Universitat Autònoma de Barcelona

Variability and diversity in the skin of healthy dogs: microbiome, genome and environment

Anna Maria Cuscó Martí

Variability and diversity in the skin of healthy dogs: microbiome, genome and environment

Anna Maria Cuscó Martí

Departament de Ciència Animal i dels Aliments

Facultat de Veterinària, Universitat Autònoma de Barcelona

A thesis submitted for the degree of Doctor (PhD)

Director: Dr. Olga Francino Martí

Tutor: Dr. Armand Sánchez Bonastre

June 2017

La **Dra. Olga Francino Martí**, investigadora del Departament de Ciència Animal i dels Aliments de la Universitat Autònoma de Barcelona

Certifica que **Anna Maria Cuscó Martí** ha dut a terme sota la meva direcció el treball de recerca realitzat al Departament de Ciència Animal i dels Aliments de la Facultat de Veterinària de la Universitat Autònoma de Barcelona, que ha portat a l'elaboració d'aquesta Tesi Doctoral, titulada «Variability and diversity in the skin of healthy dogs: microbiome, genome and environment».

Bellaterra, 2 de juny de 2017

Dra. Olga Francino Martí

Anna Maria Cuscó Martí

This work was supported by a grant awarded by **Pla de Doctorats Industrial** (2013 DI 011) provided by the Agencia de Gestió d'Ajuts Universitaris i de Recerca (AGAUR); Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement de la Generalitat de Catalunya

This Industrial PhD project was performed in collaboration with **Vetgenomics, SL**

Anna Maria Cuscó Martí was funded by Vetgenomics, SL.

Resum

En la darrera dècada han incrementat molt els estudis del microbioma –conjunt de microbis que habiten un microhàbitat comú– gràcies a la millora de les tècniques de seqüenciació massiva. Aquesta nova manera d'analitzar les comunitats microbianes ha permès detectar una gran diversitat abans ignorada en infinitat d'ambients diferents, entre ells diferents parts del cos.

El microbioma associat a certes parts del cos presenta un gran interès per la seva interacció amb les cèl·lules i la immunitat de l'hoste. En un estat normal, aquestes interaccions microbioma – hoste promouen la salut de l'individu. En canvi, en una patologia, aquestes interaccions es troben alterades.

L'objectiu d'aquesta tesis és definir la variabilitat pròpia del gos sa, principalment a nivell de microbioma cutani però també en immunitat innata. Per tal d'aconseguir aquest objectiu utilitzem tècniques de seqüenciació massiva i procedim a: i) caracteritzar la immunitat innata dels gossos sans a nivell de polimorfisme en *Toll-like Receptors* –primers sensors de patògens; ii) caracteritzar el microbioma de la pell en dues cohorts de gossos sans; iii) provar la seqüenciació de tercera generació *(single-molecule seqüencing)* per nanopors aplicada a la caracterització del microbioma.

S'ha caracteritzat la diversitat genètica del *Toll-like Receptors* en una població d'estudi amplia per representar la gran diversitat genètica pròpia de les espècies de cànids (set races diferents de gossos i dues poblacions de llops). S'han descobert noves variants genètiques, i s'han detectat d'altres prèviament identificades. S'ha dissenyat i validat un xip de genotipatge individual amb 64 sondes per detectar algunes de les variants amb possible efecte a la proteïna (mutacions no sinònimes), ja que són les que poden estar afectant a la resposta immunitària.

El microbioma de la pell en gossos sans s'ha avaluat en 8 regions diferents de la pell i en dues cohorts de gossos sans. S'han inclòs diferents regions representatives de la diversitat cutània i se n'han pogut determinar certes característiques comunes. Entre totes les regions de la pell incloses, la cara interna del pavelló auricular és la que presenta valors més grans de diversitat mentre que la regió perianal i nasal les que presenten valors més baixos. Tot i això, hem detectat que les diferents regions cutànies són més semblants en un mateix individu, tan en gossos d'ambient diferents com en gossos vivint en un mateix espai i interaccionant entre ells. Per tant, l'efecte individual és el factor més determinant tan de l'estructura com de composició del microbioma de la pell dels gossos. Hem d'entendre aquest efecte individual, com un conjunt entre l'hoste, el seu ambient i el seu comportament.

Finalment, hem provat una nova estratègia per caracteritzar el microbioma basada en la seqüenciació de tercera generació per nanopors. L'estratègia més usual en l'estudi de microbioma està basada en la seqüenciació d'un gen marcador de bacteris i arqueobacteris (16S rRNA) utilitzant fragments curts i seqüenciadors de segona generació. La seqüenciació de

tercera generació *(single-molecule seqüencing)* permet seqüenciar fragments més llargs d'ADN, permeten analitzar tot el gen marcador sencer. Tot i la baixa precisió de la tècnica escollida, hem detectat encara més diversitat en la pell dels gossos i hem pogut assignar la taxonomia a nivells inferiors. La tecnologia avaluada té molt de potencial, però probablement altres aproximacions experimentals donarien resultats més satisfactoris en l'estudi del microbioma mitjançant gens marcadors.

En conclusió, s'han pogut ampliar els coneixements en la variabilitat pròpia del gos sa sobretot a nivell de microbioma de la pell, però també en immunitat innata. Per altra banda, aquest projecte de Doctorat Industrial conjunt amb Vetgenomics ha permès obrir i posar a punt una nova línia estratègica a l'empresa.

Summary

The community of microorganisms that inhabits a specific environment is named as the microbiome. In the last decade, microbiome studies have boosted due to the apparition and development of next-generation sequencing (NGS) techniques that allow massive sequencing of DNA. This new way of analyzing microbial communities allowed detecting a wide range of diversity previously ignored in many different environments, including the body.

The microbiome associated to different body sites, interacts to host cells and immunity to contribute to the host functions as well as promote health. In contrast, in a pathology these microbiome-host interactions are altered.

The aim of this thesis is to define the intrinsic variability of the healthy dog, mainly at skin microbiome level, but also at innate immunity. To achieve this goal we use massive sequencing techniques and proceed to: i) characterize the innate immunity of healthy dogs at polymorphism level in Toll-like Receptors, which are the first sensors of pathogens; ii) characterize the skin microbiome of two cohorts of healthy dogs; iii) test the third generation sequencing (single-molecule sequencing) by nanopores applied to microbiome studies.

We characterized the genetic diversity of the Toll-like receptors in a wide population to represent the genetic diversity of the canid species (seven different breeds of dogs and two populations of wolves). New genetic variants have been discovered and others previously identified have also been detected. We designed and validated a genotyping chip with 64 individual probes to detect some of the variants with a likely effect in the protein (non-synonymous mutations), since they may be affecting the immune response.

The skin microbiome in healthy dogs has been evaluated in eight different regions of the skin and in two cohorts of healthy dogs. Different regions representing the skin diversity have been included and some common trends identified. Among all the skin regions, the inner part of the ear presented higher alpha diversity values, whereas nasal and perianal regions presented the lowest ones. However, we found that different skin regions on the same dog resemble more than the same skin region in different dogs, either in dogs living in different environments or in dogs living in the same space and interacting with each other. Therefore, the individual effect is the main factor driving the structure and composition of the skin microbiome in dogs. We must understand this effect as the individual, together with its environment and its behavior.

Finally, we tested a new strategy to characterize the microbiome based on third generation sequencing by nanopores. The usual approach in microbiome studies is targeting of a marker gene for bacteria and archaea (16S rRNA) with short primers and second-generation sequencers. The third-generation sequencing (single-molecule sequencing) can sequence long DNA fragments, and hence analyze the full-length marker gene (1,500 bp). Despite the low accuracy of the technique chosen, we were able to detect even more diversity in the skin of dogs and we could assign sequences to lower taxonomic levels. The technology assessed has a

great potential, but probably other experimental approaches would give better results in the study of the microbiome by marker genes.

In conclusion, we have been able to expand the knowledge on the variability of the healthy dog, especially at the skin microbiome level, but also at innate immunity one. Moreover, this Industrial Doctorate project together in collaboration with Vetgenomics allowed us to open and consolidate a new strategic business area for the company.

Contents

List of Boxes

List of Tables

List of figures

1. Introduction

In the present section we will provide an extensive overview of the topics covered in this thesis.

The first part will provide the technical framework needed to fully understand a microbiota study, from the experimental design, to the sample collection and extraction, further processing and sequencing to the final bioinformatics analyses and interpretation of the results.

The second part will provide the biological background to understand the object of this study, which is the skin. It will review skin anatomy and ecology on dogs, as well as genetic variability in genes responsible for skin innate immunity.

The final part of the introduction will review the present state of the art of skin microbiota on human and dogs, both in health and in disease.

1.1. Conducting a microbiota study

Microbiome studies have boosted since the apparition of next-generation sequencing (NGS) techniques that allow sequencing DNA massively and in parallel to obtain a huge amount of data (Kuczynski et al., 2011) (Section 1.1.2). The microbiome is defined as the microbial community (bacteria, virus, fungi, etc.) that inhabits a specific environment (Box 1) (Marchesi and Ravel, 2015). Thus, a microbiome sample is synonymous of thousands of different DNA targets and NGS is the key tool to understand all this complexity. See Box 1 for a collection of the main key concepts on microbiome research.

Microbiome studies can follow either an amplicon-based approach or a shotgun whole genome sequencing (WGS) one (Figure 1a). On one hand, amplicon-based approaches aim to describe the taxonomic composition and diversity of a community amplifying some regions of 16S rRNA marker gene, which is highly conserved among prokaryotes but also has hypervariable regions that differ among taxa (Section 1.1.1.1.3). On the other hand, WGS approach aims to sequence all the DNA present within an ecological niche, thus seeing the entire community (bacteria, archaea, fungi, virus, etc.) with all their genes (Kuczynski et al. 2011; Grice and Segre 2012). This approach gives both taxonomical and functional information although it is more expensive and computationally challenging (Luo et al., 2013).

Once the design of the study is clear and well defined, samples need to be collected using a consistent method from the defined skin site/s. Then, researchers should choose the most adequate DNA extraction protocol and PCR primer sets for their specific environment, if following an amplicon-based approach. After that, DNA is ready to be sequenced using NGS techniques. The last step is analyzing the output data using several bioinformatics tools (Figure 1b).

In this thesis we have chosen an amplicon-based approach (16S rRNA gene) to study skin microbiota in healthy dogs. In the following sections we will provide a more detailed overview of the workflow to follow when performing an amplicon-based microbiota study of skin samples. We will review step by step the process, from the experimental considerations and the main NGS techniques to the data analysis.

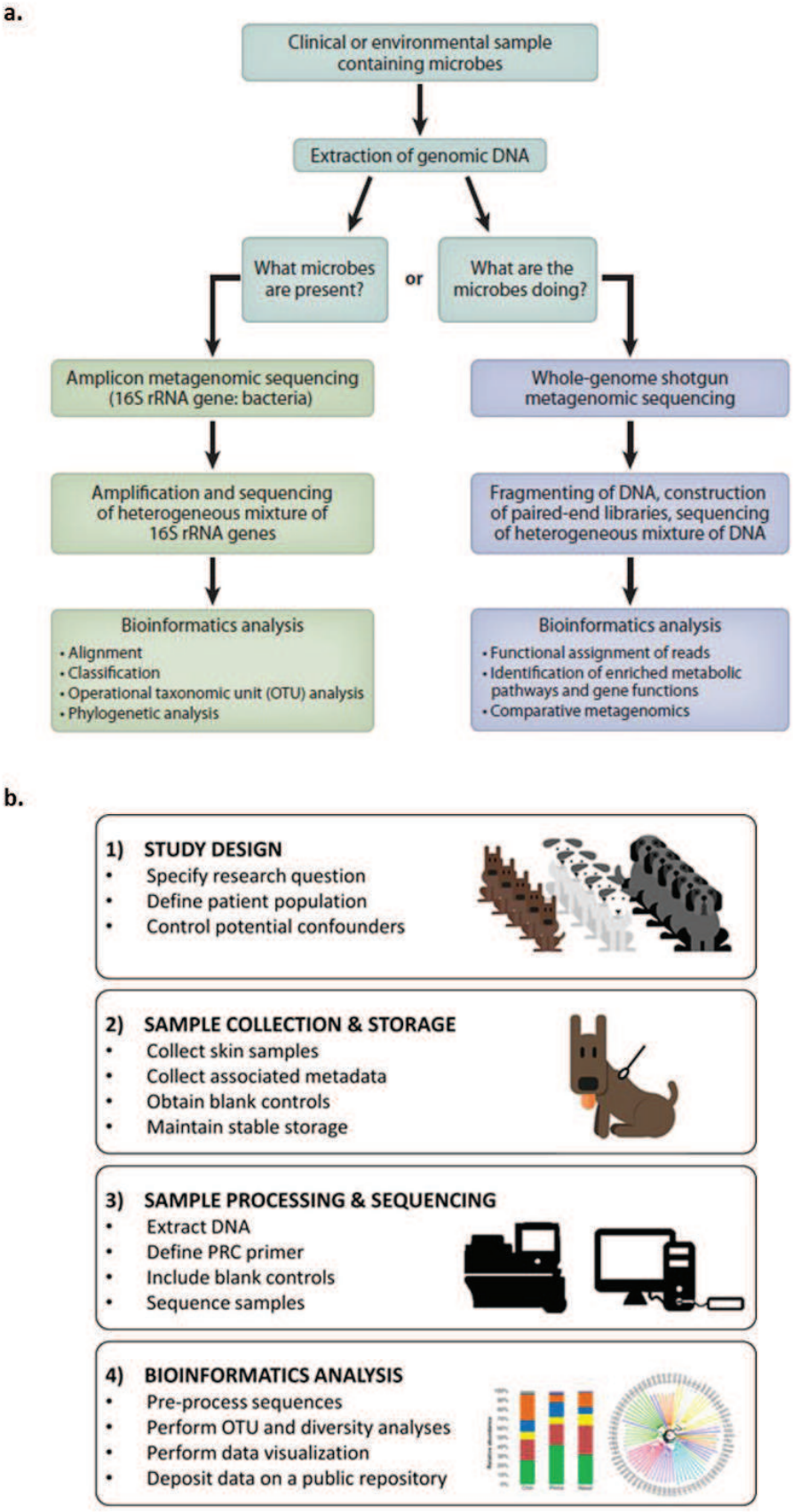


Figure 1. Performing a microbiota study. In a) main approaches depending on the objective; (from Grice and Segre 2012). in b) main steps of a skin microbiota study (adapted from Kong et al. 2017).

Box 1. Summary of the main nomenclature used on microbiota studies.

Microbiome	Collection of microorganisms together with their genes, in a defined microenvironment
Microbiota	Collection of microorganisms in a defined microenvironment
Metagenome	Collection of genes and genomes present in a defined microenvironment obtained with WGS approaches.
Whole genome shotgun sequencing (WGS)	From each defined microenvironment DNA is collected, which can contain bacteria, fungi, viruses and even host DNA. This DNA is randomly sheared into small pieces and sequenced. Afterwards, they are assembled into continuous longer sequences.
16S rRNA sequencing	16S rRNA gene is universal among prokaryotes and can be used to taxonomically classify bacteria. This taxonomic resolution comes from its gene structure: nine hypervariable regions surrounded by highly conserved regions.
Operational Taxonomic Unit (OTU)	Cluster of microorganisms that have similar DNA sequences on a taxonomic marker gene at specific threshold (97%, 99%, etc.). They are usually representing a specific bacterial species or taxon.
Taxonomic microbial composition	Relative abundances of different microorganisms of the microbiota.
Alpha diversity	Measures the number and distribution of OTUs (~taxa) within a unique sample.
Richness	Alpha diversity metrics that consider only the number of OTUs (e.g. Observed species)
Evenness	Alpha diversity metrics that consider the number of OTUs and their relative abundances (e.g. Shannon index)
Beta diversity	Measures the similarity (~shared OTUs) between bacterial communities. It computes distance matrices that can be plotted.
Unweighted UniFrac	Beta diversity metrics to assess divergences in microbiota composition (phylogeny and presence/absence of OTUs).
Weighted UniFrac	Beta diversity metrics to assess divergences in microbiota structure (phylogeny, presence/absence of OTUs and relative abundances).

1.1.1. Experimental considerations

Microbiome studies using next-generation sequencing need to be designed considering many technical variables (Kuczynski et al., 2011; Rogers and Bruce, 2010). Researchers have seen divergences on microbiome composition depending on: sample collection and storage (Dominianni et al., 2014); DNA extraction protocols (Wagner Mackenzie et al., 2015; Wesolowska-Andersen et al., 2014); primers chosen for amplification (Kuczynski et al., 2011; Meisel et al., 2016); sequencing platform used (Castelino et al., 2017; Clooney et al., 2016; Fouhy et al., 2016) or even clustering method employed on the bioinformatics analyses (Kopylova et al., 2016; Schmidt et al., 2014, 2015).

Skin microbiota samples are more challenging to process than those from other body sites (Kong et al., 2017). Thus, researchers have to deal with some extra factors such as: low microbial biomass linked with a higher risk of sample contamination (Salter et al., 2014); or the fact that skin has different microenvironments and harbors site-specific microbiota (Costello et al. 2009; Grice et al. 2009). At least, sample collection methods (Chng et al. 2016; Grice et al. 2008) or storage conditions (Lauber et al., 2010) seem not to be affecting significantly skin microbiota structure or composition.

In this dissertation most of the experimental factors were chosen following the recommended procedures of the Human Microbiome Project consortium (Human Microbiome Project Consortium., 2012).

1.1.1.1. Microbiota study design

The first step when performing a microbiota study is the study design, which needs to be appropriate to answer the research question. On one hand, cross-sectional studies including different skin sites allow describing and assessing skin microbiota variability and diversity on healthy individuals. Cross-sectional studies can also be case-control studies for assessing skin microbiota changes during a dermatological disease or any other disruption that affects the skin. On the other hand, longitudinal studies will allow assessing skin microbiota stability through time, either under normal conditions or under any disruption (disease, environmental change).

Cross-sectional studies targeting healthy skin are found in the early stages of the microbiota research and allow creating a background and determining the normal variability of the skin microbiota. In human skin microbiota, first studies assessed the normal variability on healthy individuals using a cross-sectional approach (Grice et al. 2009; Costello et al. 2009). On dog skin microbiota, only one study aimed to assess the normal variability of different skin sites in healthy dogs (Rodrigues Hoffmann et al., 2014). Later on, we can find more case-control and longitudinal studies aimed to assess altered states.

Besides choosing an appropriate study design, one should also define the target population. To do that, the researcher needs to define which metadata to collect, inclusion and exclusion criteria, and sample size.

Assessment of the metadata is really important on a microbiota study, to account for possible confounder factors. Commonly collected metadata include age, sex, antibiotic use, and sampling sites, but also other factors may influence the skin microbiome, such as cohabitation, hygiene, season, time of day, country of birth, mode of delivery, or diet (Kong et al., 2017). Moreover, inclusion and exclusion criteria that define the study population should be very well defined and also the appropriate cohort sample size to reach significant results should be previously estimated (Kong et al., 2017). Some approaches and tools have been proposed to assess the suitable sample size (Kelly et al., 2015; La Rosa et al., 2012) (https://github.com/biocore/Evident).

1.1.1.2. DNA extraction

Microbiota harbors a wide range of microorganisms with different characteristics, so a robust protocol to extract DNA of all the representatives is needed. For example, gram-positive bacteria present a cell wall with thick layers of peptidoglycan difficult to break. These differences in cell wall composition can cause bacterial cell lysis to be less efficient on gram-positive bacteria, and that in turn can distort the apparent microbiota composition and diversity (Kong et al., 2017; Yuan et al., 2012).

Thus, to avoid this bias different researchers have evaluated the effect of adding a mechanical disruption step by bead beating to the conventional DNA extraction procedure, which had led to improved results (Albertsen et al., 2015; Santiago et al., 2014; Sergeant et al., 2012; Walker et al., 2015; Yuan et al., 2012). After the lysis, DNA can be purified by alcohol precipitation or by binding to affinity columns.

1.1.1.3. Amplicon-based approach: 16S RNA marker gene

The 16S small ribosomal subunit gene (16S rRNA) is the most widely used marker gene for microbiota studies(Kuczynski et al., 2011), due to: (1) it is a gene ubiquitously found in bacteria and archaea; (2) its structure includes both conserved regions, which can be used for designing "universal" amplification primers, as well as nine hypervariable regions (V1-V9), which can be effectively used to distinguish among taxa (Clarridge, 2004) (Figure 2); (3) it has several large databases of reference sequences and taxonomies, such as Greengenes (DeSantis et al., 2006; McDonald et al., 2012), SILVA (Quast et al., 2013) or the Ribosomal Database Project (Cole et al., 2014).

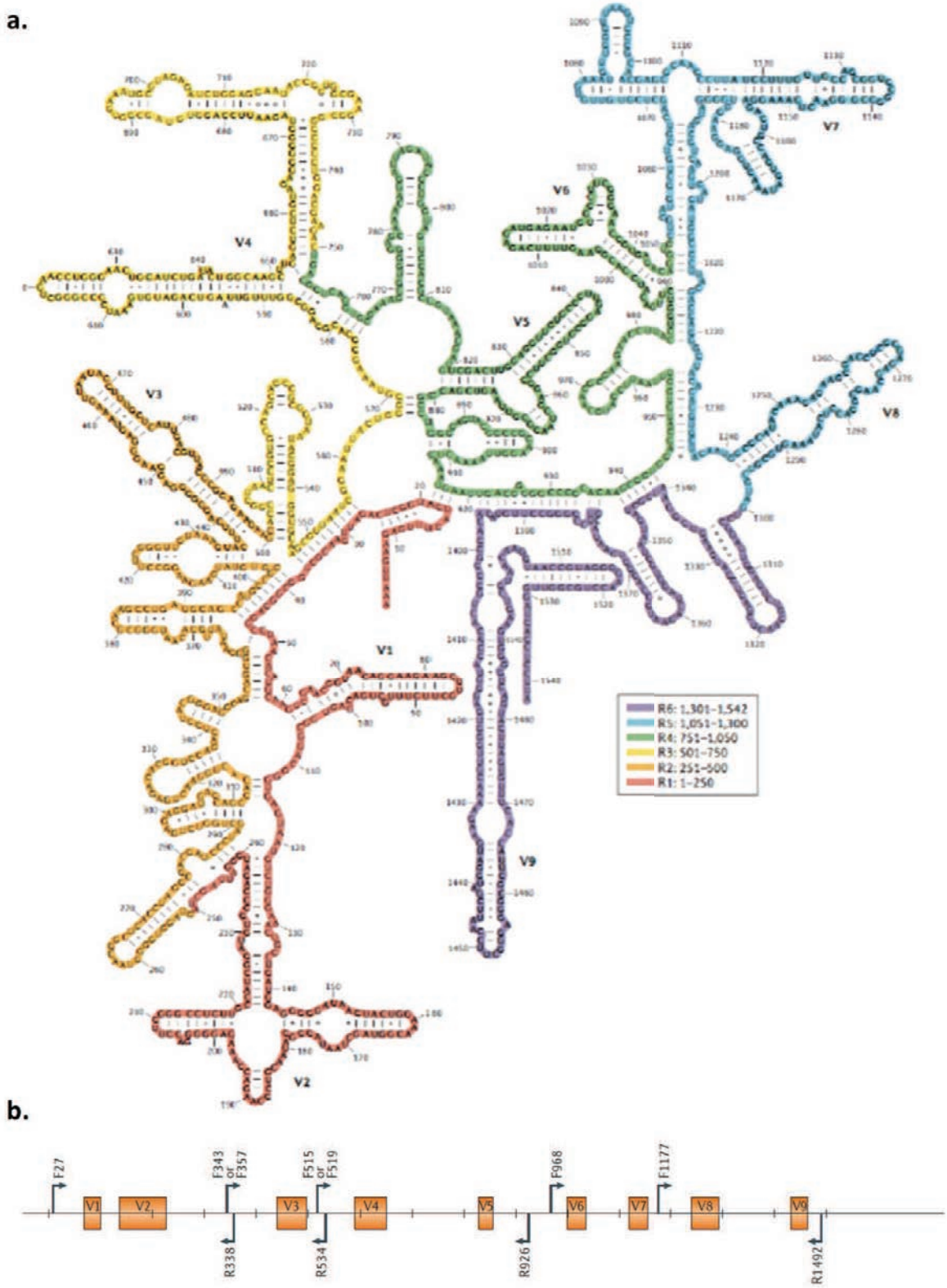


Figure 2. 16S rRNA marker gene. a) 16S rRNA gene secondary structure with hypervariable regions indicated (from Yarza et al. 2014); and b) 16S rRNA gene scheme with the most widely used primers for microbiota assessment (modified from Kuczynski et al. 2011).

When working with 16S rRNA gene to assign taxonomy, different similarity thresholds are applied to define a bacterial species. Archaeal and bacterial species were classically defined as a group of strains with: (a) certain degree of phenotypic consistency, (b) 70% of DNA–DNA hybridization and (c) at least 97% of gene-sequence identity on their 16S rRNA (Gevers et al., 2005). Later, another characteristic was added: (d) 94 - 96% of average sequence identities of shared genes (Richter and Rossello-Mora, 2009).

When focusing on 16S rRNA gene to define a species, the 97% identity cut-off was not enough within some bacterial genera to discriminate among different species and several researchers proposed that a 98.7% similarity threshold was more adequate (Stackebrandt and Ebers, 2006; Yarza et al., 2014). This identity cut-off should probably be adapted within each bacterial genus (Rossi-Tamisier et al., 2015).

The most commonly used strategy to assess microbiota composition is sequencing specific hypervariable regions of 16S rRNA gene and clustering the sequences using a similarity threshold of 97%. Thus, many authors assessed the specificity and universality of several 16S primers sets (Chakravorty et al., 2007; Jumpstart Consortium Human Microbiome Project Data Generation Working Group, 2012; Klindworth et al., 2013; Kuczynski et al., 2011; Mizrahi-Man et al., 2013). These 16S hypervariable regions are sequenced using NGS platforms, which are classified as 2nd or 3rd generation platforms (reviewed in Section 1.1.2), depending on the sequencing technology they use.

On one hand, 2nd generation sequencers have sequencing length limitations, so researchers have to choose primers targeting specific hypervariable regions to obtain short amplicons. Among the different options, the most universal primer set is F515-R806 (region V4) that captures both bacteria and archaea. However, it fails to amplify *Propionibacterium*, which is an abundant genus on the skin, thus in this dissertation we have worked with F27-F338 (region V1-V2), which was more suitable for our skin microbiota samples (Kuczynski et al., 2011; Walters et al., 2011).

On the other hand, 3rd generation sequencers can sequence long reads, so researchers choose amplifying the full (or almost full) length 16S rRNA gene using several universal primers (Klindworth et al., 2013). This approach gives a better and more accurate assessment of taxonomic diversity (Yarza et al., 2014). Here in this dissertation we used mainly primer set F27-R1492 (region V1-V9) and also F27-R1391 (region V1-V8) when testing Oxford Nanopore MinION™ 3rd generation sequencer. These two primer sets have low non-coverage rates even at phylum level being a good choice to assess microbiota diversity (Mao et al., 2012).

1.1.2. Next generation sequencing (NGS)

Once the sample has been successfully processed, the next step is transforming the DNA into data. The development of NGS has allowed obtaining huge amount of data in short periods of time, although with higher associated error rates (~0.1–15%) when compared to classical Sanger sequencing (Goodwin et al., 2016).

Main aspects to consider when choosing the appropriate sequencing platform for a specific study are: length of the reads; number of the reads (sequencing depth); and error rate and type (Vincent et al., 2016). The length of the reads is linked to the amount of information obtained from a single molecule, which is particularly significant when the aim is to phylogenetically differentiate 16S rRNA gene fragments (Jumpstart Consortium Human Microbiome Project Data Generation Working Group, 2012; Schloss et al., 2016a). The total number of reads is the most important parameter for quantitative applications (Vincent et al., 2016), such as determining the microbiota diversity and composition. The combination of read length and number of reads defines the throughput of an instrument in number of bases per run. Both the error rate and the different types of errors are specific of each sequencing technology and can be minimized with an increased coverage (Vincent et al., 2016)(Goodwin et al., 2016).

Conventionally NGS platforms have been divided in two groups: 2nd and 3rd generation platforms. 2nd generation sequencing aim is to obtain large amounts of data by massive sequencing in parallel millions of short reads at an affordable cost. The sequencing process requires a previous PCR-amplification of the DNA. In contrast, 3rd generation platforms use a single-molecule real-time sequencing approach, which means that they can sequence individual DNA molecules avoiding the necessity of PCR amplification and its associated bias. Another advantage is their ability to sequence long fragments (Glenn, 2011; Goodwin et al., 2016; Mardis, 2017).

Some of the most commonly used platforms of 2nd generation sequencing are 454, Illumina and Ion Torrent, which have some common and distinctive traits (Figure 3). The first step in the sequencing process is the library preparation, were the DNA is fragmented and the sequencing adaptors ligated. The following step is the clonal amplification of the individual DNA fragments that can be bead-based (454 and Ion Torrent) or solid-state (Illumina). The last step is the sequencing process itself.

Both 454 and Ion Torrent perform a bead-based amplification prior to sequencing, where each DNA molecule is immobilized in a single bead and clonally amplified using emulsion PCR (emPCR) (Figure 3b.1) (Dressman et al., 2003). In the case of Illumina a solid-state amplification is used with covalently bound forward and reverse primers where DNA molecules attach and clonally amplify forming clusters of sequences (Figure 3c.1).

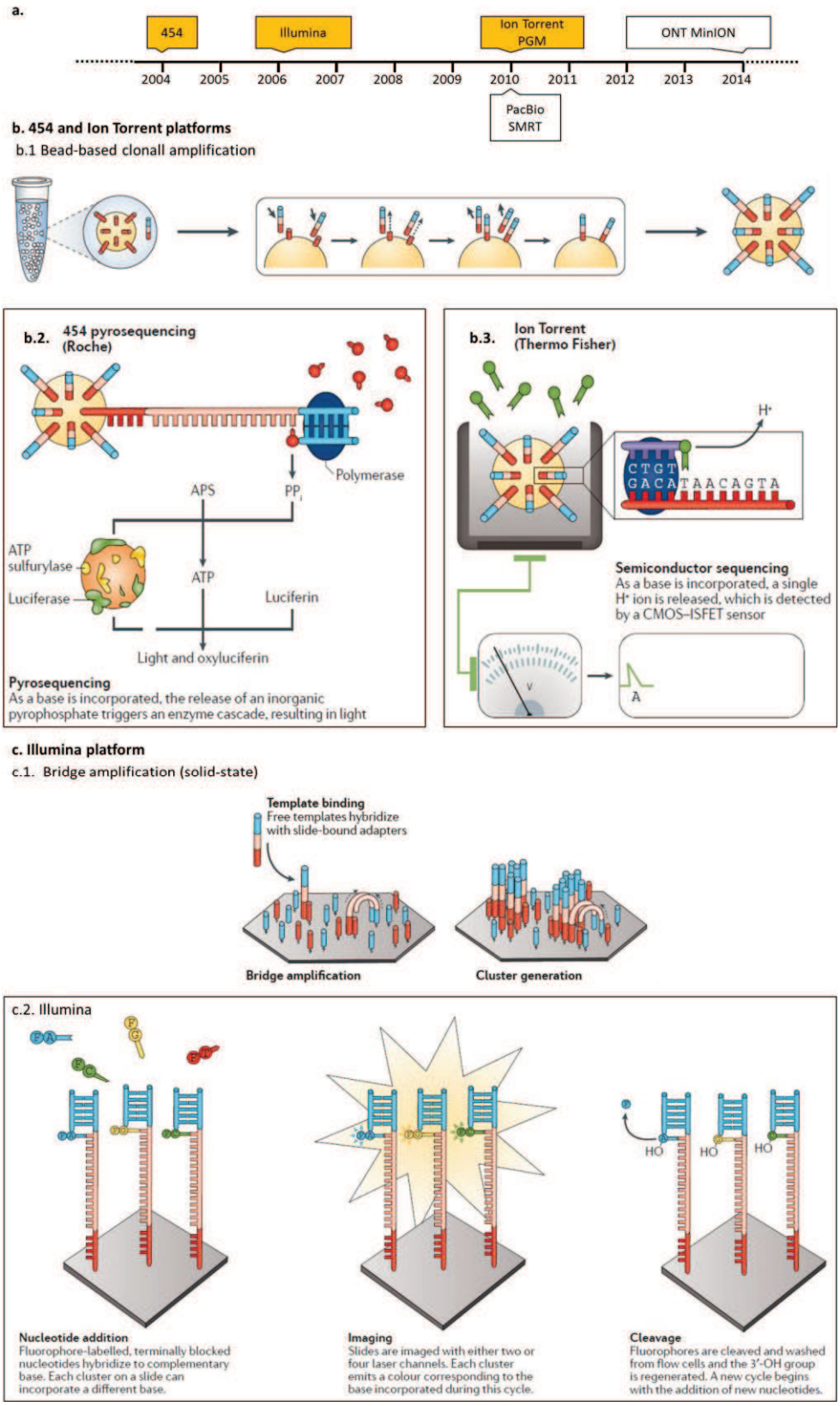


Figure 3. 2nd generation sequencing platforms and chemistries. a) Chronological timeline with the launching dates of the main platforms. In b) 454 and Ion Torrent sequencing technologies: b.1) bead-based clonal amplification, prior to b.2) 454 sequencing or b.3) Ion Torrent sequencing. In c) Illumina sequencing technology: c.1) solid-state clonal amplification (bridge amplification) prior to c.2) sequencing with Illumina and its sequencing strategy (Figure modified from Goodwin, McPherson, and McCombie 2016)

The three 2nd generation platforms reviewed here use sequencing by synthesis (SBS) approaches, so they rely on DNA polymerases. They differ mainly on the sequencing chemistry and the produced secondary signals.

- 454 pyrosequencing (Figure 3b.2) was the first 2nd generation platform that successfully achieved commercial introduction, and it was released on 2004 (Mardis, 2017). Once a polymerase incorporates a dNTP into a strand, it releases a pyrophosphate molecule that produces a bioluminescence signal through an enzymatic cascade. Differences on light signal intensity will indicate how many dNTPs are incorporated, which sometimes can lead to homopolymer errors (Goodwin et al., 2016).

- IonTorrent PGM (Figure 3b.3) was the first 2nd generation platform to sequence without optical sensing (Rothberg et al., 2011). Once a polymerase incorporates a nucleotide to a sequence, a proton (H^+) is liberated producing a measurable change in the pH. The pH change is detected through a sensor and it is proportional to the number of nucleotides incorporated. However, it is not perfectly linear and presents certain problems with homopolymers (Goodwin et al., 2016; Morey et al., 2013).

- In Illumina platforms (Figure 3c.2), each nucleotide is labeled with a base-specific cleavable fluorophore and then blocked, which allows incorporating one nucleotide per cycle. The posterior identification is achieved through total internal reflection fluorescence (TIRF) microscopy (Goodwin et al., 2016). Among the different Illumina platforms, the most commonly used in microbiota studies is MiSeq.

Main limitations of 2nd generation platforms in a microbiota project are: i) the short read length, which makes difficult the correct taxonomic classification up to species level using specific hypervariable regions of 16S rRNA (Schloss et al., 2016a); and ii) the need of a PCR-amplification step previous to sequencing, which can bias the original microbiota composition of an specific environment. 3rd generation sequencers emerged to overcome some of these limitations. Nowadays, the main technologies and sequencers in the market are: Single Molecule Real Time Sequencing from Pacific Biosciences (PacBio RS II platform) and Single-Molecule Nanopore sequencing from Oxford Nanopore Technologies (ONT: MinION, GridION and PromethION platforms) (Glenn, 2011; Goodwin et al., 2016; Heather and Chain, 2016; Mardis, 2017). The main aim of these platforms is sequencing single molecules at real-time, without short-length limitation and high-throughput, although lower than 2nd generation platforms.

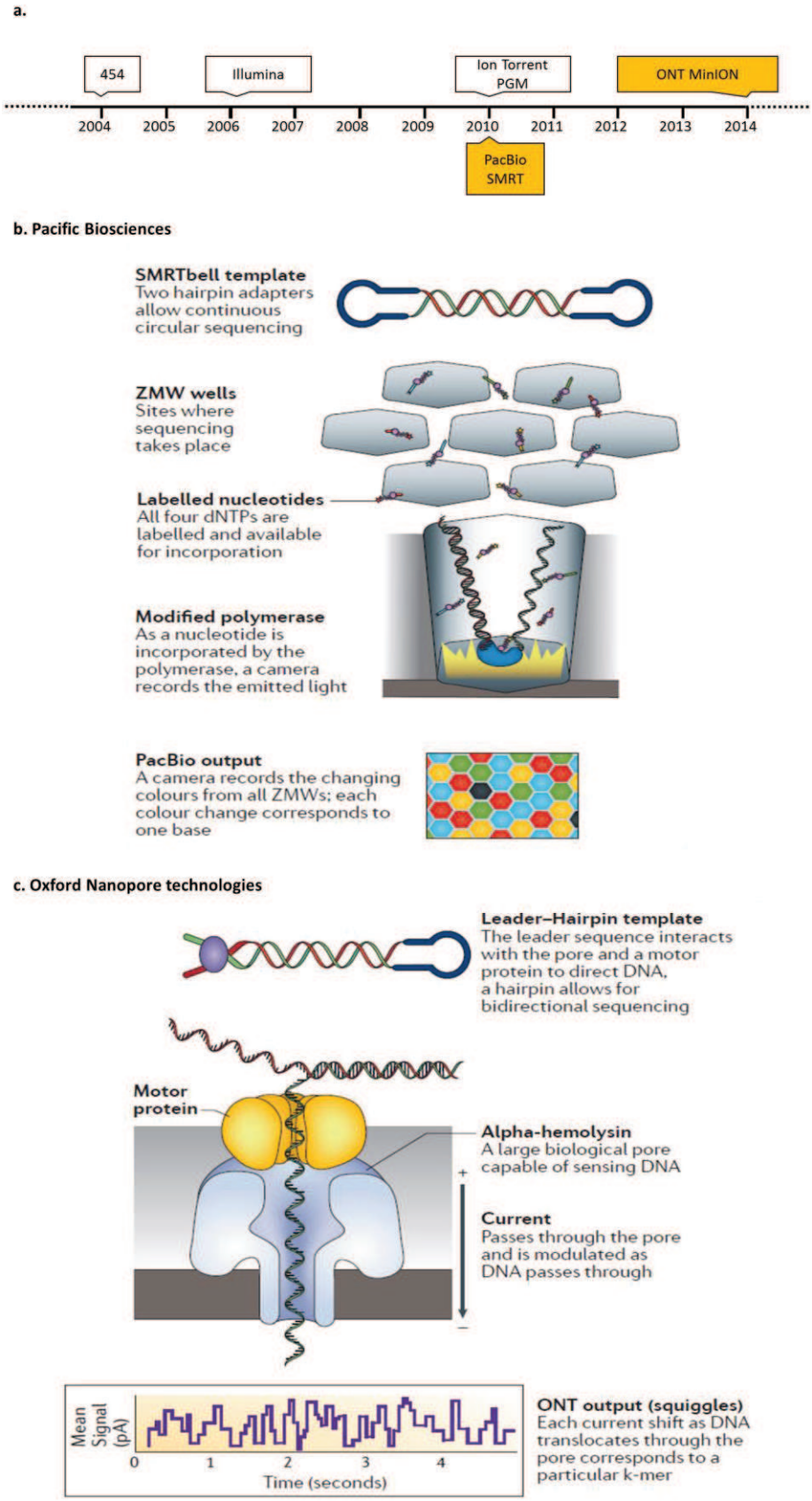


Figure 4. 3rd generation sequencing platforms and chemistries. a) Chronological timeline with the launching of the main platforms. In b) Pacific Biosciences and c) Oxford Nanopore Technologies sequencing. (Figure modified from Goodwin, McPherson, and McCombie 2016)

Sequencing technologies behind these two sequencers are very different: while PacBio continues using sequencing by synthesis, Oxford Nanopore Technology relies on direct-sequencing through nanopores. More detailed:

- Pacific Biosciences (PacBio, http://www.pacb.com/) platform is based on the single-molecule real-time (SMRT) sequencing approach (Eid et al., 2009). DNA molecules will be sequenced in SMRT-bell structure, which consists on a double-stranded region (DNA insert of interest) ligated with a single-stranded hairpin loop on both ends that provides a binding site for the primer. This structure enables multiple passes of the polymerase per DNA molecule, increasing the accuracy by creating a circular consensus sequence (CCS) (Travers et al., 2010). Sequencing reactions take place in picoliter wells called zero-mode waveguides (ZMW) (Levene et al., 2003) that have in their transparent bottom a fixed polymerase. By having a constant location of nucleotide incorporation, the optical system can focus on a single molecule and once a fluorophore is cleaved it diffuses away from the sensor. Thus, each dNTP incorporated on a single-molecule template emits a light that is continuously recorded and later basecalled (converted) to a DNA read (Goodwin et al., 2016).

- Oxford Nanopore Technologies (ONT, https://nanoporetech.com) launched in 2014 a small, USB-like, portable sequencer based on nanopore sequencing: MinION™ (Ip et al., 2015). Unlike other technologies, nanopore sequencing directly detects the DNA composition of a native single strand DNA molecule rather than a secondary signal (light, pH, colour, etc.) derived from the synthesis of a complementary strand (Goodwin et al., 2016). A DNA strand is sequenced as it passes through the nanopore while electrical current passes through the pore (Clarke et al., 2009). DNA molecules passing through the nanopore produce blockages in electrical conductivity that can be used to discriminate individual nucleotides (Olasagasti et al., 2010; Wang et al., 2015). Each subset of nucleotide bases (k-mer) translocating through the nanopore produces a unique shift of voltage called *squiggle*, which is the raw signal that will be basecalled to a DNA read. DNA template used in ONT platforms presents a leader sequence that interacts with the nanopore to direct the DNA through it, and a hairpin that joins the template and complement DNA strands when performing bidirectional sequencing with 2D chemistry (Goodwin et al., 2016)(Goodwin et al., 2016). Since May 2017, 2D sequencing chemistry has been discontinued and $1D^2$ is replacing it. This new sequencing strategy does not require a hairpin anymore, avoiding secondary structure problems and improving accuracy and throughput (Brown, 2017).

Several reviews aimed to compare different sequencing technologies available on the market from different perspectives (van Dijk et al., 2014; Glenn, 2011; Goodwin et al., 2016; Loman et al., 2012; Morey et al., 2013; Reuter et al., 2015; Vincent et al., 2016). In general, 3rd generation sequencers allow obtaining longer reads but they present higher error rates when compared to 2nd generation sequencers. Some other divergences can be found in Table 1.

Table 1. Main sequencing platforms used in microbiota studies and their main characteristics.

Platform	Type of instrument	Amplification	Sequencing chemistry	Secondary signal	Read length (bp)
Ion Torrent PGM	Benchtop	emPCR	Synthesis	H^+ (pH change)	200-400
Illumina MiSeq	Benchtop	solid-state	Synthesis	Light	150-300
454 Roche	Benchtop	emPCR	Synthesis	Light	400-650
PacBio RS	High-end machine	-	Synthesis	Light	up to 60,000
MinION	Portable	-	-	-	ultra-long reads*

*Nanopore sequencing read length is only limited by high-molecular weight DNA extraction, the longest read obtained at date 16/05/2017 is close to 1 Mb by Josh Quick & Nick Loman (Quick and Loman, 2017).

In this dissertation, we work with data from Ion Torrent PGM™ to assess skin microbiota composition (Chapter 3.2 and 3.3). Moreover, we assess the potential of nanopore sequencing at technical level for microbiota studies using MinION™ by ONT (chapter 3.4).

1.1.3. Bioinformatics analysis

The starting point of a microbiota bioinformatics analysis is the raw data obtained from the sequencing platform. In this dissertation, we have mainly worked with output from Ion Torrent PGM™ (as a 2nd generation sequencer). Moreover, we have also tested nanopore sequencing using MinION™ (as a 3rd generation sequencer). For more details on nanopore data processing see Chapter 3.4.

In this section, we will cover the main analytical steps performed in a microbiota study illustrating them with the bioinformatics tools and the metrics used in this thesis.

Usually microbiota studies rely on pipelines that gather together many different bioinformatics tools. In this thesis, we mainly used Quantitative Insights Into Microbial Ecology (QIIME v1.9.1) (Caporaso et al., 2010a) and in Chapter 3.3 also VSEARCH in some steps (Rognes et al., 2016a). Another equivalent and widely used pipeline is *mothur* although we have not used it (Schloss et al., 2009).

The analytical workflow is divided in three main steps: (i) pre-processing of the sequences, (ii) operational taxonomic unit (OTU) picking and OTU table building and (iii) downstream analysis (Figure 5).

This taxonomic-based analysis can be completed with tools predicting the functional potential of the microbial community. We have used Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) (Langille et al., 2013), but other tools such as Tax4fun (Aßhauer et al., 2015) and PanFP (Jun et al., 2015) can be used for the same aim. Finally, Linear Discriminant Analysis Effect Size (LEfSe) (Segata et al., 2011)) software is usually used to assess statistically which taxa is under or overrepresented between different groups of samples.

