



Article

Genome-Wide Pathway Exploration of the *Epidermidibacterium keratini* EPI-7^T

Yunseok Oh ^{1,2,†} , Seyoung Mun ^{2,3,†} , Young-Bong Choi ⁴, HyungWoo Jo ^{2,5}, Dong-Geol Lee ^{2,5,*} and Kyudong Han ^{1,2,3,6,*}

¹ Department of Bioconvergence Engineering, Dankook University, Jukjeon, Yongin 16890, Republic of Korea; 72210364@dankook.ac.kr

² Department of Microbiology, College of Science & Technology, Dankook University, Cheonan 31116, Republic of Korea; 12200281@dankook.ac.kr (S.M.); chohw@cosmax.com (H.J.)

³ Center for Bio Medical Engineering Core Facility, Dankook University, Cheonan 31116, Republic of Korea

⁴ Department of Chemistry, College of Science & Technology, Dankook University, Cheonan 31116, Republic of Korea; chem0404@dankook.ac.kr

⁵ R&I Center, COSMAX BTI, Pangyo-ro 255, Bundang-gu, Seongnam 13486, Republic of Korea

⁶ R&D Center, HuNBiome Co., Ltd., Gasan Digital 1-ro, Geumcheon-gu, Seoul 08507, Republic of Korea

* Correspondence: leedg@cosmax.com (D.-G.L.); jim97@dankook.ac.kr (K.H.); Tel.: +82-41-550-1240 (K.H.)

† These authors contributed equally to this work.

Abstract: Functional cosmetics industries using skin microbiome screening and beneficial materials isolated from key microorganisms are receiving increasing attention. Since *Epidermidibacterium keratini* EPI-7^T was first discovered in human skin, previous studies have confirmed that it can produce a new pyrimidine compound, 1,1'-biuracil, having anti-aging effects on human skin. Therefore, we conducted genomic analyses to judge the use value of *E. keratini* EPI-7^T and provide up-to-date information. Whole-genome sequencing analysis of *E. keratini* EPI-7^T was performed to generate new complete genome and annotation information. *E. keratini* EPI-7^T genome was subjected to comparative genomic analysis with a group of closely-related strains and skin flora strains through bioinformatic analysis. Furthermore, based on annotation information, we explored metabolic pathways for valuable substances that can be used in functional cosmetics. In this study, the whole-genome sequencing (WGS) and annotation results of *E. keratini* EPI-7^T were improved, and through comparative analysis, it was confirmed that the *E. keratini* EPI-7^T has more metabolite-related genes than comparison strains. In addition, we annotated the vital genes for biosynthesis of 20 amino acids, orotic acid, riboflavin (B2) and chorismate. In particular, we were able to prospect that orotic acid could accumulate inside *E. keratini* EPI-7^T under uracil-enriched conditions. Therefore, through a genomics approach, this study aims to provide genetic information for the hidden potential of *E. keratini* EPI-7^T and the strain development and biotechnology utilization to be conducted in further studies.

Keywords: *Epidermidibacterium keratini* EPI-7^T; skin anti-aging; de novo whole-genome sequencing; comparative genomic analysis; anti-aging; biosynthetic gene; 20 amino acids; orotic acid; functional cosmetic industry



Citation: Oh, Y.; Mun, S.; Choi, Y.-B.; Jo, H.; Lee, D.-G.; Han, K. Genome-Wide Pathway Exploration of the *Epidermidibacterium keratini* EPI-7^T. *Microorganisms* **2023**, *11*, 870. <https://doi.org/10.3390/microorganisms11040870>

Academic Editor: Matthias Noll

Received: 25 October 2022

Revised: 27 March 2023

Accepted: 27 March 2023

Published: 28 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The *Epidermidibacterium keratini* EPI-7^T was first introduced by Lee DG et al. in 2018. In a previous study, a novel bacterial strain, EPI-7^T, was isolated from skin samples (human epidermal keratinocytes) and characterized by taxonomic identification [1]. In addition, it has been proven that *E. keratini* EPI-7^T culture solutions have a positive effect on the improvement of anti-aging-related gene expression in Hs68 cells to UV-irradiated fibroblast cells [2]. Moreover, a recent study explored anti-aging compounds produced by *E. keratini* EPI-7^T (orotic acid, 1,1'-biuracil) depending on the precursor added to the culture medium [3]. Through previous studies, *E. keratini* EPI-7^T is regarded as a hidden

beneficial microbe with potential applicabilities in skin anti-aging, skincare, cosmetics, the human beauty industry, and the skin microbiome [1–3]. However, fundamental genomics analysis and genomic resource research based on the *E. keratini* EPI-7^T genome has not progressed much since 2021. The lack of genomic resource research makes it difficult to understand the organism's potential molecular and biological processes. Furthermore, the closely-related microorganisms are still not defined, and thus the study of the microbial genome and classification of functional genes is important in ensuring bacterial secondary metabolites and socioeconomic benefits that are not yet known in the *E. keratini* EPI-7^T.

With the advancement of next-generation sequencing (NGS) technology and the reduction in sequencing costs, many microbial genome data have been published and used for further examination, such as genetic screening, functional characterization, metabolism pathway prediction, and resource search [4,5]. NGS technology has lowered barriers to entry for the genetic approach at the genome level and allowed deeper evolutionary analysis. In this study, bacterial whole-genome sequencing (WGS) has been performed using Illumina NovaSeq 6000 platform to generate paired-end (PE) libraries (short-read). In addition, long-read sequencing data generated by the SMRTbell library and PacBio RSII platform is also often used in the construction of a bacterial genome [6,7]. These PE and long-read sequencing data have been involved in hybrid de novo assembly, depending on the research purposes. Here, we also choose the de novo whole-genome assembly strategy for the construction of the precise *E. keratini* EPI-7^T genome.

Recently, broadly-used and systematic web-based tools that can perform a comparative analysis of strains along with increasing WGS data of bacteria are continuously developed [8,9]. PATRIC enables an accurate bioinformatic analysis using publicly and quickly available microbial genome sequences by supporting eight services, such as web-based sophisticated comparative analysis and comprehensive genome analysis [10,11]. In addition, OrthoVenn2 efficiently performs functional and evolutionary analysis of proteins across multiple species, comparative genome studies, and visualization of taxonomic evidence, through genome-wide comparative analysis of orthologous clusters with free access to a user-friendly web server [12]. EggNOG-mapper also provides several functional annotation sources, such as the KEGG pathway, Gene Ontology labels, EC numbers, and COG functional categories [13]. In addition, the large metagenomic data sets in EggNOG software make possible faster annotation than Prokka, which is considered to be the fastest annotation tool for annotating prokaryotes [14]. Therefore, we performed the pan-genome analysis pipeline (PGAP) using the seven most closely-related species with *E. keratini* EPI-7^T (*Sporichthya polymorpha*, *Modestobacter marinus*, *Modestobacter roseus*, *Modestobacter versicolor*, *Blastococcus aggregatus*, *Geodermatophilus obscurus*, and *Frankia alni*) and two species belonging to the class *acिनomycetia* (*Streptomyces klenkii* and *Streptomyces griseocarneus*) to identify the genetic features that are fundamentally essential and specific to *E. keratini* EPI-7^T. To establish the dermatophilic status of EPI-7 in the microbiome, we initially investigated the microbial taxonomy between *E. keratini* EPI-7^T and 20 species of the prominent skin microbiota in human using the UBCG (up-to-date bacterial core gene) pipeline and thus selected 7 major species [15–17]. Additionally, PATRIC, OrthoVenn2, and EggNOG-mapper were performed with *E. keratini* EPI-7^T and seven major species (*Cutibacterium acnes*, *Staphylococcus epidermidis*, *Staphylococcus aureus*, *Staphylococcus warneri*, *Micrococcus luteus*, *Streptococcus pyogenes*, and *Streptococcus mitis*) known as human skin flora [18,19]. We confirmed the genetic ability of skin-derived *E. keratini* EPI-7^T to have similarities with the skin flora species and its unique genetic abilities.

Bacterial cellulose, known to have biocompatible, hydrophilic, and biodegradable biopolymer properties, is used as an emulsion stabilizer and skin treatment in cosmetics [20]. In addition, biomolecules such as secondary metabolites of bacterial origin, enzymes, pigments, and exopolysaccharides are attracting attention for commercial and cosmetic applications [21,22]. Particularly, it was confirmed that the gene cluster involved in the orotic acid biosynthesis pathway mentioned in the previous study was present in the *E. keratini* EPI-7^T chromosomal loci. Furthermore, the genes involved in the biosynthesis

pathway of 20 amino acids and metabolites that can affect the skin while inhabiting the skin were identified.

Hence, we perform a WGS and comparative genomic analysis of *E. keratini* EPI-7^T. This study predicts the evolutionary relationship of *E. keratini* EPI-7^T and the effect of substrate added to the culture medium on the pathway and compound accumulations. By providing a more accurate and up-to-date complete genome following its basic information, we provide bioinformatical clues to further study *E. keratini* EPI-7^T and encourage sharper demonstrations and industrial applications on the development of biomaterials and cosmetic compounds.

2. Materials and Methods

2.1. Bacterial Materials and DNA Extraction

E. keratini EPI-7^T used for the WGS and hybrid de novo assembly was provided by COSMAX BTI R&I Center. *E. keratini* EPI-7^T was cultured on an R2A agar medium at optimal conditions (R2A agar at 25 °C, pH 6.0, absence of NaCl) for 3–5 days. Then, a single colony was inoculated in an R2A broth medium and cultured for 3–5 days under optimal conditions. After five days of incubation, stock and DNA extraction were executed. The stock was made so that the final glycerol was 25% and stored at –80 °C. This genomic DNA of *E. keratini* EPI-7^T from cultured in an R2A broth medium was extracted using the Qiagen DNeasy UltraClean Microbial Kit (Qiagen, Hilden, Germany), and all experimental processes were performed following the optimal protocols provided by the DNA extraction kit. The quality check of all extracted DNA was conducted using the NanoDrop One (ThermoFisher Scientific, Waltham, MA, USA) equipment.

2.2. Whole-Genome Sequencing and Hybrid De Novo Assembly

E. keratini EPI-7^T genomic DNA was sequenced at the Teragen Bio (Theragen Bio, Seongnam, Republic of Korea) by Illumina NovaSeq 6000 (Illumina Inc., San Diego, CA, USA). Illumina PE libraries were prepared using the TruSeq Nano DNA Prep kit (Illumina, San Diego, CA, USA) in terms of the included instructions. Subsequently, we prepared the PacBio SMRTbell raw sequence data downloaded from COSMAX R&I Center, Seongnam, Republic of Korea.

With the Illumina PE libraries and PacBio SMRTbell raw sequence data, the *E. keratini* EPI-7^T genome was assembled with the de novo hybrid strategy. First, the Canu v1.7 generated long-read assembly contigs from PacBio SMRTbell raw sequence data with default option [23]. Next, to correct and improve accuracy, Illumina PE reads were integrated onto the draft genome using BWA mem v.0.7.17 (Wellcome Trust Sanger Institute, Cambridge, UK) [24]. Then, a polishing round was performed five times using pilon v.1.22 (Broad Institute of MIT and Harvard, Cambridge, MA, USA) [25].

