseases and syndromes, *P. acnes* is considered a contaminant of blood products (11, 38), surgical wounds (12, 39), and heart valve tissue cultures used in the diagnosis of infection during valve replacement surgery (10). The fact that *P. acnes* is often detected at similar frequencies in surgical wounds and samples from patients with various diseases could imply that the detection of *P. acnes* is incidental as a result of its widespread presence. Studies investigating the role of *P. acnes* or other ubiquitous bacteria in infections or diseases should therefore be designed to allow researchers to distinguish between contamination and true infection. Approaches suggested to help distinguish between these include the performance of multiple cultures (11), histopathological analysis (7), distinguishing of single bacterial cells from biofilm-related infections by immunofluorescence microscopy (12), and investigation of the ability of *P. acnes* isolates to form biofilm (40). The invasiveness of *P. acnes* isolates may also hint to a possible disease-causing role of *P. acnes* in some cases (41). MLST schemes have revealed three major phylogenetic groups of *P. acnes*, types I, II, and III (42, 43), with type I being further divided into types IA, IB, and IC. Clustering analysis found acne vulgaris to be predominantly associated with certain clonal complexes within type
...while types IB and II were more frequently recovered from soft tissue and retrieved medical devices. Certain P. acnes sequence types from types II and III were recovered only from healthy skin (44).

In our samples, recurring P. acnes strains predominated across sample types, including in the nontemplate controls (see Table S4 in the supplemental material). Two of these strains were also ubiquitously detected in skin microbiomes (3). These findings could suggest that some strains are widespread among individuals or that they arise from reagent or lab contamination, as P. acnes or propionibacteria are well-known contaminants of next-genera-
ing sequencing data (45–47). Our nontemplate controls contain-
ing PBS run in parallel with the enriched samples contained high
relative levels of P. acnes. This could occur as a result of cross-
contamination during sample processing, contamination from lab personnel or the reagents used, or bleeding over from libraries with high P. acnes levels sequenced on the same lanes of the flow cell (34, 48). With the exception of two, all the nontemplate con-
trol libraries contained DNA at levels below the detection limit of both the Qubit and Bioanalyzer instruments (data not shown), but the libraries were nevertheless sequenced. Comparison of the nontemplate controls data sets to the data sets for the other en-
riched samples showed that the controls contained an average of 12% human reads, while the biological samples contained an av-
erage of 91% human reads (see Table S2 in the supplemental ma-
terial). This dissimilar composition of the data sets suggests that
the sequence reads in the controls, including the P. acnes reads, did
not mainly originate from cross contamination or bleeding over.
Studies investigating contamination in next-generation sequenc-
ing data show that more dilute samples contain higher fractions of contaminating reads (46, 47) and that negative controls contain
reads mapping to a variety of bacterial genera, including propi-
onibacteria (47). This could be explained by the more extensive
PCR amplification of contaminant templates from samples con-
taining nearly no DNA than of the same amount of contaminant
templates from complex samples containing human DNA as well
competing for the same quantity of primers and other PCR re-
agents. Considering this skewed amplification and the circum-
stance that the biological data sets are diluted with a large amount
of endogenous human reads, the relative level of P. acnes reads in
the controls becomes less astounding. The detection of P. acnes in
the controls can thus be interpreted as further evidence of how
easily samples are contaminated with this ubiquitous bacterium,
regardless of the source.

We believe that the high relative P. acnes levels detected, par-
ticularly in enriched skin tissue and breast cancer biopsy samples,
reflect patient-derived contamination. Similar patterns showing
traces of P. acnes in most sample types and higher levels in samples
from P. acnes-rich body sites have been seen in other microbiome
studies (49). Our hypothesis is supported by the high diversity of
P. acnes strains seen in these samples compared to the low diversity
seen in other samples, including the controls (see Fig. S1 in the
supplemental material). The low strain diversity in the controls
supports the hypothesis that overamplification occurs in these
samples. More adequate negative controls for P. acnes levels in
our study were the blood samples from colon cancer patients, as
these are, in principle, sterile samples containing high levels of endogenous DNA. In these samples, we detected practically no P.
acnes reads (median, 5 in absolute numbers, or 0.05 ppm; see Table S2 in the supplemental material).