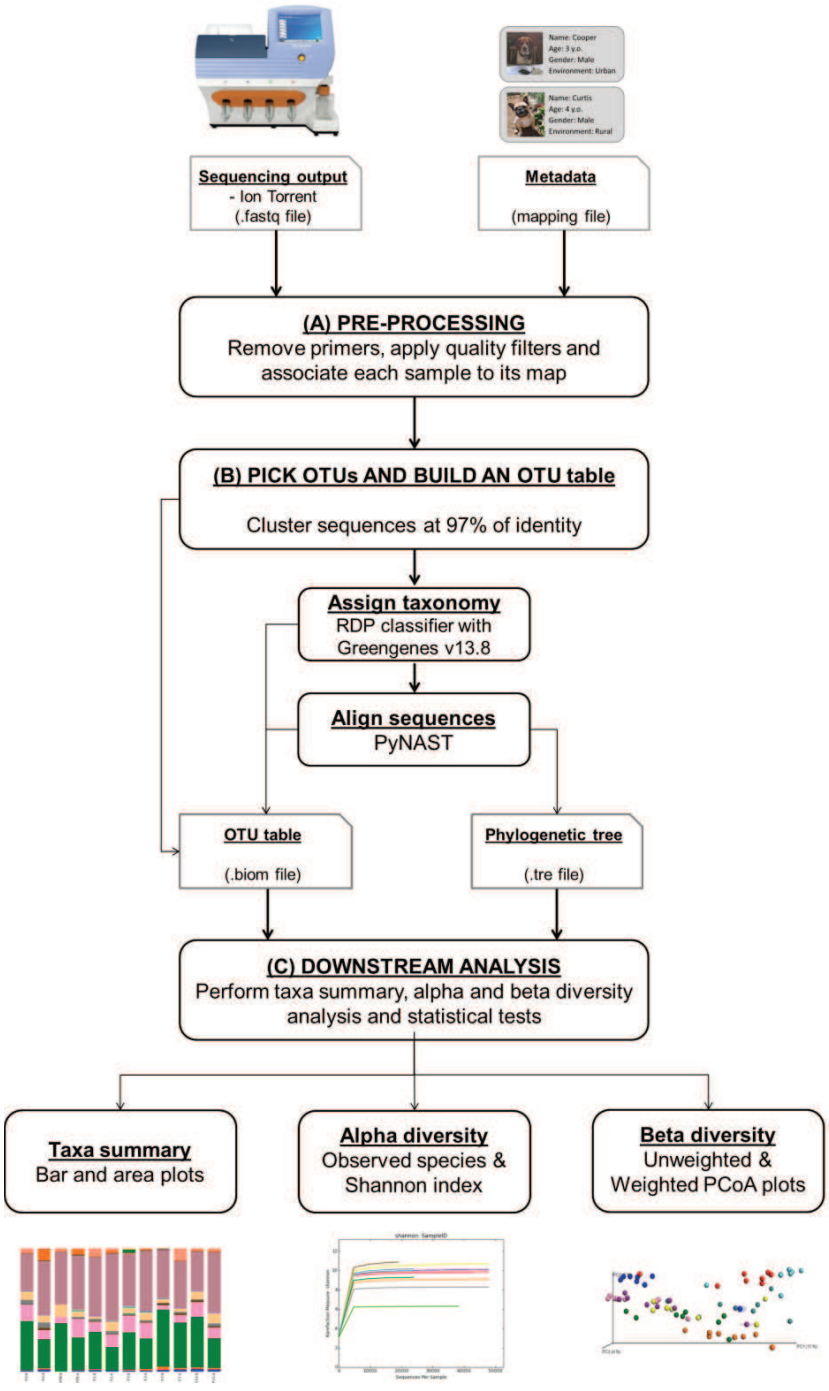


Figure 5. Overview of the bioinformatics workflow used in this thesis to analyze the microbiota.

1.1.3.1. Pre-processing step

Each Ion Torrent PGM™ run contains a pool of several barcoded samples, with an average length of 350 bp for primers 27F-R338R targeting V1-V2 regions of 16S rRNA gene. Raw data from Ion Torrent PGM™ is given demultiplexed, which means separated by barcodes, thus we have a unique file per sample included in the run.

The pre-processing steps are performed using Quantitative Insights Into Microbial Ecology (QIIME v1.9.1) software (Caporaso et al., 2010a). We need two input files, one containing the DNA sequences and the other its associated metadata. Whereas the DNA sequences are the output of Ion Torrent, the mapping file with all the metadata needs to be created by the researcher. This should contain all the technical information and metadata about the samples: sample ID, forward and reverse primers, treatment, description, age, or any other relevant information.

The following step is merging the sequencing file with the mapping file for each sample. Reading the information on the mapping file, the script trims the primers and labels the sequences with their sample ID.

Moreover, this script performs the initial quality control of the sequences discarding those ones shorter than 300 bp, with quality phred scores below 25, with mismatches on the primer and other default parameters (script: split_libraries.py).

1.1.3.2. Operational taxonomic unit (OTU) picking and OTU table building

The basic unit when working with 16S data is the Operational Taxonomic Unit (OTU), which is constituted by a group of sequences that share a certain percentage of similarity. Usually sequences with 97% of similarity are clustered together constituting an OTU, but other percentages have been proposed.

There are two main strategies to pick OTUs: *de novo*, which clusters the sequences among them considering a specific percentage of similarity (97% of similarity in our studies); and *closed*, which clusters the sequences against a reference database (previously clustered at the similarity threshold chosen). The first approach allows all the sequences to be clustered, although it is usually computationally expensive. The second approach is faster, but sequences that are not in the database are excluded for further analysis. QIIME preferred option is a combined one, which is called *open reference*: It first clusters the sequences against a database and the ones that do not match are clustered *de novo* among them (Rideout et al., 2014). We have used both open and closed reference approaches in Chapter 3.2, see further details in material and methods section (script: pick_open_references_otus.py). When working with VSEARCH (Rognes et al., 2016b) we have used the *de novo* strategy because it runs faster than its homologous in QIIME. See further details in Chapter 3.3, in the material and methods section.

Chimeras are amplification artifacts really common in microbiota surveys (Ashelford et al., 2005), due to the large amount of PCR targets within a single reaction. Thus, an aborted

extension can work as a primer for another target forming a chimeric artifact. If chimeras are not removed from the data, they are incorrectly identified as novel taxa and that leads to an inflated microbial diversity (Ashelford et al., 2005; Haas et al., 2011). Chimeric artifacts can be minimized by the correct choice of 16S rRNA primers and an adjusted number of PCR cycles among others (Schloss et al., 2011). Moreover, once you have already sequenced the data, several bioinformatics tools have been designed to identify chimeric sequences and remove them. Here we have used ChimeraSlayer (Haas et al., 2011) after OTU picking on QIIME and UCHIME (Edgar et al., 2011) in VSEARCH pipeline.

Once we have picked OTUs and removed the invalid and low-quality sequences from the analysis, we assign the taxonomy to our sequences as well as align them to create a phylogenetic tree. Taxonomic assignment of the OTUs is performed using the RDP Classifier (Wang et al., 2007) against Greengenes v13.8 database (DeSantis et al., 2006; McDonald et al., 2012). Alignment of sequences is performed using PyNast (Caporaso et al., 2010b) as default in QIIME pipeline.

1.1.3.3. Downstream analysis

1.1.3.3.1. Taxa summary

The simplest way to describe a microbial community is with a list of the bacteria together with their abundances, which are usually plotted in bar graphs.

V1 and V2 hypervariable regions of 16S allow classification of the sequences mostly up to family level, some of them even up to the genus level. Taxa analysis can be summarized at a specific taxonomic level or can be global.

1.1.3.3.2. Microbial ecology measures: alpha and beta diversity

Microbial diversity is assessed using metrics that can take into account: OTU counts, relative abundances and/or phylogenetic information. When bacterial community diversity is assessed within a community, it is called alpha diversity. On the other hand, when bacterial diversity is compared and assessed among different communities, it is called beta diversity.

Many different ecological indexes have been described for assessing both alpha and beta diversity, which can be classified depending on which properties they take into account: (1) qualitative, when only presence/absence data is considered; or (2) quantitative, when also relative abundances are taken into account; they can also be (3) species-based, when different taxa are treated as equally related; or (4) divergence-based, when phylogenetic distances among each pair of taxa are considered (Lozupone and Knight, 2008, 2007).

In this dissertation we choose two species-based methods to assess alpha diversity: Observed species (qualitative) and Shannon index (quantitative) (Shannon and Weaver, 1949). On the other hand, for beta diversity analysis we choose two divergence-based methods: Unweighted

UniFrac (qualitative) (Lozupone and Knight, 2005) and Weighted UniFrac (quantitative) (Lozupone et al., 2011).

Alpha diversity

Alpha diversity assesses diversity within a community. As a qualitative index we used Observed Species metric, which assesses richness of the sample by simply counting the unique OTUs found. As a quantitative index we used Shannon index (Shannon and Weaver, 1949), which assesses the evenness of a sample. Shannon index value gets larger as the number of species increases and as the distribution of species becomes even.

Alpha diversity is usually highly dependent on sequencing depth. For that reason, it is common to use rarefaction curves to show the cumulative number of species as a function of sampling depth (Figure 6a and 6b)(Hughes and Hellmann, 2005). Consequently alpha diversity should always be given at a specific sequencing depth. Alpha diversity values can be represented in tables, or in plots similar to those shown in Figure 6.

Finally, as an example to better understand the concepts, we should now have a look at Figure 6. We have a study cohort of nine samples that belong to two groups (A in red and B in blue). In Figure 6a we can see the rarefaction curves of all the samples included in one study, and despite presenting uneven sequencing depths, all reach the sequencing depth of 8,300 reads per sample. So, we decide to work at this sequencing depth and now we group the samples from the same group together and create a second rarefaction plot (Figure 6b). Finally, we can visually compare groups using boxplots (Figure 6c) and even assess statistical significance using non-parametric tests, e.g. Monte Carlo permutation test.

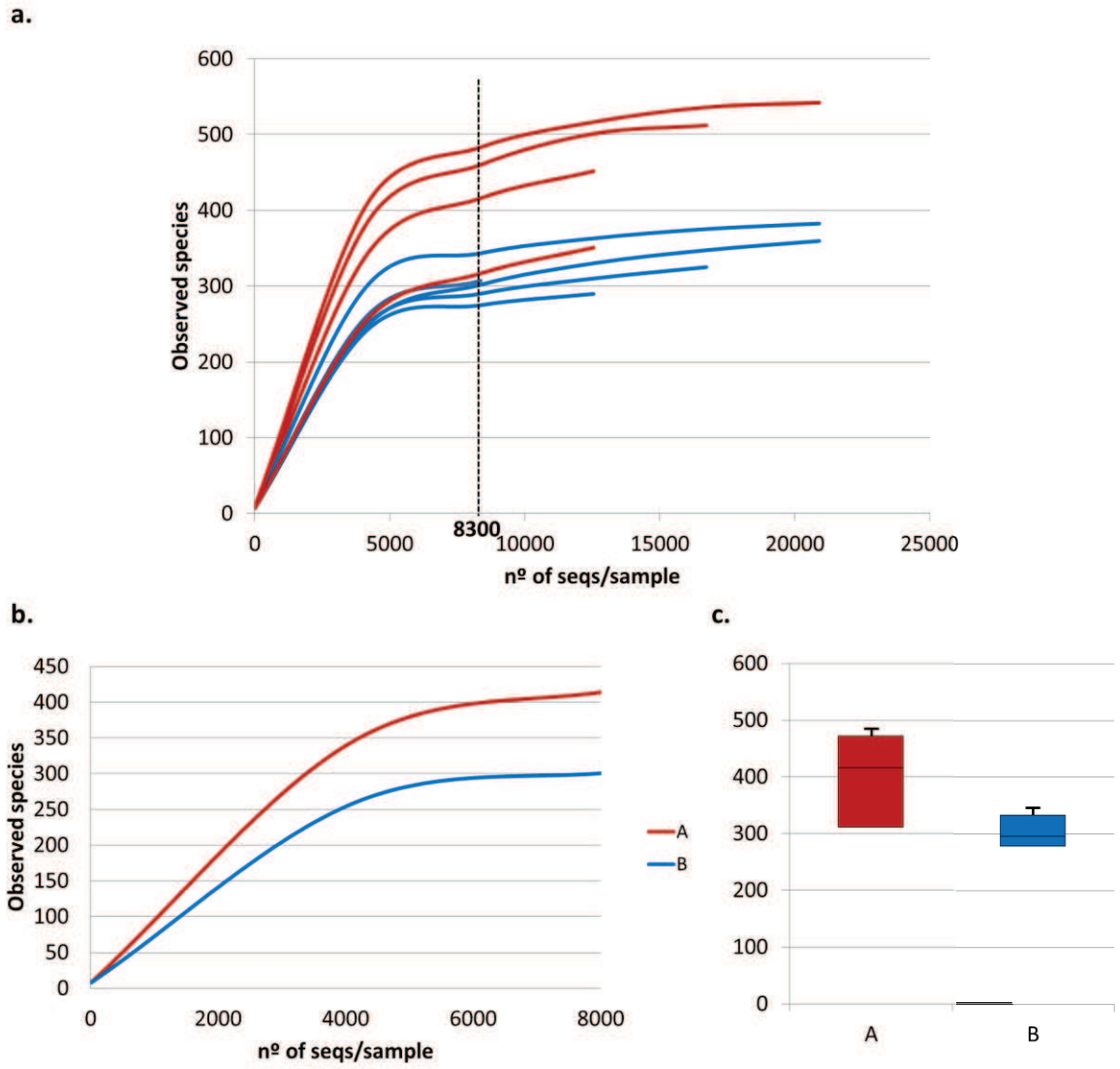


Figure 6. Alpha diversity plots. Red and blue represent two different biological categories. a) Alpha diversity rarefaction plot of all the samples at their own sequencing depth; all the samples reach 8,300 sequences per sample so further analyses are performed at this depth. b) Alpha diversity rarefaction plot of the two groups. c) Boxplots representing alpha diversity values distribution within each group.

Beta diversity

Beta diversity assesses differences among bacterial communities computing distance matrices that can be combined with multivariate statistical techniques, such as principal coordinate analysis (PCoA) and hierarchical clustering. These statistical analyses allow plotting the results graphically as well as detecting patterns and clusters of samples.

Because sample depth can affect beta diversity results and to prove its robustness, it is always recommended to apply a jackknifing protocol. Jackknifing consists on subsampling randomly and repeatedly the initial sample using a specific sequencing depth (Zahl, 1977).

In this dissertation, we choose UniFrac metrics that takes into account the phylogenetic distance among bacteria. Thus, two communities sharing similar phylogenetic lineages will be more similar, despite not sharing the same exact species.

UniFrac distance matrix (Lozupone and Knight, 2005) measures the phylogenetic distance between two collections of sequences as the fraction of branch length in a phylogenetic tree that leads to descendants of one sample or the other, but not both. Besides phylogenetic information, when these metrics consider only the presence and absence of the different bacteria found in the community it is called Unweighted UniFrac, whereas when they take into account relative abundances it is called Weighted UniFrac (Figure 7) (Lozupone et al., 2007).

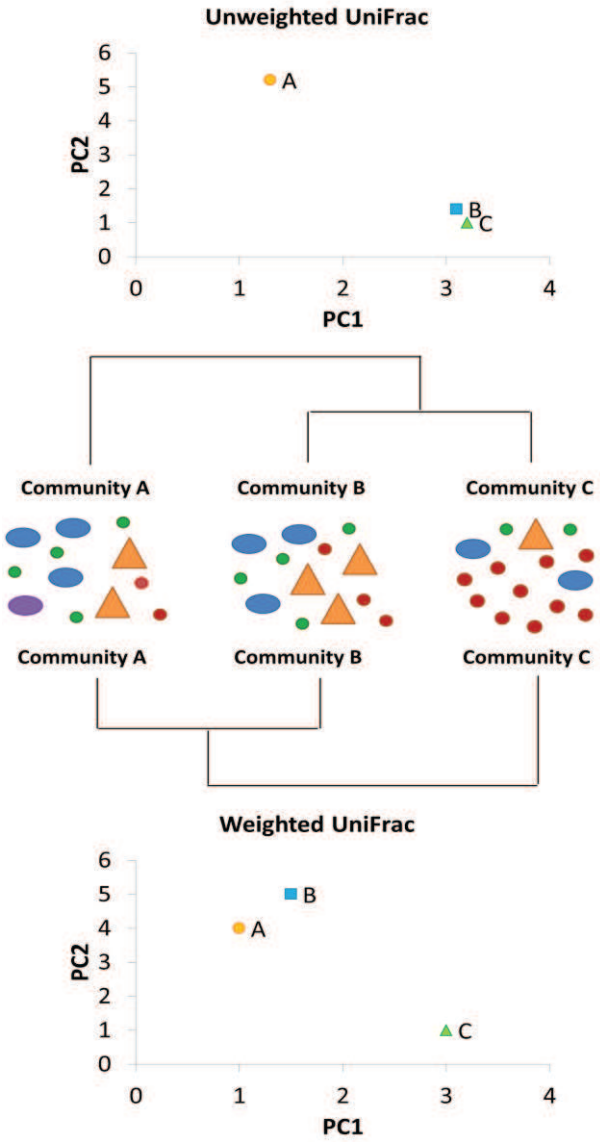


Figure 7. Unweighted and Weighted UniFrac metrics plots. Beta diversity represents distances among the bacterial communities using either UPGMA trees or PCoA plots. Bacterial community B has the same bacterial species as bacterial community C, A presents one more species. On the other hand A and B present similar relative abundances of the main bacteria, whereas C presents a predominant bacterium. a) Unweighted UniFrac, which only considers composition and phylogeny, B and C will be more similar and b) Weighted UniFrac, which also considers relative abundances, A and B will be more similar.

To assess the extent and significance of the clustering of the samples regarding a specific grouping variable, non-parametric statistical tests such as ANOSIM (analysis of similarities)(Clarke, 1993) and adonis are commonly used. ANOSIM test give an R value that when it is closer to +1 means that the dissimilarity among groups is high, whereas when it is closer to 0 means that there is no grouping based in that variable. Adonis test computes an R^2 value (effect size) that shows the percentage of variation explained by the grouping category. Both tests also output a p-value to determine the statistical significance.

1.1.3.3.3. Functional prediction using PICRUSt

Profiling 16S rRNA gene is a widely used method to provide insights into microbial community, although a more complete view would include information about community's functional capabilities. This can be achieved with metagenomics approaches, but when only 16S data is available it is also possible by predictive tools such as PICRUSt (Langille et al., 2013).

PICRUSt needs an OTU table obtained through a closed reference approach, since the prediction is based on a reference database. The first step is normalizing the OTU table by dividing the OTU counts by predicted marker gene copy number. The second step is inferring metagenomes, multiplying the inferred number of OTUs per sample by the predicted gene content. The final output is a matrix of the gene count with samples as columns and functional pathways as rows (Figure 8).

To assess the reliability of the software, Langille and colleagues used this tool with the Human Microbiome Project dataset (Human Microbiome Project Consortium, 2012) obtaining sufficiently accurate results, even for skin samples (Langille et al., 2013).

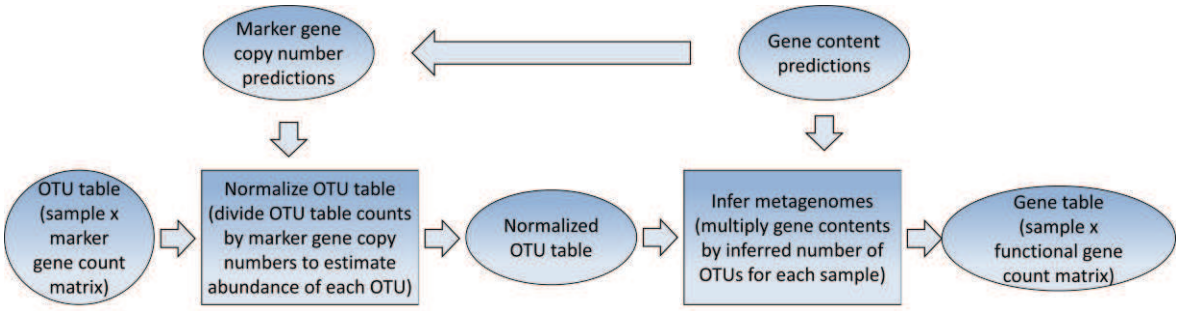


Figure 8. The PICRUSt workflow. The metagenome inference workflow takes a closed-reference OTU table, as well as the copy number of the marker gene in each OTU and the gene content of each OTU and outputs a metagenome table (i.e., counts of functional pathways per sample) (Figure adapted from Langille et al. 2013).

1.1.3.3.4. Detecting differentially distributed features using LEfSe

In both taxonomic and predicted functions analyses, we usually want to assess which features are statistically different among groups. To assess that, here in this dissertation we choose to use Linear Discriminant Analysis (LDA) Effect Size (LEfSe) algorithm (Segata et al., 2011).

LEfSe algorithm first uses the non-parametric factorial Kruskal-Wallis (KW) sum-rank test (Kruskal and Wallis, 1952) to identify features that are statistically different among biological classes. Biological significance is subsequently assessed using a set of pairwise tests among subclasses (when provided). As a last step, LEfSe uses Linear Discriminant Analysis (Fisher, 1936) to estimate the effect size of each differentially abundant feature.

1.1.3.3.5. Building ecological networks using CoNet

The microbiome is a complex ecosystem where microbes compete and cooperate. These microbial interactions can support health or promote disease. Thus identifying and characterizing them will allow us to better understand microbial dynamics.

We used CoNet software (Faust and Raes, 2016) in Chapter 3.3, which is implemented as an application in Cytoscape (Shannon et al., 2003), to obtain an overview of the microbial dynamics underlying canine skin microbiota.

The input file needed is an OTU table that will be first pre-processed normalizing the data and removing the taxa with too many zero values, which can lead to spurious interactions. In the next step, you can select the methods used to compute the networks that can be correlations, similarities and/or dissimilarities metrics. Most of these association measures allow assigning a positive or negative sign to a predicted relationship, which will be plotted on the network using a green or red edge respectively. The recommended approach is to choose multiple metrics and then combine networks, and if in any edge measures disagree on the sign, it is discarded (Faust and Raes, 2016).

1.2. Overview of the skin

Skin is an anatomical barrier that separates the animal from the outer environment. However, it is not only an anatomical barrier but also a living barrier, covered by thousands of microorganisms that cross-talk with the host cells as well as the immune system, maintaining the homeostasis and equilibrium (Naik et al., 2012a). Besides being a physical and immunological barrier, skin also percepts different kind of stimuli; facilitates motion and gives shape; produces adnexa (such as sweat and sebaceous glands, claws or hairs); regulates body temperature; stores fat, proteins, vitamins or water among others; may indicate the general health status or sexual identity; gives pigmentation; has antimicrobial properties; produces vitamin D; and secretes and excretes (Miller et al., 2013).

The object of study in this dissertation is the skin, so we will delve into its anatomy and physiology as well as into its cellular composition and histology. Finally we will provide some background on skin innate immune functions and specifically Toll-like Receptors as the first sensors of microbes.

1.2.1. Anatomy and physiology of the canine skin

On dogs, a dense hair coat, also named fur, covers the skin. The hair coat insulates the skin thermally and sense stimuli, as well as protects the skin against chemical, physical and microbial damage (Miller et al., 2013). Breeds have been classified regarding their hair coat length in four main groups: short, wire, long, and curly. These phenotypes and some combinations and variations have been associated to alleles located in three genes: FGF5, RSPO2 and KRT71 (Figure 9) (Cadieu et al., 2009).

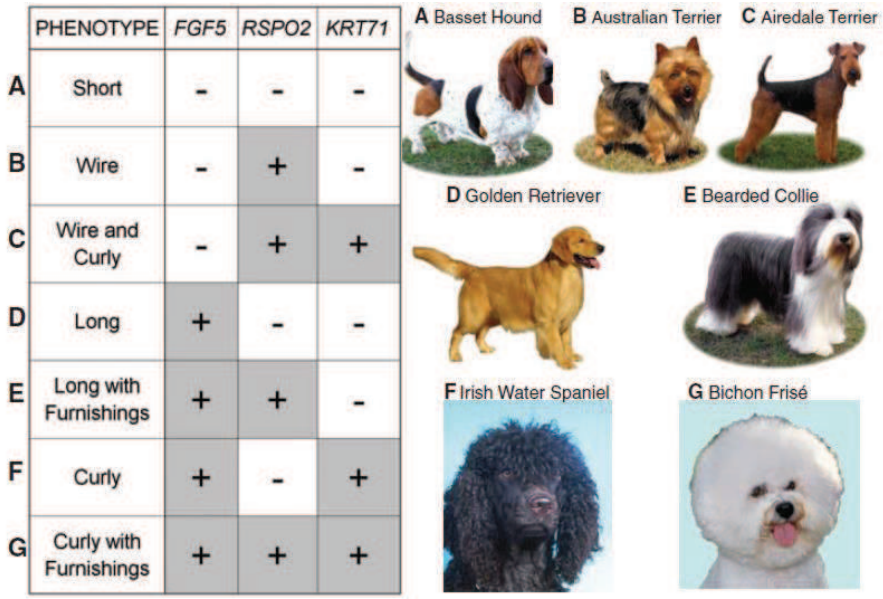


Figure 9. Hair coat phenotypes on dogs. Different combinations of alleles located on FGF5, RSPO2 and KRT71 genes produce most of the hair coat phenotypes on dogs (Figure from Cadieu et al. 2009)

Both the skin and the hair coat vary dynamically within individual canine species in quantity and quality depending on many factors such as the age, the sex, the skin site, the breed and even the individual (Miller et al., 2013). Daily growth rate of the hair shaft is both season dependent –being greater during the colder season– and site dependent (Al-Bagdadi, 2013).

In general, skin thickness decreases dorsally to ventrally on the trunk and proximally to distally on the limbs. Thus, the hair coat is usually thickest over the dorsolateral regions of the body and thinnest ventrally, on the concave part of the pinnae and on the undersurface of the tail (Miller et al., 2013; Scott et al., 1995).

The skin of haired mammals is usually acidic, however normal pH values on dog skin have been reported to range from 4.84 to 9.95 (Miller et al., 2013). A dynamic study assessing pH on healthy dogs (Ruedisueli et al., 1998) reported that: 1) pH values differ day to day and among skin sites; 2) males and spayed females had significantly higher pH values on all skin sites than females and intact females, respectively; 3) black Labrador retrievers had significantly higher pH values than yellow ones; and 4) some breeds had significantly different pH values when compared to other breeds.

1.2.2. Histology of the canine skin

Three main cellular layers constitute the dog skin: the epidermis, the dermis and the hypodermis (or subcutis). Moreover, the skin presents invaginations and appendages that go from the deep dermis to the skin surface, being the main ones sweat and sebaceous glands and hair follicles (Miller et al., 2013).

1.2.2.1. The epidermis

The epidermis is mostly constituted by keratinocytes (~85% of the epidermal cells), followed by Langerhans cells (~3-8%), melanocytes (~5%) and Merkel cells (~2%) (Figure 10) (Miller et al., 2013). Keratinocytes and Langerhans cells are members of the skin immune system, being the keratinocytes the first line of defense against invading pathogens (Suter et al., 2009) whereas Langerhans cells are antigen-presenting cells (White and Yager, 1995). Melanocytes are the pigment (melanin) producers and Merkel cells are mechanoreceptors involved in light touch sensation (Miller et al., 2013).

The epidermis of the dog is usually subdivided into four to five cell layers anchored to the dermis: the *stratum basale*, the *stratum spinosum*, the *stratum granulosum*, the *stratum lucidum* and the *stratum corneum* (Figure 10). Each epidermic layer is mainly representing different maturation stages of the keratinocytes (Miller et al., 2013).

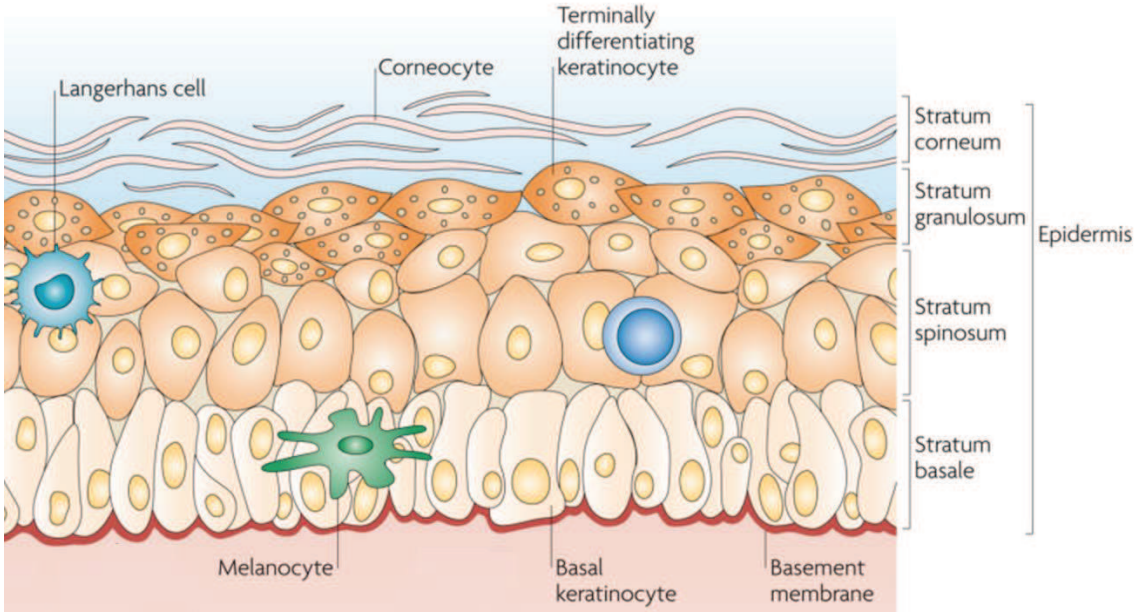


Figure 10. Skin anatomy and main cells. The epidermis is constituted by keratinocytes at different stages and they form layers called the stratum basale, the stratum spinosum, the stratum granulosum and the outermost layer, the stratum corneum. Specialized cells in the epidermis include melanocytes, which produce pigment (melanin), and Langerhans cells (Figure modified from Nestle et al. 2009).

Stratum basale is made of a single layer of columnar cells, mostly of them dividing and non-dividing keratinocytes. As basal keratinocytes reproduce and mature, they move towards the

outer layer of skin, initially forming the *stratum spinosum*. In this layer, keratinocytes are connected one to other through intercellular bridges. Continuing their transition to the surface the keratinocytes of the *stratum granulosum* flatten, their nuclei are reduced and their cytoplasm appears granular (Miller et al., 2013). *Stratum lucidum* is only present in footpads and nasal planum, and it is a fully keratinized compact layer of dead cells (Schwarz et al., 1979). Finally, the *stratum corneum* is made of non-viable terminally differentiated keratinocytes known as corneocytes. These cells (bricks) are embedded in a lipid-rich matrix (mortar) which helps to hold them together (Leigh and Watt, 1994), as well as retain the water.

1.2.2.2. Appendages and invaginations of the dog skin

In dog skin, the main cutaneous invaginations and appendages are sebaceous glands and sweat glands (apocrine and eccrine), and the hair follicle. These different cutaneous structures with their secretions allow creating different microenvironments, which likely influence the microbiota they harbor as suggested in human skin (Figure 11) (Grice and Segre 2011).

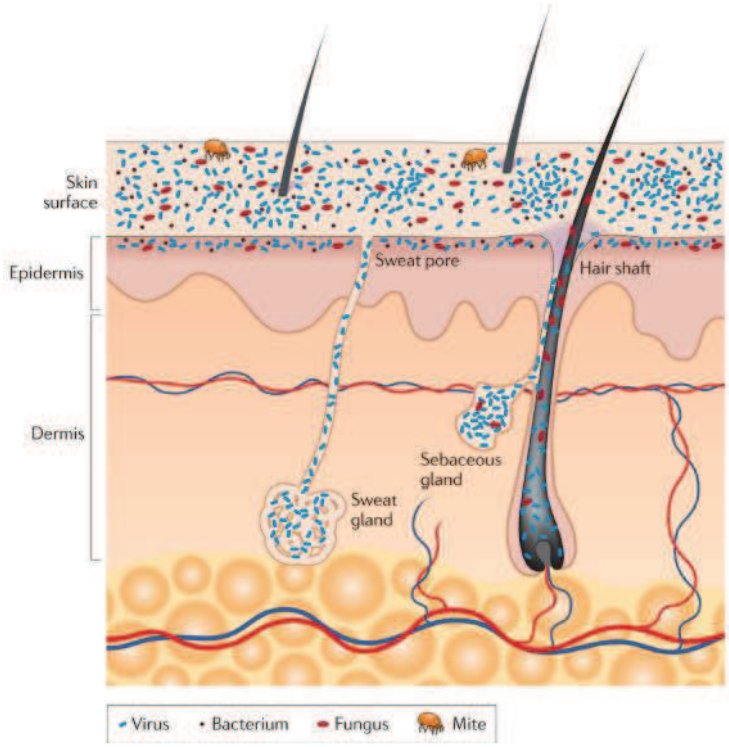


Figure 11. Skin structure: main layers, appendages and associated microbiota. (Figure from Grice and Segre 2011).

Sebaceous glands are simple or branched alveolar glands distributed throughout all haired skin in mammals with association to hair follicles. However they are absent in the dog's footpads and nasal planum. They are larger and more abundant near mucocutaneous junctions, in the interdigital spaces, on the dorsal part of the neck and rump, on the chin and on the dorsal tail (Miller et al., 2013). They secrete sebum that is an oily substance composed mainly by triglycerides and wax esters (Clarys and Barel, 1995), which bacteria metabolize to convert them to free fatty acids. The sebum provides a hydrophobic coating of the hair and skin that

protects from overwetting and may have a role on heat insulation (Smith and Thiboutot, 2008).

Sweat glands can be *epitrichal (apocrine)* or *atrichial (eccrine)* and, unlike humans, they do not play a significant role in thermoregulation (Cotton et al., 1975). *Epitrichal* sweat glands are distributed throughout all haired skin in exception of footpads or nasal planum. *Epitrichal* sweat has probably pheromonal and antimicrobial properties (Miller et al., 2013). *Atrichial* sweat glands release a watery secretion and are only found in the nasal planum and footpads placed deeply in the dermis and subcutis (Scott et al., 1995).

The hair follicles of the adult dog are complex and consist of bundles of hairs sharing a common opening on the skin surface. Usually a primary larger hair is surrounded by few secondary and thinner hairs forming a hair follicle complex. Each hair has its own sebaceous gland, but only the primary hairs are associated with an *epitrichal* sweat gland and an arrector pili muscle. The contraction of this muscle empties the contents of the glands on the skin surface (Al-Bagdadi, 2013). There are breed differences regarding both the number of hairs using a same hair follicle orifice and the density of hair follicles in skin surface.

1.2.3. Skin immunity

In skin we can find complex immunological processes of both innate and adaptive immune system (Bangert et al., 2011; Pasparakis et al., 2014). Keratinocytes are the first active participant in the skin immune response, but also Langerhans cells (Nestle et al., 2009) and even mast cells seem to have a role in skin immunity (Kumar and Sharma, 2010).

The main functions of the innate immune system are detecting pathogens and presenting their antigens to the adaptive immune system, as well as differentiating self *vs* non-self cells and microorganisms. All microorganisms –commensals and pathogens– have structures with conserved molecular patterns called microbial-associated molecular patterns (MAMPs), when these ones are specific from a pathogen they are called pathogen-associated molecular patterns (PAMPs). PAMPs are present in bacteria, fungi, viruses and other parasites and can be nucleic acids or structural components of the membrane or the envelope. Sensors of pathogens of the innate immune system are called PAMPs Recognition receptors (PRRs) and major families include Toll-like Receptors (TLRs), RIG-I like Receptors (RLRs) and Nod-like Receptors (NLRs) (Kawai and Akira, 2009).

The innate immunity of a healthy individual is capable to defend against pathogens as well as not reacting against its own cells. However, immunological dysregulations occur and create inflammation. On one hand, innate immune pathways can be improperly activated by host cells or endogenous DNA, causing an autoimmune response (Beutler, 2009; Fischer and Ehlers, 2008). On the other hand, a defect on innate immunity can make someone more susceptible to infections or even can transform a normal infection to a chronic one (Kumar et al., 2013) (Figure 12).

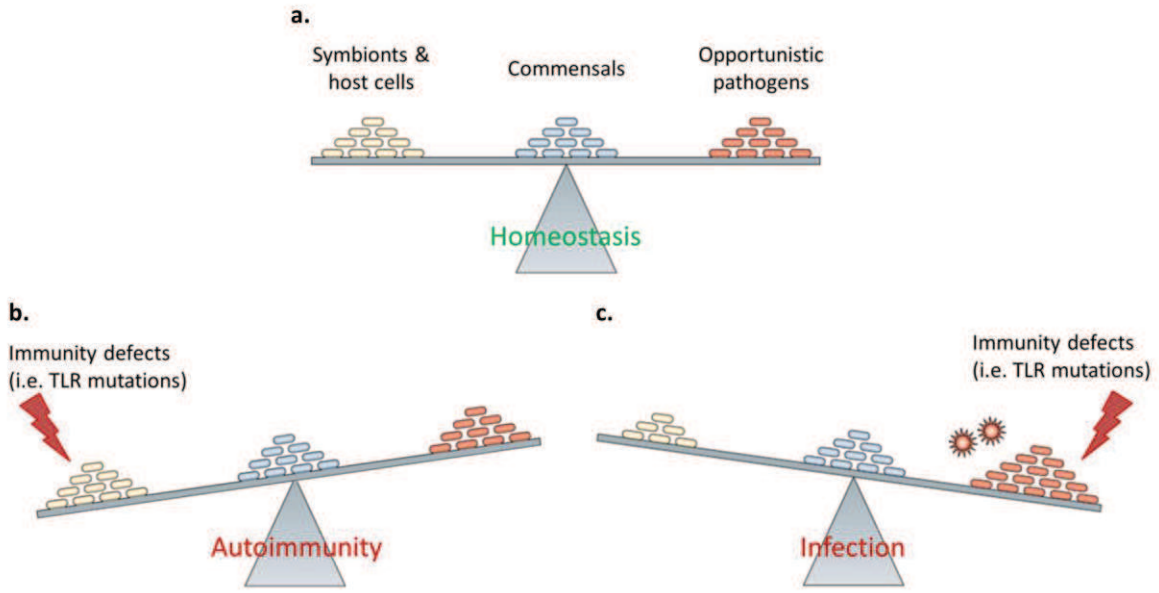


Figure 12. Skin in homeostasis and inflammation states. a) In a healthy individual, all the microorganisms are in equilibrium and pathogens are controlled with the innate immune system. b) When immunity defects exist or occur, innate immune system can detect own cells and microbiota as pathogens and overreact against them producing an autoimmunity reaction. c) When immunity defects exist or occur, pathogens can easily invade and colonize the skin, innate immune system fails to control the overgrowth and that produces an infection.

This dissertation is focused on microbiota analyses, not innate immunity; however they are two sides of the same coin. Therefore, we also describe innate immunity of dog skin by screening polymorphisms on Toll-like Receptor genes and here we present a little overview of these innate immune receptors.

1.2.3.1. Toll-like Receptors (TLRs)

Toll-like receptors are considered to be the first sensors of microbes, through their microbial associated molecular patterns (MAMPs) (Werling and Jungi, 2003). The pool of microbes they have to sense is huge and comes both from the environment and from the own microbiota. Despite the high variable microbial world, some molecular structures (MAMPs and PAMPs) are highly conserved across them and are sensed through TLRs (Beutler, 2004; Kumar et al., 2009) (Figure 13).

TLRs are transmembrane proteins constituted by leucine-rich repeat (LRR) domains, a unique intramembrane domain and a Toll/Interleukin-1 receptor (TIR) domain. MAMPs and PAMPs are sensed through the LRR domain, and signals are transduced through the TIR domain, which is always located in the cytoplasm, in order to activate innate immunity response. Several crystallography studies elucidated TLRs structure: TLRs act in dimers and create structures together with their ligands (Gay and Gangloff, 2007; Kang and Lee, 2011; O'Neill et al., 2013).

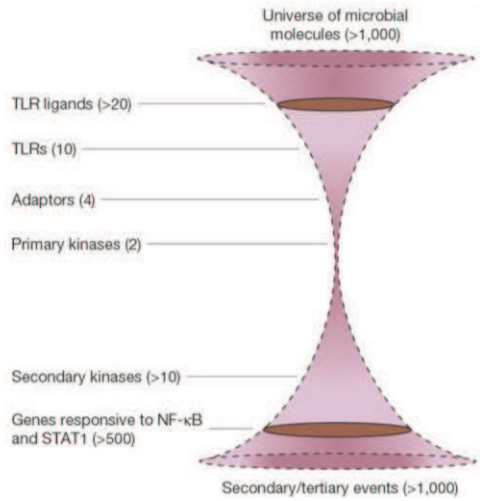


Figure 13. The hourglass shape of innate immune response. Ten Toll-like receptors (TLRs), four TIR adaptors and two protein kinases are required for most microbial perception (Figure excerpted from Beutler 2004).

Up to date, 10 functional TLRs have been described in humans and 12 in mice (Kawai and Akira, 2009). Public repositories have sequences for 10 TLRs in dogs.

TLRs have been classically classified depending on their cellular localization and their ligands: extracellular or intracellular TLRs. TLR1, 2, 4, 5, 6 and 10 are located in the cellular membrane and usually recognize components of the microorganism cell wall or membrane, whereas TLR 3, 7, 8 and 9 are usually located in intracellular vesicles and recognize mainly nucleic acids (Kawai and Akira, 2010). Once TLRs detect PAMPs, they send a signal trough an adapter molecule (MyD88 or TRIF) to initiate in the nucleus the production of transcription factors (NFκB, IRF3 or IRF7) that will activate pro-inflammatory cytokines and type 1 interferons (Figure 14) (Kumar et al., 2011; Netea et al., 2012) .

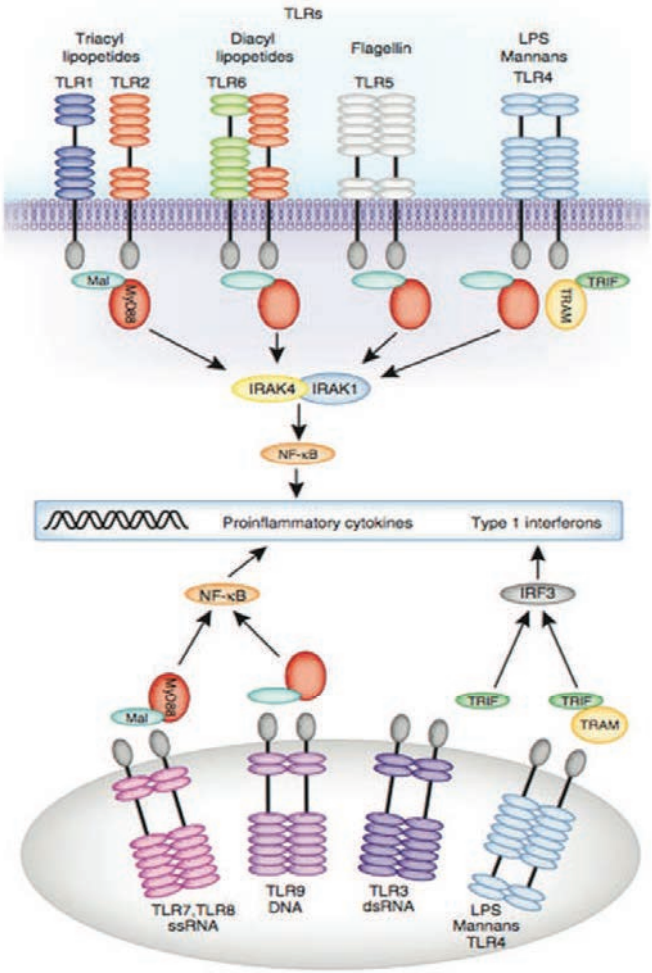


Figure 14. TLRs and their main ligands. Extracellular TLRs (above) and intracellular TLRs (below) act in dimeric form to detect their ligands and initiate the innate immune response that leads to the stimulation of proinflammatory cytokines.

1.2.3.1.1. Evolution and variability of TLR genes

TLRs are innate immune genes that are highly conserved among different vertebrates (Roach et al., 2005)(Song et al., 2012). As all the other PRRs, TLR genes evolve under both non-adaptive and adaptive forces. The non-adaptive evolution is common to all genes within a species and is produced through genetic drifts, bottlenecks and migration routes. The adaptive evolution is stronger on PRRs genes and is mainly caused by infectious pressures (Netea et al., 2012). In general, when novel DNA variants appear in a population they can be under three types of selection: purifying, positive or balancing (Harris and Meyer, 2006; Quintana-Murci, 2016) (Box 2).

When looking at TLR evolution throughout different vertebrates, it was seen that purifying selection is mostly acting in intracellular domains that have to give the signal that triggers immunity; and positive selection is mostly acting in extracellular domains that have to recognize different pathogens and adapt to different infectious pressures (Werling et al., 2009;

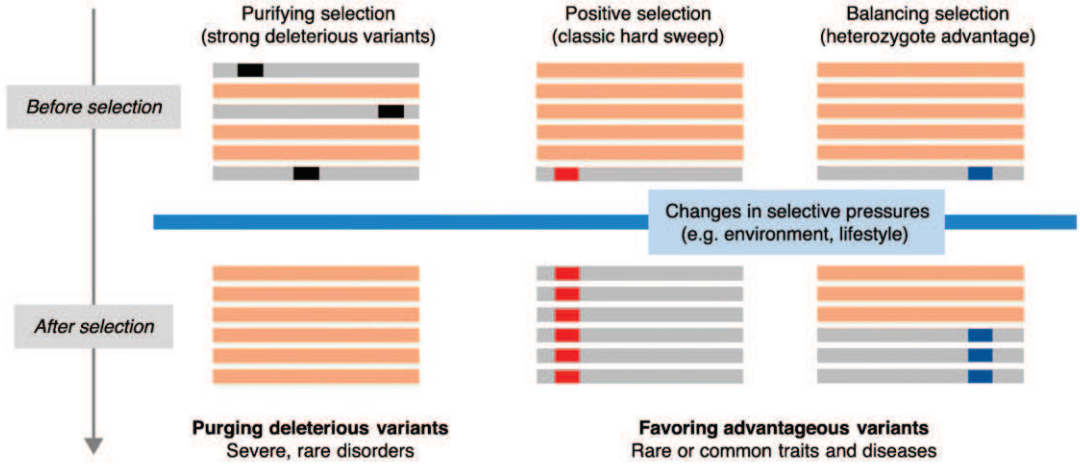
Zhou et al., 2007). Despite being highly conserved, signatures of different types of selection have been detected on TLR genes when comparing different species or even populations that show the infectious pressures they have undergone (Barreiro et al., 2009; Netea et al., 2012; Quach et al., 2013).