2.3. Genomic Gene Prediction and Genome Annotation

The newly-assembled genome was used for gene prediction and annotation using Rapid Prokaryotic Genome Annotation (Prokka) ver1.14.6 with the RNAmmer and Adgenes [26,27]. Structural annotation was conducted by using Prodigal, Aragorn, Barnap, and MinCED were used to predict coding sequence (CDS), tRNA, rRNA, and CRISPRs assay, respectively [28–31]. For functional annotation, genes were searched against the UniProt/Swissprot and NCBI non-redundant proteins RefSeq databases using BLASTP v2.2.29+ with an E-cutoff value of 1.0×10^{-10} [32–34]. Then, protein domains were also searched against the Pfam databases using InterProScan v.5.19-58.0 [35,36].

2.4. Improvement of Whole-Genome Sequencing and Gene Annotation for *E. keratini* EPI-7^T

For comparison between the newly-assembled genome and the previous genome in NCBI, the direction of the two genomes was aligned based on the newly-assembled genome using Chromeister (<https://usegalaxy.org/> (accessed on 12 September 2021)), PATRIC 3.6.12 (University of Málaga, Málaga, Spain) with genome alignment (Mauve)

services [11,37]. Moreover, the difference between the two assembled genomes was compared using BioEdit 7.2 (North Carolina State University, Raleigh, NC, USA) [38]. In addition, gene annotation was improved by excluding hypothetical protein and domain of the unknown function (DUF) through manual inspection between the newly-annotated data and the previous one registered in NCBI.

2.5. Analysis of Pan-Genome and Comparative Whole-Genome Analysis

With the hybrid de novo assembled genome for pan-genome and comparative analysis, the comparable bacterial genome data of nine comparison strains were downloaded from the NCBI genome datasets (<https://www.ncbi.nlm.nih.gov/datasets/> (accessed on 14 October 2021)) (Table S1). Furthermore, pan-genome analysis was conducted with a group of the given strains closely related to *E. keratini* EPI-7^T. The pan-genome analysis pipeline (PGAP) 1.2.1 was used for pan-genome analysis [39]. PGAP has five analysis modules: (i) cluster analysis of functional genes, (ii) pan-genome profile analysis, (iii) genetic variation analysis, (iv) species evolution analysis, and (v) function enrichment analysis. The multi-paranoid (MP) method was used for cluster analysis of functional genes and showed the orthologous genes presented in the comparison strains. Moreover, the phylogenetic analysis based on the core genome was performed in the MEGA 11 program with 1000 bootstrap replications for checking an evolutionary distance [40]. Phylogenetic analysis was conducted using the neighbor-joining method with the maximum composite likelihood model.

Genome-based phylogeny analysis between *E. keratini* EPI-7^T and 20 human skin microbiota (*Cutibacterium granulosum*, *Cutibacterium avidum*, *Cutibacterium acnes*, *Staphylococcus aureus*, *Staphylococcus hominis*, *Staphylococcus haemolyticus*, *Staphylococcus lugdunensis*, *Staphylococcus warneri*, *Staphylococcus epidermidis*, *Micrococcus luteus*, *Streptococcus pyogenes*, *Streptococcus agalactiae*, *Streptococcus mitis*, *Corynebacterium jeikeium*, *Corynebacterium aurimucosum*, *Pseudomonas aeruginosa*, *Acinetobacter pittii*, *Acinetobacter baumannii*, *Roseomonas mucosa*, *Neisseria meningitidis*) was performed using UBCG pipeline v3.0 (<https://www.ezbiocloud.net/tools/ubcg>, (accessed on 15 October 2021)), which is based on 92 core genes that exist as a single copy [17,41,42] (Table S1). A phylogenetic tree was visualized using the MEGA 11 program (Tokyo Metropolitan University, Tokyo, Japan) [40]. Comparative whole-genome analysis was also performed with a group of seven major species (*Cutibacterium acnes*, *Staphylococcus epidermidis*, *Staphylococcus aureus*, *Staphylococcus warneri*, *Micrococcus luteus*, *Streptococcus pyogenes*, and *Streptococcus mitis*) known as skin flora (Table S1). Web-based genome comparison tools, such as PATRIC (<https://www.patricbrc.org> (accessed on 15 October 2021)), OrthoVenn2 (<https://orthovenn2.bioinfotoolkits.net> (accessed on 15 October 2021)) and EggNOG-mapper (<http://eggnog-mapper.embl.de> (accessed on 14 March 2022)) were used [11–13]. PATRIC confirmed the amino acid (AA) identity between *E. keratini* EPI-7^T and the comparison strains using proteome analysis services with the default setting. OrthoVenn2 also confirmed similarity and common gene clusters by comparing orthologous genes using the AA sequence between *E. keratini* EPI-7^T and the comparison strain; E-value: 1.0×10^{-2} , inflation value: 1.5. The EggNOG-mapper was used for functional annotation and orthology assignments by submitting the amino acid sequence data; minimum hit e-value: 0.001, minimum hit bit-score: 60, percentage identity: 40, minimum % of query coverage: 20, minimum % of subject coverage: 20.

2.6. Prediction of the Particular Biosynthesis Pathway in *E. keratini* EPI-7^T

Functional analysis was performed through manual inspection based on the annotation data in this study. Based on Table S2, the annotation-based manual inspection resulted in identification of biosynthetic pathways of amino acids or compounds in bacteria and necessary gene discovery concerning MetaCyc (MetaCyc.org), KEGG pathway (<http://www.genome.jp/kegg/> (accessed on 11 October 2021)), and other studies [21,43–45].

3. Results

3.1. Hybrid De Novo Assembly and Genomic Characteristics of *E. keratini* EPI-7^T

The overall experiment was performed in a serial process (Figure 1). Before the NGS data production, the result of the 16S rRNA gene full-length Sanger sequencing of prepared bacterial gDNA matched 99.50% identity to the *Epidermidibacterium keratini* strain EPI-7^T chromosome (Table S3). An Illumina PE libraries generated 14,057,360 reads covering 2,122,661,360 bp (2.12 Gb) of sequence data, the Q30 (more base rate, nucleotide accuracy >99.9% rate) rate was 92.13%, and the GC content was 67.70%. PacBio SMRTbell sequence data comprised 75,720 reads covering 680,622,771 bp (0.68 Gb), and mean length was 8989 bp, N50 was 12,811, and average GC content was 65.17% (Table S4). As a result, a single chromosome (a contig) of *E. keratini* EPI-7^T was obtained through hybrid de novo assembly and polishing using Canu, BWA mem, and Pilon. The total *E. keratini* EPI-7^T genome length was 4,018,778 bp, and the average GC content was 67.34% (67.34%) (Table 1).

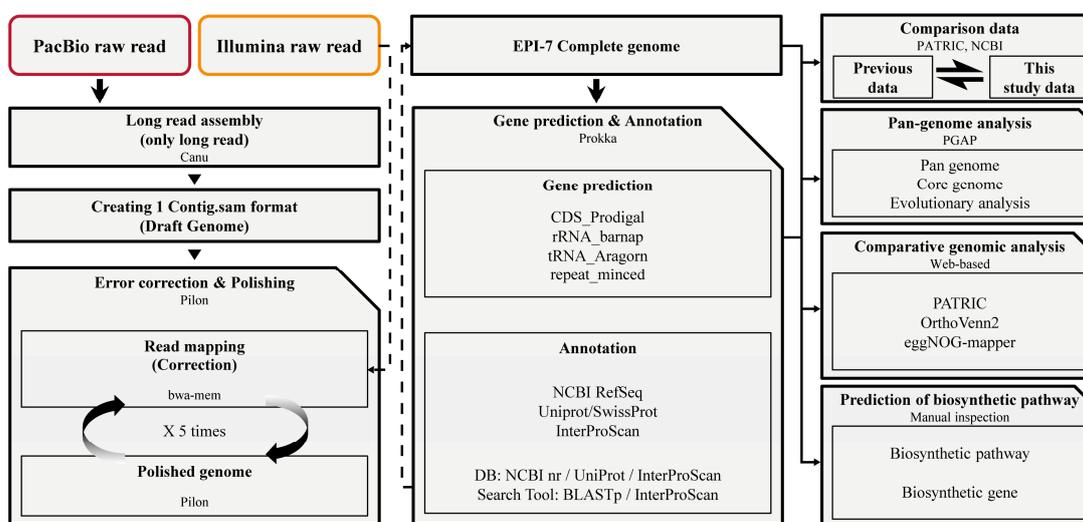


Figure 1. A comprehensive analysis was performed at each stage to understand *Epidermidibacterium keratini* EPI-7^T. Step 1. Construction of new *E. keratini* EPI-7^T genome using PacBio raw read and Illumina raw read through hybrid de novo assembly strategy. Step 2. Gene prediction and annotation using newly-assembled *E. keratini* EPI-7^T genome. Step 3. Comparison between this study data and previous data. Step 4. Pan-genome analysis with strains closely related to *E. keratini* EPI-7^T and comparative genomic analysis with skin flora. Step 5. The discovery of functional genes and biosynthesis genes that can affect skin health.

3.2. Genomic Gene Prediction and Annotation

The *E. keratini* EPI-7^T genome had 3851 predicted genes (3799 CDSs, 3 rRNAs, 48 tRNAs, and 1 tmRNA) and 1 CRISPR array (Table S5). Of the 3851 predicted genes of *E. keratini* EPI-7^T, except for 13 CDSs, 51 RNA genes, and 1 CRISPR array, 3786 CDSs had annotated in the UniProt, NCBI nr, InterProScan (respectively, 2754, 3779, 3509, accounting for 72.49%, 99.47%, 92.37% of the total number of annotated genes) (Table S6). Among the 3786 annotated CDSs, 3169 genes (83.70%) were annotated functional genes, and 617 (113 + 504) genes (16.30%) were unknown (DUF) or hypothetical genes.

3.3. Comparison between Newly-Generated and Previous Genome and Annotation

The size of the genome in the NCBI used for comparison with the newly-assembled genome was 4,018,808 bp, which is 30 bp longer than the newly-assembled genome. Moreover, we identified significant differences in the comparison results between the two genomes, which were aligned with the sequence orientation using the genome alignment service (Mauve) in PATRIC. First, in the case of sequence orientation, when the newly-assembled genome was referenced, it was found that the sequences of strands with different

senses were represented (Figure S1). Second, the alignment results of the two genomes show that the region of 1 bp to 2,090,211 bp of the newly-assembled genome complementarily matched 2,090,242 bp to 1 bp of the previous genome, and the remaining 2,090,212 bp to 4,018,778 bp region complementarily matches 4,018,808 bp to 2,090,243 bp (Figure 2). As a result, the difference between the two genome sequences is about 70 bp, which shows that the newly-assembled genome is 99.998258% similar. Five single nucleotide gaps and one small gap were filled. The newly-assembled genome corrected two single nucleotide gaps and one duplication error. We also found four SNV mismatches and corrected them in this study (Table S7).

Table 1. Characteristics of *E. keratini* EPI-7^T genome.