Considering our findings, it is possible that contamination de-

erived from the patient, the hospital personnel, or the lab personnel
or that arises during sample processing and analysis contributes to
the presence of P. acnes DNA in samples processed in clinical and
molecular laboratories. The risk of contamination should there-
fore be considered carefully when interpreting molecular and se-
quencing data showing the presence of P. acnes, and negative con-
trols should always be included. Typing of P. acnes can be a
relevant tool when tracing the origin of P. acnes, but it does not
necessarily provide clinically relevant information, as presumed
cutaneous P. acnes strains are readily detected in diverse sample
types. Efforts should be made to exclude the possibility that P.
acnes originates from contamination and, importantly, to prove
the causal relationship between this bacterium and the disease in
question.

ACKNOWLEDGMENTS
We thank Eske Willerslev for support and the use of laboratories. We
thank Robert Gniadecki (Bispebjerg Hospital), Christian von Buchwald
(Rigshospitalet), Ewa Rajpert-De Meyts (Rigshospital), Line Groth-Pe-
dersen (Rigshospital), Torben Steindke (Aarhus University Hospital),
Hans Johnsen and Karen Dybkær (Aalborg University Hospital), Estrid
Hogdal (The Danish Cancer Biobank), Peter Hokland (Aarhus University
Hospital), Ulrik Baandrup (Aalborg University Hospital), Jacob Rosen-
berg (Herlev Hospital), Zoltan Szallaszi (Technical University of Den-
mark), Ildikó Verczeczy (National Institute of Oncology, Budapest, Hun-
gary), Imre Pete (National Institute of Oncology, Department of
Gynecology, Budapest, Hungary), and Zsolt Baranyai (1st Department of
Surgery, Semmelweis University, Budapest, Hungary) for providing clini-
cal samples. We thank BGI Europe for sequencing for the samples.

L.P.N., A.J.H., S.M., T.M., L.V., and S.B. conceived the concept and
designed the experiments. S.M., L.V., S.R.R., and H.F. performed the
experiments. T.M., J.F.-N., J.M.G.I., and T.A.H. conducted the computa-
analyzed the computational data. S.M. prepared the manuscript. All au-

tors provided critical revisions to and approved the manuscript.

We declare no conflicts of interest.

The work was supported by the Innovation Fund Denmark (The
GenomeDenmark platform, grant no. 019–2011–2), the Danish National
Research Foundation (grant no. DNRF94), and the Lundbeck Foundation.

The funders had no role in study design, data collection and inter-
pretation, or the decision to submit the work for publication.

REFERENCES
1. Brook I, Frazier EH. 1991. Infections caused by Propionibacterium spe-
819.
2. Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, NISC
Comparative Sequencing Program, Bouffard GG, Blakesley RW, Mur-
ray PR, Green ED, Turner ML, Segre JA. 2009. Topographical and
temporal diversity of the human skin microbiome. Science 324:1190–
3. Oh J, Byrd AL, Deming C, Conlan S, NISC Comparative Sequencing
Program, Kong HH, Segre JA. 2014. Biogeography and individuality
shape function in the human skin metagenome. Nature 514:59–64. http:
//dx.doi.org/10.1038/nature13786.
D, Erfe MC, Loncaric A, Kim J, Modlin RL, Miller JF, Sodergren E,
Craft N, Weinstock GM, Li H. 2013. Propionibacterium acnes strain
populations in the human skin microbiome associated with acne. J Inves-
6. Achermann Y, Goldstein EJC, Coenye T, Shirliff ME. 2014. Propi-
onibacterium acnes: from commensal to opportunistic biofilm–associated


33. McDowell A, Gao A, Barnard E, Fink C, Murray PI, Dowson CG, Nagy I, Lambert PA, Patrick S. 2011. A novel multilocus sequence typing scheme for the opportunistic pathogen Propionibacterium acne and...