On human TLRs, different selection forces were described to drive genetic variation. Purifying selection has been detected in intracellular TLRs (TLR3, 7, 8 and 9) and not in extracellular TLRs (TLR1, 2, 4, 5, 6 and 10) suggesting that intracellular TLRs have an essential non-redundant role in host survival (Barreiro et al., 2009; Wlasiuk and Nachman, 2010). On the other hand, extracellular TLRs evolved under less evolutionary pressure and tolerate more damaging mutations, which suggests their likely immunological redundancy (Barreiro et al., 2009). In contrast, Mukherjee and colleagues found that also TLR2 and 4 were under purifying selection (Mukherjee et al., 2009, 2014). Moreover, another study also reported that human TLR genes were unequally polymorphic: TLR1, 5, 6 and 10 had a greater number of alleles when compared to TLR2, 3, 4, 7, and 9, which is in line with the previous works (Georgel et al., 2009). Finally, balancing selection was also detected on TLR1, 6 and 10 (Ferrer-Admetlla et al., 2008). These divergent results are elucidating the differences among populations.

Box 2. Types of natural selection. (Figure adapted from Quintana-Murci 2016)

Positive, purifying and balancing selection are the three forms of selection that shape genetic variation within and between species.

Natural selection	Definition	Consequence
Purifying selection (or negative selection)	A novel DNA variants have a selective disadvantage with respect to others, and tend to remain at low frequencies or be removed.	Increase the proportion of low-frequency variants within a population.
Positive selection	A novel DNA variant has a selective advantage over others, so rises in frequency	Overall decrease in the genetic variation, as consequence of increase of the advantageous DNA variant (also known as selective sweep).
Balancing selection	This selection increases genetic variation within a species through different biological processes.	Increase the proportion of intermediate-frequency variants within a population (also known as heterozygote advantage).



Thus, variability of the TLRs should be assessed within the same host species, or even breed, because it has been seen that the type of selection and the number of genetic variants differed on certain populations from different geography (Ferrer-Admetlla et al., 2008)(Barreiro et al., 2009); and TLRs responses differ in a species-specific manner (Werling et al., 2009).

In Chapter 3.1, we describe the genetic polymorphisms of TLRs on canids. Both adaptive and non-adaptive evolution should be considered in dogs. Non-adaptive evolution has probably a more significant role than in other species due to dogs were under a first bottleneck with domestication and a second one with the artificial selection of breeds (Figure 15) (Lindblad-Toh et al., 2005). For this reason it should be taken into account the need for dealing with different breeds, and even with other wild canids such as the wolf for the analysis of canine TLR polymorphism.

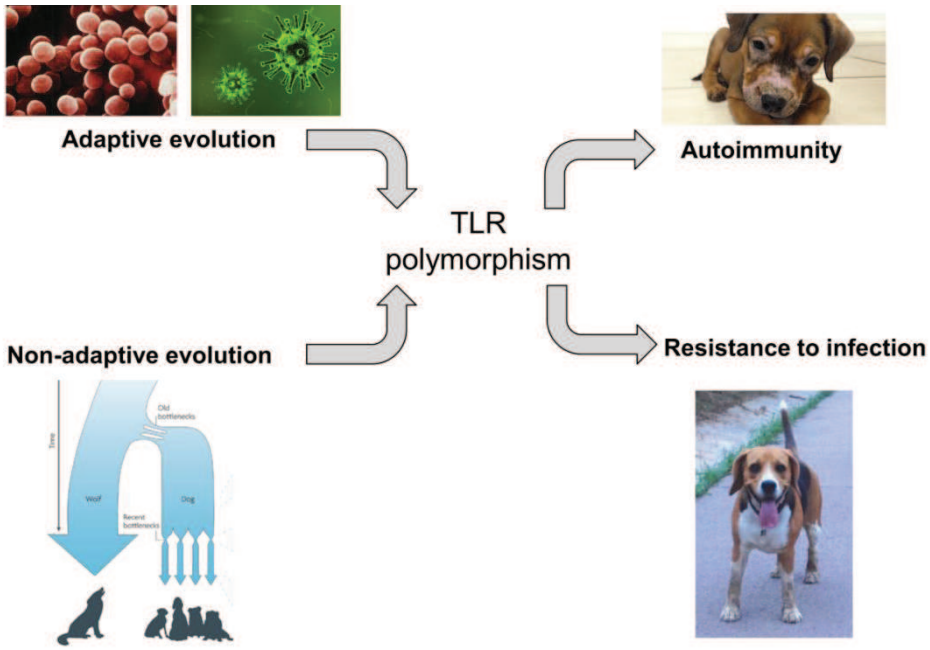


Figure 15. Evolution forces acting on TLR genes. Adaptive evolution caused by infections, combined with non-adaptive evolution caused by genetic drift, population bottlenecks and migration routes, contribute to TLR polymorphisms. Some polymorphisms can lead to resistance to infections, and other to autoimmunity phenomena. (Figure adapted from Netea, Wijmenga, and O'Neill 2012).

1.2.3.1.2. TLRs in skin diseases

Toll-like receptors are innate immune sensors that have a role in both identifying commensal microbiota and maintaining the homeostasis, and in disease establishment. On one hand, TLRs promote mutually beneficial commensal-host interactions (Kubinak and Round, 2012). On the other hand, they are responsible for some diseases, including cutaneous ones (Table 2).

TLRs are expressed not only at innate immune cells, but also at skin cells such as keratinocytes, mast cells, stromal cells and adipocytes (Nestle et al., 2009). Some relationships between skin diseases and TLRs have been described. Thus, for specific skin diseases either certain TLR polymorphisms are correlated to increased susceptibility or resistance or TLR expression is increased or decreased (Table 2). Due to this clear link TLR-disease, therapies targeting TLRs are being assessed for dermatological diseases, using for example molecules that have the ability to modulate TLR expression (TLR agonists and antagonists) (Matin et al., 2015).

Specifically on dog skin, only *Leishmania* infected dogs have reported an altered expression of TLR2, 3, 4 and 9 (Figueiredo et al., 2013; Hosein et al., 2015; Melo et al., 2014). When assessing other canine diseases some links have been reported in both differential TLR expression and polymorphisms.

In gastrointestinal pathologies, differential expression of TLR2 has been described in Inflammatory Bowel Disease (IBD) (McMahon et al., 2010); TLR2, 4, 5 and 9 in chronic enteropathies in German Shepherd (Burgener et al., 2008)(Allenspach et al., 2010); and TLR2 and 4 in inflammatory colorectal polyps (Yokoyama et al., 2017). Moreover, genetic polymorphisms in TLR4 and TLR5 have been associated with IBD in German Shepherd dogs (Kathrani et al., 2010), but only protective SNPs from TLR5 have been associated with IBD in other 38 dog breeds (Kathrani et al., 2011). Other pathologies presented differential expression of TLRs, such as TLR4 in osteoarthritis (Kuroki et al., 2010) and in infected canine endometrium (Chotimanukul and Sirivaidyapong, 2011); TLR2 and 7 in arthritis (Riggio et al., 2014); and TLRs 1-4, 6-10 in sino-nasal aspergillosis and idiopathic lymphoplasmacytic rhinitis (Mercier et al., 2012).

Table 2. Dermatological diseases associated with an alteration on Toll-like Receptors.

Skin disease	TLRs alterations	References
Infectious diseases		
Acne vulgaris	Increased expression of TLR2, correlated with severity	(Bakry et al., 2014; Kim et al., 2002)
Candidiasis	TLR9-deficient mice resistant to candidiasis	(Kasperkovitz et al., 2011)
Leishmaniosis	Increased expression of TLRs 2, 4 and 9	(Tuon et al., 2010)
Leprosy	Hypo-functional TLR1 variant protected against leprosy	(Wong et al., 2010)
	Polymorphisms on TLRs 1, 2 and 4 associated with susceptibility	(Bochud et al., 2009)
	Increased expression of TLR1 and 2	(Krutzik et al., 2003)
Rosacea	Increased expression of TLR2	(Yamasaki et al., 2011)
Staphyloccocal infections	Increased expression of TLR2	(Hilmi et al., 2014)
Immunological diseases		
Atopic dermatitis	Polymorphisms on TLR2 associated with the disease	(Potaczek et al., 2011)
	Decreased expression of TLR2 and 4	(Lesiak et al., 2012)
Psoriasis	Increased expression of TLR1, 2, 4, 5, 9	(Begon et al., 2007)
Lichen planus	Increased expression of TLR4 and reduced expression of TLR2	(Janardhanam et al., 2012)
	Increased expression of TLR4 and 9	(Siponen et al., 2012)
	Decreased expression of TLR1 and 2	(Salem et al., 2013)
Pemphigus	Increased expression of TLRs 2, 3 and 4	(Abida et al., 2013)
Systemic autoimmune diseases (with skin manifestations)		
Systemic Lupus Erythematosus	Pathogenic role for TLR7 and protective role for TLR9	(Celhar et al., 2012)
Sarcoidosis	Increased expression of TLRs 2, 3, 4, 5, 6, 7, and 8	(Huizenga et al., 2015)
Systemic sclerosis	Increased expression of TLR3	(Agarwal et al., 2011; Farina et al., 2010)
Cancer		
Basal cell carcinoma	Higher expression of TLRs 1, 2, 3, 5, 6, 7, 8	(Muehleisen et al., 2012)
Squamous cell carcinoma	Higher expression of TLRs 1, 2, 3, 5, 6, 7, 8	(Muehleisen et al., 2012)
Other skin diseases		
Stevens–Johnson syndrome	Polymorphisms on TLR3 associated with the disease	(Ueta et al., 2007)

1.3. Skin microbiota

The microbiota is defined as a collection of microorganisms that inhabit a specific environment (Marchesi and Ravel, 2015), thus the canine skin microbiota is the collection of microorganisms that inhabits the skin of dogs.

Until today (June 2017), 637 out of 846 studies found in *pubmed* regarding skin microbiota are human-based, so we will provide an extensive overview of skin microbiota focusing first on what is known for human and then specifically for dog in health (Chapter 1.3.1, 1.3.2 and 1.3.3) and disease (Chapter 1.3.4). We will finish this chapter reviewing the clinical potential of the microbiota (Chapter 1.3.5).

1.3.1. What is living on healthy skin?

The human skin microbiota is predominantly inhabited by bacteria, followed by fungi and viruses with lower abundances (Oh et al., 2014). Microbiota members mostly inhabit the skin surface, hair follicles and other appendages (Figure 16)(Grice and Segre 2011). Moreover, Nakatsuji and colleagues also detected bacteria in the deep dermis and subcutaneous tissues of healthy individuals (Nakatsuji et al., 2013).

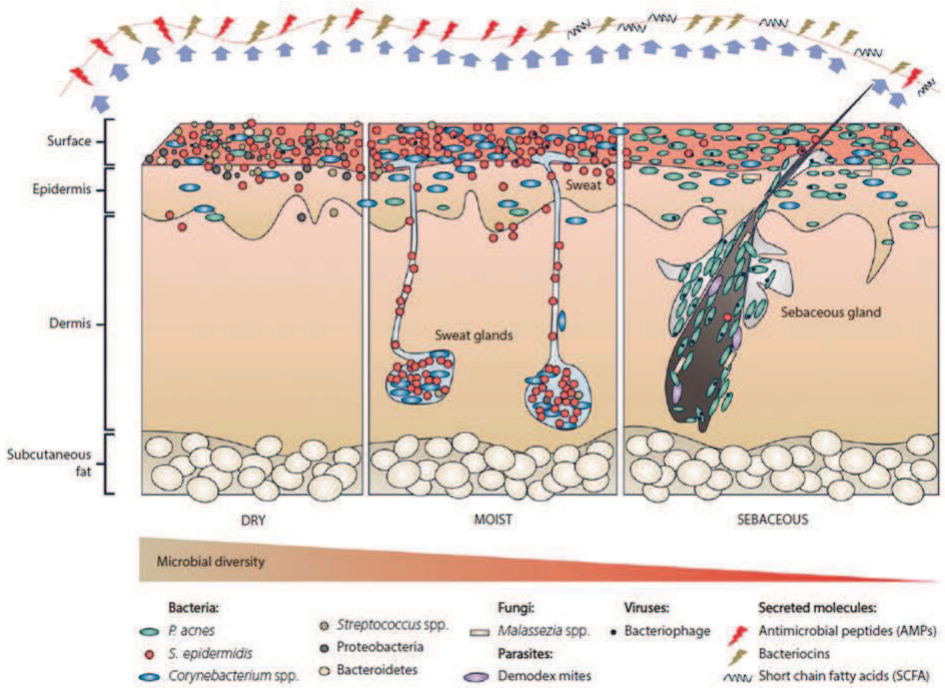


Figure 16. Human skin microhabitats and their associated microbiota. Human skin has three main microhabitats: dry, moist and sebaceous. Each of them has its different skin appendages and specific physicochemical properties that define the microbiota they harbor (Figure from Barnard and Li 2017).

First human microbiota studies using next-generation sequencing techniques elucidated in skin that the three main microhabitats harbor specific bacteria creating bacterial signatures. So, sebaceous sites are inhabited by *Propionibacterium spp*; moist sites, by *Staphylococcus* and *Corynebacterium spp*; and dry sites with gram-negative microorganisms (Grice et al. 2009; Costello et al. 2009; Grice and Segre 2011) (Figure 17a). Dry skin sites present higher diversity values when compared to the others. Similarly, when compared to other body site microbiotas, skin presents higher diversity values despite the lower microbial load (Belkaid and Segre, 2014a).

Dog skin is almost entirely covered by a dense fur creating a more uniform microenvironment, so no clearly defined microhabitats have been identified. Bacteria from the Proteobacteria phylum are the most abundant all over the dog skin (Figure 17b) (Song et al. 2013; Rodrigues Hoffmann et al. 2014), contrasting with the microhabitat-specific taxa on human. Within the skin sites, haired skin regions usually present higher alpha diversity values when compared to mucosal or mucocutaneous regions (Rodrigues Hoffmann et al., 2014).

Besides skin site divergences, a large diversity between individuals has also been reported (Hoffmann et al., 2014).

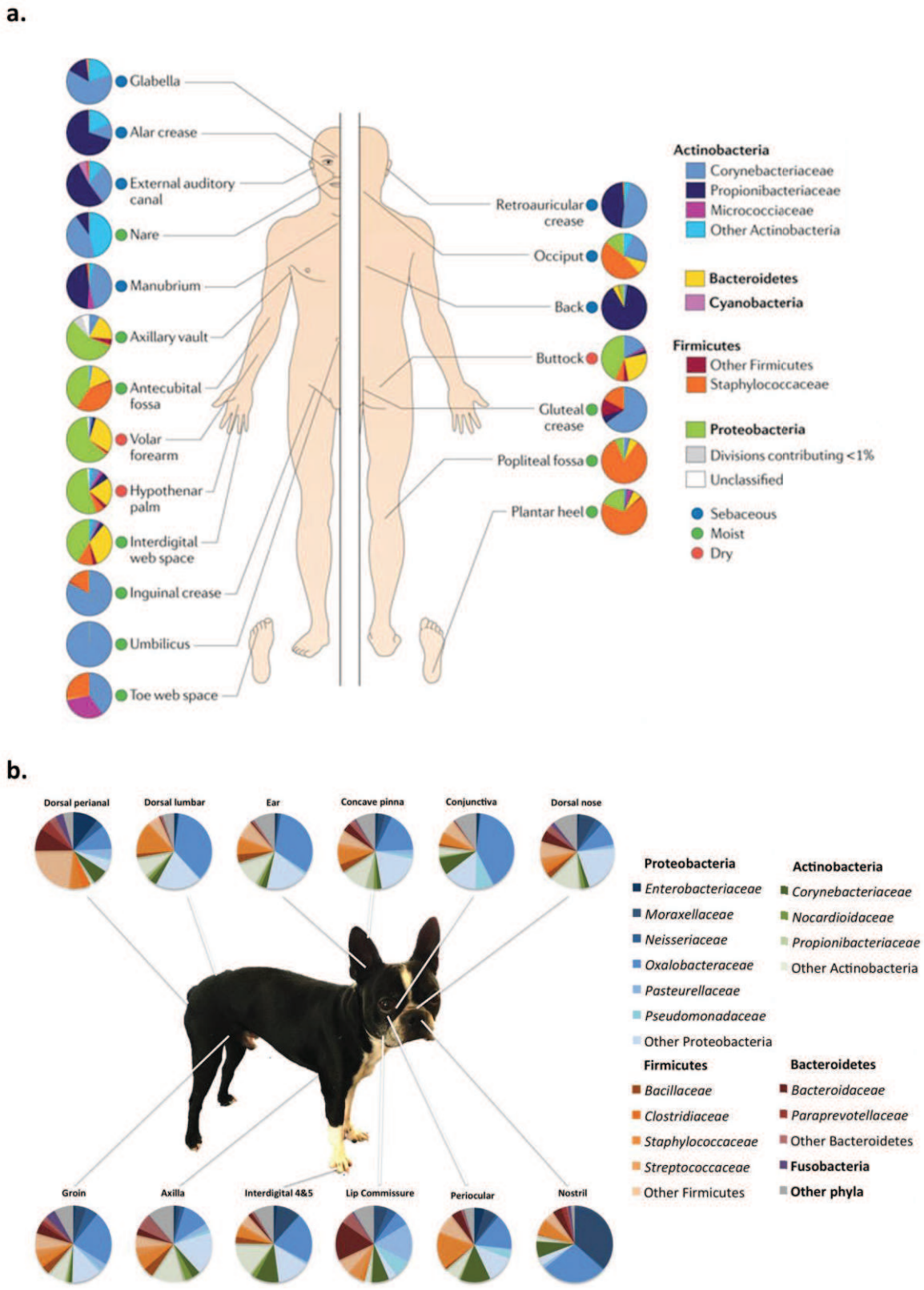


Figure 17. Skin microbiota composition per skin site. In a) human skin microbiota classified by microhabitat (Figure from Grice and Segre 2011) and in b) dog skin microbiota (Figure from Rodrigues Hoffmann et al. 2014).

Human skin microbiota presents high variability, not only between skin sites or microhabitats –intra-individual variability–, but also between individuals –inter-individual variability (Grice and Segre, 2011; Human Microbiome Project Consortium, 2012; Oh et al., 2014).

Individual signatures of the skin microbiota have the ability to identify items that an individual came in contact with (Fierer et al., 2010; Lax et al., 2015; Meadow et al., 2014). Despite the temporal stability is a personalized feature (Flores et al., 2014; Oh et al., 2016), skin microbiota in sebaceous and dry sites is relatively constant over time (Figure 18) (Oh et al., 2016). Finally strain-level differences in *Propionibacterium acnes* were individual-specific, whereas those in *Staphylococcus epidermidis* were more site-specific (Oh et al., 2014). For all these reasons demonstrating individual signatures in skin, microbiota analyses have even been proposed as a new tool for forensic science (Hampton-Marcell et al., 2017).

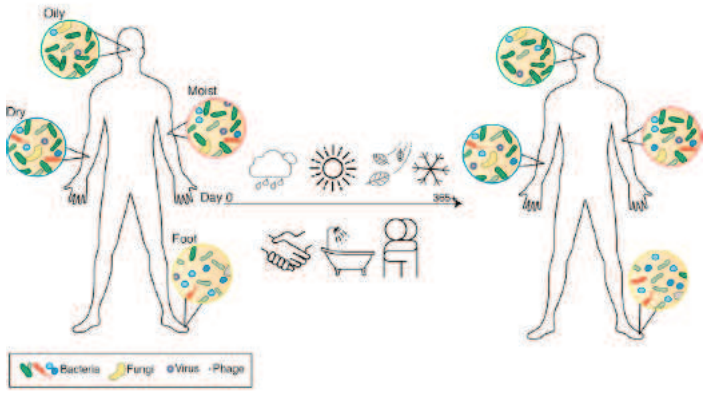


Figure 18. Temporal stability of the human skin microbiota. Healthy adults maintain their skin microbial communities at sebaceous and dry sites over time despite the constant exposure to external environment. (Excerpted from Oh et al. 2016).

Microbial diversity is not limited to bacteria; microorganisms such as archaea, fungi, and viruses also have major roles in human health and disease (Eckburg et al., 2003; Handley, 2016; Peleg et al., 2010) and whole genome sequencing approaches have allowed deciphering this global picture. Depending on the habitat, the abundance of fungi and viruses varies from <0.1% of microorganisms in the gastrointestinal tract to up to 10% and 40% on skin, respectively (Oh et al. 2014; Belkaid and Segre 2014). Archaea can be found on gastrointestinal tract with <1% (Arumugam et al., 2011) and are nearly absent on skin (Oh et al., 2014).

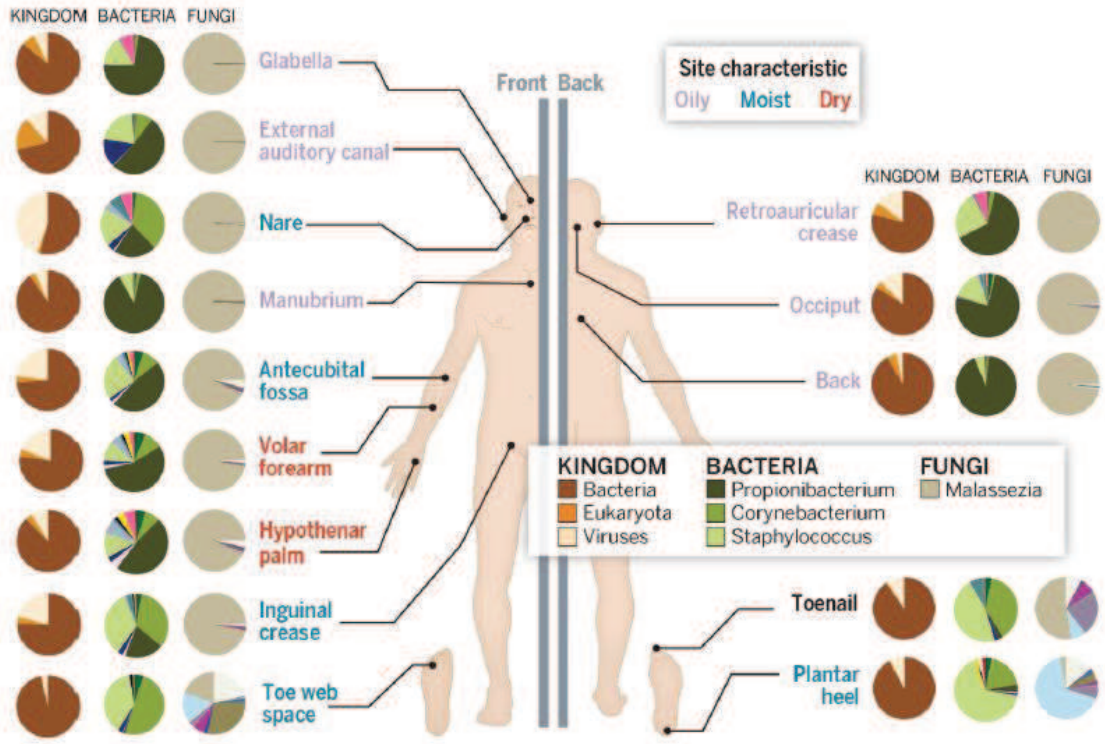


Figure 19. Microbial community of the skin using shotgun metagenomics. Relative abundances of viral, bacterial, and fungal components per skin site. Sites represent three skin microhabitats: sebaceous (blue), dry (red), and moist (green) (excerpted from Belkaid and Segre 2014)

Fungal microbiota can also be retrieved using amplicon-based approaches targeting their rRNA gene to amplify the Internal Transcribed Spacer (ITS) regions or 18S (Halwachs et al., 2017).

Malassezia species were detected as the main colonizers of the human skin in forearm and face (Paulino et al., 2006; Zhang et al., 2011). Findley and colleagues found that in fact most of the skin sites in healthy individuals were colonized by *Malassezia*, detecting site signatures only at species level. By contrast, the three foot sites analyzed presented more diversity being colonized by a combination of *Malassezia*, *Aspergillus*, *Cryptococcus*, *Rhodotorula*, *Epicoccum*, and others (Findley et al., 2013). This pattern seemed to be specific to adult population, whereas pre-pubertal skin presented higher diversity (Jo et al., 2016). When analyzed with metagenomics approaches, fungi presented high abundances near the ears and forehead, but low representation in the feet (<1%) despite the high diversity observed in the previous amplicon-based studies (Oh et al., 2014).

On dog skin using an amplicon-based approach, Meason-Smith and colleagues found that the most abundant fungi were *Alternaria* and *Cladosporium* independently from the skin site, which was not an influencing factor on mycobiota composition and structure (Figure 20). In fact, different skin sites tended to be similar within a dog (Meason-smith et al., 2015). *Alternaria* and *Cladosporium* were also the most abundant fungi on cat skin mycobiota (Meason-Smith et al., 2017).

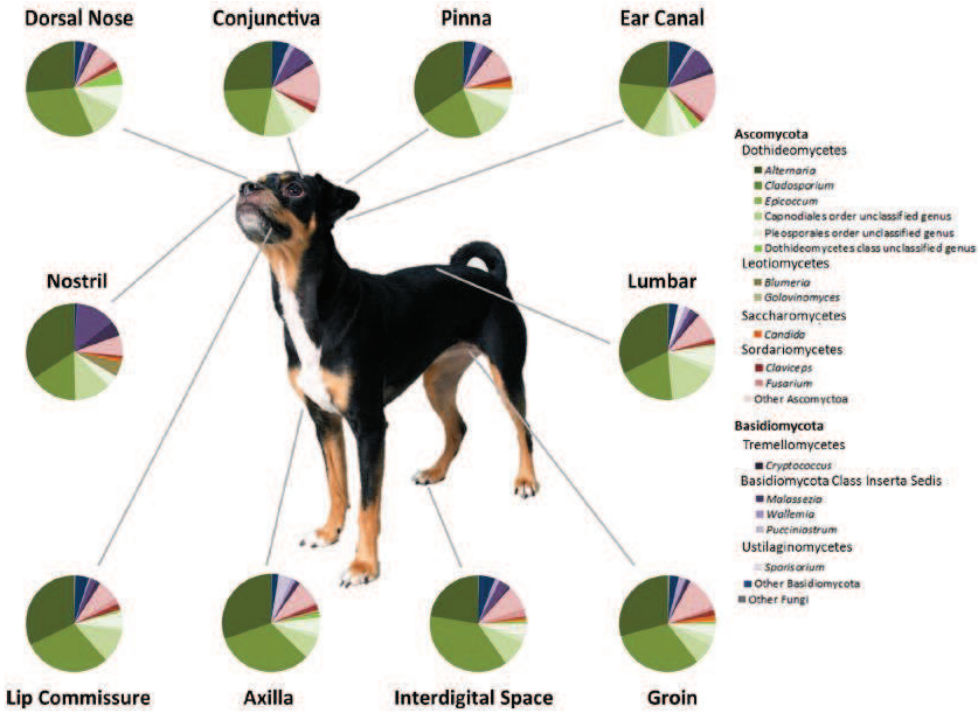


Figure 20. Skin mycobiota composition on healthy dogs per skin site. (Excerpted from Meason-smith et al. 2015).

The viruses are the most forgotten microbiota members, and main insights on human skin come from whole shotgun metagenomics approaches (Foulongne et al., 2012; Oh et al., 2014, 2016; Wylie et al., 2012). Only one study enriched viral particles and assessed cross-sectional diversity within different skin sites, which found that skin virome was highly site-specific and modulated by occlusion and exposure, in addition to sebum and moisture (Hannigan et al., 2015).

The virome can be classified within two main groups: prokaryotic virome, which includes viruses affecting prokaryotes, such as bacteriophages; and eukaryotic virome, which includes viruses affecting eukaryotic cells (Virgin, 2014). Among the prokaryotic virome the most abundant members were *Propionobacterium* and *Staphylococcus* phages (Hannigan et al., 2015; Oh et al., 2014, 2016) and *Pseudomonas* and *Bacillus* phages (Hannigan et al., 2015). In eukaryotic skin virome, *Papillomaviruses* as well as *Polyomavirus* were the most commonly found (Foulongne et al., 2012; Oh et al., 2014, 2016; Wylie et al., 2012); however the most abundant viruses were specific to single individuals (Oh et al., 2016). On dog skin, the virome has not been assessed yet.

1.3.2. Factors shaping bacterial skin microbiota

The key variables driving human skin microbiota structure and composition are skin site followed by the individual, as seen in the previous section. Besides, many other variables are affecting its composition and structure, which can be globally classified as: intrinsic factors (or host-specific) and extrinsic factors (or environment-specific) (Grice and Segre 2011).

On one hand, host factors include those relative to the individual such as the age, gender, racial origin or immune system and all of them seem to have an effect on skin microbiota (Sanmiguel and Grice, 2015).

Regarding the age, skin microbiota evolves over the first year of life, showing an initial colonization by *Staphylococci* that decreases with age as diversity increases (Capone et al., 2011). Moreover, progressive microbial shifts in skin and nares have been detected in different sexual maturation stages and when comparing children to adults (Oh et al., 2012). Cohorts usually include people from both genders, but only few studies have found a strong link between gender and skin microbiota structure and composition in hands (Fierer et al., 2008), upper buttock (Zeeuwen et al., 2012) and axillary vault (Callewaert et al., 2013). Racial origin or ethnicity presented a strong effect on skin microbiota when linked to clinical metadata, as it was shown by Human Microbiome Project that included around 300 individuals from different genetic backgrounds (Asian, Black, Mexican, Puerto Rican and White) (Human Microbiome Project Consortium, 2012). Other studies detected differences in skin microbiota of Chinese people when compared to other racial groups (Leung, Wilkins, and Lee 2015). Also, a recent study including six different ethnic groups living in New York City detected racial origin as a secondary factor shaping skin microbiota after skin site (Perez Perez et al., 2016). Finally, the immune system is the main responsible of recognizing the microorganisms and it definitely affects skin microbiota, as well as skin microbiota affects immunity (Naik et al. 2012; Naik et al. 2015). This part is covered in the following section (Chapter 1.3.3).

On the other hand, extrinsic or environmental factors include those derived from the surroundings (geography, urbanization or cohabitation) or personal habits (hygiene, cosmetics use or diet).

Tribal populations that spend more time outdoors presented higher abundances of environmental-derived taxa and increased alpha diversity values when compared to other populations (Clemente et al., 2015; Hospodsky et al., 2014). Skin microbiota varied also when comparing different environments of a delimitated region (Hanski et al., 2012; Ying et al., 2015). In line with this, individuals cohabiting together –even when including pets– shared a larger proportion of skin microbiota if compared to other individuals (Song et al. 2013; Misic et al. 2015). Regarding personal habits, Fierer and colleagues found that skin hand microbiota was influenced not only by gender, but also by the time since the last hand washing. In fact, the gender differences they reported could also have been due to extrinsic factors such as the use of cosmetics or moisturizers (Fierer et al., 2008). A recent study found that soap and shampoo practices were secondary factors shaping skin microbiota on New Yorkers (Perez Perez et al., 2016).

Finally some of these microbial drivers are both intrinsic and extrinsic (racial origin – different environment; or sex – different hygiene practices), and even some of them are a combination of both. For example, Lehtimäki and colleagues recently reported that geography shapes microbiota in an age-dependent way: environmental effects decrease as age increase (Lehtimäki et al., 2017).

In canine skin, all of these factors have the potential to drive skin microbiota composition and structure. Moreover, those derived from the environment are likely having a major impact in dogs than in humans. On one hand, as host factors we could include sex, age, skin site, individual, immune system, genetic background or even hair coat. On the other hand, as environmental factors we could include cohabitation, lifestyle, environment, season, diet, hygiene or geography (Figure 21).

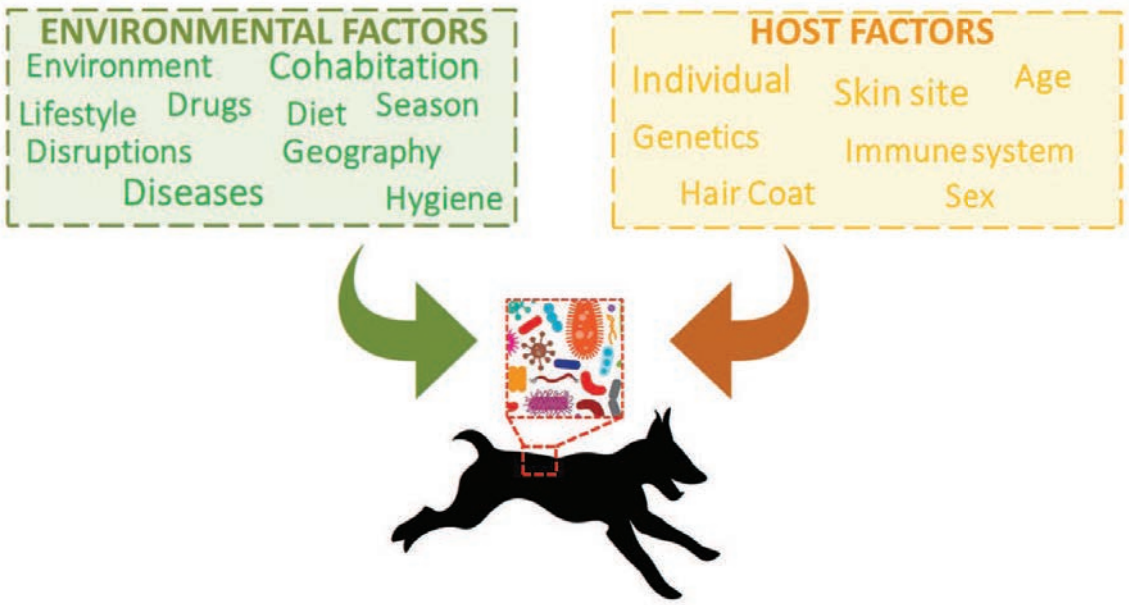


Figure 21. Canine skin microbiota is potentially shaped by both host and environmental factors.

Few of these factors have been tested as potential drivers of skin microbiota structure and composition in healthy dogs.

Rodrigues-Hoffmann and colleagues described the skin site effect on dog skin microbiota, showing mucosal and muco-cutaneous regions less diversity when compared to haired skin sites. They also detected an individual effect, despite not assessing it directly. Other factors such as the presence of fleas, time spend outdoors *vs* indoors, the sex and age were assessed as potential drivers of canine skin microbiota without finding any relationship. This could have been due to they worked with a small (N=12) and heterogeneous cohort including dogs from different households, genetic backgrounds and ages, so true effect could have been obscured (Rodrigues Hoffmann et al., 2014). Besides, animals cohabiting together (similar to humans) share more skin microbiota and moreover cohabitation was affecting skin microbiota of the pet owner, increasing its bacterial diversity (Song et al. 2013).

1.3.3. Commensal microbiota functions on the skin

Skin is one of the main interfaces with the environment and it is constantly exposed to external microorganisms, which some of them can be potential pathogens. Besides that, skin is not only exposed but also colonized by many microorganisms that constitute the skin microbiota.

This living barrier that is the skin microbiota contributes to the skin immunity mainly in three different ways: directly inhibiting pathogen growth; enhancing host innate immunity; and educating and priming adaptive immunity (Figure 22) (Sanford and Gallo, 2013). We will not review how the skin microbiota shapes the adaptive immunity, as it is beyond the scope of this dissertation.

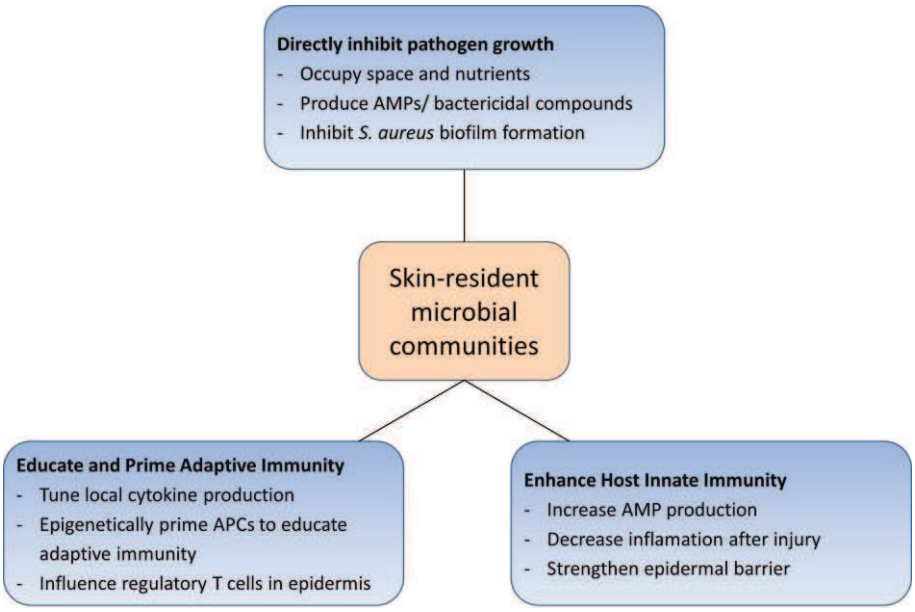


Figure 22. Skin microbiota immune functions on health. Host cells and skin microbiota are associated and their cross-talks help skin immunity through several mechanisms. (Excerpted from Sanford and Gallo 2013).

Skin microbiota members are inter-communicated (bacteria-bacteria interactions) and moreover interact with host cells establishing cross-talks (bacteria-host cells interactions). These interactions can be classified in three main categories: commensalism, when one benefits from the other; mutualism, when both benefit; and detrimental relationship, when one harms the other (Figure 23) (Schommer and Gallo, 2013).

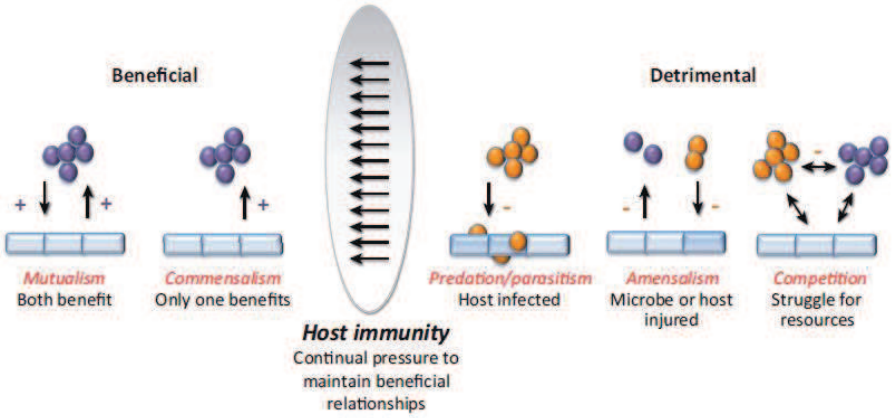


Figure 23. Dynamics of microbial interaction at the skin surface. Members of the microbiota form complex interaction networks, with both microbe-microbe and microbe-host interactions**. (**Excerpted from Schommer and Gallo 2013)..

Usually these interactions and their functions are studied in individual commensal members of the microbiota. Among all the bacteria inhabiting skin microbiota, we delve into the functions of the two most well-characterized microbiota members on human skin: *Propionibacterium acnes* and *Staphylococcus epidermidis* (Table 3).

Table 3. *Propionibacterium acnes* and *Staphylococcus epidermidis* contributions to the innate immunity functions.

Commensal bacteria	Functions in skin health
Propionibacterium acnes	Producing SCFAs that inhibit growth of other microorganisms
	Maintaining acidic skin pH
	Producing bacteroicins
	Promoting commensal growth
Staphylococcus epidermidis	Producing AMPs and bacteroicins
	Promoting host immune to produce host AMP
	Promoting host immune responses via TLR signaling

Propionibacterium acnes mechanisms that contribute to skin immunity are mostly based on directly inhibiting pathogen growth, which is achieved by: 1) metabolizing triglycerides from the sebum to short chain fatty acids, which have antimicrobial properties and contribute to acidic pH of the skin (Shu et al., 2013; Ushijima et al., 1984); and 2) producing bacteroicins (Faye et al., 2011), which are toxins that inhibit the growth of similar bacteria. Moreover, *P. acnes* strains associated with healthy skin –not those ones associated with acne–, carry genes that synthesize thiopeptides, which are antimicrobial compounds that inhibit the growth of gram-positive species (Christensen and Brüggemann, 2014).

Staphylococcus epidermidis contributes to skin immunity by both enhancing and modulating innate immunity via TLR signaling (Lai et al., 2009; Wanke et al., 2011) and by directly inhibiting pathogen growth producing antimicrobial peptides (Bastos et al., 2009; Cogen et al., 2010). For example, *Staphylococcus epidermidis* was able to amplify the innate immune response against the invading pathogen *Staphylococcus aureus* by increasing antimicrobial peptides expression and abolishing the inhibition of NF-κB signaling asserted by the pathogen (Figure 24) (Wanke et al., 2011).

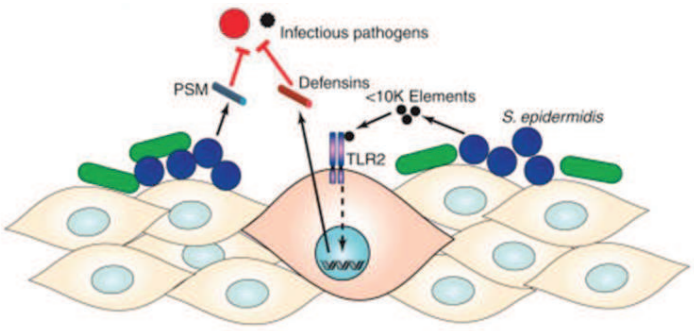


Figure 24. *Staphylococcus epidermidis* cross-talk with innate immunity. *Staphylococcus epidermidis* produces antimicrobial peptides to act as a barrier against colonization of potentially pathogens. In addition, it secretes a small molecule, which increases expression of defensins through TLR2 signaling. (Excerpted from Gallo and Nakatsuji 2011).

1.3.4. Skin microbiota and dermatological diseases

In a healthy skin, homeostasis is maintained through a cross-talk between immune system and microbiota and when dysregulated produces inflammation (Belkaid and Segre 2014). In humans, several studies have associated skin diseases to dysbiosis, which is an imbalance of microbiota (Table 4 and Figure 25). However, if dysbiosis is the cause or the consequence of the disease in most cases still remains to be clarified. On dogs, only few studies have been performed and are focused on canine atopic dermatitis (Table 4).

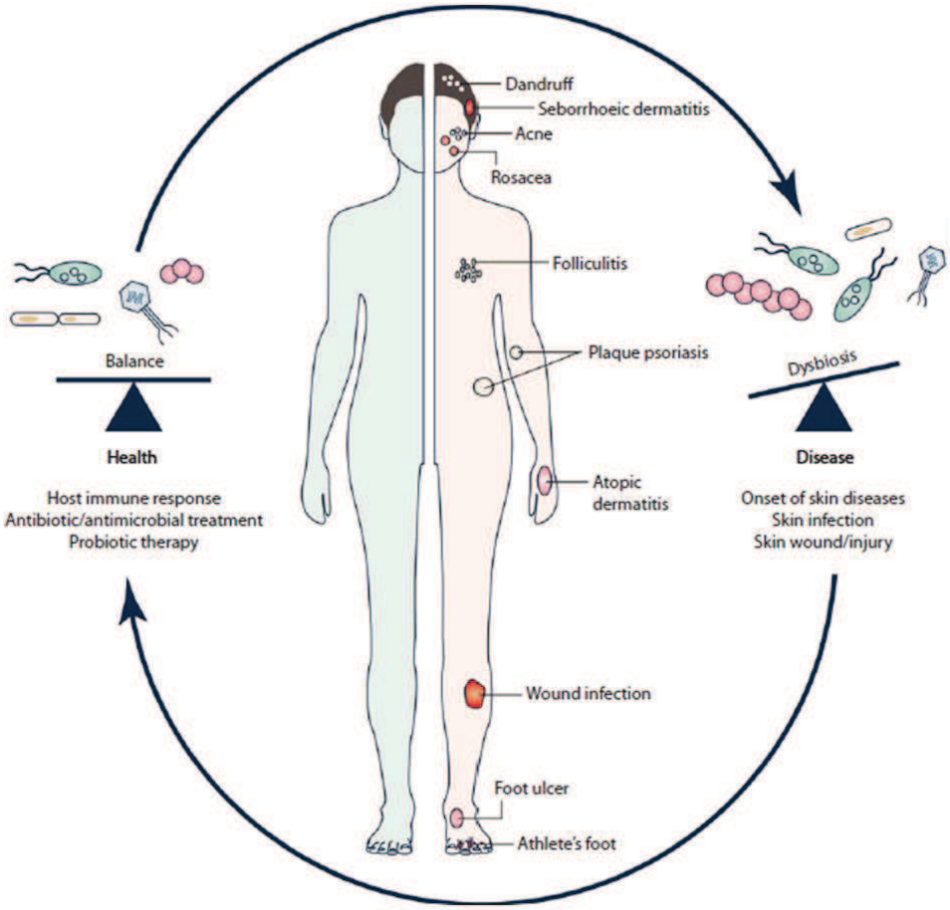


Figure 25. Dermatological diseases associated with dysbiosis on skin microbiota. (Excerpted from (Barnard and Li, 2017).