Assembly Characteristics	Value
Number of contigs	1
Total length (bp)	4,018,778
N50 (bp)	4,018,778
N90 (bp)	4,018,778
G + C content (%)	67.34%
Genome Characteristics	Value
Genome size (bp)	4,018,778
DNA coding (bp)	3,719,702
Number of total gene	3851
Number of coding sequence gene	3799
Number of RNA genes	52
Number of CRISPR array	1
Total length of gene (bp)	3,719,702
Total length of coding sequence gene (bp)	3,707,562
RNA gene total length (bp)	8814
CRISPR total length (bp)	3326
Gene/Genome (%)	92.558

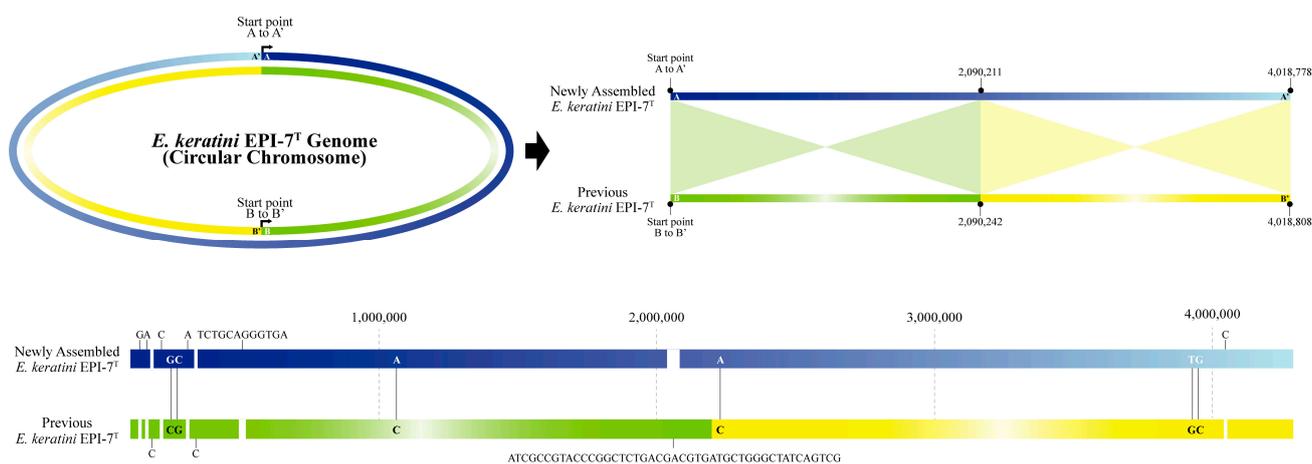


Figure 2. Representation of the newly-assembled and previous genome. A comparison between the newly-assembled genome and the previous genome was performed using Chromeister and the genome alignment (Mauve) services of PATRIC. In the case of the base sequence direction, when the newly-assembled genome of *E. keratini* EPI-7^T is expressed as a dark blue line to a light blue line (A to A'), the genome sequence of the previous *E. keratini* EPI-7^T is expressed as a green line to a yellow line (B to B'). It was found that the genome sequences were expressed in different orientation.

Using the corrected *E. keratini* EPI-7^T genome, we were able to identify a total of 3851 genes, compared to the previous genome data. Among a total of annotated genes, 3799 CDSs were implicated in the gene annotation process. As a result, the number of

annotated genes increased from the previous 3753 (98.79%) to 3786 (99.66%). Through an accurate gene annotation and the better-assembled genome data, we finally mined 3169 functional genes, a higher number than the 3012 previous annotated genes. The annotated data produced in this study showed more information on 157 functional genes. The number of unknown (DUF) and hypothetical genes were 113 and 504, respectively. A total of 3 rRNA, 48 tRNA, 1 tmRNA, and 1 CRISPR cluster were searched out using Aragorn, Barrnap, and MinCED software (Table 2) [28–30].

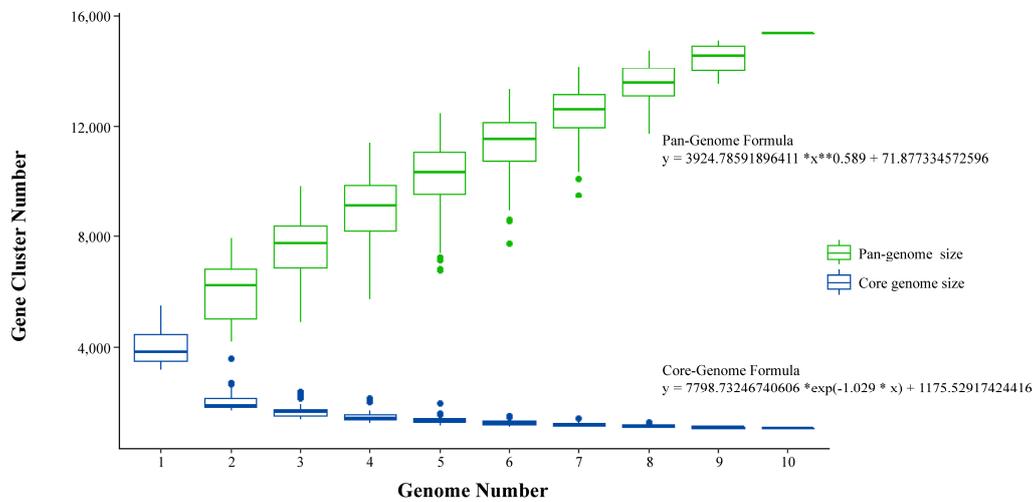
Table 2. Comparison of annotations between newly-generated and previous *E. keratini* EPI-7^T.

Newly-Assembled <i>E. keratini</i> EPI-7 ^T		Previous <i>E. keratini</i> EPI-7 ^T	
Total Genes	3851	Total Genes	3851
CDS	3799	CDS	3799
rRNA	3	rRNA	3
tRNA	48	tRNA	47
tmRNA	1	tmRNA	3
Repeat Region	0	Repeat Region	0
CRISPR Array	1	CRISPR Array	1
CDSs	3799	CDSs	3799
Annotated Genes	3786	Annotated Genes	3753
Unannotated Genes	13	Unannotated Genes	46
Annotated Genes (%)	99.66%	Annotated Genes (%)	98.79%
Annotated Genes	3786	Annotated Genes	3753
Functional Genes	3169	Functional Genes	3012
Unknown (DUF)	113	Unknown (DUF)	178
Hypothetical Genes	504	Hypothetical Genes	563

3.4. Core and Pan-Genome Analysis with Closely-Related Species Group

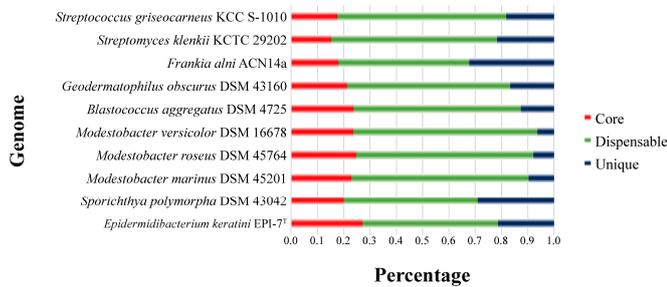
Pan-genome analysis showed that the overall orthologous gene cluster of *E. keratini* EPI-7^T and the 9 comparison strains had 15,363 pan-genomes. A comparative pan-genome analysis was conducted to study genomic diversity and evolutionary relationships, and the comparison strains were chosen by considering taxonomy relations based on the 16s RNA gene. The pan-genome and core genome size, the number of accumulated genes about the number of genomes, could be predicted by Heaps' law ($y = A \cdot x^{**} B + C$, where x is the number of genomes, y is the size of the pan-genome, $*$ is the multiplication sign (\times), $**$ is the square sign (\wedge), and A , B , and C fit parameters) [46,47]. Therefore, the pan-genome size steadily increased as the comparison strains genomes were added, and it was confirmed that the pan-genome formula ($y = 3924.78591896411 \cdot x^{**} 0.589 + 71.877334572596$) was $0 < B < 1$ (Figure 3A). Correspondingly, the orthologous gene cluster shared by *E. keratini* EPI-7^T and all comparison strains had 1032 core genomes. Then, the core genome size was maintained after a drastic decrease as the comparison strains genomes were added, and the core genome formula was $y = 7798.73246740606 \cdot \exp(-1.029 \cdot x) + 1175.52917424416$ (Figure 3A). However, the pan-genome of 15,363 orthologous gene clusters was divided into the core genome of 1032 orthologous gene clusters (6.72%) present in all 10 strains, the accessory genome of 5236 orthologous gene clusters (34.08%) partially present in 10 strains, and 9095 strain-specific orthologous gene clusters (59.20%) (Figure 3B). Furthermore, these results not only show that very few orthologous gene clusters were shared between *E. keratini* EPI-7^T and the comparison strains, but also suggest that *E. keratini* EPI-7^T has very different genetic repertoires, such as metabolic ability, production of secondary metabolites, and host cell communications. A phylogenetic tree was constructed based on the core genome to confirm further the evolutionary relationship between *E. keratini* EPI-7^T and the comparison strain. The branch of *E. keratini* EPI-7^T shows a significant evolutionary distance from other comparative strains (Figure 3C). Thus, the pan-genome analysis indicates that *E. keratini* EPI-7^T branched independently from other strains in the phylogenetic tree

constructed based on the core genome and that *E. keratini* EPI-7^T had low genetic similarity with the other strains.

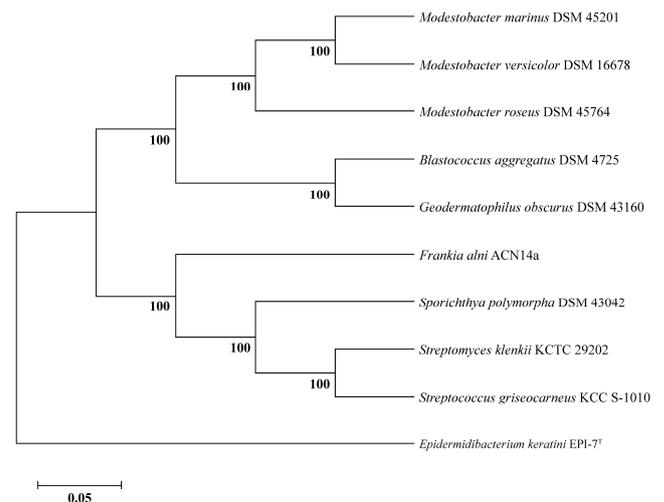


(a)