Table 4. Dermatological diseases associated with skin microbiota alterations.

Disease	Skin microbiota alterations	References
Acne vulgaris	Strain-level differences in *P. acnes* are associated with acne.	(Fitz-Gibbon et al., 2013)
	Balance between acne- and health-associated species determines health status.	(Barnard and Li, 2017)
	Increased *Staphylococci*, correlated with severity.	(Dreno et al., 2017)
Atopic dermatitis (AD)	Increased *Staphylococcus*, correlated with severity. Decreased diversity during flares.	(Kong et al., 2012a)
	Increased *Staphylococcus* and *Corynebacterium*, correlated with severity (in AD and PID).	(Oh et al., 2013)
	Differences btw affected *vs* unaffected skin. *Stenotrophomonas* prevalent in therapy responders.	(Seite and Bieber, 2015)
	1-year-old infants with AD were not colonized with S. aureus before developing AD.	(Kennedy et al., 2017)
	Signature in AD-prone skin: enriched in *Streptococcus* and *Gemella* but depleted in *Dermacoccus*.	(Chng et al., 2016)
Bacterial infection (*H. ducreyi*)	Resolvers and pustule formers have distinct skin bacterial communities	(Rensburg et al., 2015)
Primary immunodeficency	Decreased site specificity and temporal stability. Colonization with species not found in controls.	(Oh et al., 2013)
	Increased Gram-negative bacteria, especially *Acinetobacter*. Decreased *Corynebacterium*.	(Smeekens et al., 2013)
Psoriasis vulgaris	Increased Firmicutes and Actinobacteria.	(Gao et al., 2008)
	Decreased *Staphylococci* and *Propionibacteria*.	(Fahlén et al., 2012)
	Increased *Corynebacterium, Propionibacterium, Staphylococcus*, and *Streptococcus*.	(Alekseyenko et al., 2013)
Vitiligo	Decreased diversity. Altered networks: Firmicutes as central nodes (Actinobacteria in health)	(Ganju et al., 2016)
Wounds (non-chronic)	Troughout healing process, wound microbiota became increasingly similar to adjacent skin.	(Hannigan et al., 2014)
Wounds (chronic)	Prevalent anaerobic bacteria (from Clostridiales family XI)	(Price et al., 2009)
	Prevalent uncharacterized Bacteroidales, anaerobes and *Staphylococcus, Corynebacterium,* and *Serratia*	(Wolcott et al., 2009)
	Increased anaerobic bacteria and *Corynebacterium* and *Staphylococcus*.	(Gontcharova et al., 2010)
Canine atopic dermatitis	Decreased species richness on allergic dogs.	(Rodrigues Hoffmann et al., 2014)
	Decreased bacterial diversity, increased *Staphylococcus* (*S. pseudintermedius*) and *Corynebacterium*.	(Bradley et al., 2016a)
	Increased abundance of *S. pseudintermedius* at the site of lesion induction.	(Pierezan et al., 2016)

In most of the previous studies, affected skin showed lower diversity when compared to healthy one, independently from the skin pathology. This lower diversity was usually linked to an overgrowth of specific taxa, which sometimes was also correlated to severity. *Staphylococcus* is a genus that was overrepresented in acne vulgaris (Dreno et al., 2017), atopic dermatitis (Kong et al. 2012; Oh et al. 2013), primary immunodeficiencies (Oh et al., 2013), psoriasis (Alekseyenko et al., 2013) and chronic wounds (Gontcharova et al., 2010; Wolcott et al., 2009) and it was correlated with severity for acne (Dreno et al., 2017), atopic dermatitis (Kong et al. 2012; Oh et al. 2013) and primary immunodeficiencies (Oh et al., 2013). However, another study in Psoriasis detected a decrease on this genus (Fahlén et al., 2012). *Corynebacterium* was also a common genus altered in several pathologies: increased in atopic dermatitis and primary immunodeficiencies (Oh et al., 2013), psoriasis (Alekseyenko et al., 2013), and chronic wounds (Gontcharova et al., 2010; Wolcott et al., 2009); and also decreased in primary immunodeficiencies (Smeekens et al., 2013). Besides these highly common trends, each skin disease presented its own microbial characteristics and even some studies lead to contradictory results (Table 4).

In dogs, the three studies targeting atopic dermatitis found also that affected skin presented lower diversity when compared to healthy one (Bradley et al., 2016a; Pierezan et al., 2016; Rodrigues Hoffmann et al., 2014). *Staphylococcus pseudointermedius* increased in affected skin, either when compared to healthy dogs (Bradley et al. 2016) or the same dog in a non-affected site (Pierezan et al., 2016). Also *Corynebacterium* increased in affected skin (Bradley et al., 2016).

1.3.5. Skin microbiota as a clinical tool

As skin microbiota differs in health and disease, some researchers have suggested that microbiota analyses have the potential to be used in clinics as a diagnosis, prognosis and even therapeutic tool (Grice 2014; Grice 2015).

All these potential clinical approaches are based on microbiota property to sense its environment and respond against alterations. These alterations can be harmful or beneficial. They are harmful when they create a dysbiotic state, which can be produced by pathogens, environmental stimuli or immune defects among others. Beneficial alterations are usually intentioned and aim to recover the homeostasis, which can be achieved using products such as pre- and pro-biotics or even antibiotics (Egert and Simmering, 2016).

As previously seen, some skin diseases have a kind of bacterial signature, with some overrepresentation of specific bacteria (Table 4). Moreover, as seen in acne vulgaris, sometimes a skin disease is not due to the presence of a specific set of bacteria but to an altered balance between health- and disease-specific bacteria (Barnard and Li, 2017). So, microbiota analysis could be used as a diagnosis tool for certain skin diseases identifying the microbial balance and signature.

Regarding prognosis, higher proportions of specific bacteria are associated to increased severity of certain dermatological diseases, such as atopic dermatitis or acne vulgaris (Dreno et al., 2017; Kong et al., 2012b; Oh et al., 2013). Moreover, certain skin microbiota

compositions make you a resolver or a pustule former in an infection with *Haemophylus ducreyi* (Rensburg et al., 2015) and also people prone to atopic dermatitis have a different skin microbiota when compared to healthy individuals (Chng et al., 2016). In non-chronic wounds, a specific microbiota profile was correlated with future complication (Hannigan et al., 2014). Thus, different researchers have confirmed the potential to use the microbiota as a prognosis tool in several skin conditions.

Finally, the most promising tool in clinics is using microbiota manipulation as a therapeutic approach to resolve certain diseases (Grice, 2014; Reid et al., 2011). These manipulations can be divided in two main groups: the classical antimicrobial approach, to reduce the number of pathogenic microorganisms; and the novel pre- and pro-biotic approach, to increase a specific subset of beneficial bacteria (Egert and Simmering, 2016). Despite the function of antibiotics is to restore the healthy status through pathogen elimination, it also affects microbiota and has detrimental effects such as killing beneficial microorganisms and creating antibiotic resistances with its long-term use. Moreover, antibiotics perturb the original microbiota and the altered version is more likely to present overgrowth of opportunistic pathogens (for a review on microbiota alterations due to antibiotics use, see Ferrer et al. 2017). All these concerns about antibiotics use led researchers to investigate the novel approach of using pre- and probiotics as a therapy or complement for certain diseases.

Prebiotics are non-digestible food ingredient that affects the host by stimulating the growth or activity of specific beneficial bacteria, whereas probiotics are products containing living microorganisms in sufficient numbers, so they can alter the microbiota and produce beneficial health effects in the host (Schrezenmeir and de Vrese, 2001) (Fig 26A). When developing pre- and pro-biotics, researchers need to target specific bacterial products that are beneficial for the skin (Fig 26B) and therefore find the bacterial strain that metabolizes that product or the substrate that promotes the growth of that specific strain (Lew and Liong, 2013). Microbiota studies are invaluable to understand the mechanisms and dynamics of these novel products and to ensure their safety. Most of pre- and pro-biotics aim to manipulate the gastrointestinal microbiota rather than the skin, even when treating dermatological problems (Baquerizo Nole et al., 2014; Krutmann, 2009). Some topical probiotics have shown an effect until the date: *Lactobacillus rhamnosus* for atopic dermatitis (Hoang et al., 2010; Viljanen et al., 2005) ; *Bifidobacterium longum* for sensitive skin (Guéniche et al., 2009); and kefir –uncharacterized probiotic mixture– for wound healing (Huseini et al., 2012). Also sphingomyelinase produced by *Streptococcus thermophilus* was shown to increase skin-ceramide levels in aged subjects (Dimarzio et al., 2008).

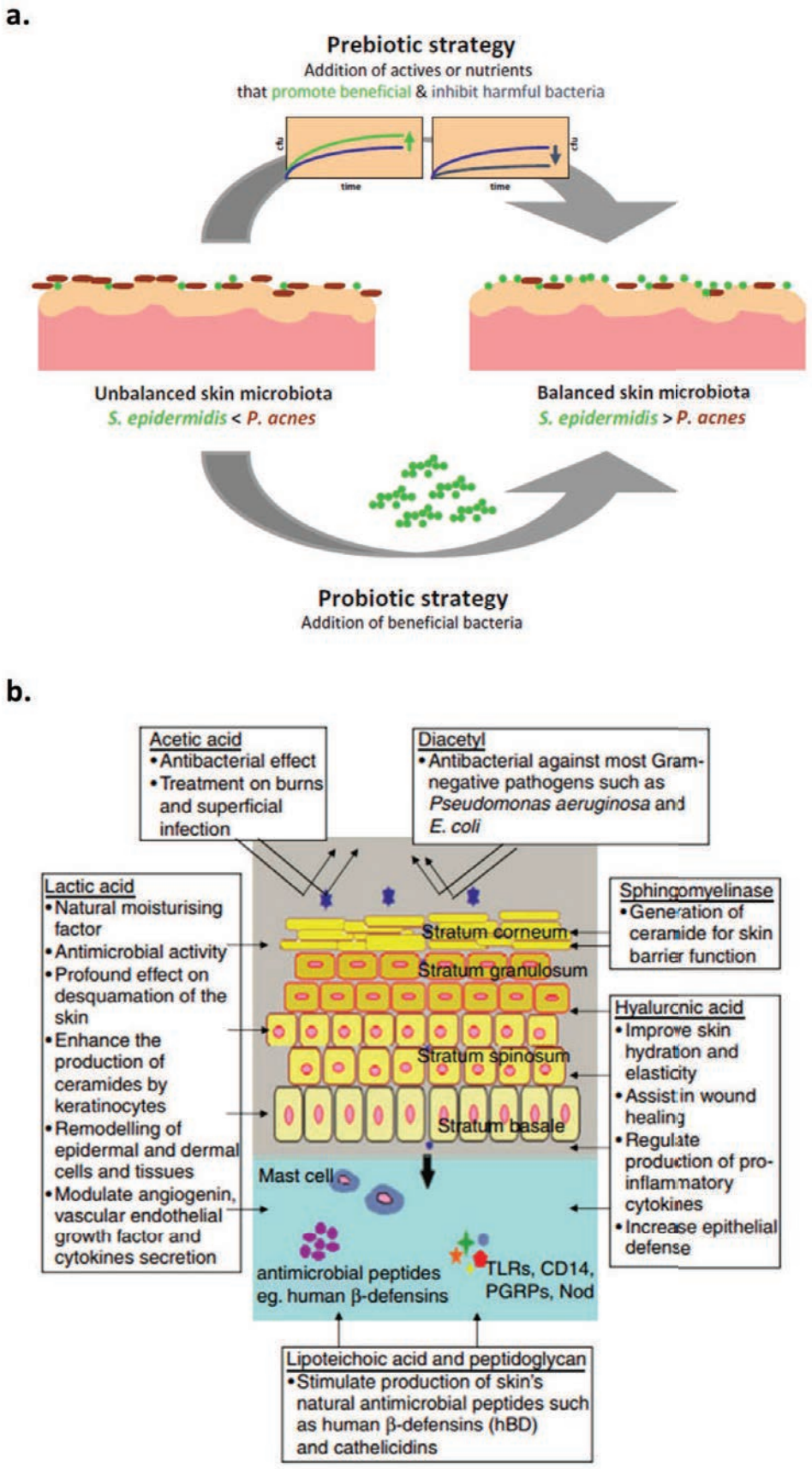


Figure 26. Pre- and probiotics for the skin. a) An example to illustrate the concepts of pre- and pro-biotics for skin and b) bacterial products with beneficial effects on the skin (Figures from (Egert and Simmering, 2016) (a) and (Lew and Liong, 2013) (b)).

When speaking about microbiota used as a therapy, we must include microbiota transplantation from healthy donors to affected receptor. Fecal transplants have already been used in human to treat *Clostridium difficile* recurrent infection since 40 years ago with a success rate up to 92%-100% (Gough et al., 2011), but the underlying microbiota dynamics are just beginning to be understood (Song et al. 2013; Fuentes et al. 2014; Leber et al. 2015; Seekatz et al. 2016; Staley et al. 2017). On the skin of a murine model for atopic dermatitis, Myles and colleagues tried some kind of "microbiota transplantation" –only culturable gram-negative bacteria– that led the mice to an improved outcome (Myles et al., 2016).

Most pre- and probiotics products worked without understanding its dynamics and mainly in the gut. Networks dynamics aim to give light in these mechanisms, and are classified in three types: individual dynamics; group dynamics; and universal dynamics (Figure 27) (Bashan et al., 2016). Nowadays, network studies on microbiota dynamics of healthy individuals elucidated that some body sites presented universal dynamics (similar among the individuals), whereas others were highly personalized (Bashan et al., 2016). In fact, the healthy gut microbiota presented universal dynamics, which could be the reason why most of the pre- and pro-biotics work on this body site. In contrast, among the several skin sites included in this study only the retroauricular crease presented universal dynamics, whereas forehead, palm, and antecubital fossa presented individual dynamics. Thus, generic microbiota manipulations may result ineffective or even detrimental in some body sites (Bashan et al., 2016).

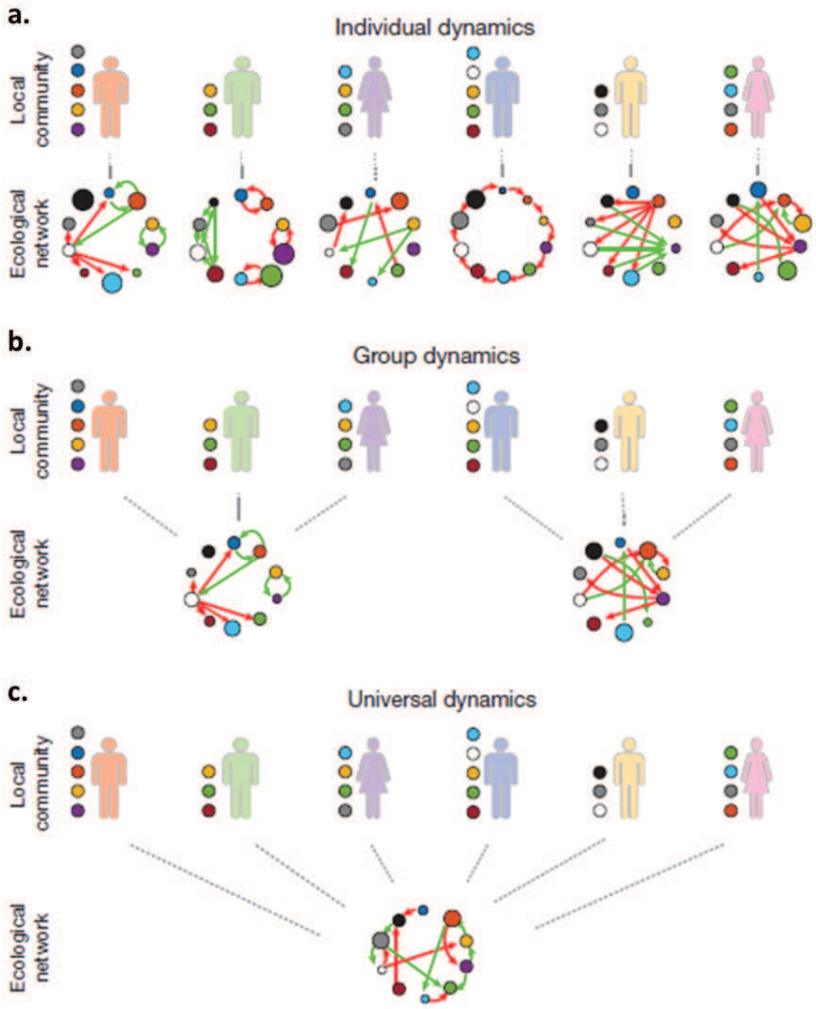


Figure 27. Alternative scenarios of microbial dynamics across different healthy individuals. Nodes represent microbial species and edges represent interspecies interactions (green and red edges represent co-ocurrence or mutual exclusion interactions, respectively). The three main scenarios are: a) individual dynamics, where the underlying dynamics is unique for each individual; b) group dynamics, where dynamics are shared across different groups; and c) universal dynamics, where different subjects have the same underlying dynamics. (Excerpted from Bashan et al. 2016)

To conclude, microbiota analyses have the potential to diagnose, prognose and even treat several skin diseases. To design safe products and accurate diagnoses and prognoses tests, healthy and altered microbiota should be better defined, as well as microbial dynamics should be fully understood.

2. Objectives

The overall goal of this thesis is to expand the knowledge on what defines the healthy status of skin in dogs, first at innate immunity level but mainly characterizing skin microbiota. Moreover, this is an industrial PhD project in collaboration with Vetgenomics, SL., so this knowledge will provide the scientific background and the know-how needed to develop a new strategic business area on Vetgenomics focused on microbiota analyses.

To provide insights on dog innate immune system (Chapter 3.1), we aim to:

- Design and validate an individual genotyping array of non-synonymous polymorphisms on canine Toll-like Receptor (TLR) genes; creating a tool to characterize the innate immune system at TLR level.
- Describe the normal variability of TLR genes on different cohorts of healthy dogs and wolves.

To provide insights on skin microbiota in healthy dogs, we aim to:

- Assess the effect of breed, individual or skin site on shaping skin microbiota in a pilot and heterogeneous cohort (Chapter 3.2).
- Assess the effect of either host or environmental factors (such as skin site, surgery, sex, date of birth date, pigmentation, and time spent in the kennel) on shaping skin microbiota (Chapter 3.3).
- Finally, we aim to assess the potential of the 3rd generation single-molecule sequencing technology for microbiota studies using Nanopore sequencing (Chapter 3.4).

3. Results

This section, divided in four chapters, encompasses a description of the results obtained in four different studies I carried out during the last three years.

Chapter 3.1 describes the genetic variability in canine TLR genes focusing on non-synonymous mutations, by massive sequencing and individual genotyping.

Chapter 3.2 depicts the skin microbiota structure and composition on a heterogeneous cohort of nine healthy dogs, from three different breeds inhabiting different households.

Chapter 3.3 depicts the skin microbiota structure and composition on a homogeneous cohort of thirty-five healthy dogs, from the same cross-breed sharing a common environment.

Chapter 3.4 assesses the potential of sequencing the full-length 16S rRNA gene with MinION™ (3rd generation sequencer) to perform a microbiota study.

3.1. Non-synonymous genetic variation in exonic regions of canine Toll-like receptors

This chapter consists of the article entitled "Non-synonymous genetic variation in exonic regions of canine Toll-like receptors" published in Canine Genetics and Epidemiology journal in October 2014 (1:1).

Supplementary material of this article is available at Canine Genetics and Epidemiology online: https://cgejournal.biomedcentral.com/articles/10.1186/2052-6687-1-11

Cuscó et al. Canine Genetics and Epidemiology 2014, 1:11
http://www.cgejournal.org/content/1/1/11

Canine Genetics and Epidemiology

RESEARCH Open Access

Non-synonymous genetic variation in exonic regions of canine Toll-like receptors

Anna Cuscó[1,2], Armand Sánchez[1], Laura Altet[2], Lluís Ferrer[3] and Olga Francino[1*]

Abstract

Background: Toll-like receptors (TLRs) are pattern recognition receptors (PRRs) considered to be the primary sensors of pathogens in innate immunity. Genetic variants could be associated to differences in breed innate immune response to pathogens and thus to susceptibility to infections or autoimmune diseases. There is therefore great interest in the characterization of canine TLRs.

Results: Polymorphisms in canine TLRs have been characterized by massive sequencing after enrichment of their exonic regions. DNAs from 335 dogs (seven different breeds) and 100 wolves (two different populations) were used in pools. The ratio of SNP discovery was 76.5% (in relation to CanFam 3.1); 155 out of 204 variants identified were new. Functional annotation identified 64 non-synonymous variants (43 new), 73 synonymous variants (56 new) and 67 modifier variants (57 new). 12 out of 64 non-synonymous variants are breed or wolf specific. TLR5 has been found to be the most polymorphic among canine TLRs. Finally, a TaqMan OpenArray® plate containing 64 SNPs with a possible functional effect in the protein (4 frameshifts and 60 non-synonymous codons) has been designed and validated.

Conclusions: Non-synonymous genetic variation has been characterized in exonic regions of canine Toll-like Receptors. The TaqMan OpenArray® plate developed to capture the individual variability that affects protein function will allow high-throughput genotyping either to study association to infection susceptibility or even TLR evolution in the canine genome.

Keywords: TLRs, Toll-like receptor, Polymorphism, SNPs, Non-synonymous SNPs, Canine, Dog, Innate immunity

Lay summary

Toll-like receptors (TLRs) are pattern recognition receptors (PRRs) and are the primary sensors of pathogens in the body. Genetic variants could be associated with differences in breed response to pathogens and also to susceptibility to infections and/or autoimmune diseases. There is great interest in the characterization of canine TLRs.

Genetic variation in canine TLRs has been characterized using massive parallel sequencing. DNA from 335 dogs (seven breeds: Beagle, German Shepherd dog, Yorkshire terrier, French bulldog, Boxer, Labrador and Shar Pei) plus 100 wolves (two populations: Iberian and Russian) were sequenced in 16 pools of 25 dogs or 50 wolves. In total, we found 204 variants, of which 155 were new. Comparison of these variants with the published dog genome sequence (called CanFam 3.1) Functional annotation identified 64 non-synonymous variants (43 new), 73 synonymous variants (56 new) and 67 modifier variants (57 new). Twelve of 64 non-synonymous variants were breed or wolf specific. TLR5 has been found to be the most polymorphic among canine TLRs. Finally, a TaqMan OpenArray(R) plate containing 64 SNPs with a possible functional effect in the protein (4 frameshifts and 60 non-synonymous codons) has been designed and validated.

Non-synonymous genetic variation has been characterized in exonic regions of canine Toll-like Receptors. The TaqMan OpenArray(R) plate developed to capture the individual variability that affects protein function will allow high-throughput genotyping either to study association to infection susceptibility or even TLR evolution in the canine genome.

* Correspondence: Olga.Francino@uab.cat
[1]Molecular Genetics Veterinary Service. Veterinary School, Universitat Autònoma de Barcelona, Barcelona, Spain
Full list of author information is available at the end of the article

Background

Toll-like receptors (TLRs) are the most widely studied pattern recognition receptors (PRRs) and are considered to be the primary sensors of pathogens in innate immunity. These molecules are constituted by leucine-rich repeat (LRR) domains, a unique intramembrane domain and a Toll/Interleukin-1 receptor (TIR) domain. Pathogen-associated molecular Patterns (PAMPs) are sensed through LRR domain, and signals are transduced through TIR domain, which is always located in the cytoplasm, in order to activate innate immunity response (for a review, see [1]).

Ten TLRs have been identified in dogs. TLRs can be classified into groups, depending on the PAMPs detected and their cellular location. TLR 1, 2, 4, 5 and 6 detect pathogen extracellular components. TLRs 3, 7, 8 and 9 target nucleic acids. The ligand for TLR10 is unknown [2].

Another way to classify TLRs is their cellular location. TLRs 1, 5, 6 and 10 are expressed at the cell surface and mainly recognize bacterial products. On the other hand, TLRs 3, 7, 8 and 9 are located almost exclusively in intracellular compartments and are specialized in recognition of nucleic acids, with self versus non-self discrimination provided by the exclusive localization of the ligands rather than their different molecular structure from that of the host. TLRs 2 and 4 can be located both on the cell surface and intracellular [2,3]. In this study, TLRs will be divided in two groups: TLRs 1, 2, 4, 5, 6 and 10 as extracellular TLRs and TLRs 3, 7, 8 and 9 as intracellular TLRs and nucleic acid sensors.

TLRs are conserved through evolution, from Drosophila to mammals (reviewed at [4]), because of its essential role in innate immunity. However, there are significant distinctions between intracellular and extracellular TLRs. Intracellular TLRs do not accept much variability, because they have evolved under strong purifying selection [5]. Viruses can only be detected through their nucleic acids; therefore intracellular TLRs have an essential non-redundant role in host survival. Moreover, mutations in those TLRs could end up with an autoimmune disease against own nucleic acids or with high susceptibility to some viral infections. On the other hand, membrane or extracellular TLRs have evolved under less evolutionary pressure, due to they can recognize one pathogen through different PAMPs (immunological redundancy). So they show a higher rate of damaging non-synonymous and STOP mutations.

Although infective pressure that has reached these molecules is one of the main mechanisms of evolution, it is not the only one. Non-adaptative evolution has also an important role, through genetic drift, bottlenecks and migratory routes [6]. This kind of evolution should be seen in dogs, due to a first bottleneck with domestication and a second one for the artificial selection of breeds [7]. For these reason it should be taken into account the need for dealing with different breeds, and even with the wolf, for the analysis of canine TLR polymorphism.

In humans, many studies are addressed to find out possible links between some TLR polymorphism and susceptibility or resistance to disease (for a review see [6]). Some genetic variants in TLRs in dogs could be associated to differences in breed innate immune response to pathogens and thus to susceptibility to infections or autoimmune diseases. So far, polymorphisms in TLR4 and TLR5 have been associated with Inflammatory Bowel disease (IBD) in German Shepherd dogs (GSD) [8], but only protective SNPs from TLR5 have been associated with IBD in other 38 dog breeds [9] There is therefore great interest in the characterization of canine TLRs. TLR5 risk-associated haplotype for canine IBD confers hyper-responsiveness to flagellin [10]. Moreover, dogs with spontaneous IBD exhibit alterations in the enteric microbiota, which bear resemblance to dysbiosis reported in humans with chronic intestinal inflammation [11].

Although no other polymorphisms have been associated to illness in dogs until date, some studies have reported differential expression of some TLRs related to inflammatory or infectious diseases, such as TLR2 in IBD [12], TLRs 2, 4, 5 and 9 in chronic enteropathies in German Shepherd [13,14]; TLR4 in osteoarthritis [15] and in infected canine endometrium [16]; TLRs 1-4, 6-10 in sino-nasal aspergillosis and idiopathic lymphoplasmacytic rhinitis [17]; and TLR2 and TLR9 in *Leishmania* infected dogs [18,19].

So our aim is the analysis of genetic variation in exonic regions of canine TLRs by massive sequencing, focusing in non-synonymous substitutions and their segregation in different dog breeds and wolf populations. A second objective is to design and validate a TaqMan OpenArray® plate of SNPs with a possible functional effect in the protein (STOP, frameshift and non-synonymous codons). High-throughput genotyping of canine TLRs with this TaqMan OpenArray® plate will allow studying the association of non-synonymous variants with individual differences in immune response, their relationship with either the commensal or the disease associated microbiota and TLR evolution in the canine genome.

Results

We have identified 156 new variants in canine TLRs by massive sequencing after the enrichment of exonic regions. DNAs from 335 dogs (seven breeds) and 100 wolves (two populations) were pooled in 16 pools and sequenced in 2 lanes of Illumina Hiseq, with a mean coverage value of 15,162.23. Dog breeds included were Beagle, Labrador, German Shepherd, Yorkshire, French Bulldog, Boxer and Shar Pei. Wolves included were Iberian (*Canis lupus signatus*) and Russian (European grey wolf, *Canis lupus lupus*). A total of 204 variants were detected: 193 SNP and

11 insertions or deletions (1 to 18 bases). Only one of the indels (insertion/deletion) mapped to an exonic region (TLR7 3′ UTR), meanwhile the others were mapping to intronic regions (5 out of 11) and intergenic regions upstream or downstream a TLR gene (5 out of 11). The SNPs identified were classified by functional annotation from ENSEMBL [20] (effect and effect impact): 73 synonymous variants, 64 non-synonymous variants and 67 modifier variants which include intergenic (upstream and downstream a TLR gene), intronic and 3′ UTR (untranslated region) variants (see Table 1). None of the variants detected in the pools analyzed had a high effect (STOP codon, frameshift mutation or splicing) on the protein function. The ratio of SNP discovery was 76.5% (in relation to CanFam 3.1); 156 out of 204 variants identified were new: 43/64 non-synonymous variants (nsSNP), 56/73 synonymous variants (synSNP) and 57/67 modifier variants.

Genetic variation differs among all TLRs. Variants detected in either extracellular or intracellular canine TLRs by massive sequencing and its classification according their effect in the protein are shown in Table 1. TLR5 gene presents the highest polymorphism, with 28 synonymous changes and 23 non-synonymous changes (Additional file 1, Table 1), although it also codifies for the longest annotated protein (1422 aminoacids).

Table 2 shows the aminoacid (AA) change ratio, which are AA changes caused by nsSNPs or frameshift mutations divided by total number AA for each one of the TLRs. The AA change ratio confirms that indeed TLR5 and TLR4 are the most polymorphic ones. On the other hand, TLR3 seem to be the most conserved receptor, just presenting one AA change in 905 AA.

Non-synonymous SNPs

A more exhaustive analysis was performed for the 64 nsSNP detected through massive sequencing, because they are expected to have a greater effect on the protein function. First, a glimpse on allelic frequencies of the nsSNP was performed. The frequencies of the alternative allele for all 64 nsSNPs are shown for each breed and wolf pools in Additional file 2.

Allelic frequencies for alternative variants in nsSNPs differ among breeds. Beagle and Russian wolf are the most variable pools, with 35 out of 64 nsSNPs segregating. Some of the variants identified are breed-specific (8 out of 64) or wolf-specific (4 out of 64). Most of the breed specific variants are found in TLR5 and TLR4, which as seen before, are the most polymorphic TLRs. German Shepherd dogs (GSD) and wolf share 3 nsSNPs, all located in TLR4. The same happens with Shar Pei and wolf, they share 3 nsSNPs located in TLR2, TLR5 and TLR6.

SNPs with a MAF (Minor allele frequency) <0.05 have been considered to be fixed in the cohort [21]. Usually it is the reference allele the one which is fixed, but in some cases (perhaps due to bad annotation of the SNP) is the alternative one. Iberian wolves' cohort is the one with more fixed variants, with only 24 out of 64 that are segregating, followed by Yorkshire and Boxer, with 25 out of 64 segregating variants.

Predicted impact of canine TLRs amino acid substitutions

Polyphen-2 [22,23], SIFT [24,25], and PROVEAN [26,27] tools were used in order to predict the effect of each nsSNP in the protein structure. Each of these tools uses a different algorithm to predict the consequence of the aminoacid change on the protein and classifies it as benign/tolerated/neutral or damaging/affect protein function/deleterious (for more detail, see Methods). 28 out of 64 nsSNPs were predicted to have an effect on the protein structure by at least one of the tools used (Table 3). When frequency of the alternative variant was high for all the cohorts tested, the alternative allele was exchanged with the reference allele in ENSEMBL sequences [20] in order to perform the Polyphen-2 analysis with the less frequent variant as the "alternative variant". Therefore, SNPs with

Table 1 Variants detected in canine TLRs by massive sequencing

		Extracellular TLRs						Intracellular TLRs				
Effect impact	**SNP effect**	**TLR1**	**TLR2**	**TLR4**	**TLR5**	**TLR6**	**TLR10**	**TLR3**	**TLR7**	**TLR8**	**TLR9**	**Total**
Low	Syn coding	2	5	4	28	2	5	6	6	8	7	73
Moderate	Non-syn coding	4	3	12	23	4	3	1	3	4	7	64
Total coding SNP (cSNP)		***6***	***8***	***16***	***51***	***6***	***8***	***7***	***9***	***12***	***14***	***137***
Modifier	Downstream	1	4	0	0	2	3	0	0	2	1	12
	Intron	0	0	6	3	0	0	10	7	0	2	28
	Upstream	0	6	0	0	2	1	1	1	1	0	12
	UTR 3′	0	0	4	0	0	0	0	10	0	0	14
Total non coding SNP (ncSNP)		***1***	***10***	***10***	***3***	***4***	***4***	***11***	***18***	***3***	***3***	***66***
Total SNP		***7***	***18***	***26***	***54***	***10***	***12***	***18***	***27***	***15***	***17***	***204***

Variants are classified according to their effect on the protein and their spread along cell surface or intracellular TLRs.

Table 2 Total number of variants affecting protein in extracellular and intracellular TLRs

Canine gene	ENSEMBL protein ID	Protein length (aa)	AA change ratio[a]
Extracellular TLRs			
TLR1	ENSCAFP00000032660	790	1/113
TLR2	ENSCAFP00000012269	785	1/196
TLR4	ENSCAFP00000031395	833	1/69
TLR5	ENSCAFP00000016726	1422	1/53
TLR6	ENSCAFP00000023836	797	1/199
TLR10	ENSCAFP00000023840	807	1/269
Intracellular TLRs			
TLR3	ENSCAFP00000011004	905	1/905
TLR7	ENSCAFP00000017193	1121	1/374
TLR8	ENSCAFP00000031505	1038	1/260
TLR9	ENSCAFP00000030804	1032	1/129

Variants from CanFam 3.1 have been added to variants identified by massive sequencing in this table. [a]AA change ratio: aminoacid changes caused by nsSNPs or frameshift mutations divided by the length of the protein in aminoacids.

frequencies greater than 0.25 for the alternative allele were tested also for the annotated reference allele. Then, 3 more SNPs were predicted to affect the protein structure (indicated as reference on the column dbSNP ID in Table 3). When considering also these ones, 31 out of 64 nsSNPs (48%) were predicted to have an impact on the protein structure. Results from Polyphen-2, SIFT and PROVEAN were convergent in predicting damaging effects for 8 out of 31 nsSNPs (27%). On the other hand, 6 out of 64 nsSNPs were not correctly predicted, giving unknown or low confidence results, because they were not aligning to enough similar sequences to give a reliable result. Curiously most of this nsSNPs were located on the N-terminal region of TLR5.

Protein structure of the canine TLRs was assessed using SMART [28,29], which predicts domains taking into account aminoacid sequences: 6 out of 31 nsSNPs predicted to be damaging in canine TLRs were found to be in a Leucine Rich Repeat C-terminal (LRRCT) or really close to it. Only 1 out of 31 was found to affect TIR domain in TLR 5, other 2 were found to be really close to this domain in TLR5 and 10. With the exception of these last ones, nsSNPs were in most cases located in the sensor domain of TLRs (Table 3).

Frequencies in Table 3 are an average of all dog pools tested and both wolf populations respectively, so all variants are polymorphic (MAF > 0.05) at least in one breed. 17 out of 31 show a MAF > 0.05 when considering the average frequencies in all the pools together (15 out of 31 with MAF > 0.05 in wolf populations). However, as mentioned above, frequencies of nsSNPs differ among breeds (see Additional file 2). It's worthy to note the differences on the alternative allele frequency observed for the 8 nsSNPs that were predicted to affect protein function by the three tools used (Figure 1).

TaqMan open array design and SNP validation

A TaqMan OpenArray® plate has been developed for the validation of the nsSNPs by individual genotyping (Table 4). This panel contains (i) 27 out of 31 nsSNPs that were predicted to have an impact on the protein structure (4 wolf-specific SNPs were not considered for the array: TLR1 A525V, TLR5 N833K, TLR6 P579L and TLR10 F787L; see Table 3); (ii) 28 out of the 33 remaining nsSNPs segregating in dogs (5 SNPs that were not suitable for a correct primer design were rejected for posterior analysis: TLR4 T36A, TLR4 T36I, TLR5 T243A, TLR5 Q213R and TLR9 A442V); and (iii) 4 frameshift and 4 non-synonymous TLR polymorphisms described on CanFam 3.1 but not detected in our cohorts (see Table 4). One of the non-synonymous variants added (rs23572381, TLR1 N634K) was designed with two different TaqMan assays due to the presence of other variants close to the interrogated SNP.

A total of 99 DNA samples of the first massive sequencing pools were chosen to be individually genotyped in order to validate the SNPs with the TaqMan Open Array® designed: 15 Beagle, 15 Boxer, 14 French bulldog, 15 Labrador, 15 German Shepherd dog, 13 Yorkshire and 12 Shar Pei were used. One Shar-Pei and 2 Yorkshires do not pass the quality control for samples (call rate > 0.9) and were removed from the posterior analysis. Finally, analysis was performed with a total of 96 individuals. Fifty-nine out of the 64 SNPs (92%) included in the OpenArray have been successfully validated and all of them had a call rate greater than 0.9.

Some downstream analyses have been performed with the individual genotypes. However, it should be taken into account that these are just preliminary results, which need to be validated with larger cohorts of dogs.

All the TLR SNPs were in Hardy-Weinberg Equilibrium (HWE) on Beagle, Boxer, German Shepherd, Labrador and

Table 3 Non-synonymous SNPs predicted to impact protein function either by Polyphen-2, SIFT or PROVEAN

Canine gene	Position	SNP	dbSNP ID	AA Subst	Protein domain[a]	Polyphen-2 result	SIFT result	Provean result	Variant freq (dog)[b]	Variant freq (wolf)[b]
EXTRACELLULAR TLRs										
TLR1	3:73542337	G/T	rs23585044	S29I	ncp	Pos. damaging	Tolerated	Neutral	0,36	0,77
	3:73543092	T/G	new	S281A	ncp	Pos. damaging	Tolerated	Neutral	0,06	0,00
	3:73543825	C/T	new	A525V	LRRCT[2]	Pos. damaging	Tolerated	Deleterious	0	0,11
TLR2	15:51463020	C/A	rs22410121	S46Y	ncp	Pos. damaging	Tolerated	Neutral	0,10	0,00
	15:51464430	***C/T***	***new***	***S516L***	***ncp***	***Prob. damaging***	***Aff. function***	***Deleterious***	***0,14***	***0***
TLR3	16:44623632	C/G	new	E176D	ncp	Pos. damaging	Tolerated	Neutral	0,16	0,12
TLR4	11:71356420	C/T	reference[1]	A8V	ncp	Prob. damaging	Tolerated	Neutral	0,77	0,57
	11:71360887	G/A	new	V82M	ncp	Pos. damaging	Tolerated	Neutral	0,09	0,15
	11:71364581	***T/C***	***rs22145736***	***L167P***	***ncp***	***Prob. damaging***	***Aff. function***	***Deleterious***	***0,15***	***0***
	11:71364681	A/C	reference[1]	Q200H	ncp	Pos. damaging	Aff. function	Neutral	0,88	0,23
	11:71365810	A/G	new	T577A	LRRCT[3]	Pos. damaging	Aff. function	Neutral	0,01	0
TLR5	38:23702837	C/T	rs9070447	R269C	ncp	Prob. damaging	Aff. function	Neutral	0,19	0,01
	38:23702918	G/A	new	V296I	ncp	Pos. damaging	Tolerated	Neutral	0,05	0,26
	38:23703629	G/A	new	G533S	ncp	Prob. damaging	Tolerated	Neutral	0,02	0
	38:23704331	G/T	new	D767Y	ncp	Prob. damaging	Tolerated	Deleterious	0,04	0
	38:23704531	C/G	new	N833K	LRRCT	Pos. damaging	Tolerated	Deleterious	0	0,06
	38:23704562	C/T	new	R844C	LRRCT	Pos. damaging	Tolerated	Neutral	0,04	0,39
	38:23704581	***C/T***	***reference[1]***	***S850L***	***LRRCT***	***Prob. damaging***	***Aff. function***	***Deleterious***	***0,68***	***0,02***
	38:23704695	***T/G***	***new***	***F888C***	***low complexity***	***Prob. damaging***	***Aff. function***	***Deleterious***	***0,04***	***0***
	38:23705081	C/T	new	H1017Y	TIR	Benign	Aff. function	Neutral	0,02	0
	38:23705264	G/A	new	A1078T	TIR[4]	Pos. damaging	Aff. function	Neutral	0,07	0,00
TLR6	3:73521250	A/G	new	Y182C	ncp	Prob. damaging	Tolerated	Deleterious	0,01	0,09
	3:73522074	***C/T***	***new***	***L457F***	***ncp***	***Prob. damaging***	***Aff. function***	***Deleterious***	***0,03***	***0***
	3:73522242	G/A	rs23570247	D513N	ncp	Pos. damaging	Tolerated	Neutral	0,73	1,00
	3:73522441	***C/T***	***new***	***P579L***	***LRRCT***	***Pos. damaging***	***Aff. function***	***Deleterious***	***0,01***	***0,07***
TLR10	***3:73569402***	***C/T***	***rs23518574***	***T361M***	***ncp***	***Prob. damaging***	***Aff. function***	***Deleterious***	***0,13***	***0,12***
	3:73570681	T/A	new	F787L	TIR[5]	Pos. damaging	Low confidence	Neutral	0,00	0,39
INTRACELLULAR TLRs										
TLR8	***X:9397240***	***T/C***	***new***	***V157A***	***ncp***	***Pos. damaging***	***Aff. function***	***Deleterious***	***0,06***	***0***
TLR9	20:37544129	G/A	new	V87I	ncp	Benign	Aff. function	Neutral	0,02	0
	20:37546230	C/T	new	P787L	ncp	Pos. damaging	Tolerated	Neutral	0,22	0,24
	20:37546454	C/T	new	R862W	ncp	Prob. damaging	Tolerated	Neutral	0,2	0

In italics, SNPs that are predicted to have an effect on protein function by the three algorithms. [a]ncp, no confident prediction. [b]Observed frequency by massive sequencing. [1]reference allele tested as the alternative in the SNP. [2]Leucine Rich Repeat C-terminal (LRRCT) domain predicted from aminoacid 528 to 582. [3]LRRCT domain predicted from aminoacid 579 to 629. [4]TIR domain predicted from aminoacid 927 to 1074. [5]TIR domain predicted from aminoacid 641 to 784.

Shar-Pei. In Yorkshire, TLR10 has two SNPs in linkage disequilibrium which are not in HWE, one of them is predicted to affect protein function by the algorithms tested (Table 3). French Bulldog was the breed that had more SNPs that did not follow HWE proportions, with 3 SNPs in TLR4 and one SNP in TLR5 (Table 5). TLR7 and 8 were not included because they are both located in chromosome X.

Principal components analysis (PCA) combined data from the individual genotypes obtained for the subset of SNPs which were not in linkage disequilibrium. It was used to illustrate if dogs cluster by breed for genetic variants in TLRs. The first two components from the PCA have been plotted in Figure 2. Visual examination of this plot shows overlapping for most breeds, excluding Labrador

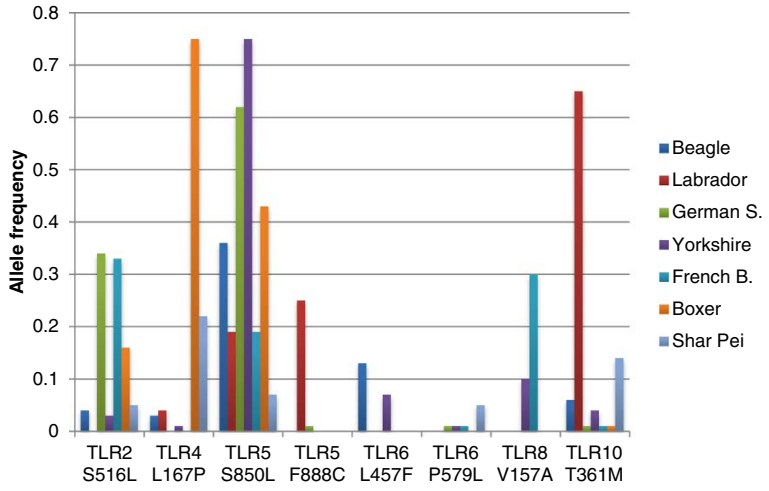


Figure 1 Breed allelic frequencies for the 8 nsSNP with a damaging prediction from Polyphen-2, SIFT and PROVEAN.

and perhaps German Shepherd, which seem to be more differentiated for these receptors.

Discussion

Canine breed specific variants in TLRs could be associated to differences in breed innate immune response to pathogens and thus to susceptibility to infections or autoimmune diseases. So far, polymorphisms in TLR4 and TLR5 have been associated with IBD in German Shepherd dogs [8], but only protective SNPs from TLR5 have been associated with IBD in other 38 dog breeds [9]. There is therefore great interest in the characterization of canine TLRs. Different dog breeds and 2 different populations of wolves (Iberian and Russian) were included in the analysis to represent some of the major phylogenetic radiations: Wolves, Ancient&Spitz breeds, Scent hounds, Working dogs, Mastiff-like dogs, Small Terriers and Retrievers [30]. A total of 204 variants have been discovered and functionally annotated in exonic regions of canine TLRs by massive sequencing: 155 of the variants were new in relation to the most recent annotation of the canine genome (CanFam 3.1; September 2012). Variants have been functionally annotated and correspond to 64 non-synonymous variants (43 new), 73 synonymous variants (56 new) and 67 modifier variants (57 new). None of the variants detected in the pools analyzed had a high effect (STOP codon, frameshift mutation or splicing) on the protein function, although 4 frameshift mutations are annotated in CanFam 3.1.