Strain	Total genes	Dispensable	Unique	Core
<i>Epidermidibacterium keratini</i> EPI-7 ^T	3799	1958 (51.5%)	809 (21.3%)	1032 (27.2%)
<i>Sporichthya polymorpha</i> DSM 43042	5148	2622 (50.9%)	1494 (29.0%)	1032 (20.0%)
<i>Modestobacter marinus</i> DSM 45201	4506	3038 (67.4%)	436 (9.7%)	1032 (22.9%)
<i>Modestobacter roseus</i> DSM 45764	4178	2815 (67.4%)	331 (7.9%)	1032 (24.7%)
<i>Modestobacter versicolor</i> DSM 16678	4358	3053 (70.1%)	273 (6.3%)	1032 (23.7%)
<i>Blastococcus aggregatus</i> DSM 4725	4346	2764 (63.6%)	550 (12.7%)	1032 (23.7%)
<i>Geodermatophilus obscurus</i> DSM 43160	4875	3033 (62.2%)	810 (16.6%)	1032 (21.2%)
<i>Frankia alni</i> ACN14a	5743	2859 (49.8%)	1852 (32.2%)	1032 (18.0%)
<i>Streptomyces klenkii</i> KCTC 29202	6779	4273 (63.0%)	1474 (21.7%)	1032 (15.2%)
<i>Streptomyces griseocarneus</i> KCC S-1010	5868	3770 (64.2%)	1066 (18.2%)	1032 (17.6%)



(b)



(c)

Figure 3. Analysis of pan-genome with species closely related to *E. keratini* EPI-7^T. Pan-genome analysis with strains closely related to *E. keratini* EPI-7^T indicated that *E. keratini* EPI-7^T had low genetic similarity to the comparison strains, which meant that *E. keratini* EPI-7^T was also genetically different from the most closely-related strains. (a) Pan-genome size increased, and core genome decreased as the comparison strains genomes were added. (b) The pan-genome of 15,363 orthologous gene clusters was divided into the core-genome of 1032 orthologous gene clusters (6.72%) present in all 10 strains, the accessory genome of 5236 orthologous gene clusters (34.08%) partially present in 10 strains, and 9095 strain-specific orthologous gene clusters (59.20%). (c) The phylogenetic analysis based on core-genome was performed with 1000 bootstrap replications for checking an evolutionary distance.

3.5. Comparative Genomic Analysis with Skin Flora Group

To identify suitable comparison candidates, we conducted a taxonomic analysis of the *E. keratini* EPI-7^T strain and 20 species of the human skin microbiota using the UBCG pipeline. *Micrococcus luteus* has the closest genetic correlation to *E. keratini* EPI-7^T (Figure 4). Based on these results, an in-depth comparative genetic characterization and specificity analysis was performed between the seven major skin microbiota and *E. keratini* EPI-7^T. The

integrated and comparative analysis of genomic and gene associations was investigated using web-based PATRIC, OrthoVenn2, and EggNOG-mapper. Proteome comparison analysis based on amino acid sequence using PATRIC showed that *E. keratini* EPI-7^T had a low protein sequence identity with the skin flora group (Figure 5). The hit ratio to the amino acid sequence of *E. keratini* EPI-7^T was less than 50%, except for *Micrococcus luteus* (50.73%). In addition, most amino acid identity (AAI) with *E. keratini* EPI-7^T was less than 50% in all strains (1110 to 1646). However, *Cutibacterium acnes* (244) and *Micrococcus luteus* (298) had a relatively high ratio of AAI that was more than 90% and less than 50% with *E. keratini* EPI-7^T (Table S8).

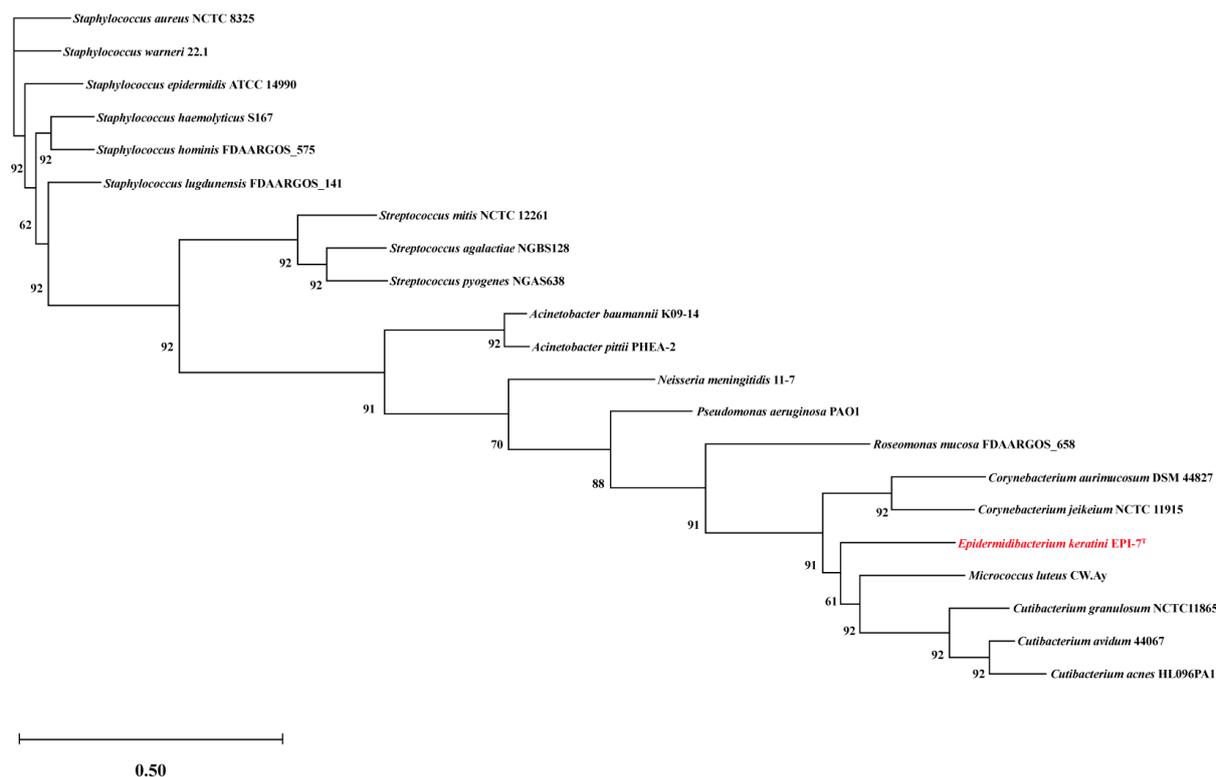


Figure 4. Genome-based phylogeny analysis with *E. keratini* EPI-7^T and the 20 prominent skin microbiota in humans using UBCG. Evolutionary distance based on the 92 core genes selected by UBCG shows the relevant relationship between *E. keratini* EPI-7^T and 20 human skin microbiotas. Bootstrap analysis was performed at 100 replications. Bootstrap values are given at branching points. Bar, 0.50 substitution per position. It can be confirmed that the *E. keratini* EPI-7^T (in red) is the most relevant relationship with *Micrococcus luteus*. On the contrary, the most independent branch is formed at *Staphylococcus* and *Streptococcus* genera.

OrthoVenn2 was used to confirm orthologous gene clusters and the pairwise heatmap based on the amino acid sequence. The pairwise heatmap showed a correlation based on the number of overlapping clusters of *E. keratini* EPI-7^T and the skin flora group. As a result, *C. acnes* and *M. luteus* showed the highest associations with *E. keratini* EPI-7^T in amino acid identity. This result corresponds with PATRIC analysis. Although the correlation with other strains was low, the correlation between strains belonging to the same genus was close to each other (Figure 6A). Then, in GO enrichment, GO categories and GO terms of the clusters were shared by the seven comparison strains. The 311 orthologous gene clusters across all strains are involved in 266 biological process (BP), 41 molecular function (MF), and 1 cellular component (CC) GO term. Of those GO terms, “Translation” (GO:0006412, 46 clusters); “SOS response” (GO:0009432, 8 clusters); “‘de novo’ IMP biosynthetic process” (GO:0006189, 6 clusters); “Translation elongation factor activity” (GO:0003746, 4 clusters); “Translation initiation factor activity” (GO:0003743, 3 clusters) were shown to have significant associations (Figure S2). As for the closest microorganisms, *C. acnes* and *M. luteus*, we

screened that 240 orthologous gene clusters between 3 species were associated with 162 BP, 58 MF, and 10 CC GO terms. In addition, the GO enrichment shared by the three strains were significantly associated with “Terpenoid biosynthetic process” (GO:0016114), including four genes: 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (HMBPP reductase), HMBPP synthase, 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MEcPP), and 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXP reductoisomerase) (Figure 6B). The unique gene clusters in *E. keratini* EPI-7^T are implicated in the function annotation and prediction. A total of 143 unique clusters are related to 61 BP, 26 MF, and 3 CC GO terms. Those GO terms were highly enriched into “Aromatic compound catabolic process” (GO:0019439); “Oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen” (GO:0016705) (Figure 6C). Cytochromes P450 (CYPs) and luciferase-like monooxygenases (LLM) class F420-dependent oxidoreductase were involved in BP: GO:0019439. Helix-turn-helix (HTH) domain-containing protein, 2Fe-2S iron-sulfur cluster binding domain-containing protein, and Biphenyl-2,3-diol 1,2-dioxygenase 3 were involved in MF: GO:0016705.

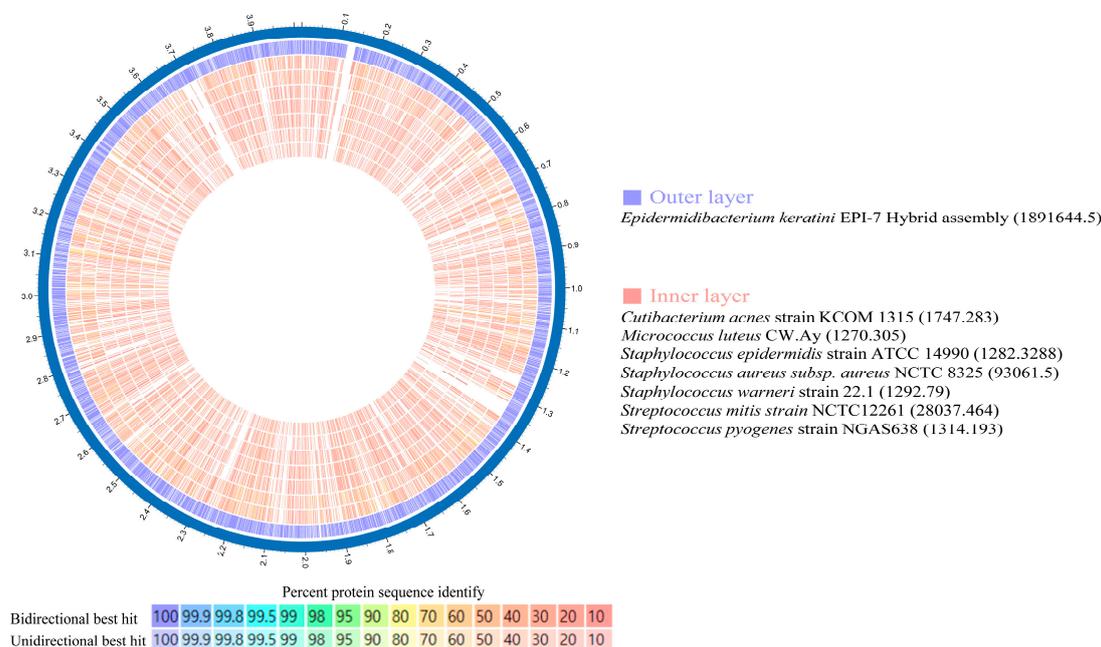
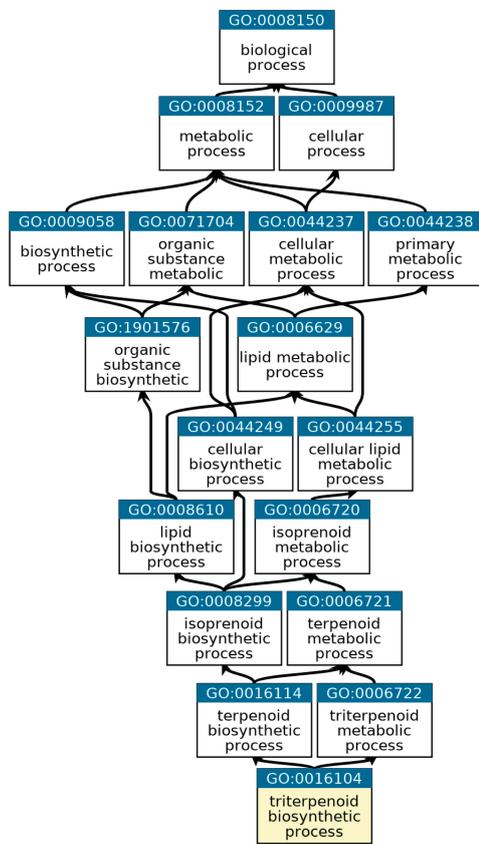
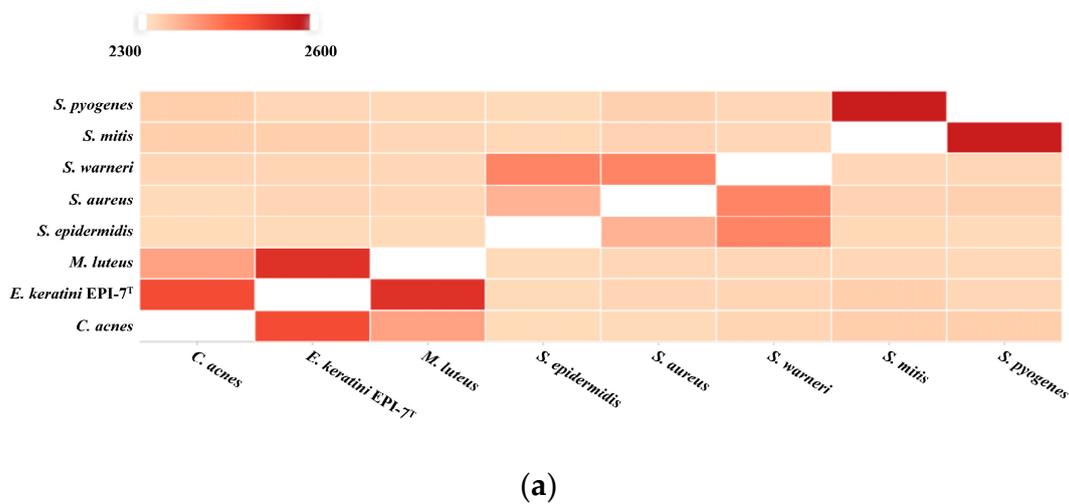
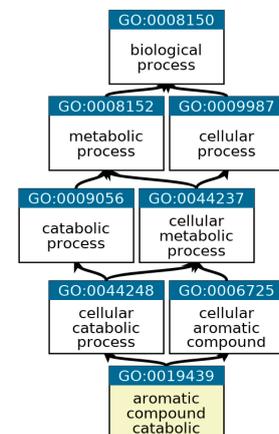
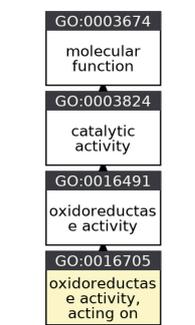


Figure 5. Amino acid identity (AAI) using Proteome comparison (PATRIC). Proteome comparison analysis based on amino acid sequence using PATRIC showed that *E. keratini* EPI-7^T had a low AAI with those of skin flora. However, *Cutibacterium acnes* (244) and *Micrococcus luteus* (298) had a relatively high AAI ratio of more than 90%, less than 50%, with *E. keratini* EPI-7^T.

We also checked that CDSs aligned to Clusters of Orthologous Groups (COGs) families comprising 18 functional categories (Table S9). Among them, except for function unknown, metabolism categories that existed significantly in *E. keratini* EPI-7^T than in other skin flora were “Energy production and conversion” (C, 237 genes); “Amino acid transport and metabolism” (E, 399 genes); “Carbohydrate transport and metabolism” (G, 210 genes); “Coenzyme transport and metabolism” (H, 147 genes); “Lipid transport and metabolism” (I, 275 genes); “Transcription” (K, 331 genes); “Replication, recombination and repair” (L, 167 genes); “Inorganic ion transport and metabolism” (P, 245 genes); “Secondary metabolites biosynthesis, transport, and catabolism” (Q, 153 genes); and “Signal transduction mechanisms” (T, 101 genes) (Figure 7).



(b)



(c)

Figure 6. Pairwise heatmap based on amino acid sequence and GO enrichment. (a) In the pairwise heatmap, each cell is the overlapping cluster number between the two genomes in contact with the horizontal and vertical axes, and the darker the color, the higher the overlapping cluster number. Pairwise heatmap based on amino acid sequence confirmed that *C. acnes* and *M. luteus* showed the highest association with *E. keratini* EPI-7^T. (b) GO enrichment based on GO terms showed that *C. acnes*, *M. luteus*, and *E. keratini* EPI-7^T shared “Terpenoid biosynthetic process”, but (c) “Aromatic compound catabolic process” and “Oxidoreductase activity” was unique to *E. keratini* EPI-7^T.

3.6. Genome-Wide Prediction of the Particular Biosynthesis Pathway in Accumulation of Orotic Acid in *E. keratini* EPI-7^T

It is well known that most microorganisms are auxotrophs, relying on external nutrients from host or syntrophic microorganisms for growth and survival [48–50]. Nevertheless, *E. keratini* EPI-7^T is regarded as one autotrophic microorganism with genes necessary to synthesize 20 amino acids. Representatively, asparagine (Asn), glutamine (Gln), and serine (Ser) are essential biofactors that play the essential role of pH control and nitrogen donor through continuous bacterial metabolisms [51]. As shown in Table S10, the biosynthetic pathway-related genes of Asn, Gln, and Ser, typically consumed the most in microorganisms, were coded in the *E. keratini* EPI-7^T genome. Two copies of genes (gene01609 and gene01999) can convert aspartate (Asp) to Asn. Furthermore, the other two copies of the gene (gene01563 and gene01580) encode an enzyme capable of converting Glutamate (Glu) to Gln. It also has genes encoding enzymes necessary for the process of converting 3-phosphate-D-glycerate to Ser. In addition, genes and gene clusters necessary for the biosynthesis of 20 amino acids from various substrates or other amino acids are preserved in *E. keratini* EPI-7^T.

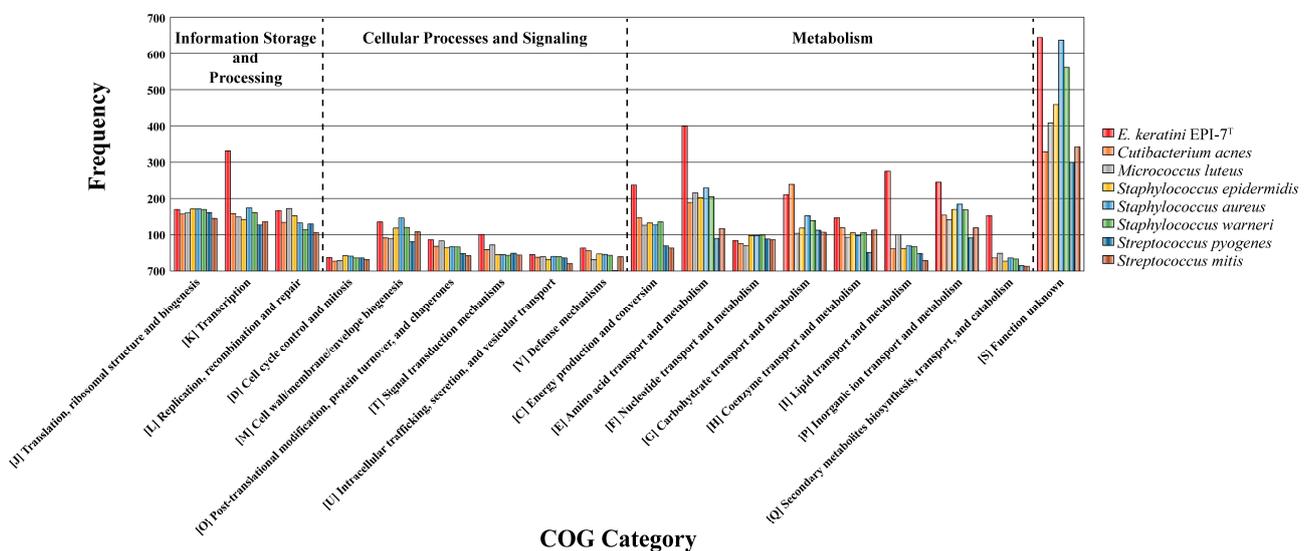


Figure 7. Bar chart plot showing the frequency of genes per COG category. EggNOG-mapper based on COG category frequency showed that *E. keratini* EPI-7^T contained significantly more genes related to the metabolism category than other strains.

In addition, in the previous study, it was confirmed that *E. keratini* EPI-7^T used uracil as a substrate to make orotic acid, an anti-aging functional substance [3]. Based on the new annotated data information through this study, we reorganized and demonstrated gene clusters associated with orotic acid biosynthesis in the *E. keratini* EPI-7^T genome; (gene01830 encoding aspartate carbamoyltransferase [EC 2.1.3.2], gene01831 encoding dihydroorotase [EC 3.5.2.3], gene01832 encoding carbamoyl-phosphate synthase small chain [EC 6.3.5.5], gene01833 encoding carbamoyl-phosphate synthase large chain [EC 6.3.5.5], gene01834 encoding dihydroorotate dehydrogenase (quinone) [EC 1.3.5.2]) (Figure 8A). The biosynthesis pathway of substances that can affect skin health (riboflavin [vitamin B2], chorismate [shikimate], anthranilate, phytoene) mentioned in the previous study and related genes were confirmed to be possessed by *E. keratini* EPI-7^T. Furthermore, they had genes involved in the production of enzymes related to skin antioxidants, such as antioxidants (trehalose), superoxide dismutase (SOD), and peroxidase (catalase, glutathione peroxidase) (Table S11).