SNPs functionally annotated as non-synonymous are expected to have a greater effect on protein function, and therefore a more exhaustive analysis was performed on them. Although allelic frequencies for nsSNPs differ among breeds and 12.5% of them are breed-specific (6.25% are wolf specific), dogs from different breeds share most non-synonymous variants in TLRs.

A TaqMan OpenArray® plate containing 64 SNPs with a possible functional effect in the protein (4 frameshifts and 60 nsSNPs) has been designed and validated. 55 out of 64 SNPs contained in the OpenArray® plate have been identified in this work through massive sequencing by HISEQ; the remaining 9 were obtained from CanFam 3.1.

As shown in Figure 2, the individual genotypes are not clustering by breed, with the exception of Labrador and German Shepherd dogs.

The functional impact of non-synonymous variants in dog TLRs was predicted using Polyphen-2, SIFT and PROVEAN. Knowing that TLRs are highly conserved receptors, it is not unexpected that half non-synonymous mutations in dogs have a benign effect, which agrees with results from similar approaches in other non-primate species such as bovine [31]. In dogs, TLR5 is the one that presents more damaging non-synonymous mutations (possibly damaging + probably damaging), followed by TLR4, both of them extracellular receptors.

Results from SIFT and Polyphen-2 from some nsSNPs located in TLR5 returned no output and no prediction ("unknown" or "low confidence"). In dogs, TLR5 was described as a longer protein compared to their homologs in other species. In CanFam 3.1 TLR5 has 1422 aminoacids, however other species like human, cow and pig have 858 aa, 858 aa and 856 aa, respectively. A protein BLAST was performed with the extra 5′ and 3′ TLR5 fragments, but no result was obtained. Furthermore, the 5′ sequence begins with ATG codon in the same phase as the initial coding ATG in other species, whereas the 3′ sequence eliminates the STOP codon due to some repeats in tandem (data not shown). So, bad annotation of this gene in CanFam3.1 is suggested. However, SNPs have been found in this region. In fact, there are 2 SNPs that had already been wrongly described as an aminoacid change

Table 4 Non-synonymous SNPs and frameshift mutations of canine TLRs in the TaqMan Open Array plate

Canine gene	SNP	Chr:bp position	dbSNP ID	AA Subst	Previous detected[a]	Validated?
TLR1	G/T	3:73542337	rs23585044	S29I	Massive seq	YES
	T/G	3:73543092	new	S281A	Massive seq	YES
	G/A	3:73543185	new	V312I	Massive seq	YES
	T/A	3:73544153	rs23572381	N634K[1]	CanFam 3.1	YES[2]
	T/A	3:73544153	rs23572381	N634K[1]	CanFam 3.1	NO
	G/A	3:73544221	rs23572380	S657N	CanFam 3.1	YES[2]
TLR2	C/A	15:51463020	rs22410121	S46Y	Massive seq	YES
	A/0	15:51464076	rs8958543	A398-	CanFam 3.1	YES[2]
	C/T	15:51464430	new	S516L	Massive seq	YES
	C/T	15:51464700	new	T606M	Massive seq	YES
TLR3	C/G	16:44623632	new	E176D	Massive seq	YES
TLR4	T/C	11:71356420	rs22120766	V8A	Massive seq	NO
	G/C	11:71360743	rs22157966	A34P	Massive seq	YES
	G/A	11:71360887	new	V82M	Massive seq	YES
	T/C	11:71364581	rs22145736	L167P	Massive seq	YES
	C/A	11:71364681	rs22189454	H200Q	Massive seq	YES
	A/G	11:71364769	rs22189456	K230E	Massive seq	YES
	G/A	11:71365120	new	A347T	Massive seq	YES
	A/T	11:71365652	rs22124023	E524V	Massive seq	YES
	A/G	11:71365810	new	T577A	Massive seq	YES
	G/A	11:71365888	rs22123995	E603K	Massive seq	YES
TLR5	G/A	38:23702193	rs24029590	G54E	CanFam 3.1	NO
	0/C	38:23702251	rs9070448	-74C	CanFam 3.1	YES[2]
	A/G	38:23702514	rs9070450	Y161C	CanFam 3.1	NO
	A/C	38:23702539	new	E169D	Massive seq	YES
	G/A	38:23702562	new	S177N	Massive seq	YES
	G/C	38:23702640	rs9070451	R203P	Massive seq	YES
	T/C	38:23702684	rs9070452	W218R	Massive seq	NO
	C/T	38:23702837	rs9070447	R269C	Massive seq	YES
	G/A	38:23702918	new	V296I	Massive seq	YES
	T/C	38:23703180	new	L383S	Massive seq	YES
	G/A	38:23703237	new	R402Q	Massive seq	YES
	G/A	38:23703279	new	R416Q	Massive seq	YES
	T/0	38:23703591	rs9125247	T520-	CanFam 3.1	YES[2]
	G/A	38:23703629	new	G533S	Massive seq	YES
	G/A	38:23704233	new	R734Q	Massive seq	YES
	G/T	38:23704331	new	D767Y	Massive seq	YES
	C/T	38:23704562	new	R844C	Massive seq	YES
	T/C	38:23704581	rs24029975	L850S	Massive seq	YES
	T/G	38:23704695	new	F888C	Massive seq	YES
	G/A	38:23704718	new	A896T	Massive seq	YES
	C/T	38:23705081	new	H1017Y	Massive seq	YES

Table 4 Non-synonymous SNPs and frameshift mutations of canine TLRs in the TaqMan Open Array plate *(Continued)*

	G/A	38:23705090	new	G1020S	Massive seq	YES
	G/A	38:23705178	new	R1049Q	Massive seq	YES
	G/A	38:23705264	new	A1078T	Massive seq	YES
TLR6	A/G	3:73521250	new	Y182C	Massive seq	YES
	C/T	3:73522074	new	L457F	Massive seq	YES
	G/A	3:73522242	rs23570247	D513N	Massive seq	YES
TLR7	C/G	X:9334108	new	A16G	Massive seq	YES[2]
	C/A	X:9355727	new	F167L	Massive seq	YES
	C/T	X:9358423	new	P1066L	Massive seq	YES
TLR8	T/C	X:9397240	new	V157A	Massive seq	YES
	G/A	X:9397663	new	R298Q	Massive seq	YES
	G/A	X:9398094	rs24607342	G442S	Massive seq	YES
	G/A	X:9398827	rs24607358	R686H	Massive seq	YES
TLR9	G/A	20:37544129	new	V87I	Massive seq	YES
	0/A	20:37544851	rs9188882	-328A	CanFam 3.1	YES[2]
	A/G	20:37545011	new	K381E	Massive seq	YES
	C/A	20:37545245	new	P459T	Massive seq	YES
	A/G	20:37546031	rs22882109	S721G	Massive seq	YES
	C/T	20:37546230	new	P787L	Massive seq	ND[3]
	C/T	20:37546454	new	R862W	Massive seq	YES
TLR10	C/T	3:73569402	rs23518574	T361M	Massive seq	YES
	A/G	3:73570094	new	M592V	Massive seq	YES

[a]Massive seq indicates a SNP variant detected in our cohorts. An "rs" name is indicated in dbSNP ID if the SNP is annotated in CanFam 3.1. [1]SNP considered twice with a different surrender SNP in order to detect it. [2]Assay has been validated technically, although not genetically because all individuals have only the reference allele. [3]ND (not determined), there are incongruent results: massive sequencing showed that this SNP was present at a frequency of 0.2 in all breeds tested, whereas it has not been genotyped through TaqMan OA plate.

(in ENSEMBL) moreover in this study 5 more SNPs have been detected. So it would be interesting to either determine the existence and functionality of these extra fragments in canine TLR5 cDNA or correctly annotate it in CanFam 3.1.

Intracellular TLRs, which detect nucleic acids, have less nsSNPs (15), moreover these are predicted to be less damaging variants than those identified in extracellular TLRs, suggesting that intracellular TLRs are selectively constrained. TLR9 is the intracellular TLR that accepts more nsSNPs in dogs, but the predicted effect of these nsSNPs is usually benign.

These results agree with previously reported data revealing major differences in the intensity of selection acting upon the different members of the TLR family. Different TLRs differ in their immunological redundancy, reflecting their distinct contributions to host defense [5,32]. Intracellular TLRs act as nucleic acid sensors and have evolved under strong purifying selection, indicating their essential non-redundant role in host survival. Higher rates of

Table 5 SNPs in different breeds that are not in Hardy-Weinberg Equilibrium ($p < 0.05$)

Breed	Canine gene	AA change	SNP	Genotypes[a]	p-value	SNP prediction[b]
Yorkshire	TLR10	T361M	C/T	1/1/9	0.0416978	Prob. damaging
Yorkshire	TLR10	M592V	A/G	1/1/9	0.0416978	Benign
French B.	TLR4	V82M	G/A	4/2/8	0.0099493	Pos. damaging
French B.	TLR4	H200Q	C/A	6/1/7	0.0013535	Pos. damaging*
French B.	TLR4	K230E	A/G	6/1/7	0.0013535	Benign
French B.	TLR5	S177N	G/A	0/11/3	0.0154748	Benign

[a]genotypes, indicate genotype count for reference homozygotes, heterozygotes and alternative homozygotes. [b]SNP prediction, using Polyphen-2 classification. *possibly damaging when reference allele is tested as alternative in the SNP.

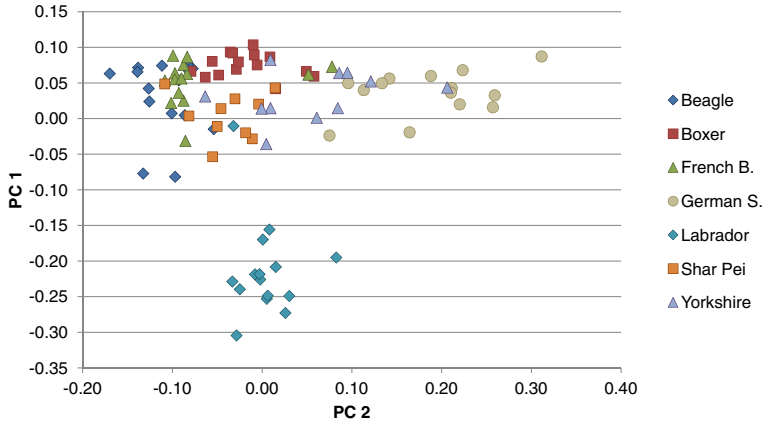


Figure 2 Principal Component Analysis (PCA) plot of the two first components for canine TLRs.

damaging non-synonymous and nonsense mutations are tolerated in cell-surface or extracellular TLRs, which recognize compounds other than nucleic acids, suggesting a higher redundancy.

Location of the SNPs in the protein was approached using the software SMART [28], which identifies TLR domains using the aminoacid sequence. The intracellular TIR domain is highly conserved between different TLRs and species due to its involvement in intracellular signaling [33]. Also in dogs, TIR domains have few SNPs; only one is present in the predicted TIR domain (TLR5 H1017Y) and another two are really close to it (TLR5 A1078T and TLR10 F787L). Extracellular domains of TLRs are those that recognize PAMPs, and they have an enhanced susceptibility to mutate adapting to different microbiologic environments [33]. It can also be seen that a high number of mutations (some with damaging effects) are located in LRR domains, which form the extracellular domain of TLRs.

So far, polymorphisms in TLRs have been associated with Inflammatory Bowel disease (IBD) in German Shepherd dogs (GSD) and in other breeds. Variants in TLR5 previously reported to be associated to IBD (*G22A, C100T and T1844C* from [8,9]) have been also detected in our cohorts and correspond with TLR5 T243A, TLR5 R269C and TLR5 L850S respectively.

SNP *G22A*, where the risk allele is A in *G22A* (corresponding to Thr in TLR5 T243A as named in this work), is found to be an additive allele. So when a GSD is homozygous for the risk allele it has more susceptibility to suffer IBD. This risk allele is not segregating in our GSD cohort. This could be due to the difference in the geographical origin of the GSD cohort between both studies. In [9] GSD are from UK, whilst our cohort is from Spain. SNPs *C100T* and *T1844C* were found to be significantly protective against canine IBD in many breeds [9]. The frequencies of the protective alleles (T in *C100T* and T in *T1844C* or Cys in TLR5 R269C and Leu in TLR5 L850S as named in this work) differ among breeds (Figure 3), with a frequency higher than 0.5 in Yorkshire and GSD.

Conclusions

Polymorphisms in the exonic regions of canine TLRs have been characterized by massive sequencing and 156 out of 204 variants identified were new: 43/64 non-synonymous variants, 56/73 synonymous variants and 57/67 modifier variants. None of the variants detected in the pools analyzed had a high effect (STOP codon, frameshift mutation or splicing) on the protein function.

A TaqMan OpenArray® plate containing 64 SNPs with a possible functional effect in the protein (4 frameshifts and 60 nsSNPs) has been designed and validated to allow the high throughput genotyping of canine TLRs.

Methods

Ethics statement

The dogs in the study were examined during routinary veterinary procedures by the veterinary clinics participating in the study. All samples were collected for routine diagnostic and clinical purposes. The samples were obtained during veterinary procedures that would have been carried out anyway and DNA was extracted from residual surplus of samples and used in the study with verbal owner consent. This is a very special situation in veterinary medicine. As the data are from client-owned dogs that underwent normal veterinary exams, there was no "animal experiment" according to the legal definitions in Spain and the United Kingdom, and approval by an ethical committee was not necessary.

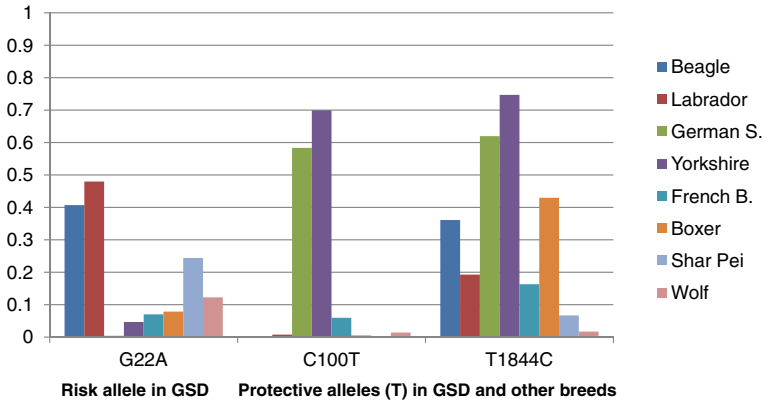


Figure 3 Observed allele frequency of the alleles related with IBD in our pools (A in G22A, and T in both C100T and T1844C).

DNA sources

Samples available from the DNA bank at the SVGM (Molecular Genetics Veterinary Service, UAB) were used. Total DNA from blood cells had been extracted either as described elsewhere [34] or using QIAamp DNA Mini Kit (Qiagen).

DNAs from 7 different dog breeds, including 50 Beagle, 50 German Shepherd, 50 Yorkshire, 35 French bulldog, 75 Boxer, 50 Labrador and 25 Shar Pei were used. All the dogs included in this study are from Spain region, and come from hospital population or normal pet dogs. Also 2 different populations of wolves, with 50 Iberian (*Canis lupus signatus*) and 50 Russian (European grey wolves, *Canis lupus lupus*), were analyzed. DNA pools were prepared with 200 ng of DNA from 25 unrelated dogs (with the exception of one pool of French bulldogs, with only 10 dogs). Two pools of each breed were analyzed, in exception of Shar Pei (only 1 pool) and Boxer (3 pools). Pools of wolves were of 50 individuals.

Some DNA samples of the first massive sequencing analyses were chosen to be individually genotyped in order to validate SNPs in the TaqMan Open Array® designed (15 Beagle, 15 Boxer, 14 French bulldog, 15 Labrador, 15 German Shepherd, 13 Yorkshire and 12 Shar Pei).

Exon capture and massive sequencing for SNP discovery

Twenty exonic regions of 10 canine TLR genes annotated in CanFam 2.0 were chosen to perform the enrichment (see Additional file 3 with corresponding coordinates in CanFam 3.1).

Oligonucleotides were first automatically designed for the enrichment of selected regions [35]. Regions rejected in the automated design, because of the presence of gaps, repeats or shorter sizes than required (at least 120 nucleotides) were manually redesigned. Finally 235 ultra-long 120-mer biotinylated cRNA baits were designed to capture the exonic regions of canine TLRs (28,200 bases) by the Agilent Sure Select technique. High-throughput sequencing was performed using 2 lanes of Illumina HISEQ, with 8-labelled pools each, at CNAG (Centre Nacional d'Anàlisi Genòmica, Barcelona, Spain).

Sequences obtained were mapped to CanFam 3.1 (released September 2012). All pools were analysed together for variant calling, for better comparison. Alternative variant frequencies were estimated for each breed pool and wolf populations. The variants were annotated with statistical information from the Genome Analysis Tool Kit (GATK) and functional annotations were added from Ensembl using snpEff [36].

Prediction of functional impact of non-synonymous SNPs

The functional impact of non-synonymous mutations detected was predicted using Polyphen-2 [22,23], SIFT [25,24] and PROVEAN [27,26]. When the mean frequency of an alternative variant on the dog population analyzed was more than 0.25, both alleles of those SNPs were tested with algorithms mentioned before as reference and alternative.

Polyphen-2 classifies mutations in three categories: *benign*, *possibly damaging* and *probably damaging*. Polyphen algorithm considers protein structure and/or sequence conservation information for each gene [23]. SIFT is based on the evolutionary conservation of the amino acids within protein families performing multiple sequencing analyses using PSI-Blast algorithm. Highly conserved positions tend to be intolerant to substitution, whereas those with a low degree of conservation tolerate most substitutions. Therefore, it classifies each non-synonymous polymorphism as *tolerated* or *affect protein function* and provides also a confidence measure [24]. PROVEAN introduced a region-based "delta alignment score" which measures the impact of an amino acid variation not only based on the amino acid residue at the position of interest but also the quality of sequence alignment derived from the neighborhood flanking sequences. It classifies variants either as neutral or deleterious [26].

SMART was used in order to identify protein domains of each TLR using their aminoacid sequence [28].

TaqMan OpenArray® design

A TaqMan OpenArray® was designed for genotyping and validating 64 SNPs with a possible functional effect in the protein. Selected SNPs and their surrounding sequences, 60 nucleotides upstream and 60 nucleotides downstream were introduced in Custom TaqMan® Assay Design Tool web site [37] from Life Technologies® to validate if the sequences were suitable for TaqMan assay design. Other SNPs in the context sequences were indicated with an "N" before the assays design. SNPs included are listed in Table 4.

Analysis was performed with the TaqMan Genotyper software v.1.3 (Applied Biosystems). Further analysis of individual genotypes was performed with SVS (version 7) of Golden Helix Inc. SNPs or samples that do not pass call rate >0.9 were removed for posterior analysis.

Additional files

Additional file 1: Total number of detected synonymous and non-synonymous SNPs for each canine Toll-like receptor.

Additional file 2: Frequencies of coding variants obtained by massive sequencing. Frequency per breed (50 individuals approximately) and mean frequency per dog and wolf species. Frequency of the SNP is represented as the frequency of the alternative allele.

Additional file 3: Coordinates of exonic regions of 10 canine TLR genes as annotated in CanFam 3.1.

Competing interests

A patent application has been filled in related to the use of some of the markers described in the manuscript.

Authors' contributions

AS, LF and OF designed the experiment. AS, LA, LF and OF supervised the project and gave conceptual advice. LA collected the samples. AC carried out the molecular genetic studies and designed the chip. AC and OF analyzed the data and drafted the manuscript. AS, LA and LF edited the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We acknowledge Lorena Serrano (Vetgenomics) for helping with the collection of the samples; Sophia Derdak and Sergi Beltran (CNAG) for the raw data processing of the massive sequencing results; and Anna Mercadé (SVGM) for technical advice and support on the OpenArray design and validation.

Author details

[1]Molecular Genetics Veterinary Service. Veterinary School, Universitat Autònoma de Barcelona, Barcelona, Spain. [2]Vetgenomics. Ed Eureka. Parc de Recerca UAB, Barcelona, Spain. [3]Department of Clinical Sciences, Cummings School of Veterinary Medicine, Tufts University, North Grafton, MA, USA.

Received: 6 June 2014 Accepted: 23 September 2014
Published: 22 October 2014

References

1. Werling D, Jungi TW: **TOLL-like receptors linking innate and adaptive immune response.** *Vet Immunol Immunopathol* 2003, **91:**1–12.
2. Kawai T, Akira S: **The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors.** *Nat Immunol* 2010, **11:**373–384.
3. Mogensen TH: **Pathogen recognition and inflammatory signaling in innate immune defenses.** *Clin Microbiol Rev* 2009, **22:**240–273.
4. Aderem A, Ulevitch RJ: **Toll-like receptors in the induction of the innate immune response.** *Nature* 2000, **406:**782–787.
5. Barreiro LB, Ben-Ali M, Quach H, Laval G, Patin E, Pickrell JK, Bouchier C, Tichit M, Neyrolles O, Gicquel B, Kidd JR, Kidd KK, Alcaïs A, Ragimbeau J, Pellegrini S, Abel L, Casanova J-L, Quintana-Murci L: **Evolutionary dynamics of human Toll-like receptors and their different contributions to host defense.** *PLoS Genet* 2009, **5:**e1000562.
6. Netea MG, Wijmenga C, O'Neill LAJ: **Genetic variation in Toll-like receptors and disease susceptibility.** *Nat Immunol* 2012, **13:**535–542.
7. Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M, Clamp M, Chang JL, Kulbokas EJ, Zody MC, Mauceli E, Xie X, Breen M, Wayne RK, Ostrander EA, Ponting CP, Galibert F, Smith DR, DeJong PJ, Kirkness E, Alvarez P, Biagi T, Brockman W, Butler J, Chin C-W, Cook A, Cuff J, Daly MJ, DeCaprio D, Gnerre S, *et al*: **Genome sequence, comparative analysis and haplotype structure of the domestic dog.** *Nature* 2005, **438:**803–819.
8. Kathrani A, House A, Catchpole B, Murphy A, German A, Werling D, Allenspach K: **Polymorphisms in the TLR4 and TLR5 gene are significantly associated with inflammatory bowel disease in German shepherd dogs.** *PLoS One* 2010, **5:**e15740.
9. Kathrani A, House A, Catchpole B, Murphy A, Werling D, Allenspach K: **Breed-independent toll-like receptor 5 polymorphisms show association with canine inflammatory bowel disease.** *Tissue Antigens* 2011, **78:**94–101.
10. Kathrani A, Holder A, Catchpole B, Alvarez L, Simpson K, Werling D, Allenspach K: **TLR5 risk-associated haplotype for canine inflammatory bowel disease confers hyper-responsiveness to flagellin.** *PLoS One* 2012, **7:**e30117.
11. Suchodolski JS, Dowd SE, Wilke V, Steiner JM, Jergens AE: **16S rRNA gene pyrosequencing reveals bacterial dysbiosis in the duodenum of dogs with idiopathic inflammatory bowel disease.** *PLoS One* 2012, **7:**e39333.
12. McMahon LA, House AK, Catchpole B, Elson-Riggins J, Riddle A, Smith K, Werling D, Burgener IA, Allenspach K: **Expression of Toll-like receptor 2 in duodenal biopsies from dogs with inflammatory bowel disease is associated with severity of disease.** *Vet Immunol Immunopathol* 2010, **135:**158–163.
13. Burgener IA, König A, Allenspach K, Sauter SN, Boisclair J, Doherr MG, Jungi TW: **Upregulation of toll-like receptors in chronic enteropathies in dogs.** *J Vet Intern Med* 2008, **22:**553–560.
14. Allenspach K, House A, Smith K, McNeill FM, Hendricks A, Elson-Riggins J, Riddle A, Steiner JM, Werling D, Garden OA, Catchpole B, Suchodolski JS: **Evaluation of mucosal bacteria and histopathology, clinical disease activity and expression of Toll-like receptors in German shepherd dogs with chronic enteropathies.** *Vet Microbiol* 2010, **146:**326–335.
15. Kuroki K, Stoker AM, Sims HJ, Cook JL: **Expression of Toll-like receptors 2 and 4 in stifle joint synovial tissues of dogs with or without osteoarthritis.** *Am J Vet Res* 2010, **71:**750–754.
16. Chotimanukul S, Sirivaidyapong S: **Differential expression of Toll-like receptor 4 (TLR4) in healthy and infected canine endometrium.** *Theriogenology* 2011, **76:**1152–1161.
17. Mercier E, Peters IR, Day MJ, Clercx C, Peeters D: **Toll- and NOD-like receptor mRNA expression in canine sino-nasal aspergillosis and idiopathic lymphoplasmacytic rhinitis.** *Vet Immunol Immunopathol* 2012, **145:**618–624.
18. Figueiredo MM, Amorim IFG, Pinto AJW, Barbosa VS, Pinheiro LDJ, Deoti B, Faria AMC, Tafuri WL: **Expression of Toll-like receptors 2 and 9 in cells of dog jejunum and colon naturally infected with Leishmania infantum.** *BMC Immunol* 2013, **14:**22.
19. Melo LM, Perosso J, Almeida BFM, Silva KLO, Somenzari MA, de Lima VMF: **Effects of P-MAPA immunomodulator on Toll-like receptor 2, ROS, nitric oxide, MAPKp38 and IKK in PBMC and macrophages from dogs with visceral leishmaniasis.** *Int Immunopharmacol* 2014, **18:**373–378.
20. Ensembl genome browser 75: **Canis lupus familiaris.** [http://www.ensembl.org/Canis_familiaris/].
21. Nelson MR, Marnellos G, Kammerer S, Hoyal CR, Shi MM, Cantor CR, Braun A: **Large-scale validation of single nucleotide polymorphisms in gene regions.** *Genome Res* 2004, **14:**1664–1668.
22. **PolyPhen-2: prediction of functional effects of human nsSNPs.** [http://genetics.bwh.harvard.edu/pph2/index.shtml]
23. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR: **A method and server for predicting damaging missense mutations.** *Nat Methods* 2010, **7:**248–249.

24. Kumar P, Henikoff S, Ng PC: **Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm.** *Nat Protoc* 2009, **4:**1073–1081.
25. **SIFT.** [http://sift.jcvi.org/]
26. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP: **Predicting the functional effect of amino acid substitutions and indels.** *PLoS One* 2012, **7:**e46688.
27. **PROVEAN.** [http://provean.jcvi.org/index.php]
28. Letunic I, Doerks T, Bork P: **SMART 7: recent updates to the protein domain annotation resource.** *Nucleic Acids Res* 2012, **40**(Database issue):D302–D305.
29. **SMART.** [http://smart.embl-heidelberg.de/]
30. Vonholdt BM, Pollinger JP, Lohmueller KE, Han E, Parker HG, Quignon P, Degenhardt JD, Boyko AR, Earl DA, Auton A, Reynolds A, Bryc K, Brisbin A, Knowles JC, Mosher DS, Spady TC, Elkahloun A, Geffen E, Pilot M, Jedrzejewski W, Greco C, Randi E, Bannasch D, Wilton A, Shearman J, Musiani M, Cargill M, Jones PG, Qian Z, Huang W, *et al*: **Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication.** *Nature* 2010, **464:**898–902.
31. Fisher CA, Bhattarai EK, Osterstock JB, Dowd SE, Seabury PM, Vikram M, Whitlock RH, Schukken YH, Schnabel RD, Taylor JF, Womack JE, Seabury CM: **Evolution of the bovine TLR gene family and member associations with Mycobacterium avium subspecies paratuberculosis infection.** *PLoS One* 2011, **6:**e27744.
32. Wlasiuk G, Nachman MW: **Adaptation and constraint at Toll-like receptors in primates.** *Mol Biol Evol* 2010, **27:**2172–2186.
33. Werling D, Jann OC, Offord V, Glass EJ, Coffey TJ: **Variation matters: TLR structure and species-specific pathogen recognition.** *Trends Immunol* 2009, **30:**124–130.
34. Francino O, Altet L, Sánchez-Robert E, Rodriguez A, Solano-Gallego L, Alberola J, Ferrer L, Sánchez A, Roura X: **Advantages of real-time PCR assay for diagnosis and monitoring of canine leishmaniosis.** *Vet Parasitol* 2006, **137:**214–221.
35. **SureDesign.** [https://earray.chem.agilent.com/suredesign/]
36. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Ruden DM, Lu X: **A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3.** *Fly* 2012, **6:**80–92.
37. **Custom TaqMan® Assay Design Tool - Life technologies.** [https://www5.appliedbiosystems.com/tools/cadt/]

doi:10.1186/2052-6687-1-11
Cite this article as: Cuscó *et al.*: Non-synonymous genetic variation in exonic regions of canine Toll-like receptors. *Canine Genetics and Epidemiology* 2014 **1**:11.

3.2. Individual Signatures Define Canine Skin Microbiota Composition and Variability

This chapter consists of the article entitled "Individual Signatures Define Canine Skin Microbiota Composition and Variability" published in Frontiers in Veterinary Science in February 2017 (4:6).

See Section 3.2.1 Annex 1**: Erratum: Figure 3** to see the newest version of Figure 3,

The Supplementary Material for this article can be found online at:

http://journal.frontiersin.org/article/10.3389/fvets.2017.00006/full#supplementary-material.

frontiers in Veterinary Science

ORIGINAL RESEARCH
published: 06 February 2017
doi: 10.3389/fvets.2017.00006

Individual Signatures Define Canine Skin Microbiota Composition and Variability

Anna Cuscó[1,2], Armand Sánchez[1], Laura Altet[2], Lluís Ferrer[3] and Olga Francino[1]*

[1] *Molecular Genetics Veterinary Service (SVGM), Veterinary School, Universitat Autònoma de Barcelona, Barcelona, Spain,* [2] *Vetgenomics, Ed Eureka, Parc de Recerca UAB, Barcelona, Spain,* [3] *Department of Clinical Sciences, Cummings School of Veterinary Medicine, Tufts University, North Grafton, MA, USA*

Dogs present almost all their skin sites covered by hair, but canine skin disorders are more common in certain skin sites and breeds. The goal of our study is to characterize the composition and variability of the skin microbiota in healthy dogs and to evaluate the effect of the breed, the skin site, and the individual. We have analyzed eight skin sites of nine healthy dogs from three different breeds by massive sequencing of 16S rRNA gene V1–V2 hypervariable regions. The main phyla inhabiting the skin microbiota in healthy dogs are Proteobacteria, Firmicutes, Fusobacteria, Actinobacteria, and Bacteroidetes. Our results suggest that skin microbiota composition pattern is individual specific, with some dogs presenting an even representation of the main phyla and other dogs with only a major phylum. The individual is the main force driving skin microbiota composition and diversity rather than the skin site or the breed. The individual is explaining 45% of the distances among samples, whereas skin site explains 19% and breed 9%. Moreover, analysis of similarities suggests a strong dissimilarity among individuals (R = 0.79, P = 0.001) that is mainly explained by low-abundant species in each dog. Skin site also plays a role: inner pinna presents the highest diversity value, whereas perianal region presents the lowest one and the most differentiated microbiota composition.

Keywords: skin, microbiota, microbiome, dog, canine, coat, skin site, 16S

OPEN ACCESS

Edited by:
Carl James Yeoman, Montana State University, USA

Reviewed by:
Eric Altermann, AgResearch, New Zealand
Suleyman Yildirim, Istanbul Medipol University International School of Medicine, Turkey
Jan Slapeta, University of Sydney, Australia

****Correspondence:***
Anna Cuscó
anna.cusco@vetgenomics.com

Specialty section:
This article was submitted to Veterinary Experimental and Diagnostic Pathology, a section of the journal Frontiers in Veterinary Science

Received: *17 October 2016*
Accepted: *17 January 2017*
Published: *06 February 2017*

Citation:
Cuscó A, Sánchez A, Altet L, Ferrer L and Francino O (2017) Individual Signatures Define Canine Skin Microbiota Composition and Variability. Front. Vet. Sci. 4:6. doi: 10.3389/fvets.2017.00006

INTRODUCTION

The skin is the living interface between an individual and the exogenous environment. It is covered with millions of microorganisms (1) interacting together with hosts' cells and immune receptors to maintain the equilibrium (2). Bacteria are the most abundant microorganisms living on skin surface (3), and their whole population is defined as the microbiota. The high variability of the microbiota in the healthy skin has been captured during the last years using next-generation sequencing techniques [for a review, see Ref. (4)]. Marker-based approaches, mainly using 16S rRNA gene, focus on detecting who is living there—bacterial composition and diversity.

Main phyla inhabiting human skin are Actinobacteria, Firmicutes, Bacteroidetes, and Proteobacteria. A feature of human cutaneous microbiota is the existence of different microhabitats, which are characterized by the predominance of a specific taxa: sebaceous sites (occiput, glabella, alar crease, and manubrium) with *Propionibacterium* spp; moist sites (nare, axilla, and inguinal crease) with *Staphylococcus* and *Corynebacterium* spp; and dry sites (palms and butlock) with gram-negative microorganisms (5). According to the first extensive study reported, dogs harbor mainly the same

phyla as human skin (6). Fusobacteria was also detected as a main phylum, when only considering the paws and the forehead (7) and also in a recent study considering the groins (8). In humans, the variation is higher among different microhabitat skin sites of the same individual than among skin sites from the same microhabitat in different individuals (5, 9). Several differences among skin sites have been described in dogs (6), but to our knowledge, no microhabitats have been defined.

Different factors such as the environment, host genetic variation, lifestyle, or hygiene cause shifts on the microbial communities of the skin (10). These shifts on the microbiota structure and composition could establish a dysbiotic state, which if not recovered could result on a dermatologic affliction. Dysbiosis of the skin microbiota has been associated with several skin afflictions in humans, such as atopic dermatitis (11, 12), psoriasis (13, 14), and acne vulgaris (15). In canine microbiota studies, association between atopic dermatitis and microbiota has been assessed showing less richness on affected animals, either when considering bacteria (6, 16) or fungal communities (17). However, in allergen-induced canine atopic dermatitis, no significant differences on diversity were reported (8). Moreover, recent studies have reported significant increases of *Staphylococcus* and *Corynebacterium* in dogs with this disease (8, 16). Nevertheless, a better characterization of the cutaneous microbiota of healthy dogs seems to be necessary before understanding its role in disease conditions.

There is much less knowledge about the potential functions of the mammals' microbiota. The potential function of a bacterial community can be assessed either directly, using shotgun metagenomics, or indirectly, using 16S data and a predictive software such as Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) (18). Langille et al. used this tool with the Human Microbiome Project dataset (19) obtaining sufficiently accurate results, even for skin samples (18). In canine intestinal microbiota studies, shotgun metagenomics has been used to study microbiota variability when feeding animals with two different diets (20), and PICRUSt was used in dogs suffering idiopathic inflammatory bowel disease (21). To our knowledge, no studies have assessed potential functions of the microbiota at the skin level.

Our aim was to characterize the composition and variability of the skin microbiota on healthy dogs, considering the breed—specially the hair coat—the skin site, and the individual. We sampled nine healthy dogs from three breeds representing the diversity of canine hair coats: French Bulldog (FB; short hair), German Shepherd (GS; long hair with undercoat), and West Highland White Terriers (WHs; wired hair) (22). These three breeds were also selected because they are among the most predisposed to suffer from atopic dermatitis (23). We also aimed to predict the functional profile of the microbiota of different skin sites using PICRUSt.

MATERIALS AND METHODS

Ethics Statement

The dogs in the study were examined during routine veterinary procedures by the veterinary clinics participating in the study. All samples were collected and used in the study with verbal owner consent. As the data are from client-owned dogs that underwent normal preventative veterinary examinations, there was no "animal experiment" according to the legal definitions in Spain, and approval by an ethical committee was not necessary.

Individuals Included and Sample Collection

A cross-sectional study was performed in nine healthy dogs to analyze skin microbiota variability in several skin sites, considering the breed, the hair coat, and the individual. They were all pure-breed dogs ranging from 3 months to 12 years of age and from different households visiting the veterinary clinic for routine procedures (Table S1 in Supplementary Material). All of them lived in urban or periurban environment. Samples from three FBs (FB1, FB2, and FB3), three GSs (GS1, GS2, and GS3), and three West Highland WHs (WH1, WH2, and WH3) were included. Skin samples were collected from eight skin regions: chin, inner pinna, nasal skin, axilla, back, abdomen, interdigital area, and perianal region. These regions are named as 1, 2, 3, 4, 5, 6, 7, and 8, respectively (**Figure 1**). Samples were obtained by firmly rubbing each area using Sterile Catch-All™ Sample Collection Swabs (Epicentre Biotechnologies) soaked in sterile SCF-1 solution (50 mM Tris buffer (pH = 8), 1 mM EDTA, and 0.5% Tween-20). To minimize sample cross-contamination, the person sampling wore a fresh pair of sterile gloves for each individual. Swabs were stored at 4°C until DNA extraction, within the following 24 h.

DNA Extraction

Bacterial DNA was extracted from the swabs using the PowerSoil™ DNA isolation kit (MO BIO) under manufacturer's conditions, with one modification. At the first lysis step, the swab tip with the sponge was cut and introduced in the beads' tube, until the first transference of the supernatant to a new tube. The remaining steps were performed as described by the manufacturer. DNA samples (100 µl) were stored at −20°C until further processing.

To assess for contaminations from the laboratory or reagents, a sterile swab tip was processed in the same conditions as the skin microbiota samples, giving negative results.

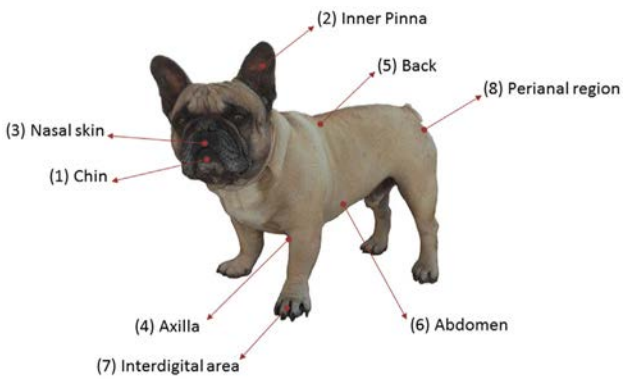


FIGURE 1 | Skin sites sampled per dog.

PCR Amplification and Massive Sequencing

V1–V2 regions of 16S rRNA gene were amplified using the widely used primer pair F27 (5′-AGAGTTTGATCCTGGCTCAG-3′) and R338 (5′-TGCTGCCTCCCGTAGGAGT-3′). PCR mixture (50 uL) contained 5 µl of DNA template (~5 ng), 5 µl of 10× AccuPrime™ PCR Buffer II, 0.2 µM of each primer, and 1 U of AccuPrime™ Taq DNA Polymerase High Fidelity (Life Technologies). The PCR thermal profile consisted of an initial denaturation for 2 min at 94°C, followed by 30 cycles for 1 min at 94°C, 1 min at 55°C, 1 min at 72°C, and a final step for 7 min at 72°C. To assess possible reagent contamination, each PCR reaction included a no template control sample, which did not amplify. For each amplicon, quality and quantity were assessed using Agilent Bioanalyzer 2100 and Qubit™ fluorometer. Both primers included sequencing adaptors at the 5′ end, and forward primers were tagged with different barcodes to pool samples in the same sequencing reaction. Each pool contained 8 barcoded samples. A total of 9 pools were sequenced on an Ion Torrent Personal Genome Machine (PGM) with the Ion 318 Chip Kit v2 and the Ion PGM™ Sequencing 400 Kit (Life Technologies) under manufacturer's conditions. The raw sequences have been deposited in NCBI under the Bioproject accession number PRJNA357691.

Quality Control of the Sequences and Operational Taxonomic Unit (OTU) Picking

Raw sequencing reads were demultiplexed, quality filtered, and analyzed using QIIME 1.9.1 (24). Reads included had a length greater than 300 bp; a mean quality score above 25 in sliding window of 50 nucleotides; no mismatches on the primer; and default values for other quality parameters. Quality-filtered reads were clustered into OTUs at 97% similarity, using UCLUST (25) in an open reference approach for taxonomy analyses and a closed reference approach for functional profiling. Taxonomic assignment of representative OTUs was performed using the RDP Classifier (26) against Greengenes v13.8 database (27). Alignment of sequences was performed using PyNast (28) as default in QIIME pipeline. Chimera checking was performed using Chimera Slayer (29).

We applied two extra filtering steps in aligned and taxonomy-assigned OTU table. First, sequences that belonged to chloroplasts class were filtered out. After that, sequences representing less than 0.005% of total OTUs were also filtered out [as previously done in Ref. (30)] from the chloroplast filtered OTU table. After these two extra filtering steps, we lost a mean of 27% of sequences (median of 25%, ranging from 2 to 77%) and a mean of 21% of sequences (median of 15%, ranging from 1 to 77%) in open and closed reference approaches, respectively (Data Sheet S1 in Supplementary Material).

Downstream Bioinformatics Analyses: Diversity, Composition, Potential Functions, and Statistical Tests

Downstream analyses were performed using QIIME 1.9.1 (24) with the filtered OTU table. Reads are clustered against a reference sequence collection, and all of the reads that do not hit a sequence in the reference sequence collection are excluded from downstream analyses in a closed reference approach or are subsequently clustered *de novo* in an open reference approach. To standardize samples with unequal sequencing depths, analyses were performed using random subsets of 25,000 sequences per sample in the open reference approach and random subsets of 10,000 sequences per sample in the closed reference approach. The perianal sample of one FB (FB1.8) failed this parameter and was discarded for posterior analyses.

Alpha diversity analysis assesses the diversity within a sample. In alpha diversity, we used two different metrics: observed species to assess richness and Shannon index to assess evenness. We assessed statistical significance with 999 permutations using the non-parametric Monte Carlo permutation test and corrected the *P* value through false discovery rate.

To assess the differences in the alpha diversity and composition at the individual level, we collapsed the eight skin samples from a dog using QIIME v1.9.1 to form a unique sample representing the individual. Therefore, the sample size for analyzing the individual effect is nine.

To assess the differences in the alpha diversity and composition when considering the breed, we used two approaches: (A) analyzing each skin site independently and (B) using the QIIME collapsed values from the eight skin site samples for each dog ($n = 9$). In the first approach, we group the three samples corresponding to a skin site from a breed and assessed differences in breeds per skin site; e.g., GS1.1, GS2.1, and GS3.1 as GS_chin and we compared them to FB_chin and WH_chin. In the second approach, we group the three collapsed individual samples from each breed; e.g., GS1, GS2, and GS3 as GS and we compared them to FB and WH.

Beta diversity analysis assesses the similarities among samples of the same community. Beta diversity was performed using both weighted and unweighted UniFrac distance metrics (31). Weighted UniFrac considers phylogeny, taxa, and relative abundances, whereas unweighted UniFrac only considers phylogeny and taxa. Those distance matrices were used to create PCoA plots and unweighted pair group method with arithmetic mean (UPGMA) trees. Trees were plotted using FigTree (REF). ANOSIM and adonis statistical methods were applied to evaluate if some variables were determining grouping and to which extent.

PICRUSt (18) was used to predict the functional profile of skin bacterial communities using 16S rRNA gene data obtained using a closed reference approach in QIIME v1.9.1. Kyoto Encyclopedia of Genes and Genomes (KEGG) (32) Ortholog (KO) hierarchy was used to make inferences of the functional gene content.

Linear discriminant analysis (LDA) effect size (LEfSe) (33) was used to compare groups and to identify differentially abundance distribution in both taxa and predicted functions ($\alpha = 0.05$ and with an LDA score >3.0).

RESULTS

To assess variability and composition of dog skin microbiota, we performed a cross-sectional study with healthy dogs from three breeds. We have analyzed 72 samples from 9 dogs: 3 FBs,

3 GSs, and 3 West Highland WH. We sampled eight skin sites: chin, inner pinna, nasal skin, axilla, back, abdomen, interdigital region, and perianal area, which are named as 1, 2, 3, 4, 5, 6, 7, and 8, respectively (**Figure 1**). These anatomic sites were selected to represent the regional diversity of the canine skin (34).

We found a total of 2,092 bacterial OTUs living on dog skin, which were taxonomically classified into 20 phyla, 51 classes, 69 orders, 132 families, and 245 genera. Data Sheet S2 in Supplementary Material contains several OTU tables: the complete OTU table for the 72 samples, the OTU table at family level obtained for all the samples, the OTU table collapsed by site, and the OTU table collapsed by individual.

The abundances of the main phyla differed on each sample (**Figure 2A**). The main phyla on skin samples were Proteobacteria (1–73%), Firmicutes (3–93%), Fusobacteria (0–58%), Bacteroidetes (0–69%), and Actinobacteria (0–35%), followed by Cyanobacteria, Tenericutes, TM7, and others with lower abundances.

Alpha diversity values were also very variable among samples (Data Sheet S3 and Figure S1 in Supplementary Material). The richness (observed species) ranged from 145.6 in the chin of WH1 to 928.8 in the inner pinna of GS1 (average of 488.42). The evenness (Shannon Index) ranged from 0.959 in the axilla of WH3 to 8.559 in the abdomen of WH2 (average of 5.8).

To assess if the variability of the dog skin microbiota depended on individual, breed, and/or skin site and to which extent, we clustered the samples using UPGMA trees and assessed statistical significance using adonis and ANOSIM tests (**Figures 2B,C**). We found that the main force driving the variability in dog skin microbiota composition is the individual, followed by the skin site and the breed.

Despite the high variability detected among samples, all of them were skin microbiota of healthy dogs and in fact shared some of their taxonomy. Thus, to assess the homogeneity of the samples, we analyzed the core microbiota. To complete the analysis, we assessed the potential functions of the bacterial community using PICRUSt.

Individual

Samples from the same individual tended to cluster together (**Figures 2B,C**). Statistical analysis using adonis test confirmed this result: the clustering of samples per individual significantly explained 40% (unweighted UniFrac) and 45% (weighted UniFrac; Figure S2A in Supplementary Material) of the distances among samples. Moreover, ANOSIM R value was close to +1 ($R = 0.79$, $P = 0.001$) in unweighted UniFrac, suggesting a strong dissimilarity among groups that was mainly explained by low-abundant species in each dog. Therefore, the individual was the variable that explained most differences among samples.