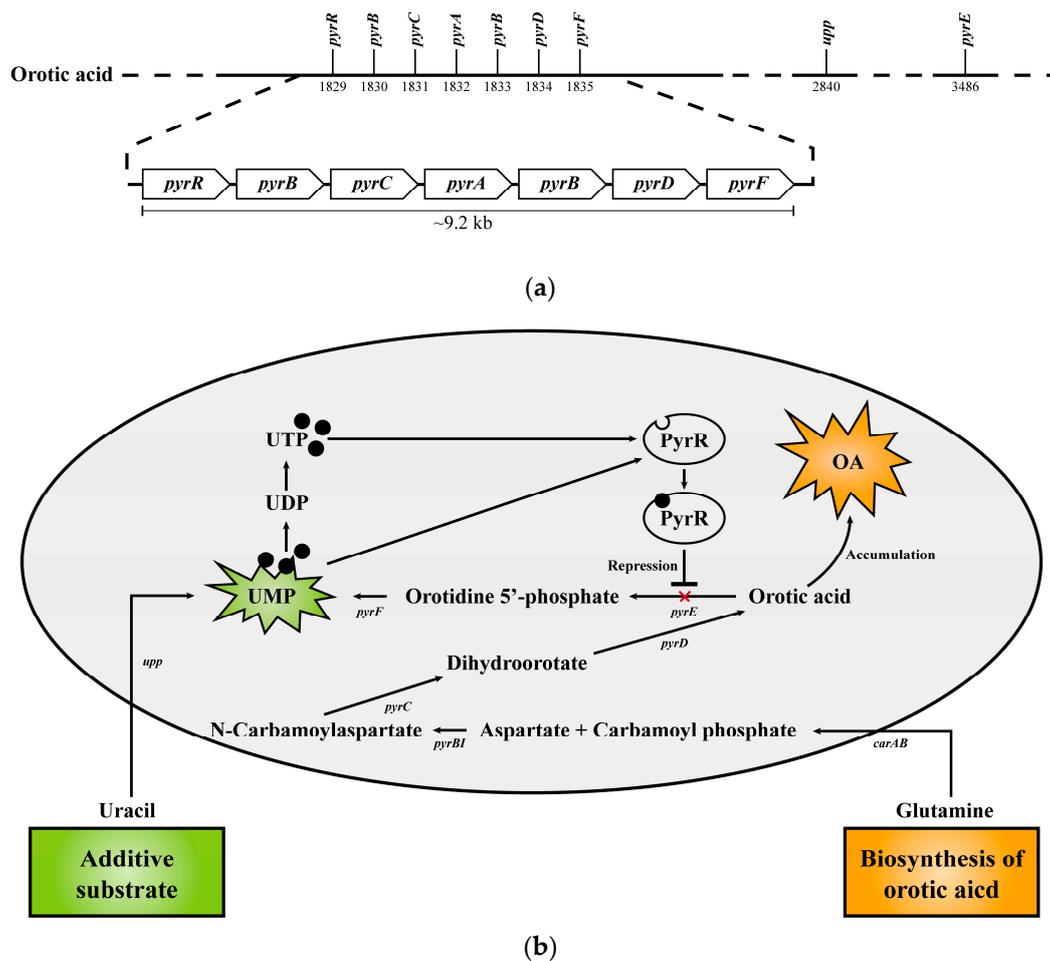


Figure 8. Biosynthetic pathway of orotic acid in *E. keratini* EPI-7^T. (a) The gene cluster for biosynthesis of orotic acid shows the distribution in the *E. keratini* EPI-7^T genome. (b) The names of genes encoding enzymes used in each step of orotic acid biosynthesis and UMP synthesis are as follows: *carAB*, carbamoyl-phosphate synthase small chain/carbamoyl-phosphate synthase large chain; *pyrBI*, aspartate carbamoyltransferase catalytic subunit/aspartate carbamoyltransferase regulatory chain; *pyrC*, dihydroorotase; *pyrD*, dihydroorotate dehydrogenase; *pyrE*, orotate phosphoribosyltransferase; *pyrF*, orotidine-5'-phosphate decarboxylase; *pyrR*, bifunctional *pyr* operon transcriptional regulator/uracil phosphoribosyltransferase *PyrR*; *upp*, uracil phosphoribosyltransferase.

4. Discussion

In this study, we conducted a comprehensive investigation based on the newly-assembled complete genome of *Epidermidibacterium keratini* EPI-7^T for a deeper understanding of *E. keratini* EPI-7^T. The results of complete genome data provide that more accurate and up-to-date comparative analysis and biosynthetic pathway prediction are possible by improving the number of functional genes and reducing the number of ambiguous genes such as unknown (DUF) or hypothetical genes compared to the previous annotation data. *E. keratini* EPI-7^T is still the only microorganism species belonging to the genus *Epidermidibacterium*. The statement means that in order to conduct a comparative analysis with *E. keratini* EPI-7^T, the nine species were chosen based on their phylogenetic features which were similar to those of *E. keratini* EPI-7^T, even if they belonged to different genera. We also tried to find common or specific features between *E. keratini* EPI-7^T and the comparison strains through whole-genome comparisons with the most genetically matched strains. By establishing the genetic association of these strains, we allowed researchers to make more meaningful comparisons with *E. keratini* EPI-7^T and provided insight into the unique characteristics and taxonomic distance of these strains.

As a result, pan-genome analysis with strains closely related to *E. keratini* EPI-7^T confirms that the pan-genome size increased whenever the genome of the comparison strain was added and that it was an open pan-genome with $0 < B < 1$. These results generally show that the pan-genome analysis, which performs a comparative analysis of different strain levels based on the same species, is in the process of environmental adaptation and evolution according to habitat [52,53]. However, our results indicate a low evolutionary correlation between *E. keratini* EPI-7^T and comparison strains with a close relationship. In addition, the proportion of orthologous gene clusters in the core genome is only 6.72%, which shows that the core genome size is small because they share only the genes (DNA replication and repair gene, ribosomal RNA gene, membrane transport gene, and regulatory genes) necessary for life and survival or add genomes of different genus strains [46,54]. Therefore, it is emphasized again through the pan-genome analysis that it is a novel strain that did not exist before at the genus level.

Genome-based phylogenetic analyses between *E. keratini* EPI-7^T and 20 human skin microbiomes revealed genetic similarities with the genera *Micrococcus*, *Cutibacterium*, and *Corynebacterium*. It also showed independent branching from the skin microbiome (especially *Staphylococcus* and *Streptococcus*), suggesting the possibility of a skin strain that could survive independently and have specific autonomous metabolic mechanisms. Therefore, to understand the hidden ability of *E. keratini* EPI-7^T derived from the skin, we tried to obtain clues by conducting comparative genome analysis with representative microbes inhabiting the skin. Among comparative genomic analyses with skin flora strains and various web-based tools, OrthoVenn2 and EggNOG-mapper showed significant results. In turn, the GO enrichment shared by all comparison strains of OrthoVenn2 is essential for the survival of organisms with functions related to biological processes and molecular functions in that essential genes are more conserved [55,56]. Exceptionally, *C. acnes* and *M. luteus*, which were most associated with *E. keratini* EPI-7^T through pairwise heatmap, share the ability of terpenoid biosynthetic processes. These terpenoids serve a variety of roles in central cellular processes, such as electron transport, photosynthesis, membrane fluidity, signaling, and cell wall formation [57,58]. Although terpenoids were generally known as metabolites of fungi or plants, it was confirmed that terpenoids of bacterial origin exist (especially *streptomycetes*). Terpene synthase is widely encoded in bacteria [59–61]. In addition, it is emphasized that the GO enrichment associated with oxidative stress and organic compound catabolism, which only *E. keratini* EPI-7^T has, can be an important ability for stress tolerance from environmental stress by rapidly catabolizing toxic substances and reactive oxygen [62,63]. Thus, it is anticipated that *E. keratini* EPI-7^T may have more exceptional resistance to external factors and environmental changes in the skin than other strains.

Amino acids are known as representative natural moisturizing factors (NMFs), although a precise mechanism has not been elucidated [64]. Collagen, which is widely applied in various fields as well as used for anti-inflammatory and beautifying purposes in the skin, has a characteristic of low permeability because it does not penetrate the skin layer well when injected from the outside due to its molecular size [65,66]. Therefore, instead of using a single method of making collagen in nano units, attempts are being made to induce collagen synthesis by mixing low-molecular-weight amino acids such as glycine, proline, hydroxyproline, alanine, and glutamic acid, which are the main components of collagen [67–69]. In addition, amino acids can be used for medical or cosmetic purposes for treating diseases and maintaining skin according to the combination of several amino acids or with other compounds as well as collagen [70,71]. Genome-wide prediction of the particular biosynthesis pathway *E. keratini* EPI-7^T confirmed that the *E. keratini* EPI-7^T genome contains genes that biosynthesize 20 amino acids. *E. keratini* EPI-7^T preserves many genes related to metabolism that can synthesize necessary amino acids from other substrates or amino acids in a nutrient-poor environment such as skin. Therefore, it is predicted that by supplying amino acids necessary for other microorganisms, it forms a symbiotic relationship, inhabits epidermal keratinocytes, and plays a role in supplying amino acids. Therefore, *E. keratini* EPI-7^T can form a syntrophic relationship and help

collagen synthesis in human skin cells while providing amino acids lacking in surrounding microorganisms and directly living in human epidermal keratinocytes [72,73].

In addition to amino acids, *E. keratini* EPI-7^T contains genes related to metabolite biosynthesis and enzymes, such as orotic acid, riboflavin, chorismite (shikimate), anthranilate, phytoene, antioxidants, superoxide dismutase, and peroxidase, which help various anti-aging, oxidative stress, wrinkle improvements, nutritional supplements, and moisturizing [74–76]. Specifically, orotic acid was well known as an intermediate in the synthesis of pyrimidines [77]. Also known as vitamin B13, orotic acid is one of the few vitamins that can help prevent skin aging [78]. In a previous study by Minsu Kang, et al., it was confirmed that the content of orotic acid in the *E. keratini* EPI-7^T solution was increased when uracil was added as a substrate [3]. Previously, orotic acid was predicted to be biosynthesized using uracil as a substrate. Although uracil cannot be converted to orotic acid, it can ultimately be converted to uridine monophosphate (UMP) in common with orotic acid. In addition, *pyrR* encodes an mRNA-binding attenuator, and UMP/UMT bind to *pyrR* protein and negatively regulates *pyr* expression [79,80]. Therefore, although uracil addition is not directly involved in the biosynthesis of orotic acid, we predict that the PyrR is inhibited as uracil is converted to UMP and the conversion of orotic acid to UMP is inhibited, resulting in orotic acid accumulation (Figure 8B). Indeed, Christopherson RI, et al. showed that in *Escherichia coli* k12, an exogenous donation such as adenine or uracil prevented the use of aspartate molecules in the de novo biosynthesis of pyrimidine nucleotides, resulting in an increase in aspartate concentration and a significant increase in orotic acid synthesis [81]. However, the computational in silico genome analysis only is limited in demonstrating the potential capabilities of *E. keratini* EPI-7^T. Therefore, we suggest the need for abundance screening of *E. keratini* EPI-7^T in different ethnic, gender, and age groups by microbial qRT-PCR quantitative analysis utilizing EPI-7^T-specific probes for the terms of skin microbiome research.