Seven of nine dogs had a taxonomic profile with the main bacterial phyla: Firmicutes and Proteobacteria with higher abundances than Actinobacteria or Bacteroidetes (**Figure 2A**). From these dogs, FB2, FB3, and WH1 presented also Fusobacteria as one of the main phyla if not the greatest one, whereas in GS1, GS2, GS3, and FB1, this phylum was almost absent. Two of nine dogs presented a predominant phylum (>50% of the total abundance) over the others, WH2 with Proteobacteria and WH3 with Firmicutes. The abundances of these two phyla and others were differentially distributed (Figure S2B in Supplementary Material) ($\alpha = 0.05$, LDA score >3).

The abundances differed in each individual, not only at the phylum level but also at the deeper taxonomic levels, such as family level (Data Sheet S2 in Supplementary Material). We can detect some individual-specific families, when looking at the most abundant families (**Table 1**): *Listeriaceae* representing a 22.5% of total microbiota composition for GS2; *Porphyromonadaceae* with a 26.1% for WH1; and or *Enterobacteriaceae* with a 12% for FB1. On the other hand, *Streptococcaceae* was present in all the individuals with low percentages, in exception of WH3 with 59% of the total composition that making it the individual with the lowest evenness value (3.71 of Shannon Index, Data Sheet S3 in Supplementary Material). Depending on the individual, families representing more than 5% (**Table 1**) were describing from 36.3 to 78.6% of total microbiota composition.

Skin Site

Clustering samples per skin site significantly explained 19% of the distances, when considering composition, phylogeny, and relative abundances (weighted UniFrac; **Figure 3A**). Visually inspecting the beta diversity plot, we found that perianal samples cluster together.

Inner pinna presented the highest diversity value with an average richness of 610.82 observed species and an average evenness of 6.85 of Shannon index, whereas the perianal region presented the lowest diversity, with only 323.1 observed species and a Shannon index of 4.41 (Data Sheet S3 in Supplementary Material). These two skin regions were significantly different between each other when considering evenness (**Figure 3B**; $P = 0.028$).

Most of the skin sites had Proteobacteria and Firmicutes as the most abundant phyla, adding up to more than 55% of the total microbiota composition (**Figure 3C**). Chin had also Bacteroidetes as an abundant phylum with a 14.6% of *Porphyromonadaceae*, becoming the main family of this skin site. The perianal region was the exception and presented the most different composition profile (**Figure 3** and **Table 2**; Data Sheet S2 in Supplementary Material). In perianal region, Bacteroidetes was the main phylum, followed by Firmicutes and Fusobacteria. Moreover, Proteobacteria, which is one of the main phyla inhabiting dog skin, was almost absent. The three main families inhabiting perianal region were *Bacteroidaceae* with 32.5% (Bacteroidetes), *Fusobacteriaceae* with 25.6% (Fusobacteria), and *Lachnospiraceae* with 6.4% (Firmicutes).

We detected differentially distributed abundances on skin sites with LEfSe analyses ($\alpha = 0.05$, LDA score >3) at the phylum and class level (**Figure 3D**) and up to the family level (Figure S3 in Supplementary Material). At the phylum and class level, Proteobacteria was significantly overrepresented at inner pinna, mainly due to the members of Alphaproteobacteria class; and Actinobacteria were overrepresented at the back. Some of the lowest abundant phyla were significantly more represented in a specific skin site: GN02 and TM7 at chin and [Thermi] (mainly from the *Deinococci* class) at inner pinna.

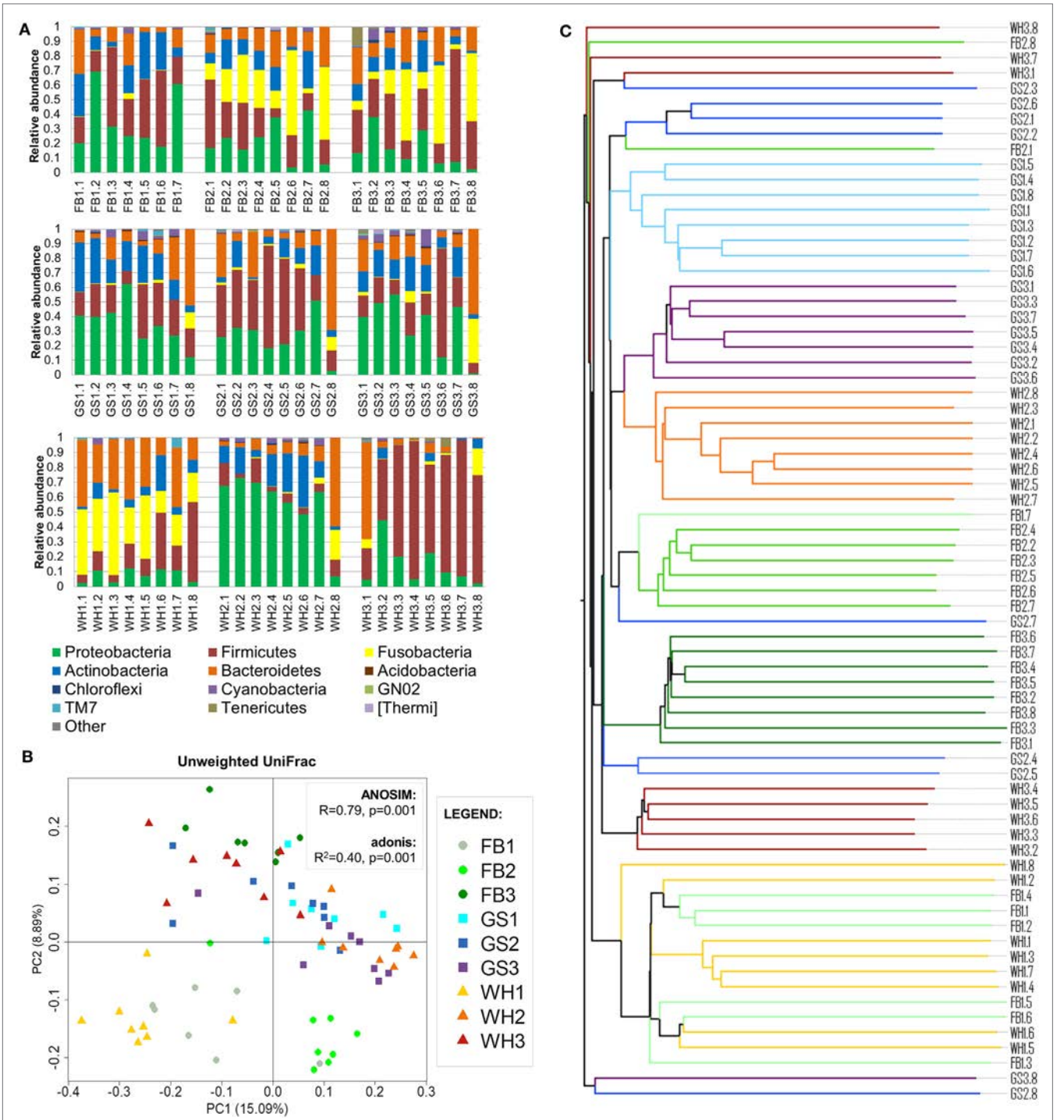


FIGURE 2 | **Individual signatures on dog skin**. **(A)** Taxonomy bar plot of each sample at phylum level. The first two letters with the number represent the breed and individual; and the numbers represent each skin site. **(B)** PCoA plot using unweighted UniFrac metrics colored by individual with values of ANOSIM and adonis statistical tests. **(C)** Unweighted pair group method with arithmetic mean tree associated with **(B)**; branches are colored by individual.

Breed

The breed explained fewer differences among the samples, but it did explain some differences. Clustering samples per breed significantly explained 10% (unweighted UniFrac) and 9% (weighted UniFrac) of the distances among samples (Figures S4A,B in Supplementary Material).

To assess the effect of the breed in diversity and composition, we used two approaches: (A) analyzing each skin site separately

TABLE 1 | Skin microbiota composition at family level for each individual.

Phylum	Family	FB1	FB2	FB3	GS1	GS2	GS3	WH1	WH2	WH3
Proteobacteria	*Rhodospirillaceae*	0.4%	1.5%	1.7%	2.4%	1.2%	5.5%	0.2%	10.9%	0.5%
Proteobacteria	*Sphingomonadaceae*	2.5%	3.3%	1.5%	3.0%	1.2%	1.9%	0.7%	9.0%	0.7%
Proteobacteria	*Enterobacteriaceae*	12.0%	0.2%	0.1%	0.2%	0.3%	0.1%	0.1%	0.0%	0.2%
Proteobacteria	*Pasteurellaceae*	6.4%	1.2%	0.1%	0.2%	0.9%	0.4%	0.4%	0.1%	0.2%
Proteobacteria	*Moraxellaceae*	0.8%	0.2%	0.3%	0.3%	2.0%	0.4%	0.5%	6.6%	0.4%
Firmicutes	*Listeriaceae*	0.1%	0.0%	0.0%	0.1%	22.5%	0.5%	0.2%	0.1%	0.0%
Firmicutes	*Staphylococcaceae*	10.2%	4.6%	7.0%	5.6%	1.2%	10.6%	0.8%	3.9%	0.3%
Firmicutes	*Streptococcaceae*	2.2%	2.6%	1.1%	1.0%	3.3%	1.6%	0.2%	0.1%	59.1%
Firmicutes	*Clostridiaceae*	0.5%	0.5%	8.2%	4.9%	1.4%	1.6%	0.9%	0.4%	2.6%
Firmicutes	*Lachnospiraceae*	0.4%	0.3%	6.9%	1.8%	0.6%	1.0%	2.9%	0.9%	0.7%
Fusobacteria	*Fusobacteriaceae*	0.6%	21.2%	23.9%	1.9%	2.3%	5.4%	32.2%	3.4%	3.6%
Fusobacteria	*Leptotrichiaceae*	0.3%	5.9%	0.0%	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%
Bacteroidetes	*Bacteroidaceae*	0.4%	3.1%	6.6%	6.2%	9.0%	8.6%	1.4%	5.7%	0.3%
Bacteroidetes	*Porphyromonadaceae*	1.8%	1.6%	3.3%	0.2%	4.7%	1.0%	26.1%	1.4%	8.5%
Bacteroidetes	*Weeksellaceae*	5.4%	0.8%	0.3%	0.8%	3.5%	1.8%	1.7%	1.0%	0.2%
Actinobacteria	*Corynebacteriaceae*	4.8%	0.5%	1.4%	6.7%	1.0%	2.6%	2.1%	1.4%	1.3%
Actinobacteria	*Intrasporangiaceae*	5.2%	1.4%	0.4%	0.9%	0.3%	2.6%	2.5%	0.3%	0.0%
% of microbiota explained by taxa >5%		**53.9%**	**48.9%**	**62.8%**	**36.3%**	**55.6%**	**45.6%**	**72.6%**	**45.0%**	**78.6%**

List of the taxa that represent >5% of the total microbiota composition and their abundances considering the individual.

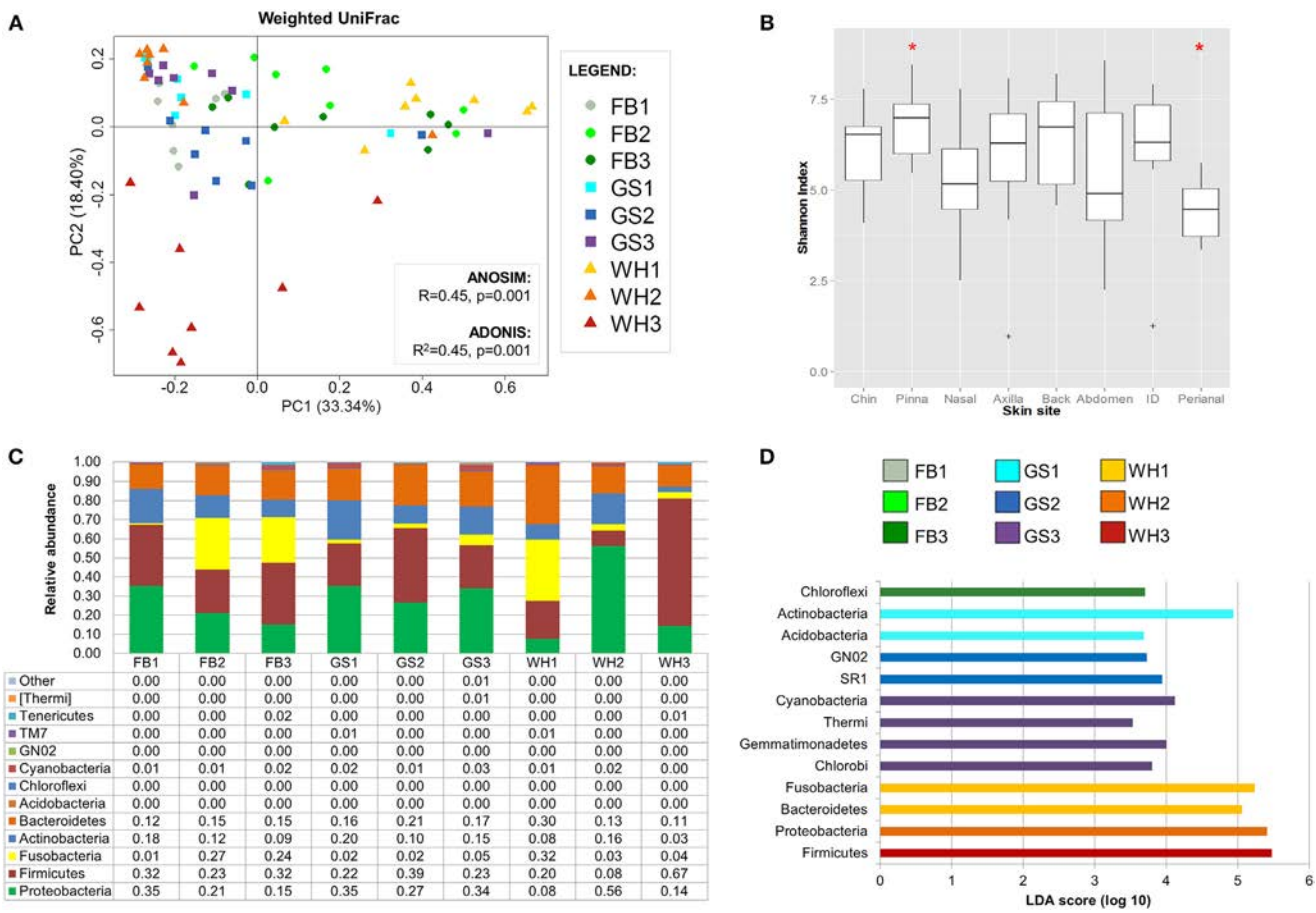


FIGURE 3 | Dog skin microbiota analysis considering site. **(A)** PCoA plot using weighted UniFrac metrics colored by skin site with values of ANOSIM and adonis statistical tests. Perianal region is circled in brown. **(B)** Boxplots of alpha diversity values. Marked with a red asterisk the two comparisons that were statistically different when using Monte Carlo permutation test ($P < 0.05$). **(C)** Bar plot representing skin microbiome composition at phylum level per skin site; each bar represents the mean values of the nine dogs per each skin site. **(D)** Histogram of linear discriminant analysis (LDA) effect size scores for differentially abundance distribution ($\alpha = 0.05$, LDA score >3) of bacterial phyla and classes among individuals.

considering the breed and (B) analyzing each collapsed dog sample per breed, adding up together all the values of the eight skin sites to form an individual dog value and grouping the three dogs from the same breed.

At taxonomic composition level, when analyzing each skin site per breed, we saw some differences ($\alpha = 0.05$, LDA score >3) (Figure S4C in Supplementary Material). At phylum level, Tenericutes were overrepresented at nasal skin of FB. At family level, GS had an overrepresentation of *Dermabacteraceae* at axilla and *Corynebacteriaceae* and *Williamsiaceae* at the interdigital region, whereas FB had *Burkholderiaceae* and *Bacillaceae* at axilla, *Gemellaceae* at the interdigital region, and *Gordoniaceae* at back and chin. When collapsing all the eight skin sites to obtain an individual sample, we only detected three families with differentially distributed abundances: *Sphingobacteriaceae* and *Dermabacteraceae* in GS and *Enterococcaceae* in FB (Figure S4D in Supplementary Material). All of these taxa had low relative abundances (Data Sheet S2 in Supplementary Material).

In alpha diversity analysis, we detected no statistical differences, both when analyzing each skin site separately (Figure S5A in Supplementary Material) and when analyzing the collapsed dog samples (Figure S5B in Supplementary Material).

Core Skin Microbiota

Each dog had its own microbiota profile, but there were also taxa shared among all samples even at low-abundant level, which we can define as the skin *core* microbiota of our cohort of individuals.

Families found in all the skin samples analyzed in this study were *Corynebacteriaceae* (Actinobacteria); *Streptococcaceae* and *Lachnospiraceae* (Firmicutes); *Fusobacteriaceae* (Fusobacteria); and *Comamonadaceae*, *Oxalobacteraceae,* and *Neisseriaceae* (Proteobacteria) (**Table 3**).

TABLE 2 | Skin microbiota composition at family level for each skin site.

Phylum	Family	Chin	Inner pinna	Nasal skin	Axilla	Back	Abdomen	Interdigital	Perianal
Proteobacteria	*Rhodospirillaceae*	2.3%	6.6%	2.9%	3.1%	3.0%	0.8%	2.9%	0.1%
Proteobacteria	*Neisseriaceae*	4.9%	1.3%	3.3%	1.5%	0.6%	1.1%	5.1%	0.0%
Proteobacteria	*Enterobacteriaceae*	0.3%	6.7%	2.1%	0.6%	0.2%	0.1%	0.1%	0.5%
Proteobacteria	*Moraxellaceae*	1.0%	0.2%	6.7%	0.6%	0.5%	0.6%	0.3%	0.1%
Firmicutes	*Listeriaceae*	1.7%	2.9%	0.2%	6.9%	5.7%	2.7%	0.8%	0.1%
Firmicutes	*Staphylococcaceae*	4.5%	2.2%	12.9%	1.7%	5.1%	9.6%	1.6%	0.5%
Firmicutes	*Streptococcaceae*	5.2%	4.3%	10.6%	11.1%	6.2%	9.5%	11.0%	5.8%
Firmicutes	*Clostridiaceae*	1.6%	2.6%	1.1%	0.8%	1.5%	1.4%	5.4%	4.6%
Firmicutes	*Lachnospiraceae*	1.1%	0.7%	0.6%	0.8%	1.0%	0.7%	3.0%	6.4%
Fusobacteria	*Fusobacteriaceae*	7.8%	6.7%	11.6%	10.6%	7.4%	13.5%	3.5%	25.6%
Bacteroidetes	*Bacteroidaceae*	0.4%	0.2%	0.4%	3.1%	0.6%	2.5%	0.5%	32.5%
Bacteroidetes	*Porphyromonadaceae*	14.6%	3.3%	7.3%	5.2%	4.4%	2.8%	4.3%	1.3%
Bacteroidetes	*Weeksellaceae*	6.0%	0.8%	2.4%	0.8%	0.6%	1.1%	1.5%	0.0%
Actinobacteria	*Corynebacteriaceae*	6.2%	2.6%	1.2%	2.0%	1.8%	1.6%	0.9%	3.1%
% of microbiota explained by taxa >5%		**57.5%**	**41.2%**	**63.3%**	**48.8%**	**38.3%**	**48.1%**	**40.8%**	**80.6%**

List of the taxa that represent >5% of the total microbiota composition and their abundances considering the skin site.

TABLE 3 | Skin core microbiota at family level for each individual and skin site.

Phylum	Family	GS1	GS2	GS3	FB1	FB2	FB3	WH1	WH2	WH3
Actinobacteria	*Corynebacteriaceae*	6.7%	1.0%	2.6%	4.8%	0.5%	1.4%	2.1%	1.4%	1.3%
Firmicutes	*Streptococcaceae*	1.0%	3.3%	1.6%	2.2%	2.6%	1.1%	0.2%	0.1%	59.1%
	Lachnospiraceae	1.8%	0.6%	1.0%	0.4%	0.3%	6.9%	2.9%	0.9%	0.7%
Fusobacteria	*Fusobacteriaceae*	1.9%	2.3%	5.4%	0.6%	21.2%	23.9%	32.2%	3.4%	3.6%
Proteobacteria	*Comamonadaceae*	3.3%	1.5%	1.8%	1.6%	1.1%	0.6%	0.9%	1.9%	0.8%
	Oxalobacteraceae	2.8%	3.4%	2.2%	1.1%	0.9%	0.5%	0.6%	3.5%	3.9%
	Neisseriaceae	1.6%	3.9%	1.5%	4.6%	0.8%	1.0%	1.9%	4.9%	0.4%
% of microbiota explained by core taxa		**19.1%**	**16.1%**	**16.2%**	**15.3%**	**27.5%**	**35.5%**	**40.7%**	**16.0%**	**69.9%**

Phylum	Family	Chin	Inner pinna	Nasal skin	Axilla	Back	Abdomen	Interdigital	Perianal
Actinobacteria	*Corynebacteriaceae*	6.2%	2.6%	1.2%	2.0%	1.8%	1.6%	0.9%	3.1%
Firmicutes	*Streptococcaceae*	5.2%	4.3%	10.6%	11.1%	6.2%	9.5%	11.0%	5.8%
	Lachnospiraceae	1.1%	0.7%	0.6%	0.8%	1.0%	0.7%	3.0%	6.4%
Fusobacteria	*Fusobacteriaceae*	7.8%	6.7%	11.6%	10.6%	7.4%	13.5%	3.5%	25.6%
Proteobacteria	*Comamonadaceae*	1.7%	1.4%	1.9%	2.0%	1.9%	1.8%	1.2%	0.1%
	Oxalobacteraceae	1.5%	3.2%	1.6%	2.0%	2.6%	2.1%	3.6%	0.2%
	Neisseriaceae	4.9%	1.3%	3.3%	1.5%	0.6%	1.1%	5.1%	0.0%
% of microbiota explained by core taxa		**28.4%**	**20.2%**	**30.8%**	**30.0%**	**21.4%**	**30.4%**	**28.2%**	**41.2%**

List of the taxa shared in all samples included in the study, their abundances, and distributions by individual and skin site.

The skin core microbiota at family level explained from 15.3 to 40.7% of the individual composition and from 20.2 to 41.2% of the skin site composition. It reached 69.9% for WH3 (with 59.1% of *Streptococcaceae*). Although a group of families constituted the core microbiota, their abundances were specific for each individual and site.

When we consider that the skin core microbiota is defined by taxa present in 85% of the samples (61 of 71 samples; to exclude some specific site or specific individual), the skin core microbiota explained a mean of 78% of the skin composition at both the individual and the skin site level, and we found 39 different families (Data Sheet S4 in Supplementary Material).

Predicted Functions

We used 16S rRNA gene sequencing data to predict the functional profile of dog skin microbiota samples, applying PICRUSt. PICRUSt developers (18) and more recently Meisel et al. (35) reported strong correlations between human metagenomic data sets and 16S-based functional prediction in skin microbiota.

We found up to 41 predicted functions for the dog skin microbiota, when considering the second level of KO hierarchy. Membrane transport (environmental information processing); replication and repair (genetic information processing); and amino acid, carbohydrate, and energy metabolism (metabolism) are the functions more spread and represented, with a mean relative abundance of 12, 7.9, 10.3, 10.4, and 5.6%, respectively (Data Sheet S6 in Supplementary Material).

Taxa composition profiles became more uniform when converting them to predicted functions (**Figures 4A,B**). However, we found some differentially distributed abundances in predicted functions at breed, individual, and skin site level ($\alpha = 0.05$, LDA score >3). We focused on assessing differences on the functional prediction among skin sites (**Figure 4C**).

Some predicted functions were overrepresented in back, chin, perianal region, and inner pinna and differentially distributed from all other sites. In pinna, we found overrepresentation of cellular processes and cell motility (cellular processes) and also signal transduction (environmental information processing). In perianal region, three metabolism pathways were increased: carbohydrate metabolism, glycan biosynthesis and metabolism, and nucleotide metabolism. In chin, we found overrepresentation of genetic information processing and its sublevel pathways—replication and repair and translation. In the back, three metabolism pathways were increased: xenobiotics biodegradation and metabolism, lipid metabolism, and metabolism of terpenoids and polyketides. Figure S6 in Supplementary Material contains LEfSe plots of differentially abundant predicted functions at level 3 of KEGG Orthology for skin site.

DISCUSSION

Our results suggest that the main force driving the variability in microbiota composition in dogs is the individual, rather than the breed—hair coat—or the skin site. This is true both considering the community structure (weighted UniFrac), but mainly when looking at the less abundant species (unweighted UniFrac). Several human studies have reported that interindividual variation is high and defines a "personal microbiome" (9, 19, 36). These low abundant bacterial signatures have been even used to identify individuals (37).

Meason et al. found recently this same pattern for canine skin mycobiome (fungal community) (17). Moreover, Rodrigues-Hoffmann et al. observed great differences on individuals, although they focused on detecting skin site differences and not on assessing the effect of the individual directly (6). On the other hand, human skin has three main microhabitats or ecological niches, depending on the physiological properties: sebaceous, dry, and moist areas and different microbiota is associated with each microhabitat (5). Conversely, dogs present almost all their skin sites covered by hair that creating a more uniform habitat.

Previous research had detected Proteobacteria (6, 7) or Firmicutes (38) as the main phyla inhabiting dog skin microbiota. Our results suggest that either Proteobacteria or Firmicutes or a combination of both can be the main phyla, depending on the individual. We also found Fusobacteria as one of the most abundant phyla for three of nine dogs, and when it was present, it spread over all the skin sites. Rodrigues-Hoffmann et al. detected Fusobacteria as one phylum specific to perianal regions (6); other studies also found them in groins (8) and paws and forehead (7), but with lower abundances than those seen here.

At the family level, taxa found in our cohort resemble those found in other canine skin microbiota studies (6–8, 18). Rodrigues-Hoffmann et al. found that *Oxalobacteraceae*, specifically *Ralstonia* spp., was the most abundant and extended taxa on dog skin (6), specially on healthy dogs; however, none of our *Oxalobacteraceae* sequences were from *Ralstonia* spp. Pierezan et al. have suggested that this could be due to the use of different supplies for the collection of samples, modifications in sample storage, extraction methods, and/or changes in the high-throughput sequencing platform used (8). *Ralstonia spp.* had been also detected in "blank" controls in microbiota studies and could be contaminants from the laboratory or the kits and reagents used (39).

Among all the individuals included, WH3 was very different with its skin mostly inhabited by *Streptococcaceae* that suggesting a colonization event. The representative sequence of the most abundant *Streptococcaceae* OTU in WH3 corresponds to *Streptococcus canis*, which are considered opportunistic pathogens inhabiting healthy dog skin. Their overgrowth has been associated to dermatitis (40) and even necrotizing fasciitis (41). Moreover, WH3 was the less diverse individual. Low alpha diversity values were characterizing skin microbiota in dogs affected by atopic dermatitis (6, 18), and in humans, they had been linked to elderly people (42). Therefore, we have two hypotheses for WH3: although considered healthy by the clinicians, the dog was beginning to develop some skin affliction; or the effect could be due to its advanced age. Further studies would be needed to assess the effect of age on healthy dog skin microbiota. Reanalyzing results excluding this sample have shown similar results for both ANOSIM and adonis tests (data not shown), confirming that the individual is the main force driving microbiota structure and composition and that the inclusion of this dog does not interfere with the results obtained.

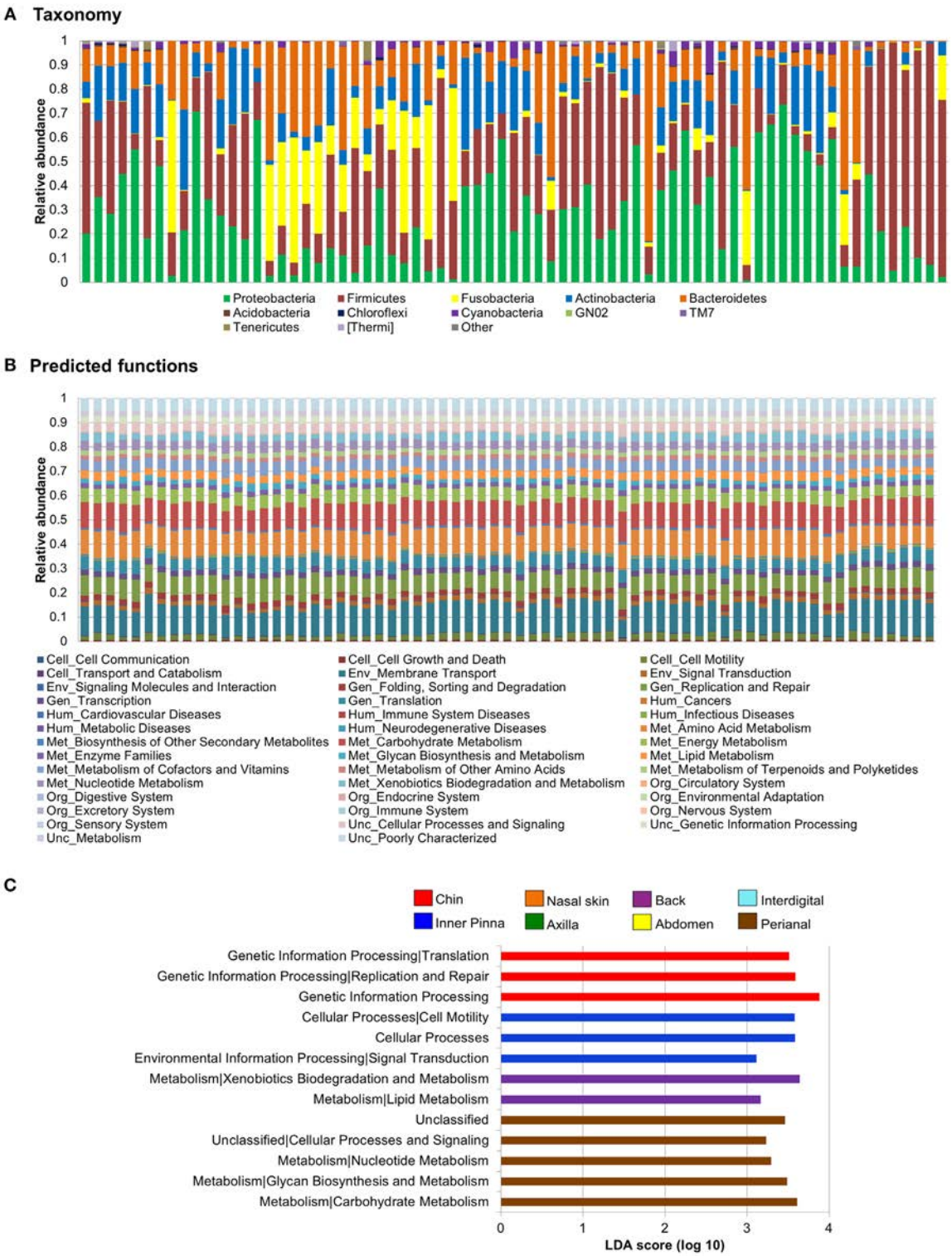


FIGURE 4 | **Skin microbiota relative abundances: taxa vs predicted functions**. Bar plots obtained through a closed reference approach to be comparable between them (see Section "Materials and Methods"). Each stacked bar represents relative abundances of each sample included in the study. Relative abundances of **(A)** bacteria at phylum level and **(B)** predicted functions (second level of the Kyoto Encyclopedia of Genes and Genomes Ortholog hierarchy) based on Phylogenetic Investigation of Communities by Reconstruction of Unobserved States data set. **(C)** Histogram of linear discriminant analysis (LDA) effect size scores for differentially abundance distribution ($\alpha = 0.05$, LDA score >3) of predicted functions. Complete list of predicted functions on dog skin and their relative abundances is available in Data Sheet S6 in Supplementary Material.

Despite the major force driving microbiota composition and variability was the individual, skin site also plays a role explaining the variability observed. The variability of the skin microbiota regarding the site could be due to the influences of other body site microbiota, such as the perianal region with the gastrointestinal microbiota or the chin with the oral microbiota, or due to the specific physiological properties of each skin site, such as the back with higher sebum production. Perianal region presented the most different composition profile: Bacteroidetes followed by Firmicutes and Fusobacteria were the main phyla, whereas Proteobacteria presented lower abundances. Moreover, Erysipelotrichi and Clostridia classes were overrepresented. This phyla pattern and taxa are more similar to that seen on canine gastrointestinal microbiota than that from the skin (20, 43). At the functional level profiling, some metabolic pathways were significantly overrepresented in the perianal region. Swanson et al. detected carbohydrate metabolism as one of the main pathways in intestinal microbiota of dogs, with values similar to those detected here (20) that are differentially higher than the other skin sites included. In chin, the most abundant phyla were Bacteroidetes, followed by Proteobacteria and Firmicutes. Sturgeon et al. have detected those same three phyla as the most abundant ones on canine oral microbiota (44). Moreover GN02 and TM7, two of the lowest abundant phyla, were overrepresented and differentially distributed on that region. These two phyla have been previously detected in canine oral microbiota (45). We also found that *Porphyromonadaceae* and *Fusobacteraceae* are the most abundant families in chin, coinciding with Bradley et al. who detected *Porphyromonas* and *Fusobacterium* (among others) as abundant genera in canine oral microbiome (16). On the other hand, physiological properties of the back skin could be influencing the microbiota function of that region. The dorsal parts of the neck, the trunk, and the tail have larger sebaceous glands than other skin regions (46). Moreover, the dorsal region has the densest hair coat, so a larger number of sebaceous glands associated with the hair follicles (46). Consequently, more sebum is produced than in other skin sites, which is mainly composed of lipid compounds. The higher abundance of this substrate could be explaining the increased lipid metabolism and fatty acid metabolism pathways in microbiota inhabiting back.

Even when our results show that the main force driving skin microbiota structure and composition is the individual, we cannot rule out the influence of the environment and lifestyle. The individual should be understood as the dog, its lifestyle, and its environment. In fact, the chloroplasts sequences that we detected and discarded for the ulterior analysis were not evenly distributed, but more represented in three dogs (Data Sheet S1 in Supplementary Material), suggesting that these individuals had a greater or more recent exposure to outdoor environment and may have more transient bacterial members detected as skin microbiota. On the other hand, despite being the human skin constantly exposed to extrinsic factors, healthy adults have shown to maintain their skin microbial communities over time (36). This last hypothesis should be assessed in dogs, because they are exposed to extrinsic factors, such as environment or human contact. In this study, we cannot distinguish whether this individual factor is solely host specific or it also includes extrinsic properties from the environment.

Some of the differences when comparing our results to previous studies could be due to differences in the methodologies chosen such as the 16S region analyzed or the sequencing platform used. We are amplifying 400 bp of the V1–V2 hypervariable regions that had been suggested to be a better choice for skin microbiota in humans among others (47). Hypervariable regions V1–V3 are the most commonly used on dog skin microbiota studies (6, 18), but only V2 region has also been used (7). Recently, Pierezan et al. used V4 (8). On the other hand, Clooney et al. found that the factor responsible for the greatest variance in microbiota composition was the chosen methodology, when comparing Illumina HiSeq, Illumina MiSeq, and Ion Torrent PGM. This problem was larger in Illumina MiSeq rather than in Ion Torrent PGM when analyzing 16S rRNA V1–V2 region amplicons (48). In another study comparing microbial profiles using V1–V2 regions, the authors concluded that the output generated from PGM Ion Torrent and 454 yielded concurrent results (49). Finally, PICRUSt is a tool that was mainly developed for the human microbiome. However, dogs share skin microbiota with their owners [as seen in Ref. (7)]. So, using PICRUSt for skin in pets is probably a valid approach. Moreover, PICRUSt has already been used in fecal samples of dogs (21).

CONCLUSION

The individual seems to be the main force driving skin microbiota composition and diversity in dogs, and dissimilarity is mainly explained by low-abundant species in each dog. The main phyla inhabiting the dog skin in our cohort are Proteobacteria, Firmicutes, Fusobacteria, Actinobacteria, and Bacteroidetes, and their abundance patterns differ among individuals.

The skin site also plays a role: the composition and function of microorganisms inhabiting chin and perianal region could be influenced by other body site microbiota. Moreover, the specific physiological properties of the back, with higher abundance of sebum, could favor the growth of specific microorganisms. We observed distinctive taxa composition profiles for each sample, but relative abundances become more uniform when converting them to predicted functions.

As the diversity among individuals is the highest, a good choice to better assess the dog skin microbiota would probably be comparing affected vs unaffected regions from the same dog rather than comparing different dogs in case–control studies, so each dog is its own control; and an accurate assessment of the environmental factor, controlling variables such as geographical region, season, lifestyle, or cohabitation with other animals.

AUTHOR CONTRIBUTIONS

AS, LF, and OF conceived and designed the experiment. AS, LA, LF, and OF supervised the project and gave conceptual advice. AC and OF performed the experiment. AC carried out the

bioinformatics analysis. AC drafted the manuscript. AS, LA, LF, and OF edited the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We acknowledge Marc Pons (UAB) for helping at the initial bioinformatics analysis and Diana Ferreira, Xavier Roura, and Mar Bardagí (Hospital Clinic Veterinari, UAB) for the sample collection. This work was supported by a grant awarded by Generalitat de Catalunya (Industrial Doctorate program, 2013 DI 011).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://journal.frontiersin.org/article/10.3389/fvets.2017.00006/full#supplementary-material.

REFERENCES

1. Grice EA, Kong HH, Renaud G, Young AC, Bouffard GG, Blakesley RW, et al. A diversity profile of the human skin microbiota. *Genome Res* (2008) 18:1043–50. doi:10.1101/gr.075549.107
2. Naik S, Bouladoux N, Wilhelm C, Molloy MJ, Salcedo R, Kastenmuller W, et al. Compartmentalized control of skin immunity by resident commensals. *Science* (2012) 337:1115–9. doi:10.1126/science.1225152
3. Oh J, Byrd AL, Deming C, Conlan S, Barnabas B, Blakesley R, et al. Biogeography and individuality shape function in the human skin metagenome. *Nature* (2014) 514:59–64. doi:10.1038/nature13786
4. Grice EA, Segre JA. The skin microbiome. *Nat Rev Microbiol* (2011) 9:244–53. doi:10.1038/nrmicro2537
5. Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, et al. Topographical and temporal diversity of the human skin microbiome. *Science* (2009) 324:1190–2. doi:10.1126/science.1171700
6. Rodrigues-Hoffmann A, Patterson AP, Diesel A, Lawhon SD, Ly HJ, Elkins Stephenson C, et al. The skin microbiome in healthy and allergic dogs. *PLoS One* (2014) 9:e83197. doi:10.1371/journal.pone.0083197
7. Song SJ, Lauber C, Costello EK, Lozupone CA, Humphrey G, Berg-Lyons D, et al. Cohabiting family members share microbiota with one another and with their dogs. *Elife* (2013) 2:1–22. doi:10.7554/eLife.00458
8. Pierezan F, Olivry T, Paps JS, Lawhon SD, Wu J, Steiner JM, et al. The skin microbiome in allergen-induced canine atopic dermatitis. *Vet Dermatol* (2016) 5:332–e82. doi:10.1111/vde.12366
9. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. Bacterial community variation in human body habitats across space and time. *Science* (2009) 326:1694–7. doi:10.1126/science.1177486
10. Findley K, Grice EA. The skin microbiome: a focus on pathogens and their association with skin disease. *PLoS Pathog* (2014) 10:e1004436. doi:10.1371/journal.ppat.1004436
11. Kong HH, Oh J, Deming C, Conlan S, Grice EA, Beatson MA, et al. Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis. *Genome Res* (2012) 22:850–9. doi:10.1101/gr.131029.111
12. Oh J, Freeman AF, Park M, Sokolic R, Candotti F, Holland SM, et al. The altered landscape of the human skin microbiome in patients with primary immunodeficiencies. *Genome Res* (2013) 23:2103–14. doi:10.1101/gr.159467.113
13. Alekseyenko AV, Perez-Perez GI, De Souza A, Strober B, Gao Z, Bihan M, et al. Community differentiation of the cutaneous microbiota in psoriasis. *Microbiome* (2013) 1:31. doi:10.1186/2049-2618-1-31
14. Takemoto A, Cho O, Morohoshi Y, Sugita T, Muto M. Molecular characterization of the skin fungal microbiome in patients with psoriasis. *J Dermatol* (2015) 42:166–70. doi:10.1111/1346-8138.12739
15. Fitz-Gibbon S, Tomida S, Chiu BH, Nguyen L, Du C, Liu M, et al. Propionibacterium acnes strain populations in the human skin microbiome associated with acne. *J Invest Dermatol* (2013) 133:2152–60. doi:10.1038/jid.2013.21
16. Bradley CW, Morris DO, Rankin SC, Cain CL, Misic AM, Houser T, et al. Longitudinal evaluation of the skin microbiome and association with microenvironment and treatment in canine atopic dermatitis. *J Invest Dermatol* (2016) 136:1182–90. doi:10.1016/j.jid.2016.01.023
17. Meason-smith C, Diesel A, Patterson AP, Older CE, Mansell JM, Suchodolski JS, et al. What is living on your dog's skin? Characterization of the canine cutaneous mycobiota and fungal dysbiosis in canine allergic dermatitis. *FEMS Microbiol Ecol* (2015) 91:1–12. doi:10.1093/femsec/fiv139
18. Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* (2013) 31:814–21. doi:10.1038/nbt.2676
19. Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* (2012) 486:207–14. doi:10.1038/nature11234
20. Swanson KS, Dowd SE, Suchodolski JS, Middelbos IS, Vester BM, Barry KA, et al. Phylogenetic and gene-centric metagenomics of the canine intestinal microbiome reveals similarities with humans and mice. *ISME J* (2011) 5:639–49. doi:10.1038/ismej.2010.162
21. Minamoto Y, Otoni CC, Steelman SM, Büyükleblebici O, Steiner JM, Jergens AE, et al. Alteration of the fecal microbiota and serum metabolite profiles in dogs with idiopathic inflammatory bowel disease. *Gut Microbes* (2015) 6:33–47. doi:10.1080/19490976.2014.997612
22. Cadieu E, Neff MW, Quignon P, Walsh K, Chase K, Parker HG, et al. Coat variation in the domestic dog is governed by variants in three genes. *Science* (2009) 326:150–3. doi:10.1126/science.1177808
23. Jaeger K, Linek M, Power HT, Bettenay SV, Zabel S, Rosychuk RA, et al. Mueller RS. Breed and site predispositions of dogs with atopic dermatitis: a comparison of five locations in three continents. *Vet Dermatol* (2010) 21:118–22. doi:10.1111/j.1365-3164.2009.00845.x
24. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* (2010) 7:335–6. doi:10.1038/nmeth.f.303
25. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* (2010) 26:2460–1. doi:10.1093/bioinformatics/btq461
26. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* (2007) 73:5261–7. doi:10.1128/AEM.00062-07
27. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* (2006) 72:5069–72. doi:10.1128/AEM.03006-05
28. Caporaso JG, Bittinger K, Bushman FD, Desantis TZ, Andersen GL, Knight R. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* (2010) 26:266–7. doi:10.1093/bioinformatics/btp636
29. Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G, et al. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* (2011) 21:494–504. doi:10.1101/gr.112730.110
30. Navas-molina JA, Peralta-sánchez JM, González A, McMurdie PJ, Vázquez-baeza Y, Xu Z, et al. Advancing our understanding of the human microbiome using QIIME. *Methods Enzymol* (2013) 531:371–444. doi:10.1016/B978-0-12-407863-5.00019-8
31. Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* (2005) 71:8228–35. doi:10.1128/AEM.71.12.8228
32. Kanehisa M, Goto S. KEGG: Kyoto encyclopaedia of genes and genomes. *Nucleic Acids Res* (2000) 28:27–30. doi:10.1093/nar/28.1.27
33. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. *Genome Biol* (2011) 12:R60. doi:10.1186/gb-2011-12-6-r60

34. Miller WH, Griffin CE, Campbell KL, Muller GH. *Muller and Kirk's Small Animal Dermatology*. 7th ed. St Louis: Elsevier (2013).
35. Meisel JS, Hannigan GD, Tyldsley AS, Sanmiguel AJ, Hodkinson BP, Zheng Q, et al. Skin microbiome surveys are strongly influenced by experimental design. *J Invest Dermatol* (2016) 136:947–56. doi:10.1016/j.jid.2016.01.016
36. Oh J, Byrd AL, Park M; NISC Comparative Sequencing Program, Kong HH, Segre JA. Temporal stability of the human skin microbiome. *Cell* (2016) 165:854–66. doi:10.1016/j.cell.2016.04.008
37. Fierer N, Lauber CL, Zhou N, McDonald D, Costello EK, Knight R. Forensic identification using skin bacterial communities. *Proc Natl Acad Sci U S A* (2010) 107:6477–81. doi:10.1073/pnas.1000162107
38. Weese JS. The canine and feline skin microbiome in health and disease. *Vet Dermatol* (2013) 24:137–46. doi:10.1111/j.1365-3164.2012.01076.x
39. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* (2014) 12:87. doi:10.1186/s12915-014-0087-z
40. Lamm CG, Ferguson AC, Lehenbauer TW, Love BC. Streptococcal infection in dogs: a retrospective study of 393 cases. *Vet Pathol* (2010) 47:387–95. doi:10.1177/0300985809359601
41. DeWinter LM, Low DE, Prescott JF. Virulence of *Streptococcus canis* from canine streptococcal toxic shock syndrome and necrotizing fasciitis. *Vet Microbiol* (1999) 70:95–110. doi:10.1016/S0378-1135(99)00128-5
42. Ying S, Zeng D, Chi L, Tan Y, Galzote C, Cardona C, et al. The influence of age and gender on skin-associated microbial communities in urban and rural human populations. *PLoS One* (2015) 10:e0141842. doi:10.1371/journal.pone.0141842
43. Middelbos IS, Vester Boler BM, Qu A, White BA, Swanson KS, Fahey GC. Phylogenetic characterization of fecal microbial communities of dogs fed diets with or without supplemental dietary fiber using 454 pyrosequencing. *PLoS One* (2010) 5:e9768. doi:10.1371/journal.pone.0009768
44. Sturgeon A, Stull JW, Costa MC, Weese JS. Metagenomic analysis of the canine oral cavity as revealed by high-throughput pyrosequencing of the 16S rRNA gene. *Vet Microbiol* (2013) 162:891–8. doi:10.1016/j.vetmic.2012.11.018
45. Camanocha A, Dewhirst FE. Host-associated bacterial taxa from Chlorobi, Chloroflexi, GN02, Synergistetes, SR1, TM7, and WPS-2 Phyla/candidate divisions. *J Oral Microbiol* (2014) 6:25468. doi:10.3402/jom.v6.25468
46. Al-Bagdadi F. The integument. In: Evans HE, de Lahunta A, editors. *Miller's Anatomy of the Dog*. St Louis: Elsevier (2012). p. 61–79.
47. Kuczynski J, Lauber CL, Walters WA, Parfrey LW, Clemente JC, Gevers D, et al. Experimental and analytical tools for studying the human microbiome. *Nat Rev Genet* (2011) 13:47–58. doi:10.1038/nrg3129
48. Clooney AG, Fouhy F, Sleator RD, Driscoll AO, Stanton C, Cotter PD, et al. Comparing apples and oranges? Next generation sequencing and its impact on microbiome analysis. *PLoS One* (2016) 11:e0148028. doi:10.1371/journal.pone.0148028
49. Indugu N, Bittinger K, Kumar S, Vecchiarelli B, Pitta D. A comparison of rumen microbial profiles in dairy cows as retrieved by 454 Roche and Ion Torrent (PGM) sequencing platforms. *PeerJ* (2016) 4:e1599. doi:10.7717/peerj.1599

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

3.2.1. Annex 1. Erratum: Figure 3

In the production step, an older version of Figure 3 was published instead of the newer. The published Figure 3 is not incorrect, it only does not match with the legend or text associated.