5. Conclusions

In this genome-scale study of *Epidermidibacterium keratini* EPI-7^T, we conducted a comprehensive investigation based on the newly-assembled complete genome of *E. keratini* EPI-7^T for a deeper understanding of *E. keratini* EPI-7^T. The genome-wide investigations confirmed *E. keratini* EPI-7^T, *C. acnes*, and *M. luteus* shared similar terpenoid biosynthesis, and *E. keratini* EPI-7^T contained many metabolite-related genes and the abilities related to environmental stress tolerance specifically developed. Additionally, the annotation-based manual inspection speculated that the gene required for vital amino acid biosynthesis is preserved in the *E. keratini* EPI-7^T genome, interacting with various microorganisms and skin cells and surviving in an environment where nutrients are challenging to obtain. In addition, considering the accumulation of orotic acid in a uracil-rich environmental, we expect to easily collect target metabolites from *E. keratini* EPI-7^T, depending on the nutrients added together. In conclusion, our study will provide important clues for future in vitro studies of EPI-7^T function and industrial applications such as derivation of microbial-based skin improvement materials.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/microorganisms11040870/s1>, Figure S1. Alignment of genome using Chromeis-ter.; Figure S2. Summary of shared GO enrichment between *E. keratini* EPI-7^T and seven skin flora.; Table S1. List of comparison strain.; Table S2. Description of annotations in EPI-7^T.; Table S3. BLAST result using the full-length *Epidermidibacterium keratini* EPI-7^T.; Table S4. Statistics of sequencing read for the *Epidermidibacterium keratini* EPI-7^T genome.; Table S5. Summary of sequencing read.; Table S6. Gene annotation result according to database.; Table S7. Comparison between newly-assembled and previous genomes.; Table S8. Bi- and uni-directional hit counting with PATRIC.; Table S9. Annotation of *Epidermidibacterium keratini* EPI-7^T and skin flora using EggNOG-mapper.; Table S10. List of 20 amino-acid-biosynthesis-related genes.; Table S11. List of metabolite-biosynthesis-related genes.

Author Contributions: Conceptualization, D.-G.L. and K.H.; methodology, Y.O., S.M., H.J., D.-G.L. and Y.-B.C.; investigation, Y.O., S.M., H.J., D.-G.L. and Y.-B.C.; writing—original draft, Y.O.; writing—review and editing, S.M. and K.H.; visualization, Y.O.; supervision, S.M., D.-G.L. and K.H.; project administration, K.H.; funding acquisition, D.-G.L. and K.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (Grant number: HP20C0171).

Data Availability Statement: All data generated during this study are included in this article.

Acknowledgments: The authors gratefully acknowledge the Center for Bio-Medical Engineering Core Facility at Dankook University. The authors also gratefully acknowledge the Department of Microbiology through the Research-Focused Department Promotion Project as a part of the University Innovation Support Program for Dankook University in 2022. This *Epidermidibacterium keratini* EPI-7^T strain was provided by COSMAX and COXMAX BTI center of Korea Cosmetic Industry Institute, Republic of Korea.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lee, D.-G.; Trujillo, M.E.; Kang, S.; Nam, J.-J.; Kim, Y.-J. *Epidermidibacterium keratini* gen. nov., sp. nov., a member of the family Sporichthyaceae, isolated from keratin epidermis. *Int. J. Syst. Evol. Microbiol.* **2018**, *68*, 745–750. [[CrossRef](#)]
2. Lee, Y.-G.; Lee, D.-G.; Gwag, J.E.; Kim, M.; Kim, M.; Kim, H.-G.; Ko, J.-H.; Yeo, H.; Kang, S.; Baek, N.-I. A 1,1'-biuracil from *Epidermidibacterium keratini* EPI-7 shows anti-aging effects on human dermal fibroblasts. *Appl. Biol. Chem.* **2019**, *62*, 14. [[CrossRef](#)]
3. Kang, M. *Exploring Anti-Aging Compounds Produced by Epidermidibacterium Keratini EPI-7 Strain Derived from Skin Microbiome*; Gyeongsang National University Graduate School: Jinju-si, Republic of Korea, 2021.
4. Buermans, H.P.J.; Den Dunnen, J.T. Next generation sequencing technology: Advances and applications. *Biochim. Biophys. Acta BBA-Mol. Basis Dis.* **2014**, *1842*, 1932–1941. [[CrossRef](#)] [[PubMed](#)]
5. van Dijk, E.L.; Auger, H.; Jaszczyszyn, Y.; Thermes, C. Ten years of next-generation sequencing technology. *Trends Genet.* **2014**, *30*, 418–426. [[CrossRef](#)]
6. Nakano, K.; Shiroma, A.; Shimoji, M.; Tamotsu, H.; Ashimine, N.; Ohki, S.; Shinzato, M.; Minami, M.; Nakanishi, T.; Teruya, K.; et al. Advantages of genome sequencing by long-read sequencer using SMRT technology in medical area. *Hum. Cell* **2017**, *30*, 149–161. [[CrossRef](#)]
7. Rhoads, A.; Au, K.F. PacBio Sequencing and Its Applications. *Genom. Proteom. Bioinform.* **2015**, *13*, 278–289. [[CrossRef](#)]
8. Fricke, W.F.; Rasko, D.A. Bacterial genome sequencing in the clinic: Bioinformatic challenges and solutions. *Nat. Rev. Genet.* **2014**, *15*, 49–55. [[CrossRef](#)] [[PubMed](#)]
9. Gunasekara, A.W.A.C.W.R.; Rajapaksha, L.G.T.G.; Tung, T. Whole-genome sequence analysis through online web interfaces: A review. *Genom. Inform.* **2022**, *20*, e3. [[CrossRef](#)]
10. Davis, J.J.; Wattam, A.R.; Aziz, R.K.; Brettin, T.; Butler, R.; Butler, R.M.; Chlenski, P.; Conrad, N.; Dickerman, A.; Dietrich, E.M.; et al. The PATRIC Bioinformatics Resource Center: Expanding data and analysis capabilities. *Nucleic Acids Res.* **2020**, *48*, D606–D612. [[CrossRef](#)]
11. Wattam, A.R.; Abraham, D.; Dalay, O.; Disz, T.L.; Driscoll, T.; Gabbard, J.L.; Gillespie, J.J.; Gough, R.; Hix, D.; Kenyon, R.; et al. PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res.* **2014**, *42*, D581–D591. [[CrossRef](#)]
12. Wang, Y.; Coleman-Derr, D.; Chen, G.; Gu, Y.Q. OrthoVenn: A web server for genome wide comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res.* **2015**, *43*, W78–W84. [[CrossRef](#)] [[PubMed](#)]
13. Almeida, A.; Nayfach, S.; Boland, M.; Strozzi, F.; Beracochea, M.; Shi, Z.J.; Pollard, K.S.; Sakharova, E.; Parks, D.H.; Hugenholtz, P.; et al. A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat. Biotechnol.* **2021**, *39*, 105–114. [[CrossRef](#)] [[PubMed](#)]
14. Cantalapedra, C.P.; Hernández-Plaza, A.; Letunic, I.; Bork, P.; Huerta-Cepas, J. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Mol. Biol. Evol.* **2021**, *38*, 5825–5829. [[CrossRef](#)] [[PubMed](#)]
15. Byrd, A.L.; Belkaid, Y.; Segre, J.A. The human skin microbiome. *Nat. Rev. Microbiol.* **2018**, *16*, 143–155. [[CrossRef](#)]
16. Grice, E.A.; Segre, J.A. The skin microbiome. *Nat. Rev. Microbiol.* **2011**, *9*, 244–253. [[CrossRef](#)]
17. Na, S.-I.; Kim, Y.O.; Yoon, S.-H.; Ha, S.-M.; Baek, I.; Chun, J. UBCG: Up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. *J. Microbiol.* **2018**, *56*, 280–285. [[CrossRef](#)]
18. Cogen, A.L.; Nizet, V.; Gallo, R.L. Skin microbiota: A source of disease or defence? *Br. J. Dermatol.* **2008**, *158*, 442–455. [[CrossRef](#)]
19. Davis, C.P. Normal Flora. In *Medical Microbiology*, 4th ed.; University of Texas Medical Branch at Galveston: Galveston, TX, USA, 1996.
20. Bianchet, R.T.; Cubas, A.L.V.; Machado, M.M.; Moecke, E.H.S. Applicability of bacterial cellulose in cosmetics—bibliometric review. *Biotechnol. Rep.* **2020**, *27*, e00502. [[CrossRef](#)]