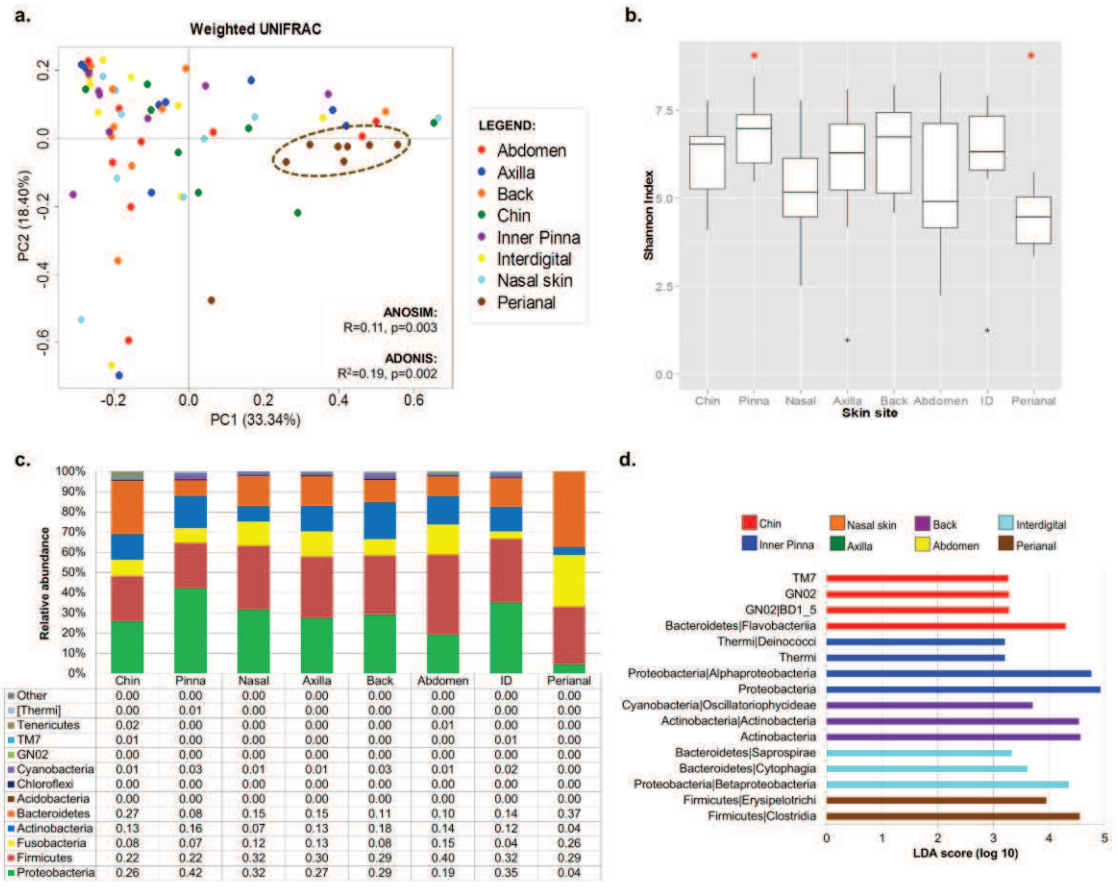


Figure 3. Dog skin microbiota analysis considering site. (A) PCoA plot using weighted UniFrac metrics colored by skin site with values of ANOSIM and adonis statistical tests. Perianal region is circled in brown. (B) Boxplots of alpha diversity values. Marked with a red asterisk the two comparisons that were statistically different when using Monte Carlo permutation test ($P < 0.05$). (C) Bar plot representing skin microbiome composition at phylum level per skin site; each bar represents the mean values of the nine dogs per each skin site. (D) Histogram of linear discriminant analysis (LDA) effect size scores for differentially abundance distribution ($\alpha = 0.05$, LDA score >3) of bacterial phyla and classes among individuals.

3.3. Individual signatures and environmental factors shape skin microbiota on healthy dogs

This chapter consists of the article entitled "Individual signatures and environmental factors shape skin microbiota on healthy dogs" recently submitted to Microbiome journal.

Supplementary material of this article is available online at the following link: https://www.dropbox.com/sh/kcd1bo4adzlh439/AADsgFQ5MwNYTieimIITZUUAa?dl=0

Individual signatures and environmental factors shape skin microbiota on healthy dogs

Anna Cuscó[1,2]*, Janelle M. Belanger[3], Liza Gershony[3], Alma Islas-Trejo[3], Kerinne Levy[4], Juan F. Medrano[3], Armand Sánchez[1], Anita M. Oberbauer[3], Olga Francino[1].

[1]Molecular Genetics Veterinary Service (SVGM), Veterinary School, Universitat Autònoma de Barcelona, Barcelona, Spain,
[2]Vetgenomics, Ed Eureka, Parc de Recerca UAB, Barcelona, Spain,
[3]Department of Animal Science, University of California, Davis, CA, USA,
[4]Canine Companions for Independence, Santa Rosa, CA, USA.

Abstract

Background. The individual, together with its environment, has been reported as the main force driving composition and structure of skin microbiota in healthy dogs. Therefore, one of the major concerns when analyzing canine skin microbiota is the likely influence of the environment. Despite the dense fur covering, certain skin diseases exhibit differential prevalence among skin sites, dog breeds and individuals.

Results. Here we aimed to assess the variability of the skin microbiota in healthy dogs cohabiting together, by analyzing eight different skin sites in a large and homogeneous cohort of Golden-Labrador Retriever crossbred dogs (N=35). We found that microbiota composition was driven by the individual, but when considering abundances, the microbiota structure was driven both by the individual and by the skin site. Network analyses elucidated bacterial interactions within and between skin sites, especially in chin, abdomen, axilla and the perianal region, with the highly shared interactions probably representing an environmental component. When analyzing each skin site independently to assess host-specific factors, we found that season of birth or the time spent in the kennel was shaping skin microbiota in all skin sites. The most abundant taxon driving this difference was Sphingomonas, which is an air-borne bacterium that cannot be cultivated at elevated temperatures. We also found some taxonomic differences linked to sex on abdomen, axilla and back. Finally, the USA and European cohorts were grouping by geographical origin in two different and well-defined clusters, even when the European individuals were very heterogeneous.

Conclusions. We observed a large inter-individual variability and effects of different host variables, such as season of birth or time spent in the kennel and sex, even in an environmental homogenous cohort.

Keywords: skin, canine, microbiota, microbiome, dog, season, skin site, pinna, 16S, environment

Introduction

Skin is a complex ecosystem inhabited by a high diversity of microorganisms, collectively referred to as the microbiota. These microbial communities not only inhabit, but also interact with the host cells impacting cellular function and immunity; likewise the host cells influence the microbes (1). This cross-talk between the host cells and the microorganisms maintains the homeostasis and the healthy status of an individual, and when disrupted usually indicates disease (2).

The dense fur that covers almost all of a dog's skin creates a homogenous microenvironment. However, some skin diseases show a preference for certain skin sites and for specific breeds (3). Previous studies have described skin microbiota on healthy dogs (4–9) , but only three of them included several skin sites to assess differences that may exist due to the anatomical location sampled (6,8,9). Results from Rodrigues-Hoffmann and colleagues showed that haired skin regions presented higher diversity values than mucosal areas and mucocutaneous junctions (6), and a similar result was reported when comparing the inner pinna and the perianal region (8). No differences among skin sites were detected when including only dorsal neck, axilla and abdomen (9).

Dog skin microbiota studies aimed at detecting differences between health and disease status have already been performed for canine atopic dermatitis (6,7,10). Skin affected with atopic dermatitis in dogs presented a less diverse microbiota (6,7) and increased proportions of *Staphylococcus* and *Corynebacterium* (7). Moreover, dogs with allergen-induced atopic dermatitis presented higher proportions of *Staphylococcus* on the challenged site compared to the contralateral site (10).

In humans, skin microbiota differs among skin sites and among individuals (11). On one hand, skin presents three main microhabitats depending on the physicochemical properties: sebaceous sites, inhabited with *Propionibacterium* spp; moist sites, with *Staphyloccocus* and *Corynebacterium* spp; and dry sites, with gram-negative microorganisms (11,12). On the other hand, individual signatures of the skin microbiota are usually driven by low abundant species (13). Following those first human studies describing skin microbiota, research then targeted key variables to ascertain if they drove skin microbiota structure and composition in the healthy individual. Variables assessed and found to have some effect on microbiota diversity, composition and structure included those related to host such as sex (14–16), age (17–19) and racial origin (20–22); or related to environment such as birth delivery mode (23), hygiene (14,22), cohabitation (5,24), geography (21,25,26) and urbanization (19,27,28).

One of the major concerns when performing skin microbiota studies on dogs is the likely influence of the environment (29). Our previous results suggest that the individual (together with its environment) was the main force driving skin microbiota composition and structure in a population of dogs from three different breeds and hair coats (8). Rodrigues-Hoffmann and colleagues assessed some environmental variables, such as presence of fleas, time spent indoors vs outdoors, sex, or age and did not detect significant associations between the microbiota and a particular environmental factor (6). However, the dog cohort assessed was very variable and included 12 individuals from different

breeds, ages, and households and therefore it is likely that any environmental effects may have been obscured. Two studies reported that dogs cohabiting together shared more skin microbiota (5,9). A recent study using a cohort of 40 healthy dogs sampled in three skin sites across seasons assessed the effects of age, sex, breed, hair type, skin site, season at time of collection and cohabitation. They found that season was the only variable significantly stratifying microbiota community structure and samples from different skin sites were more similar within the same dog (9).

Skin microbiota has been suggested as a potential clinical tool in susceptibility, diagnosis, and treatment of dermatological diseases (30), therefore characterizing the variability of skin microbiota in healthy dogs and determining which host and environmental variables are defining its structure and composition will extend the background to better design studies aimed to assess the altered skin microbiota on disease.

Here we aimed to assess the variability of the canine skin microbiota in healthy dogs cohabiting together by analyzing eight different skin sites in a large and homogeneous cohort of Golden-Labrador Retriever crossbred dogs (N=35). As most of the environmental variables were fixed, we also aimed to elucidate if any of the host factors were driving skin microbiota structure and composition in some skin sites. Finally, we aimed to assess the effect of the geographical region, thus we compared the findings from the US cohort with those obtained for a group of European dog samples.

Individuals included and sample collection

A cross-sectional study was performed in 35 healthy dogs to analyze skin microbiota variability in eight skin sites. All dogs were companion dogs, Golden-Labrador Retriever crosses, with ages ranging from 1 year and 7 months to 2 years and 3 months. They were born in different households, where they had been raised until 8 weeks of age, and then they had gone to individual puppy raisers until entering training at ~18 months of age. This cohort was composed of 20 female and 15 male dogs living and playing together in a shared environment in Santa Rosa, California. 6 females and 6 males presented a black coat, and the rest presented a yellow one. This will be named the USA cohort. Additional File 1 and Additional File 2 contain all the metadata associated with the dogs.

Skin microbiota samples were collected from eight regions taken from the right side of the dog: inner pinna, chin, nasal skin, back, axilla, abdomen, interdigital area and perianal region. These regions are named as A, B, C, D, E, F, G and H respectively (Figure 1A). Samples were obtained by firmly rubbing each area using Sterile Catch-All™ Sample Collection Swabs (Epicentre Biotechnologies) soaked in sterile SCF-1 solution (50 mM Tris buffer (pH=8), 1 mM EDTA, and 0.5% Tween-20). To minimize sample cross-contamination, the person sampling wore a fresh pair of sterile gloves for each individual. To minimize bias in sampling, only AO and AC sampled the dogs. Swabs were stored at 4ºC until DNA extraction, within the following 8 days (3 days, 2-day stop, 3 days).

On the other hand, the European cohort included 11 dogs of different breeds (Beagle, French Bulldog, German Shepherd, and West Highland white terrier), ages and

households. Nine of the dogs were previously published and described (8), whereas two of them sampled later remain unpublished. All samples were processed following the same DNA extraction protocol, PCR and sequencing approach. All samples were analyzed together following the steps explained below.

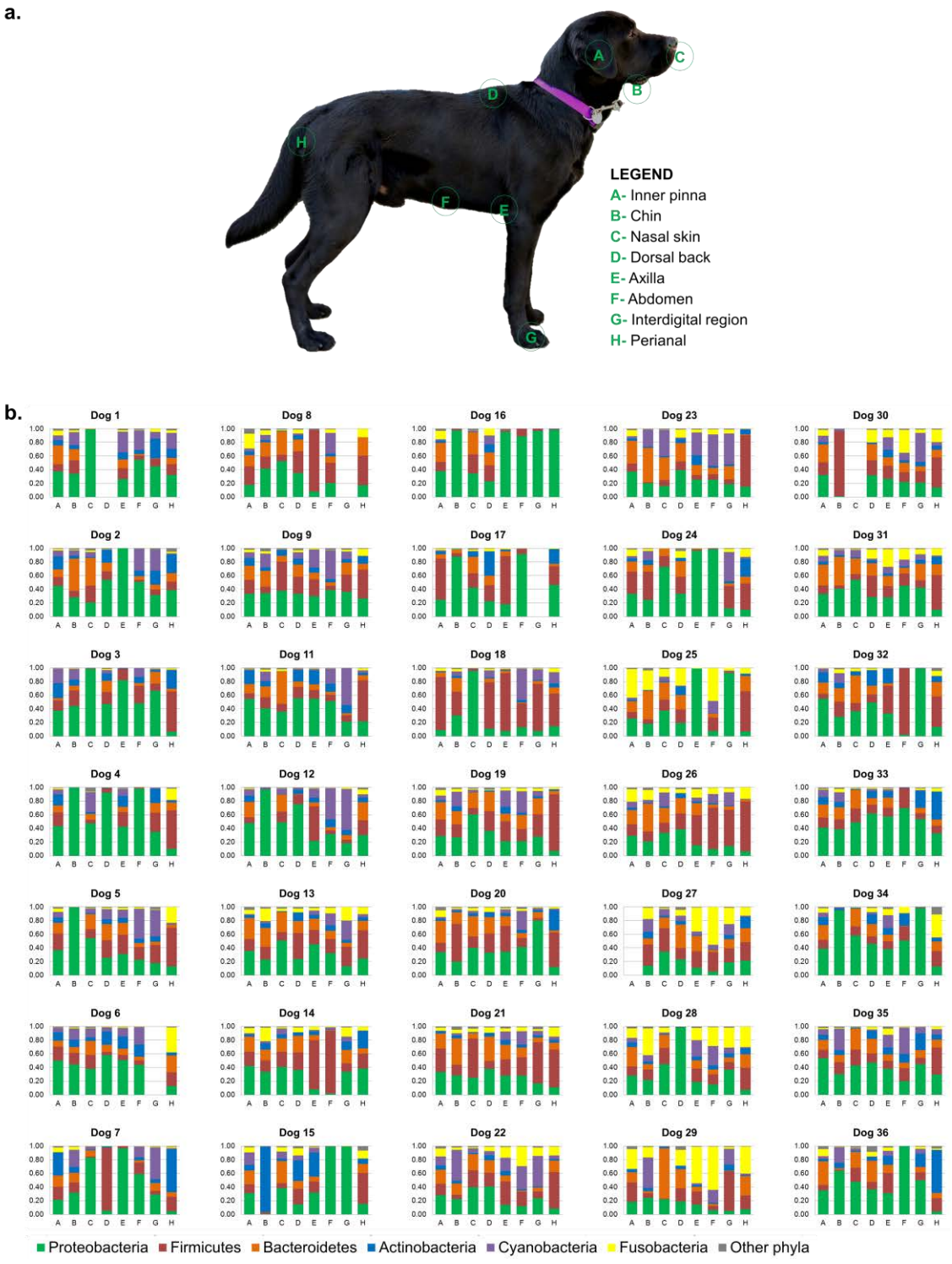


Figure 1. (A) Skin sites sampled and (B) Taxonomic composition per sample included at phylum level.

DNA extraction

Bacterial DNA was extracted from the swabs using the PowerSoil™ DNA isolation kit (MO BIO) under manufacturer's conditions, with one modification. At the first lysis step, the swab tip with the sponge was cut and placed in the bead tube, until the first transference of the supernatant to a new tube. The remaining steps were performed as described by the manufacturer in exception of the elution step, which was performed on 50µL of C6 instead of 100µL to obtain a higher concentration. Samples from different skin sites and individuals were randomly extracted to avoid confounding a batch effect with an actual effect. DNA extractions were performed within the following 8 days in random batches of samples to avoid confounding technical biases with actual ones. DNA samples (50 µl) were stored at -20°C until further processing. To assess for contamination from the laboratory or reagents two blank samples were processed: one with a sterile swab tip and the other without the sterile swab tip.

PCR amplification and massive sequencing

V1-V2 regions of 16S rRNA gene were amplified using the widely used primer pair F27 (5'-AGAGTTTGATCCTGGCTCAG-3') and R338 (5'-TGCTGCCTCCCGTAGGAGT-3'). We choose V1–V2 hypervariable regions because they had been suggested to be a better choice for human skin microbiota among others (31). PCR mixture (25 µl) contained 2 µl of DNA template, 5 µl of 5x Phusion® High Fidelity Buffer, 2.5 µL of dNTPs (2 mM), 0.2 µM of each primer and 0.5 U of Phusion® Hot Start II Taq Polymerase (Thermo Fisher).

The PCR thermal profile consisted of an initial denaturation of 30 sec at 98 °C, followed by 30 cycles of 15 sec at 98 °C, 15 sec at 55 °C, 20 sec at 72 °C and a final step of 7 min at 72 °C. Samples that did not amplify the first time were repeated increasing cycles to 33. To assess possible reagent contamination, each PCR reaction included a no template control (NTC) sample.

For each amplicon, quality and quantity were assessed using Agilent Bioanalyzer 2100 and QubitTM fluorometer, respectively. Both primers included sequencing adaptors at the 5' end and forward primers were tagged with different barcodes to pool samples in the same sequencing reaction.

Each sequencing pool included forty barcoded samples that were sequenced on an Ion Torrent Personal Genome Machine (PGM) with the Ion 318 Chip Kit v2 and the Ion PGM™ Sequencing 400 Kit (Life Technologies) under manufacturer's conditions.

Quality control of the sequences and OTU picking

Raw sequencing reads were demultiplexed and quality-filtered using QIIME 1.9.1 (32). Reads included presented: a length greater than 300 bp; a mean quality score above 25 in sliding window of 50 nucleotides; no mismatches on the primer; and default values for

other quality parameters. After that, quality-filtered reads were processed using vsearch v1.1 pipeline (33): a first de-replication step was applied, followed by clustering into operational taxonomic units (OTUs) at 97% similarity with a de *novo* approach and finally chimera checking was performed using uchime de *novo*. The raw OTU table was transferred into QIIME 1.9.1 and taxonomic assignment of representative OTUs was performed using the Ribosomal Database Project (RDP) Classifier (34) against Greengenes v13.8 database (35). Alignment of sequences was performed using PyNast (36). We sequentially applied some extra filtering steps in aligned and taxonomy-assigned OTU table to filter out: 1) sequences that belonged to Chloroplasts class; 2) sequences representing less than 0.005% of total OTUs (as previously done in (37)); 3) sequences that belonged to *Shewanellaceae* and *Halomonadaceae* families, which were highly represented in the NTC of the repetition chip (performed with an increased cycle number) and considered contamination from the reagents.

Samples 17G and 27A did not amplify and they could not be sequenced. We performed downstream analysis at a depth of 11,000 sequences per sample: 1D, 30C, 6G and 8G failed this parameter and were discarded for posterior analyses. Also, NTC and Blank with a swab tip (S-blank) presented some amplification, but failed to reach 11,000 sequences per sample; Blank without the swab tip (N-blank) could not amplify.

Downstream bioinformatics analyses

Downstream analyses were performed using QIIME 1.9.1 (32) with the filtered OTU table. To standardize samples with unequal sequencing depths, analyses were performed using random subsets of 11,000 sequences per sample.

Alpha diversity analysis assesses the diversity within a sample. Two different metrics were used for the alpha diversity: observed species to assess richness and Shannon index to assess evenness. We assessed statistical significant differences in alpha diversity values among groups with 999 permutations using the non-parametric Monte Carlo permutation test and corrected the p-value through false discovery rate.

Beta diversity analysis assesses the similarities among samples of the same community. Beta diversity was performed using both weighted and unweighted UniFrac distance metrics (38). Weighted UNIFRAC considers phylogeny, taxa and relative abundances; whereas unweighted UNIFRAC only considers phylogeny and taxa. Those distance matrices were used to create PCoA plots. ANOSIM and adonis statistical methods were applied to evaluate if some variables were determining grouping and to which extent.

Linear Discriminant Analysis (LDA) Effect Size (LEfSe) (39) was used to compare groups and to identify differentially abundance distribution in taxa (α=0.05 and with an LDA score > 3.0).

CoNet (40), which is implemented as an application in Cytoscape (41), was applied to infer networks among skin sites using bacterial families that presented a median relative abundance higher than 0.05% in each specific site. In CoNet we used a combination of five different algorithms (Pearson's correlation, Spearman's correlation, Kullback-Leibler

dissimilarity distances, Bray-Curtis dissimilarity distances and mutual information similarity). The results of the five methods were merged using Simes p-value. We performed a first permutation step, followed by a bootstrap analysis corrected for false discovery rate (α=0.05).

Results

We analyzed the variability of the canine skin microbiota in eight different skin sites from a healthy homogenous and well-controlled cohort of Golden-Labrador Retriever crossbred dogs cohabiting together in the same kennel in the United States (N=35) (see Additional File 1 and Additional File 2 for the associated metadata). We sampled microbiota from eight skin sites: inner pinna, chin, nasal skin, dorsal back, axilla, abdomen, interdigital region and perianal area, which are respectively named as A, B, C, D, E, F, G and H (Figure 1A). These anatomic sites were selected to represent the regional diversity of the canine skin (3) and to compare with our previous study (8). Samples 17G, 27A, 1D, 30C, 6G and 8G failed at some processing point and were discarded for posterior analyses (see material and methods for more detail).

We found a total of 2,216 bacterial OTUs living on dog skin (Additional File 3) that were taxonomically classified into 17 phyla, 41 classes, 62 orders, 128 families and 242 genera.

The abundance of the main phyla differed on each sample (Figure 1B) and were Proteobacteria (median: 33%; range: 0-99%), Firmicutes (median: 17%; range: 0-97%), Bacteroidetes (median: 12%, range: 0-74%) Actinobacteria (median: 5%; range: 0-95%), Cyanobacteria (median: 5%; range: 0%-5%) and Fusobacteria (median: 3%; range: 0-64%), followed by Tenericutes, TM7 and others with lower abundances.

Alpha diversity values were also very variable among samples with a median value of 5.9 for Shannon index and 414.37 for observed species. The range is broad and goes from 27.8 observed species and 0.27 Shannon index in axilla of Dog 25 to 989.9 observed species in chin of Dog 33 or 8.5 Shannon index in dorsal back of Dog 34 (Additional File 4A). None of the dogs was significantly more or less diverse than any other, because there were large differences in diversity values within the same dog. Those dogs that could seem less diverse because most of the skin sites presented less diversity, usually presented average values in inner pinna or perianal region giving no statistical significant differences among individuals (Additional File 4A).

We performed sample clustering to assess if the variability of the dog skin microbiota depended on the individual and/or the skin site and to which extent, assessing statistical significance using Adonis and ANOSIM tests. We also performed an independent analysis for each skin site to assess the effect of the individual-specific variables, such as sex, coat color, season of birth or time spent in the kennel.

Individual is driving skin microbiota structure and composition, followed by skin site

The dogs included in this study were all Golden-Labrador crossbreds with similar ages (~2 years old) and interacted and lived together in a shared environment. Moreover, in most cases a dog shared genetic background with others: 33 out of 35 dogs presented at least one half-sibling or littermate and only dogs 31, 19 and 14 were born from different sets of progenitors (Additional File 2). Grouping the samples per individual significantly explained 23% and 22% of the variation in Unweighted and Weighted UniFrac distance matrices (Table 1), suggesting that the main force driving the variability of skin microbiota in our samples was the individual. On the other hand, clustering samples per skin site explained 12% and 17% of the variation respectively and even presented an ANOSIM R value larger (R=+0.21) than the individual (R=+0.14) when considering OTUs abundance with the Weighted UNIFRAC distance matrix. Thus, composition of skin microbiota (Unweighted UniFrac) on healthy dogs was better explained by the individual, whereas structure of the skin microbiota when abundance was taken into account (Weighted UniFrac) was better explained by both individual and skin site.

When comparing individual profiles at the phylum level (Figure 1B), we detected that Proteobacteria was usually the main phylum found on the skin of our cohort: 28 out of 35 dogs presented a higher median value of Proteobacteria than any other phylum. For 5 of the sampled dogs, Firmicutes was the main phylum. Fusobacteria were most frequently found in the perianal regions, however when Fusobacteria colonized the haired-skin, the distribution was individual-specific. That is, there were a few individual dogs with a high abundance of Fusobacteria in several regions whereas other dogs had almost no Fusobacteria. Within the Fusobacteria enriched individuals, usually the highest percentages were found in the abdomen samples. Finally, Cyanobacteria phylum was mainly present with high abundances in the abdomen, interdigital region and the chin of specific individuals.

Table 1. Clustering of the samples per biological and technical variables. Beta diversity statistics ANOSIM and Adonis values. (**) p-value=0.001, (*) p-value<0.05 (-) indicates no significant clustering.

	Unweighted UniFrac		**Weighted UniFrac**	
	adonis R^2	ANOSIM R	adonis R^2	ANOSIM R
Individual	0.23**	0.20**	0.22**	0.14**
Skin site	0.12**	0.19**	0.17**	0.21**
Storage time	0.05**	0.07**	0.05**	0.05**
Chip	0.03*	0.03*	-	-
Person extracting	0.02*	-	0.02*	-
Sampler	0.01*	0.03*	0.01*	0.03*

Analysis of individual-specific variables

In order to assess if any individual-specific variable defined the skin microbiota composition or structure in any of the skin sites, we inspected the alpha and beta diversity of each skin site grouped by the different dog-specific variables such as sex, coat color, season of birth, time spent in the kennel, or recent surgery and assessed statistical significance except for the recent surgery due to the small sample size. Depending on the season of birth, the dogs can be classified in two groups: dogs born from January to May and dogs born from June to September. The time spent in the kennel coincided with the season of birth because older dogs (born from January to May) had been in the kennel for at least 6 months, whereas younger dogs (born from June to September) had been in the kennel for 3 months. We also considered the pedigree information of the dogs (Additional File 2): most of the dogs shared some genetic background to at least one other dog (half-siblings) even across the two main groups.

The season of the year when the dog was born or the time spent in the kennel significantly affected the microbiota composition (Unweighted UniFrac) and also the community structure (Weighted UniFrac) in all skin sites (Table 2; Additional File 5). This effect was especially large on the inner pinna, almost coincident with PC1 component, explaining 26% of the variation among samples and with an ANOSIM R value of 0.84 (Table 2, Figure 2A). In the other skin sites, these two variables explained more than 9% of the variation (except for nasal skin), with an R-value ranging from 0.24 to 0.38.

Either the season of birth or the time spent in the kennel was the variable that explained ubiquitously a significant amount of variation for all the skin sites. Delving into the effect of these variables on the inner pinna skin microbiota, we corroborated the pattern in the unweighted UniFrac consensus tree (Figure 2B): two clear clusters were elucidated matching with the season of birth or the time spent in the kennel (except Dog 8). Moreover, littermates were usually as similar as any other dog in the same group (except Dog 2 and 3) and sharing the sire did not make dogs resemble more in skin microbiota. Moreover, dogs born from January to May or dogs that had spent at least six months in the kennel were significantly more diverse than the other group (Figure 2C). Finally, LEFSe analysis detected 61 families differentially distributed in inner pinna when clustering in these two groups (Additional File 6) and those with higher relative abundances are plotted in Figure 2D. The greatest difference is provided by *Sphingomonadaceae* that is highly present in the inner pinna of individuals born from January to May that had been in the kennel for 6 months, whereas it is almost absent on those dogs born from June to September that had been in the kennel for 3 months.

Table 2. Individual-specific variables that cluster samples in specific skin sites.

		Unweighted UniFrac		Weighted UniFrac	
Skin site	**Variable**	ANOSIM R	adonis R^2	ANOSIM R	adonis R^2
Inner Pinna	Season of birth / Time in the kennel	0.84**	0.26**	0.41**	0.22**
Axilla	Season of birth / Time in the kennel	0.38**	0.11**	0.09*	0.07*
Dorsal back	Season of birth / Time in the kennel	0.37**	0.13**	0.28**	0.14**
Interdigital	Season of birth / Time in the kennel	0.28**	0.11**	0.09*	0.07*
Abdomen	Season of birth / Time in the kennel	0.28**	0.10**	0.09*	0.07*
Perianal	Season of birth / Time in the kennel	0.27**	0.09**	-	-
Chin	Season of birth / Time in the kennel	0.24*	0.10*	0.10*	0.08*
Abdomen	Sex	0.13*	0.05*	0.24*	0.11**
Nasal skin	Season of birth / Time in the kennel	0.11*	0.05*	0.06*	-
Back	Sex	-	0.05*	-	-
Axilla	Sex	-	-	-	0.06*

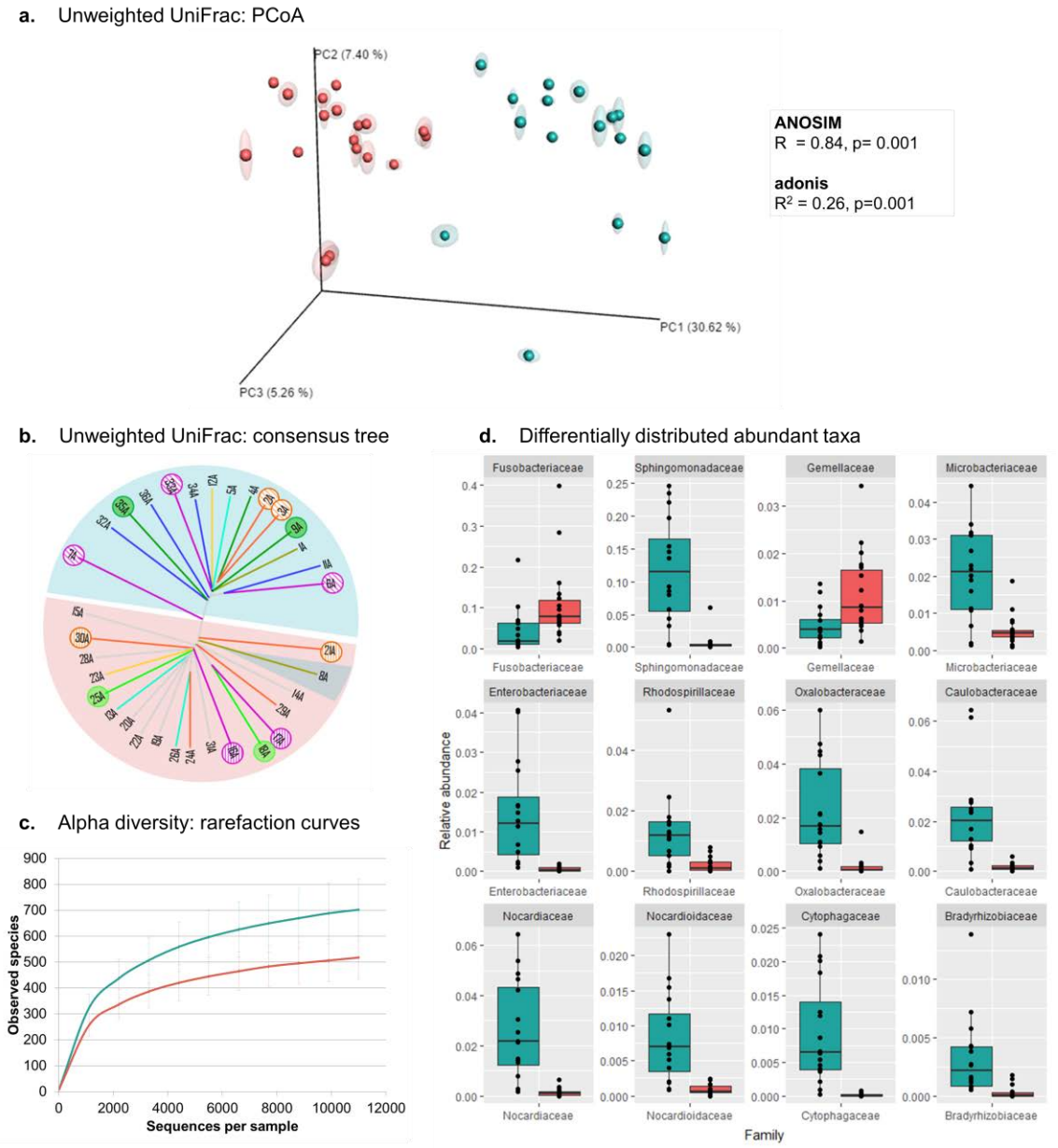


Figure 2. Season of birth or time spent in the kennel effect on inner pinna. Color blue represents Jan-May group that had been in the kennel for at least 6 months and color red, Jun-Sep group that had been in the kennel for 3 months. (A) Unweighted UniFrac PCoA beta diversity plot. (B) Unweighted UniFrac consensus tree: dogs sharing sire present same-colored branches and littermates are circled and colored with a common pattern within a group (C) alpha diversity rarefaction curves using observed species metrics, and (D) Boxplots of the main differentially distributed families: those include families with abundances > 1% in any group and also LefSe significant (LDA-score > 3.0, p-value < 0.05).

The sex of the dog also explained some variation. The microbiota community structure in the abdomen was better explained by the variable sex (11% of the variation in the weighted UNIFRAC plot and ANOSIM R value +0.24) rather than the season of birth or time spent in the kennel. This variable also explained to a lesser extent some variability of the microbiota composition in the dorsal back and the community structure in the axilla (Table 2). Considering the three skin sites affected by sex (abdomen, back and axilla), we could see that males had an overrepresentation of bacteria from Fusobacteria phylum, with *Sneathia* and Fusobacterium genera; also other genera such as *Actinomycetospora*, *Gemella*, *Parvimonas*,

Brevundimonas and phylum SR1 were also overrepresented on males. Females had an overrepresentation of *Enterobacteriaceae* family (Table 3).

Table 3. Differentially abundant taxa associated to sex.

		Abdomen		Axilla		Back	
Phylum	Family or genus	Female	Male	Female	Male	Female	Male
Fusobacteria	Fusobacteriales (order)	1.70%	21.45%	1.64%	13.44%	3.54%	9.55%
Fusobacteria	Leptotrichiaceae	0.23%	2.70%	0.24%	3.53%	0.82%	3.00%
Fusobacteria	Sneathia	0.01%	0.34%	0.05%	0.25%	0.21%	0.47%
Fusobacteria	Fusobacterium	NS	NS	1.41%	9.91%	2.72%	6.54%
Actinobacteria	Actinomycetospora	NS	NS	0.00%	0.04%	0.00%	0.15%
Firmicutes	Gemella	0.19%	3.04%	0.51%	1.61%	NS	NS
Firmicutes	Parvimonas	NS	NS	0.16%	1.82%	0.55%	1.15%
Proteobacteria	Brevundimonas	NS	NS	0.00%	0.01%	0.00%	0.01%
SR1	SR1	NS	NS	0.05%	0.19%	0.14%	0.44%
Proteobacteria	Enterobacteriaceae	14.08%	1.31%	7.78%	0.69%	NS	NS

Relative abundances of main taxa found to be differentially distributed (LDA-score >3, p-value < 0.05) between males and females in at least two out of the three skin sites affected. NS, stands for no significant differences.

We delved deeper into the five dogs that had undergone surgery, followed by a medical treatment prior to sampling (Table 4 and Additional File 4B). Dogs 14, 15, 16 and 17 presented reduced alpha diversity values in several skin sites, being chin and abdomen always affected; whereas alpha diversity values of inner pinna, nasal skin and back were not reduced in any dog. Dog 20, who underwent surgery three months before sampling, presented average alpha diversity values.

Finally, the coat color was not significantly explaining the skin microbiota structure or composition in any skin site.

Table 4. Information of the dogs that had undergone surgery prior to sampling.

Individual	Surgery date	Surgery type	Medicines	From / To	Sites w. reduced α-diversity1
Dog 14	08/04/2016	Spay	Amoxicillin (antibiotic) + Previcox (antiinflamatory)	08/04/2016 13/04/2016	Chin and abdomen
Dog 15	18/04/2016	Spay	Amoxicillin (antibiotic) + Previcox (antiinflamatory)	18/04/2016 23/04/2016	Chin, axilla, abdomen and ID region
Dog 16	30/03/2016	GI obstruction	Pepcid AC (antihistamine) + Tramadol (analgesic)	01/04/2016 06/04/2016	Chin, axilla, abdomen, ID region and perianal area
Dog 17	12/04/2016	Spay	Previcox (antiinflamatory)	12/04/2016 16/04/2016	Chin, axilla, abdomen and perianal
Dog 20	05/01/2016	Spay	Rimadyl (antiinflamatory)	05/01/2016 10/01/2016	None

1Reduced alpha diversity values include those ones that are half or less than the median alpha diversity of that specific skin site of the dogs that had not undergone recent surgery (Additional File 4B).

Influence of other variables on the skin microbiota

The main factors driving skin microbiota in our cohort were the individual and the skin site as sample associated variables and the sex, the season of birth or the time spent in the kennel as individual associated variables. However, other sample associated variables influenced albeit to a lesser extent, with all of them explaining 5% or less of the variation in the PCoA plots (Table 1). The most significant variable was storage time of the sample before DNA extraction, with extractions undertaken within the first 3 days being significantly more diverse than samples extracted in the last 3 days, both in terms of observed species and Shannon Index (Additional File 4C). Of note, the samples from dogs that had previously undergone surgery were extracted in the last 3 days, so the analyses were repeated excluding those individuals and there were still significant differences due to storage time (Additional File 4D).

We could also detect person extracting was influencing the diversity found in the samples, that was probably not a real effect though. Differences are detected among AC and AI extractions: AC was extracting the first 3-days, while AI extracted the last 3-days. So the samples from that one who extracted the last 3-days are less diverse probably due to surgery dogs included and due to the later extraction.

Skin sites: Network analysis

A network analysis detects bacterial relationships, within and among different ecological niches. The global network for all the skin sites considering the most abundant families allowed us to understand more deeply skin microbiota relationships in our cohort (Figure 3, Table 5 and Additional File S7). Some bacterial species interacted specifically in the same skin site, whereas other bacterial species interacted among different skin sites. Thus, we have different ecological niches within the skin.

Chin, abdomen, axilla and perianal region had the highest number of interactions, with 373, 226, 179 and 93 respectively, and also some extra interactions among families of other skin sites (Table 5 and Additional File 7). On the other hand, inner pinna, nasal skin, interdigital region and dorsal back presented a lower number of interactions and no inter-site interactions, as shown in Figure 2. Inner pinna had 35 family interactions; interdigital region, 23; nasal skin, 7; and dorsal back, only 2.

Table 5. Summary statistics of microbial interactions in the skin of a cohort of healthy dogs.

	Chin	Abdo-men	Axilla	Perianal region	Inner Pinna	Nasal skin	ID area	Dorsal back
Total Interactions	373	226	179	93	35	7	23	2
Common Interaction	139	103	104	43	13	4	13	2
Unique Interactions	234	123	75	50	22	3	10	0
Inter-site interaction*	3	20	10	12	0	0	0	0
% of unique interactions	63%	54%	42%	54%	63%	43%	43%	0%
% of co-occurence	92%	88%	79%	100%	100%	100%	100%	100%

ID stands for interdigital.*Inter-site interactions represent families from a specific skin site affecting other families from another skin site

In some cases, specific taxonomic interactions were found within different skin sites. We identified six interactions highly spread among different skin sites (present in 4 out of 8 skin sites): Neisseriaceae -> Weeksellaceae; Neisseriaceae -> Xenococcaceae (in chin, axilla, abdomen and perianal); Sphingomonadaceae -> Caulobacteraceae (in inner pinna, axilla, abdomen and perianal); Sphingomonadaceae -> Nocardioidaceae (in inner pinna, axilla, abdomen and interdigital region); Sphingomonadaceae -> Oxalobacteraceae (in inner pinna, chin, abdomen and interdigital); and Weeksellaceae -> Flavobacteriaceae (in chin, axilla, abdomen and interdigital). However, most interactions (517 out of 703) were exclusive from one specific skin site (Additional File 7).

This global network demonstrated that most interactions in canine skin were co-ocurrence relationships rather than mutual exclusion. Among mutual exclusion interactions, few nodes were negatively linked to many different families within a skin site (circles marked with a wider black line in Figure 3), showing an apparent invasive pattern. That was seen for Pseudomonadaceae family in axilla, chin and abdomen and also for Enterobacteriaceae family in abdomen. When blasting the most abundant OTUs from the highly connected mutual exclusion nodes, we found that the main genera were Pseudomonas (for Pseudomonadaceae) and *Erwinia* and *Pantoea* (for Enterobacteriaceae) (Additional File 8).

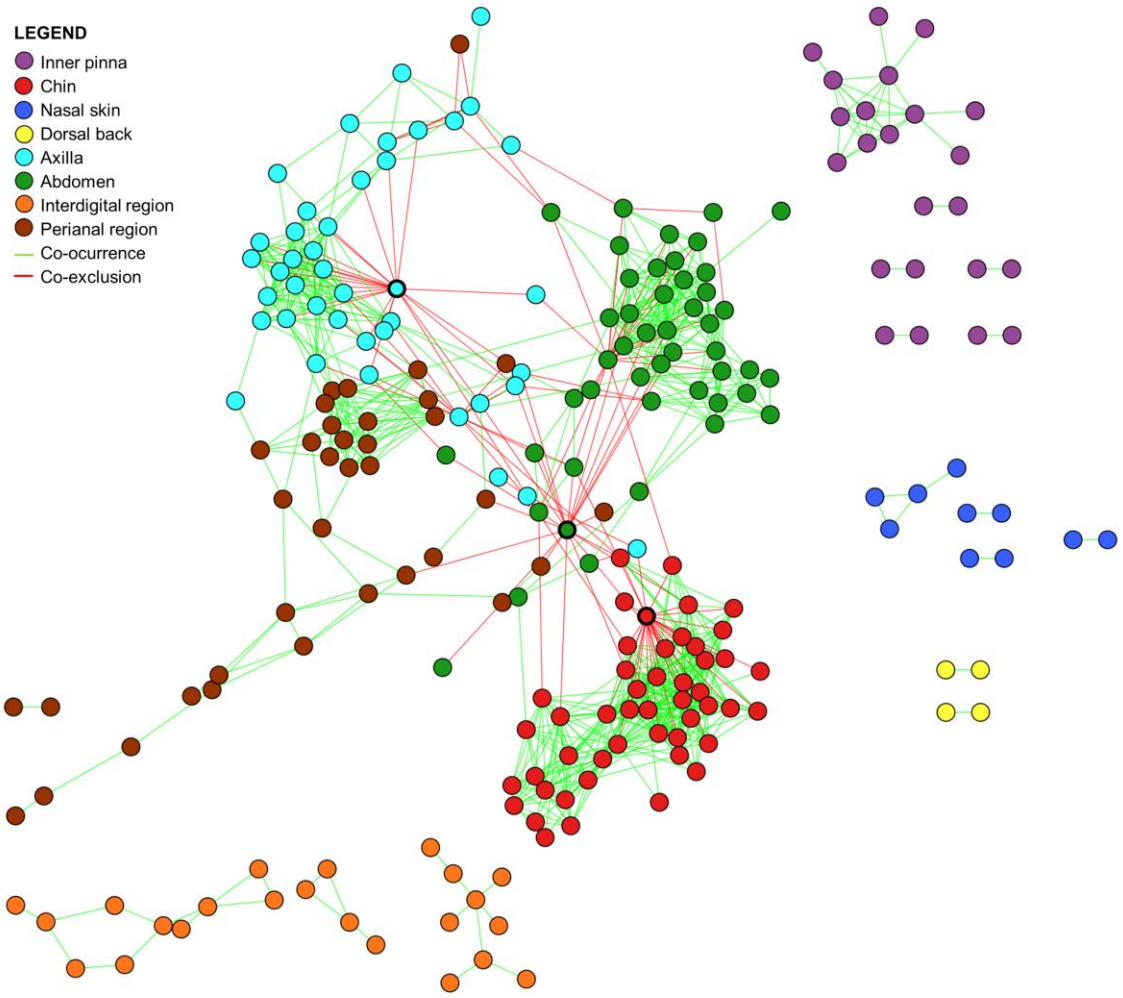


Figure 3. Significant co-occurrence and co-exclusion interactions among the abundant families (>0.005%) in the dog skin microbiota. Nodes are colored depending on the skin site they are found; nodes with a black circle are those highly connected mutual exclusion nodes; edges are green to represent co-occurrence patters and red to represent co-exclusions. Data associated with the complete network can be found at Additional File 7.