21. Gupta, P.L.; Rajput, M.; Oza, T.; Trivedi, U.; Sanghvi, G. Eminence of Microbial Products in Cosmetic Industry. *Nat. Prod. Bioprospecting* **2019**, *9*, 267–278. [CrossRef]
22. Martins, A.; Vieira, H.; Gaspar, H.; Santos, S. Marketed Marine Natural Products in the Pharmaceutical and Cosmeceutical Industries: Tips for Success. *Mar. Drugs* **2014**, *12*, 1066–1101. [CrossRef]
23. Koren, S.; Walenz, B.P.; Berlin, K.; Miller, J.R.; Bergman, N.H.; Phillippy, A.M. Canu: Scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res.* **2017**, *27*, 722–736. [CrossRef] [PubMed]
24. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997.
25. Walker, B.J.; Abeel, T.; Shea, T.; Priest, M.; Abouelliel, A.; Sakthikumar, S.; Cuomo, C.A.; Zeng, Q.; Wortman, J.; Young, S.K.; et al. Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLoS ONE* **2014**, *9*, e112963. [CrossRef] [PubMed]
26. Seemann, T. Prokka: Rapid Prokaryotic Genome Annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [CrossRef] [PubMed]
27. Lagesen, K.; Hallin, P.; Rødland, E.A.; Staerfeldt, H.-H.; Rognes, T.; Ussery, D.W. RNAmmer: Consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **2007**, *35*, 3100–3108. [CrossRef]
28. Laslett, D.; Canback, B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* **2004**, *32*, 11–16. [CrossRef] [PubMed]
29. Seemann, T. Barrnap 0.9: Rapid Ribosomal RNA Prediction. GPLv3. 2013. Available online: <https://github.com/tseemann/barrnap> (accessed on 24 October 2022).
30. Bland, C.; Ramsey, T.L.; Sabree, F.; Lowe, M.; Brown, K.; Kyripides, N.C.; Hugenholtz, P. CRISPR Recognition Tool (CRT): A tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinform.* **2007**, *8*, 209. [CrossRef]
31. Hyatt, D.; Chen, G.-L.; Locascio, P.F.; Land, M.L.; Larimer, F.W.; Hauser, L.J. Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform.* **2010**, *11*, 119. [CrossRef]
32. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef]
33. Pruitt, K.D.; Tatusova, T.; Maglott, D.R. NCBI reference sequences (RefSeq): A curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **2007**, *35*, D61–D65. [CrossRef]
34. UniProt Consortium. UniProt: A hub for protein information. *Nucleic Acids Res.* **2015**, *43*, D204–D212. [CrossRef] [PubMed]
35. Mulder, N.; Apweiler, R. Interpro and interproscan. In *Comparative Genomics*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 59–70.
36. Finn, R.D.; Coghill, P.; Eberhardt, R.Y.; Eddy, S.R.; Mistry, J.; Mitchell, A.L.; Potter, S.C.; Punta, M.; Qureshi, M.; Sangrador-Vegas, A.; et al. The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* **2016**, *44*, D279–D285. [CrossRef] [PubMed]
37. Pérez-Wohlfeil, E.; Diaz-Del-Pino, S.; Trelles, O. Ultra-fast genome comparison for large-scale genomic experiments. *Sci. Rep.* **2019**, *9*, 10274. [CrossRef] [PubMed]
38. Hall, T.A. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In *Proceedings of the Nucleic Acids Symposium Series*; Oxford University Press: London, UK, 1999; pp. 95–98.
39. Zhao, Y.; Wu, J.; Yang, J.; Sun, S.; Xiao, J.; Yu, J. PGAP: Pan-genomes analysis pipeline. *Bioinformatics* **2012**, *28*, 416–418. [CrossRef]
40. Tamura, K.; Stecher, G.; Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* **2021**, *38*, 3022–3027. [CrossRef]
41. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree 2—Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE* **2010**, *5*, e9490. [CrossRef]
42. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [CrossRef]
43. Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **2017**, *45*, D353–D361. [CrossRef]
44. Caspi, R.; Billington, R.; Keseler, I.M.; Kothari, A.; Krummenacker, M.; Midford, P.E.; Ong, W.K.; Paley, S.; Subhraveti, P.; Karp, P.D. The MetaCyc database of metabolic pathways and enzymes—A 2019 update. *Nucleic Acids Res.* **2020**, *48*, D445–D453. [CrossRef]
45. Swamy, M.K.; Akhtar, M.S. *Natural Bio-Active Compounds: Volume 2: Chemistry, Pharmacology and Health Care Practices*; Springer: Berlin/Heidelberg, Germany, 2019.
46. Tettelin, H.; Massignani, V.; Cieslewicz, M.J.; Donati, C.; Medini, D.; Ward, N.L.; Angiuoli, S.V.; Crabtree, J.; Jones, A.L.; Durkin, A.S. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial “pan-genome”. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13950–13955. [CrossRef]
47. Tettelin, H.; Riley, D.; Cattuto, C.; Medini, D. Comparative genomics: The bacterial pan-genome. *Curr. Opin. Microbiol.* **2008**, *11*, 472–477. [CrossRef] [PubMed]
48. Zengler, K.; Zaramela, L.S. The social network of microorganisms—How auxotrophies shape complex communities. *Nat. Rev. Genet.* **2018**, *16*, 383–390. [CrossRef] [PubMed]
49. Lin, L.; Du, R.; Wang, Y.; Wu, Q.; Xu, Y. Regulation of auxotrophic lactobacilli growth by amino acid cross-feeding interaction. *Int. J. Food Microbiol.* **2022**, *377*. [CrossRef]
50. Bertels, F.; Merker, H.; Kost, C. Design and Characterization of Auxotrophy-Based Amino Acid Biosensors. *PLoS ONE* **2012**, *7*, e41349. [CrossRef] [PubMed]

51. Liu, Y.-K.; Kuo, H.-C.; Lai, C.-H.; Chou, C.-C. Single amino acid utilization for bacterial categorization. *Sci. Rep.* **2020**, *10*, 12686. [[CrossRef](#)] [[PubMed](#)]
52. Kant, R.; Rintahaka, J.; Yu, X.; Sigvart-Mattila, P.; Paulin, L.; Mecklin, J.-P.; Saarela, M.; Palva, A.; Von Ossowski, I. A Comparative Pan-Genome Perspective of Niche-Adaptable Cell-Surface Protein Phenotypes in *Lactobacillus rhamnosus*. *PLoS ONE* **2014**, *9*, e102762. [[CrossRef](#)]
53. Mira, A.; Martín-Cuadrado, A.B.; D’Auria, G.; Rodríguez-Valera, F. The bacterial pan-genome: A new paradigm in microbiology. *Int. Microbiol.* **2010**, *13*, 45–57. [[CrossRef](#)]
54. Lefébure, T.; Bitar, P.D.P.; Suzuki, H.; Stanhope, M.J. Evolutionary Dynamics of Complete *Campylobacter* Pan-Genomes and the Bacterial Species Concept. *Genome Biol. Evol.* **2010**, *2*, 646–655. [[CrossRef](#)]
55. Jordan, I.K.; Rogozin, I.B.; Wolf, Y.I.; Koonin, E.V. Essential Genes Are More Evolutionarily Conserved Than Are Nonessential Genes in Bacteria. *Genome Res.* **2002**, *12*, 962–968. [[CrossRef](#)]
56. Luo, H.; Gao, F.; Lin, Y. Evolutionary conservation analysis between the essential and nonessential genes in bacterial genomes. *Sci. Rep.* **2015**, *5*, 13210. [[CrossRef](#)]
57. Tholl, D. Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. *Curr. Opin. Plant Biol.* **2006**, *9*, 297–304. [[CrossRef](#)] [[PubMed](#)]
58. Vavitsas, K.; Fabris, M.; Vickers, C.E. Terpenoid Metabolic Engineering in Photosynthetic Microorganisms. *Genes* **2018**, *9*, 520. [[CrossRef](#)] [[PubMed](#)]
59. Pollak, F.C.; Berger, R.G. Geosmin and Related Volatiles in Bioreactor-Cultured *Streptomyces citreus* CBS 109.60. *Appl. Environ. Microbiol.* **1996**, *62*, 1295–1299. [[CrossRef](#)] [[PubMed](#)]
60. Yamada, Y.; Kuzuyama, T.; Komatsu, M.; Shin-Ya, K.; Omura, S.; Cane, D.-E.; Ikeda, H. Terpene synthases are widely distributed in bacteria. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 857–862. [[CrossRef](#)]
61. Dickschat, J.S.; Martens, T.; Brinkhoff, T.; Simon, M.; Schulz, S. Volatiles Released by a *Streptomyces* Species Isolated from the North Sea. *Chem. Biodivers.* **2005**, *2*, 837–865. [[CrossRef](#)]
62. Oikawa, P.Y.; Lerda, M.T. Catabolism of volatile organic compounds influences plant survival. *Trends Plant Sci.* **2013**, *18*, 695–703. [[CrossRef](#)]
63. Juhanson, J.; Truu, J.; Heinaru, E.; Heinaru, A. Survival and catabolic performance of introduced *Pseudomonas* strains during phytoremediation and bioaugmentation field experiment. *FEMS Microbiol. Ecol.* **2009**, *70*, 446–455. [[CrossRef](#)]
64. Rawlings, A.V.; Harding, C.R. Moisturization and skin barrier function. *Dermatol. Ther.* **2004**, *17*, 43–48. [[CrossRef](#)]
65. Yamauchi, M.; Woodley, D.T.; Mechanic, G.L. Aging and cross-linking of skin collagen. *Biochem. Biophys. Res. Commun.* **1988**, *152*, 898–903. [[CrossRef](#)]
66. Subhan, F.; Hussain, Z.; Tauseef, I.; Shehzad, A.; Wahid, F. A review on recent advances and applications of fish collagen. *Crit. Rev. Food Sci. Nutr.* **2020**, *61*, 1027–1037. [[CrossRef](#)]
67. Miyahara, T.; Shiozawa, S.; Murai, A. The Effect of Age on Amino Acid Composition of Human Skin Collagen. *J. Gerontol.* **1978**, *33*, 498–503. [[CrossRef](#)] [[PubMed](#)]
68. Murakami, H.; Shimbo, K.; Inoue, Y.; Takino, Y.; Kobayashi, H. Importance of amino acid composition to improve skin collagen protein synthesis rates in UV-irradiated mice. *Amino Acids* **2012**, *42*, 2481–2489. [[CrossRef](#)] [[PubMed](#)]
69. Murakami, H.; Shimbo, K.; Takino, Y.; Kobayashi, H. Combination of BCAAs and glutamine enhances dermal collagen protein synthesis in protein-malnourished rats. *Amino Acids* **2013**, *44*, 969–976. [[CrossRef](#)] [[PubMed](#)]
70. Diaz, I.; Namkoong, J.; Wu, J.Q.; Giancola, G. Amino acid complex (AACComplex) benefits in cosmetic products: *In vitro* and *in vivo* clinical studies. *J. Cosmet. Dermatol.* **2022**, *21*, 3046–3052. [[CrossRef](#)] [[PubMed](#)]
71. Corsetti, G.; D’Antona, G.; Dioguardi, F.S.; Rezzani, R. Topical application of dressing with amino acids improves cutaneous wound healing in aged rats. *Acta Histochem.* **2010**, *112*, 497–507. [[CrossRef](#)]
72. D’Souza, G.; Kost, C. Experimental evolution of metabolic dependency in bacteria. *PLoS Genet.* **2016**, *12*, e1006364. [[CrossRef](#)]
73. Kouzuma, A.; Kato, S.; Watanabe, K. Microbial interspecies interactions: Recent findings in syntrophic consortia. *Front. Microbiol.* **2015**, *6*, 477. [[CrossRef](#)]
74. Samtiya, M.; Samtiya, S.; Badgujar, P.C.; Puniya, A.K.; Dhewa, T.; Aluko, R.E. Health-Promoting and Therapeutic Attributes of Milk-Derived Bioactive Peptides. *Nutrients* **2022**, *14*, 3001. [[CrossRef](#)]
75. Musae, A.; Sadykova, S.; Anambayeva, A.; Saizhanova, M.; Balkanay, G.; Kolbaev, M. Mare’s Milk: Composition, Properties, and Application in Medicine. *Arch. Razi Inst.* **2021**, *76*, 1125.
76. Lei, M.; Wu, X.; Huang, C.; Qiu, Z.; Wang, L.; Zhang, R.; Zhang, J. Trehalose induced by reactive oxygen species relieved the radial growth defects of *Pleurotus ostreatus* under heat stress. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 5379–5390. [[CrossRef](#)]
77. Kelley, W.N.; Fox, I.H.; Wyngaarden, J.B. Regulation of purine biosynthesis in cultured human cells: I. Effects of orotic acid. *Biochim. Et Biophys. Acta BBA-Gen. Subj.* **1970**, *215*, 512–516. [[CrossRef](#)]
78. Jastrzębska, E.; Wadas, E.; Daszkiewicz, T.; Pietrzak-Fiecko, R. Nutritional value and health-promoting properties of mare’s milk—A review. *Czech J. Anim. Sci.* **2017**, *62*, 511–518. [[CrossRef](#)]
79. Dong, H.; Zhang, D. Current development in genetic engineering strategies of *Bacillus* species. *Microb. Cell Fact.* **2014**, *13*, 63. [[CrossRef](#)] [[PubMed](#)]

80. Suzuki, H.; Murakami, A.; Yoshida, K.-I. Counterselection System for *Geobacillus kaustophilus* HTA426 through Disruption of *pyrF* and *pyrR*. *Appl. Environ. Microbiol.* **2012**, *78*, 7376–7383. [[CrossRef](#)]
81. Christopherson, R.I.; Finch, L.R. Response of the pyrimidine pathway of *Escherichia coli* K12 to exogenous adenine and uracil. *Eur. J. Biochem.* **1978**, *90*, 347–358. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.