Skin sites: taxonomy and diversity analysis

Alpha and beta diversity analyses were undertaken to create an overall characterization of the canine skin microbiota that accounted for skin sites. Differences in alpha diversity among skin sites were prevalent. The inner pinna displayed the greatest diversity when compared to all the other sites and was statistically different to all but the chin site (p-value = 0.028). The chin, when considering observed species, was significantly more diverse than

the axilla (p-value = 0.028) (Additional File 4C). As seen previously in Table 1, clustering samples per skin site explained up to 17% of the differences in beta diversity analysis when considering OTUs abundance (Weighted UniFrac). Differences in microbiota structure were also significant among almost all skin sites, with the exception of the interdigital region when compared to abdomen or axilla. We found the greatest differences when comparing any skin site to the perianal region followed by the nasal skin (Additional File 9).

Focusing on taxonomic analyses, we found that bacteria from Gammaproteobacteria class were the most abundant in dog skin microbiota, with the exception of perianal regions where Bacilli class from Firmicutes phylum were the most abundant.

Skin sites shared most of the taxa, but presented also specific taxonomic patterns: the abundance and distribution varied significantly among skin sites and unique taxa were identified in some of the sites. Figure 4A shows different bar plots, colored by the main families found in skin. The families that were differentially distributed (LDA score >3, p-value = 0.05) are shown in Additional File 10.

The inner pinna had a higher amount of Proteobacteria phylum when compared to other skin sites, with Gammaproteobacteria, Alphaproteobacteria and Betaproteobacteria classes being the main representatives. Bacilli (Firmicutes) and Flavobacteriia (Bacteroidetes) were present in similar abundances to Proteobacteria. Moreover, inner pinna presented many different and less abundant bacteria.

The chin region was enriched in Gammaproteobacteria, with Pseudomonadaceae as the main representative family. The nasal skin was also enriched in Gammaproteobacteria, but the main representative family was Pasteurellaceae. Both families were differentially distributed in their respective skin site.

Back and axilla had quite similar taxonomic patterns: the main bacterial class was Gammaproteobacteria, with *Moraxellaceae* as the main family, followed by Bacilli, with *Lactobacillaceae* as one of the main families. The greatest taxonomic difference between both sites was the higher abundance of *Staphylococcaceae* (Bacilli class) in the axilla, which was also differentially distributed when compared to the other skin sites.

The abdomen and interdigital regions had similar taxonomic patterns, where most of the bacteria were Gammaproteobacteria, specifically from *Enterobacteriaceae*, *Moraxellaceae* and *Pseudomonadaceae* families; followed by Cyanobacteria, specifically *Xenococcaceae* family. However, *Planococcaceae* was found in abdomen but not in interdigital region.

Finally, the perianal region was the skin site that presented the most differentiated pattern in dog skin microbiota. The main phylum was Firmicutes, especially Bacilli, followed by Actinobacteria. Many different families from different phyla were differentially distributed in the perianal region, indicating that it was the most divergent skin site (Additional File 10). Most of the abundant families in perianal region were also differentially distributed when compared to the other skin sites. Some of them were: *Erysipelotrichaceae*, *Lachnospiraceae*, *Lactobacillaceae* and *Veillonellaceae* (Firmicutes); *Corynebacteriaceae* (Actinobacteria); and *Bacteroidaceae* (Bacteroidetes). The perianal region was also enriched in

Fusobateriaceae, despite not being statistically differentially distributed when compared to other skin sites.

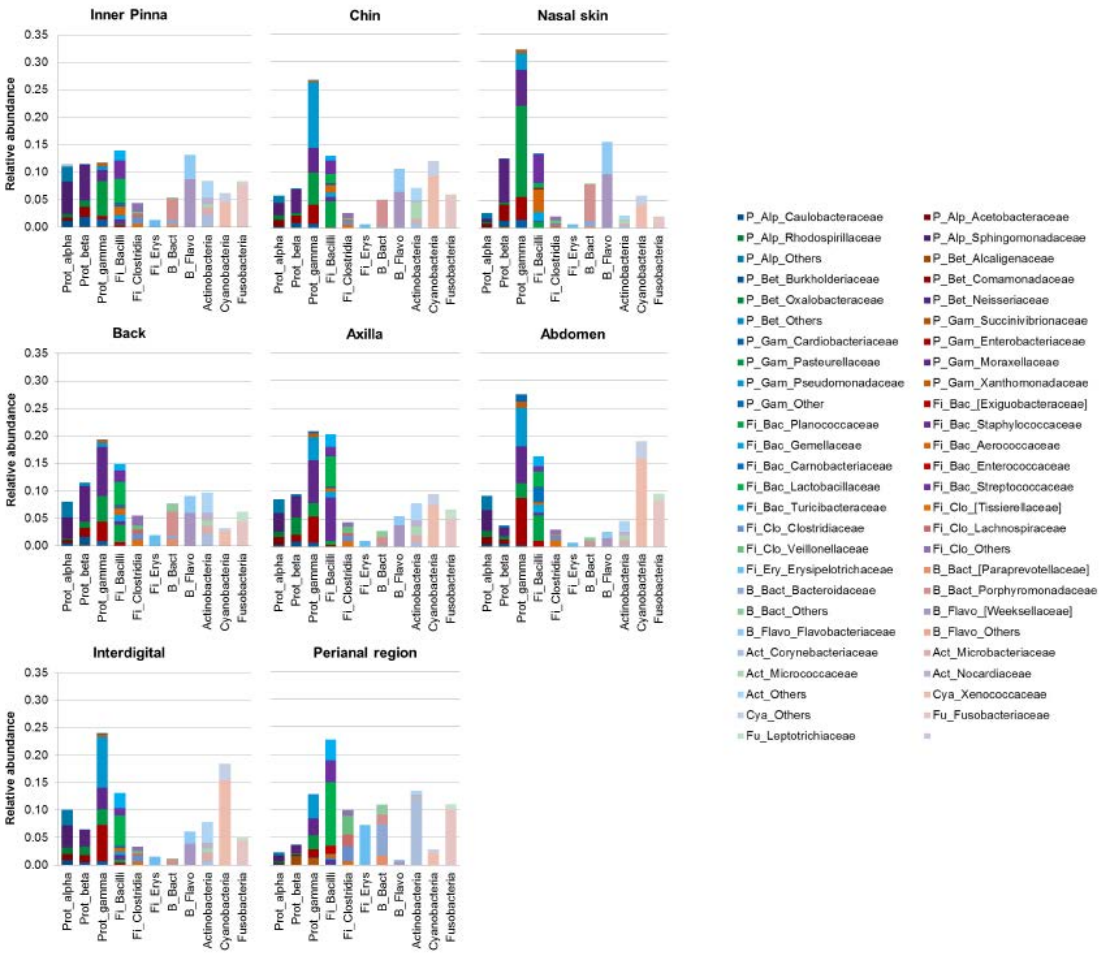


Figure 4. Taxonomic profiles per skin site. Taxa summary bar plots per class colored by main families within each skin site.

Influence of sample geographical origin: abdomen and dorsal back

To assess if the geographical origin of the samples had an effect on the microbiota, we performed an additional analysis with only abdomen and back regions, comparing this USA cohort with some European dogs (see Material and Methods for more details). While the USA cohort presented uniform characteristics (same crossbred dogs, similar ages, shared environment, etc.), the European cohort was much more variable including individuals from different breeds, ages and inhabiting in different households. Thus, we would expect a clear cluster for the USA controlled cohort and a diffuse if any clustering of Europe individuals. Samples were processed following the same protocol and the European cohort was re-analyzed together with the USA cohort, following the previously described analyses in Material and Methods section.

We visualized two clusters in both abdomen and back when considering the composition (Unweighted UniFrac), but those clusters lost power when analyzing the community structure (Weighted UniFrac). The grouping of back samples when considering the

geographical origin was strong (ANOSIM R=+0.68) and it significantly explained 12% of the variation in Unweighted UniFrac (Additional File 11A). In abdomen samples, grouping by geographical origin was lower (ANOSIM R=+0.26) and only explained 6% of the variation (Additional File 11B). Therefore, the back was better at explaining the differences in microbiota composition among geographical origin than the abdomen site. However, it is important to note that the back samples composition and structure is more uniform and less variable among individuals when compared to abdomen samples.

In conclusion, both in back and abdomen samples, the geographical origin explained some variation that was larger when looking at microbiota composition (Unweighted UniFrac). Thus, even when comparing two differently constituted cohorts, where European dogs were not homogenized by any variable, two clusters were clearly identified in beta diversity analyses representing geographical origin.

Discussion

Our results suggest that the main force driving the skin microbiota composition is the individual, rather than the skin site, even in a homogeneous cohort of dogs cohabiting and interacting together. This is in line with what we found previously in a cohort of nine healthy dogs from three different breeds, although in that study we could not elucidate whether the individual effect was real or represented an environmental influence (8). Here, we homogenized the cohort to account for different effects: same crossbreed dogs, same age and environment. Our results confirm the individual as the main factor determining skin microbiota composition in healthy dogs, when abundance of the bacterial species was not taken into account (Unweighted UniFrac). In human skin microbiota, low abundant species are those defining individuality and their signatures have the ability to identify items that the person came in contact with (13,42,43) and are relatively constant over time (44). An individual effect had also been reported as the main driver of fungal skin microbiota structure and composition in dogs from heterogeneous cohorts (45) and had been suggested to affect also bacterial skin microbiota in dogs, despite the individual was not assessed directly (6).

Our results also suggest that when relative abundances of the bacteria were taken into account (Weighted UniFrac), both skin site and individual affected the skin microbiota structure of healthy dogs. Similarly, these two factors also shaped human skin microbiota, with great variability within several skin sites of an individual and between individuals having been reported (11,46,47). Human skin has three main microhabitats (moist, dry and sebaceous) inhabited by specific taxa (12,48). Although the three microhabitats clearly identified in humans were not seen in dogs (29), Rodrigues-Hoffmann and colleagues reported significant differences between haired and mucosal or muco-cutaneous junctions (6), which coincides with our current observation. Here, we found that perianal region and, to lower extent, nasal skin presented different community structure (Weighted UniFrac) as well as lower alpha diversity values when compared to all other haired skin regions.

Globally in our cohort, Gammaproteobacteria followed by Bacilli were the most abundant classes in all regions in exception of perianal region with the same classes but the opposite order. A previous study including dorsal neck, abdomen and axilla samples from 40 domestic dogs inhabiting different households found Gammaproteobacteria and Bacilli as main classes, but also Actinobacteria (9). On the other hand, Hoffmann and colleagues (6) detected different abundant classes depending on the skin site: Betaproteobacteria was the most common in the concave pinna, dorsal lumbar and ear; Actinobacteria, in the axilla and interdigital skin; Gammaproteobacteria, in the nostril; and Clostridia and Bacteroidia, in the perianal region. Finally, in our previous study, we found Bacilli as the main class for all the skin sites with the exception of inner pinna that had Alphaproteobacteria (8). Thus, as the inter-individual variability is large, independent studies led to similar results only when a large number of individuals are included.

Network analysis elucidated the overall community organization throughout the skin of our canine cohort, with more than 40% of the interactions exclusive of each site, demonstrating a skin site signature. Back skin presented only two interactions and both of them were back-exclusive, probably other interactions remain hidden because only abundant families were included for network analysis. Among all skin sites, the inner pinna and chin were the sites that presented a higher proportion of unique interactions, suggesting stronger specialization or influences. On one hand, the inner pinna is an anatomically and environmentally isolated skin site when compared to others. On the other hand, we suggest that chin presented influences of both drinking water and oral microbiota. The most abundant families were *Xenococcaceae* and *Pseudomonadaceae*, which had been isolated in several water sources (49)(50). Moreover, the following abundant families, such as *Fusobacteriaceae*, *Moraxellaceae*, *Porphyromonadaceae*, *Neisseriaceae* and *Flavobacteriaceae*, were previously found as main taxa in canine oral microbiota (7,51).

Network analysis detected that most of the interactions among abundant families on the skin of Golden-Labrador Retriever crossbreds were co-occurrence rather than co-exclusion, which contrasts with what was previously seen for human microbiome with a balanced ratio of co-occurrence vs co-exclusion relationships within a body site (52). Despite the majority of co-occurrence interactions, few families presented a high number of mutual exclusions: *Pseudomonadaceae* and *Enterobacteriaceae* in abdomen and *Pseudomonadaceae* in axilla and chin. When blasting the OTUs that presented this apparently invasive pattern (Additional File 8), we found that the ones belonging to Enterobacteriaceae family had been mainly isolated from soil or plant surfaces (53,54); whereas those from *Pseudomonadaceae* family had been mainly isolated from soil and different sources of water (50). Thus, we suggest that this pattern is representing a recent exposure to the environment prior to sampling of some dogs. Other bacteria were suspected to have an environmental origin, despite not showing a co-exclusion pattern, such as *Xenococcaceae* with *Chroococcidiopsis* as its main genus. Bacteria from this genus had been mainly isolated from freshwater environments including lakes, soil, or inside of rocks (49). Moreover, they have already been detected on healthy dog skin (6,8). The presence of these bacteria with high abundance on the interdigital and abdomen regions may suggest

these two regions are more susceptible to environmental influences, which seems reasonable since these two skin sites have direct contact with the ground.

The skin sites could be classified based upon two patterns. The first pattern included sites having a high number of interactions among abundant families, with some interactions with other skin sites (chin, axilla, abdomen, and perianal region). The second pattern included sites having a lower number of interactions among abundant families and displayed exclusively within-site interactions (pinna, nasal skin, dorsal back and interdigital area). We suggest that the inter-site relationships could be explained due to topographical, behavioral and environmental factors. The chin is juxtaposed to the mouth, which is a main entrance for the environment through licking, eating, or drinking water. Dogs could lap the same water in which they are playing, and they usually lick themselves, which could explain some interactions among those sites. Additionally, the abdomen and axilla are anatomically continuous on the ventral side of the dog and close to the ground facilitating interactions with the environment and between the two skin sites. Furthermore, dogs may come into contact with fecal matter, which could explain shared OTUs among the abdomen, axilla and perianal regions. Main families of the second pattern, constituted by the inner pinna, dorsal back, interdigital area and nasal skin, were only interacting with other families in the same skin site, suggesting that both anatomical isolation and stronger effects of other microbiota (nostril microbiota, for nasal skin and soil microbiota for interdigital region) may account for the exclusive within-site interactions.

With this general overview, we sought to elucidate if any individual-specific variable determined the observed diversity, composition, and/or community structure in any of the skin sites. When considering the season of birth or the time spent in the kennel as variable, we observed two significantly different groups: dogs born from January to May that had spent 6 months in the kennel were different from those born from June to September that had spent 3 months in the kennel. This effect was highest on the inner pinna, with a significant ANOSIM R value of +0.84. Main taxonomic difference among inner pinna from both groups was due to *Sphingomonadaceae*, specifically *Sphingomonas*. These taxa are classically considered air- and dust-borne (55), although they had also been identified on dog skin microbiota (5,8) and in animal sheds (55,56), even specifically on dogs' (57). These classically considered air- and dust-borne bacteria are cultivable at temperatures ranging from 4–28 °C, but not at 37 °C (55). Independent studies of grapevine microbiome show a link between *Sphingomonadaceae* and lower temperatures: in leaves *Sphingobacteriaceae* was significantly overrepresented on samples from May when compared to July (58); and in grapes, samples from a colder year presented increased levels of *Sphingomonadaceae* (59). Thus, we suggest two alternate hypotheses. If the variable explaining those differences was the season of birth, maybe the maximum temperatures reached in the summer time, usually overpassing 28ºC, implied lower *Sphingomonadaceae* presence in the air, in dust and on dogs' hair coats, avoiding its establishment as a component of the dogs' skin microbiota. In human studies, infant skin microbiota resembled the maternal one up to six weeks after birth, contrasting with microbiota from stool, nares or oral cavity which significantly differed (60). Thus, applying few concepts of ecology theory (61), we suggest that the bacterial pool of the environment and the air was shaped with the season characteristics

(humidity, UV light, temperature, etc.), which in turn was shaping to some extent the skin microbiota of dogs (environmental selection). Thus, when a dog was born it had a different available bacterial pool depending on the season that will affect the invasion order: some species were more abundant, so more likely to be the first to colonize the skin. Or perhaps the dam had an increased number of season-specific bacteria on her shed, transferring them to the puppies and therefore becoming resident inhabitants of the puppy skin microbiota by historical contingency processes, where the order of invasion matters and it could randomly differ among littermates (for example, being born first vs second). If the time spent in the kennel was the variable explaining these divergences, we could be detecting again an environmental effect: those dogs that were during the autumn and winter in the kennel were exposed for a longer period of time to the environment and were more likely to harbor bacteria from it. In this case we should accept a high influence of the environment in the skin microbiota composition, entitling even a replacement of the main families depending on the time of cohabitation and the household placement conditions. Finally, we should not forget the possible role of progenitors' genetics, but in both groups there are some common sires despite different dams discarding the possible effect of sire genetics. Moreover, inner pinna skin microbiota is not more similar when comparing two littermate's than when comparing any other two dogs within the same group (in exception of Dog 2 and 3).

Although it is difficult to elucidate which bacteria are really microbiota and which others are only transient members from the environment, the pattern seems clear for *Sphingomonadaceae*. All the samples were obtained the same day and a significant difference was found regarding the season of birth or time spent in the kennel. Therefore, *Sphingomonadaceae*, classically regarded as air- and dust-borne bacteria, should be considered also a normal colonizer of dog skin microbiota. Similarly as dog skin with *Sphingomonadaceae*, researchers had found taxa classically regarded as environmental being part of human skin microbiota. For example, the genus *Enhydrobacter* was commonly found in air and surfaces of built environment of Hong Kong (62) and also presented high abundances in skin of Chinese individuals (21,63). Another example would be Amerindian individuals, which presented a very different skin microbiota profile with high proportion of bacteria commonly regarded as environmental (25).

Besides the season of birth or the time spent in the kennel, sex had a significant effect on the abdomen, back and axilla microbiota of our cohorts. Female dogs presented an overrepresentation of Enterobacteriales and *Enterobacteriaceae* families, coinciding with what was previously reported on the hands of humans (14).

Dogs that had undergone surgery within the previous month presented lower alpha diversity values, always in abdomen and chin. The surgery procedures that had undergone implied shaving the abdomen and were followed by oral medication administration (sometimes antibiotics), which would probably explain the lowered alpha diversity values.

Finally, geography also affected dog skin microbiota, most significantly its composition. The European cohort was comprised of pet dogs that did not interact with each other, were from different households, age groups and genetic backgrounds, and their samples

were collected in different seasons of the year. Even considering this heterogeneity, the European dogs clustered together in a single group differing from the environmental well-controlled USA cohort. That is in line with what has already been described for humans with geography (21,26), geographical isolation (25) or urbanization (27,28)(19) grouping differently skin microbial communities.

Focusing on the technique, our results showed that skin microbiota samples collected with swabs and stored at 4ºC for more than 5 days presented significantly less diversity when compared to those extracted sooner; this was true even when the dogs who had surgery, and were therefore less diverse, were excluded (Additional File 4D). Each day a different set of random samples was extracted and our results demonstrate that there were no impacts of extraction within the first 3-days; neither were differences detectable among the last 3-days. However, microbiota diversity values differed when extraction for weeks one and two were compared. This contrasts with the results reported by Lauber and colleagues that assessed stability of the skin microbiota samples when stored in different temperatures. They found no differences on diversity when they were stored either at 4ºC or 20ºC for up to two weeks (64). However, they worked at a sequencing depth 10 times lower than our study: 1,000 vs 11,000 sequences per sample. The lower coverage may have obscured the loss of diversity because the lower prevalent species may not have been detected at the lower sequencing depth. Thus, we recommend extracting microbial DNA from skin swab samples within 3-days of sample collection with storage at 4ºC. Finally, it is important to note that this decrease in the diversity only slightly affected skin community structure and composition (explaining ≤5% of the variation).

Conclusions

In summary, we have characterized the normal variability of dog skin microbiota in a well-controlled cohort of a large number of dogs with similar ages, related genetic background, and a shared environment. We found that microbiota composition was driven by the individual, but when considering the abundances, microbiota structure was driven both by the individual and by the skin site. Network analyses elucidated that both exclusive and shared interactions existed depending on the skin site, with the highly shared interactions probably representing an environmental origin. When analyzing each skin site independently to assess host-specific factors we found that season of birth or time spent in the kennel affected all skin sites. The most abundant taxon driving this difference was *Sphingomonas*, which is an air-borne bacterium that cannot be cultivated at elevated temperatures. We also found taxonomic differences among male and female dogs on abdomen, axilla and back. Moreover, the USA and European cohorts were grouping by geographical origin in two different and well-defined clusters, even when the European individuals were very heterogeneous. In conclusion, we observed a large inter-individual variability and effects of different host variables, even in an environmental well-controlled cohort.

Thus, to overcome the individual variability inherent to skin microbiota studies, we would recommend longitudinal studies assessing divergences between health and disease comparing affected vs unaffected regions within an individual through time; or alternatively the cohort should be large enough and well controlled if case-control studies are preferred. Understanding the skin microbiota of healthy skin will allow a better knowledge of the intrinsic variability in health and the assessment of what is an altered state. It will also provide a background to develop its clinical applications (30) such as identifying an altered skin microbiota landscape or developing personalized therapies aimed at shifting the balance toward a healthy skin microbiota, promoting beneficial bacteria growth rather than killing all bacteria.

Declarations

Ethics approval and consent to participate

All animal work was done with the approval of the University of California, Davis Institutional Animal Care and Use Committee and the scientific research oversight committee of Canine Companions for Independence.

The European dogs included in this study were examined during routine veterinary procedures by the veterinary clinics participating in the study. All samples were collected and used in the study with verbal owner consent. As the data are from client-owned dogs that underwent normal preventative veterinary examinations, there was no "animal experiment" according to the legal definitions in Spain, and approval by an ethical committee was not necessary.

Consent for publication

Not applicable.

Availability of data and material

The datasets analyzed during the current study are available in the SRA NCBI repository under the Bioproject accession number PRJNA384381.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was partly supported by a grant awarded by Generalitat de Catalunya (Industrial Doctorate program, 2013 DI 011).

Authors' contributions

AO, JM, OF, AS and AC conceived and designed the experiment. OF, AO, AS and JM supervised the project and gave conceptual advice. AO, JB, AI, KL, and AC participated in the sample collection. JB, AI, LG, and AC performed the DNA extractions. AC performed the PCRs. KL recollected all the metadata. AC carried out the bioinformatics analysis. AC drafted the manuscript. OF, AO, JB, and LG edited the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We would like to acknowledge Canine Companions for the Independence organization for providing the skin samples from their dogs. We would also acknowledge Nicolas Boulanger and Joana Ribes for the sequencing service and Gonzalo Vera for the informatics support.

Additional files

They are available at the following link:

https://www.dropbox.com/sh/kcd1bo4adzlh439/AADsgFQ5MwNYTieimIITZUUAa?dl=0

Additional File 1.xlsx - Information about the dogs included in the study. Sample were collected the 2016/04/27.

Additional File 2.docx - Pedigree chart of the dogs included in this study. Circle represent female and rectangle represent male. In blue, dogs born from January to May that had spent at least 6 months in the kennel (Jan-May group) and in red dogs born from June to September that had spent 3 months in the kennel (Jun-Sep group).

Additional File 3.xlsx - OTU table at genus level including all the samples.

Additional File 4.xlsx - Alpha diversity values and statistics.

Additional File 5.docx - Unweighted UniFrac beta diversity PCoA plot per skin site colored by season of birth or time spent in the kennel. In red, dogs born from January to May that had spent at least 6 months in the kennel (Jan-May group) and in blue dogs born from June to September that had spent 3 months in the kennel (Jun-Sep group).

Additional File 6.docx - Differentially distributed families based on season of birth or time spent in the kennel. Histogram of linear discriminant analysis (LDA) effect size (LefSe) for differentially abundance distribution ($\alpha = 0.05$, LDA score >3).

Additional File 7.xlsx - Network output. CoNet output tables with edge and node information.

Additional File 8.xlsx - Taxonomies of the highly-connected nodes obtained through BLAST.

Additional File 9.xlsx - ANOSIM R values for each pair of skin sites and both for Weighted and Unweighted UniFrac matrices.

Additional File 10.docx - Differentially distributed families based on skin site. Histogram of linear discriminant analysis (LDA) effect size (LefSe) up to family level for differentially abundant distributed taxa ($\alpha = 0.05$, LDA score >3).

Additional File 11.docx - Geographical origin effect on beta diversity for back and abdomen samples. Samples from this study (USA) were merged with previous samples (Spain) (Cuscó et al., 2017) as well as two other unpublished individuals. Unweighted UniFrac beta diversity plots of (A) dorsal back and (B) abdomen samples colored by geographical origin with their associated ANOSIM and adonis values.

Bibliography

1. Naik S, Bouladoux N, Wilhelm C, Molloy MJ, Salcedo R, Kastenmuller W, et al. Compartmentalized Control of Skin Immunity by Resident Commensals. Science 2012, 337:1115–9.

2. Cho I, Blaser MJ. The human microbiome: at the interface of health and disease. Nat Rev Genet 2012, 13:260–70.

3. Miller WH, Griffin CE, Campbell KL. Muller & Kirk's Small Animal Dermatology. 7th ed. St Louis: Elsevier; 2013.

4. Weese JS. The canine and feline skin microbiome in health and disease. Vet Dermatol 2013, 24:137–46.

5. Song SJ, Lauber C, Costello EK, Lozupone CA, Humphrey G, Berg-Lyons D, et al. Cohabiting family members share microbiota with one another and with their dogs. Elife 2013, 2:e00458.

6. Rodrigues Hoffmann A, Patterson AP, Diesel A, Lawhon SD, Ly HJ, Elkins Stephenson C, et al. The skin microbiome in healthy and allergic dogs. PLoS One 2014, 9:e83197.

7. Bradley CW, Morris DO, Rankin SC, Cain CL, Misic AM, Houser T, et al. Longitudinal evaluation of the skin microbiome and association with microenvironment and treatment in canine atopic dermatitis. J Invest Dermatol 2016, 136:1182–90.

8. Cuscó A, Sánchez A, Altet L, Ferrer L, Francino O. Individual Signatures Define Canine Skin Microbiota Composition and Variability. Front Vet Sci 2017, 4:6.

9. Torres S, Clayton JB, Danzeisen JL, Ward T, Huang H, Knights D, et al. Diverse bacterial communities exist on canine skin and are impacted by cohabitation and time. PeerJ 2017, 5:e3075.

10. Pierezan F, Olivry T, Paps JS, Lawhon SD, Wu J, Steiner JM, et al. The skin microbiome in allergen-induced canine atopic dermatitis. Vet Dermatol 2016, 5:332–e82

11. Grice EA, Segre JA. The skin microbiome. Nat Rev Microbiol 2011, 9:244–53.

12. Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, et al. Topographical and temporal diversity of the human skin microbiome. Science 2009, 324:1190–2.

13. Fierer N, Lauber CL, Zhou N, McDonald D, Costello EK, Knight R. Forensic identification using skin bacterial communities. Proc Natl Acad Sci U S A 2010, 107:6477–81.

14. Fierer N, Hamady M, Lauber CL, Knight R. The influence of sex, handedness, and washing on the diversity of hand surface bacteria. Proc Natl Acad Sci U S A 2008, 105:17994–9.

15. Zeeuwen PL, Boekhorst J, van den Bogaard EH, de Koning HD, van de Kerkhof PM, Saulnier DM, et al. Microbiome dynamics of human epidermis following skin barrier disruption. Genome Biol 2012, 13:R101.

16. Callewaert C, Kerckhof F-M, Granitsiotis MS, Van Gele M, Van de Wiele T, Boon N. Characterization of Staphylococcus and Corynebacterium Clusters in the Human Axillary Region. PLoS ONE 2013, 8:e70538.

17. Capone KA, Dowd SE, Stamatas GN, Nikolovski J. Diversity of the human skin microbiome early in life. J Invest Dermatol 2011, 131:2026–32.

18. Oh J, Conlan S, Polley E, Segre JA, Kong HH. Shifts in human skin and nares microbiota of healthy children and adults. Genome Med 2012, 4:77.

19. Lehtimäki J, Karkman A, Laatikainen T, Paalanen L, von Hertzen L, Haahtela T, et al. Patterns in the skin microbiota differ in children and teenagers between rural and urban environments. Sci Rep 2017, 7:45651.

20. The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. Nature 2012, 486:207–14.

21. Leung MHY, Wilkins D, Lee PKH. Insights into the pan-microbiome: skin microbial communities of Chinese individuals differ from other racial groups. Sci Rep 2015, 5:11845.

22. Perez Perez GI, Gao Z, Jourdain R, Ramirez J, Gany F, Clavaud C, et al. Body Site Is a More Determinant Factor than Human Population Diversity in the Healthy Skin Microbiome. PLoS One 2016, 11:e0151990.

23. Dominguez-bello MG, Costello EK, Contreras M, Magris M, Hidalgo G. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. Proc Natl Acad Sci U S A 2010, 107:11971–5.

24. Misic AM, Davis MF, Tyldsley AS, Hodkinson BP, Tolomeo P, Hu B, et al. The shared microbiota of humans and companion animals as evaluated from Staphylococcus carriage sites. Microbiome 2015, 3:2.

25. Clemente JC, Pehrsson EC, Blaser MJ, Sandhu K, Gao Z, Wang B, et al. The microbiome of uncontacted Amerindians. Sci Adv 2015, 1:e1500183.

26. Hospodsky D, Pickering AJ, Julian TR, Miller D, Gorthala S, Boehm AB, et al. Hand bacterial communities vary across two different human populations. Microbiology 2014, 160:1144–52.

27. Hanski I, von Hertzen L, Fyhrquist N, Koskinen K, Torppa K, Laatikainen T, et al. Environmental biodiversity, human microbiota, and allergy are interrelated. Proc Natl Acad Sci 2012, 109:8334–39.

28. Ying S, Zeng D, Chi L, Tan Y, Galzote C, Cardona C. The Influence of Age and Gender on Skin-Associated Microbial Communities in Urban and Rural Human Populations. PLoS One 2015, 10:e0141842.

29. Rodrigues Hoffmann A. The cutaneous ecosystem: the roles of the skin microbiome in health and its association with inflammatory skin conditions in humans and animals. Vet Dermatol 2017, 1:60–e15.

30. Grice EA. The skin microbiome: potential for novel diagnostic and therapeutic approaches to cutaneous disease. Semin Cutan Med Surg 2014, 33:98–103.

31. Kuczynski J, Lauber CL, Walters WA, Parfrey LW, Clemente JC, Gevers D, et al. Experimental and analytical tools for studying the human microbiome. Nat Rev Genet 2011, 13:47–58.

32. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 2010, 7:335–6.

33. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH : a versatile open source tool for metagenomics. PeerJ 2016, 4:e2584.

34. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 2007, 73:5261–7.

35. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol 2006, 72:5069–72.

36. Caporaso JG, Bittinger K, Bushman FD, Desantis TZ, Andersen GL, Knight R. PyNAST: A flexible tool for aligning sequences to a template alignment. Bioinformatics 2010, 26:266–7.

37. Navas-molina JA, Peralta-sánchez JM, González A, McMurdie PJ, Vázquez-Baeza Y, Xu Z. Advancing Our Understanding of the Human Microbiome Using QIIME. Methods Enzymol 2013, 531:371–444.

38. Lozupone C, Knight R. UniFrac : a New Phylogenetic Method for Comparing Microbial Communities. Appl Environ Microbiol 2005, 71:8228–35.

39. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. Genome Biol 2011, 12:R60.

40. Faust K, Raes J. CoNet app : inference of biological association networks using Cytoscape. F1000Research 2016, 5:1519.

41. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003, 13:2498–504.

42. Meadow JF, Altrichter AE, Green JL. Mobile phones carry the personal microbiome of their owners. PeerJ 2014, 2:e447.

43. Lax S, Hampton-Marcell JT, Gibbons SM, Colares GB, Smith D, Eisen JA, et al. Forensic analysis of the microbiome of phones and shoes. Microbiome 2015, 3:21.

44. Oh J, Byrd AL, Park M, Kong HH, Segre JA. Temporal Stability of the Human Skin Microbiome. Cell 2016, 165:854–66.

45. Meason-smith C, Diesel A, Patterson AP, Older CE, Mansell JM, Suchodolski JS, et al. What is living on your dog's skin ? Characterization of the canine cutaneous mycobiota and fungal dysbiosis in canine allergic dermatitis. FEMS Microbiol Ecol 2015, 91:1–12.

46. Oh J, Byrd AL, Deming C, Conlan S, Barnabas B, Blakesley R, et al. Biogeography and individuality shape function in the human skin metagenome. Nature 2014, 514:59–64.

47. Ursell LK, Clemente JC, Rideout JR, Gevers D, Caporaso JG, Knight R: The interpersonal and intrapersonal diversity of human-associated microbiota in key body sites. J Allergy Clin Immunol 2012, 129:1204-8

48. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. Bacterial community variation in human body habitats across space and time. Science 2009, 326:1694–7.

49. Cumbers J, Rothschild LJ. Salt tolerance and polyphyly in the cyanobacterium Chroococcidiopsis (Pleurocapsales). J Phycol 2014, 50:472–82.

50. Peix A, Ramírez-Bahena M-H, Velázquez E. Historical evolution and current status of the taxonomy of genus Pseudomonas. Infect Genet Evol 2009, 9:1132–47.

51. Sturgeon A, Stull JW, Costa MC, Weese JS. Metagenomic analysis of the canine oral cavity as revealed by high-throughput pyrosequencing of the 16S rRNA gene. Vet Microbiol 2013, 162:891–8.

52. Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, Raes J, et al. Microbial co-occurrence relationships in the human microbiome. PLoS Comput Biol 2012, 8:e1002606.

53. Palmer M, Steenkamp ET, Coetzee MPA, Chan W-Y, van Zyl E, De Maayer P, et al. Phylogenomic resolution of the bacterial genus Pantoea and its relationship with Erwinia and Tatumella. Antonie Van Leeuwenhoek 2017, 1–23.

54. Walterson AM, Stavrinides J. Pantoea: insights into a highly versatile and diverse genus within the Enterobacteriaceae. FEMS Microbiol Rev 2015, 39:968–84.

55. Busse H-J, Denner EBM, Buczolits S, Salkinoja-Salonen M, Bennasar A, Kämpfer P. Sphingomonas aurantiaca sp. nov., Sphingomonas aerolata sp. nov. and Sphingomonas faeni sp. nov., air- and dustborne and Antarctic, orange-pigmented, psychrotolerant bacteria, and emended description of the genus Sphingomonas. Int J Syst Evol Microbiol 2003, 53:1253–60.

56. Andersson AM, Weiss N, Rainey F, Salkinoja-Salonen MS. Dust-borne bacteria in animal sheds, schools and children's day care centres. J Appl Microbiol 1999, 86:622–34.

57. Miletto M, Lindow SE. Relative and contextual contribution of different sources to the composition and abundance of indoor air bacteria in residences. Microbiome 2015, 3:61.

58. Pinto C, Pinho D, Sousa S, Pinheiro M, Egas C, Gomes AC. Unravelling the Diversity of Grapevine Microbiome. PLoS One 2014, 9:e85622.

59. Zarraonaindia I, Owens SM, Weisenhorn P, West K, Hampton-Marcell J, Lax S, et al. The soil microbiome influences grapevine-associated microbiota. MBio 2015, 6: e02527-14.

60. Chu DM, Ma J, Prince AL, Antony KM, Seferovic MD, Aagaard KM. Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery. Nat Med 2017, 23:314–26.

61. Costello EK, Stagaman K, Dethlefsen L, Bohannan BJM, Relman DA. The application of ecological theory toward an understanding of the human microbiome. Science 2012, 336:1255–62.

62. Leung MHY, Lee PKH. The roles of the outdoors and occupants in contributing to a potential pan-microbiome of the built environment: a review. Microbiome 2016, 4:21.

63. Ling Z, Liu X, Luo Y, Yuan L, Nelson KE, Wang Y, et al. Pyrosequencing analysis of the human microbiota of healthy Chinese undergraduates. BMC Genomics 2013, 14:390.

64. Lauber CL, Zhou N, Gordon JI, Knight R, Fierer N. Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples. FEMS Microbiol Lett 2010, 307:80–6.

3.4. Using MinION™ to characterize dog skin microbiota through full-length 16S rRNA gene sequencing approach

This chapter consists of the article entitled "Using MinION™ to characterize dog skin microbiota through full-length 16S rRNA gene sequencing approach" pending to be submitted.

Using MinION™ to characterize dog skin microbiota through full-length 16S rRNA gene sequencing approach

Introduction

Bacteria, fungi, viruses and archaea are the main microorganisms constituting the microbiota, which is defined as the microbial communities inhabiting a specific environment (1). In humans, many efforts have been made to characterize the different body site ecosystems and their associated microbial communities, mainly at bacterial level (2,3), which are the most abundant microorganisms on the human-associated microbiota (4,5).

Studying host-associated microbiota has provided many insights on health and diseases for many different body sites (6,7). In human skin, alterations on skin microbiota have been associated to numerous cutaneous diseases, such as acne vulgaris (8,9), psoriasis (10–12), or atopic dermatitis (13–17). Not only humans, but also dogs presented altered microbiota states during atopic dermatitis disease (18–20).

The most common strategy to assess bacterial microbiota is amplifying and sequencing specific regions of 16S rRNA gene using 2nd generation massive sequencing technologies (for a review see (21)). This bacterial marker gene is ubiquitously found in bacteria, and has nine hypervariable regions (V1-V9) that can be used to infer taxonomy (22).

The ability to classify sequences to the genus or species level is a function of read length, sample type and the reference database (23). High-quality short-reads obtained from 2nd generation sequencers (250-350 bp) bias and limit the taxonomic resolution of this gene. The more usual regions amplified with Illumina MiSeq or Ion Torrent PGM™ for bacterial taxonomic classification are V4 or V4-V5, but these regions fail in amplifying some significant species for skin microbiota studies, such as *Propionibacterium acnes*. So, when performing a skin microbiota study the preferred choice is V1-V2 regions, although they lack sensitivity for the genus *Bifidobacterium* and poorly amplify the phylum *Verrucomicrobia* (21). On the other hand, near full-length 16S rRNA gene sequences are required for accurate richness estimations especially at higher taxa (24), which are necessary on microbiota studies. Besides, full-length reference sequences are needed for performing phylogenetic analyses or designing lineage specific primers (23), especially in species different to human or mouse, in which previous metagenomics approaches deciphered the richness of bacterial species in the great and different variety of microbiome samples analyzed.

With the launching of 3rd generation single-molecule technology sequencers, these short-length associated problems can be overcome by sequencing the full or almost full-length of 16S rRNA gene with different sets of universal primers (25). Results for full-length 16S rRNA gene have been reported for Pacific Biosciences (PacBio) platform (23,26–30). Schloss and collaborators reported the possibility of generating near full-length 16S rRNA gene sequences with error rates slightly higher, but comparable to the 2nd generation platforms (0.03%)(23). The primary limitation on the PacBio platform is the accessibility to the sequencers and the cost of generating the data.

MinION™ sequencer of Oxford Nanopore Technologies (ONT) (https://nanoporetech.com) is a 3rd generation sequencer that is portable, affordable with a small budget and offers long-read output (only limited by DNA extraction protocol). Besides, it can provide a rapid real-time and on-demand analysis very useful on clinical applications. Several studies targeting the full 16S rRNA gene have already been performed using MinION™ to: i) identify pure bacterial culture (31); ii) characterize artificial and already-characterized bacterial communities (mock community) (32–34); and to iii) characterize complex microbiota samples, from mouse gut (35), wastewater (31) and pleural effusion from a patient with empyema (34).

Here we aim to assess the potential of Nanopore sequencing in complex microbiota samples using the full-length 16S rRNA (1,500bp). First set-up step is performed using a staggered mock community (HM-783D). Then, we sequenced a pool of several skin microbiota samples previously sequenced by Ion Torrent PGM™.

Material and methods

Samples included

As simple microbial community, we used a Microbial Mock Community HM-783D kindly donated by BEI resources (http://www.beiresources.org) that contained genomic DNA from 20 bacterial strains with staggered ribosomal RNA operon counts (1,000 to 1,000,000 copies per organism per μL). This mock community allowed us to perform the MinION™ sequencing and analysis protocol set-up.

As complex microbial community, we used a sample pool from skin microbiota of healthy dogs (inner pinna), which had been previously characterized using Ion Torrent PGM™. We assessed skin microbiota from inner pinna samples: 1A, 7A, 18A, 20A and 29A from (36) (or Chapter 3.3).

Sample collection and DNA extraction

For the complex microbial community, skin microbiota samples were collected using Sterile Catch-All™ Sample Collection Swabs (Epicentre Biotechnologies) soaked in sterile SCF-1 solution (50 mM Tris buffer (pH = 8), 1 mM EDTA, and 0.5% Tween-20).

Bacterial DNA was extracted from the swabs using the PowerSoil™ DNA isolation kit (MO BIO) (for further details on sample collection and DNA extraction see (37)).

PCR amplification and barcoding

To prepare the DNA and the library we followed the Oxford Nanopore protocol 1D PCR barcoding amplicons (SQK-LSK108), however we used the Phusion Taq polymerase rather than the LongAmp Taq recommended in this protocol. Specifically, we amplified ~1,500bp fragments of the full 16S rRNA gene.

Bacterial DNA was amplified using a nested PCR with a first round to add the 16S rRNA gene primer sets and a second round to add the barcodes. In this study we used two sets of 16S universal primers. On one hand, primer set 27F-1391R (also named S-D-Bact-0008-c-S-20 and S-D-Bact-1391-a-A-17 (38)) amplified V1-V8 hypervariable regions of 16S rRNA gene. On the other hand, primer set 27F-1492R (also named S-D-Bact-0008-c-S-20 and S-D-Bact-1492-a-A-22 (38)) amplified V1-V9 hypervariable regions of 16S rRNA gene. These two sets of universal primers are the most commonly used when assessing full-length 16S rRNA gene, because they have shown a really low non-coverage rate, even at phylum level (39). The primers used in this study are listed in Table 1 and contain some ambiguous bases previously described to make the primers more universal (25).

Table 1. Primer sequences and hypervariable regions (HVR) targeted for full-length 16S rRNA gene amplification and sequencing.

Complete name	Short name	HVR	Sequence (5' --> 3')	Melting T
S-D-Bact-0008-c-S-20	27F	V1	AGRGTTYGATYMTGGCTCAG	54.4 ºC
S-D-Bact-1391-a-A-17	1391R	V8	GACGGGCGGTGWGTRCA	59.5 ºC
S-D-Bact-1492-a-A-22	1492R	V9	TACCTTGTTAYGACTT	41.6 ºC

We will distinguish among primer sets used referring to the hypervariable regions they are amplifying, so: 27F-338R will be V1-V2; 27F-1391R will be V1-V8; and 27F-1492R will be V1-V9.

We ordered the 16S rRNA gene primers with the Oxford Nanopore Universal Tag added to their 5' end. The universal tag was 5'-TTTCTGTTGGTGCTGATATTGC-3' for forward primers and 5'-ACTTGCCTGTCGCTCTATCTTC-3' for reverse primers. These universal tags will allow the second barcoding PCR using the PCR Barcoding kit (EXP-PBC001).

In the first round, PCR mixture (25 µl) contained initial DNA sample (1 µl of DNA for the mock community PCR and 5 µl of DNA for the skin microbiota PCR), 5 µl of 5X Phusion Buffer HF, 0.2 mM of dNTPs, 0.02 U/µl of Phusion High Fidelity Taq Polymerase (Thermo Scientific). Primer concentrations were adapted to each primer set: for 27F-1391R, 0.4 µM of each primer and for 27F-1492R, 0.4µM of 27F and 0.8µM of 1492R. The PCR thermal profile consisted of an initial denaturation step for 30s at 98°C, followed by

25 cycles for 15s at 98°C, 15s at primer-adjusted annealing temperature, 45s at 72°C for extension, and a final step for 7 min at 72°C. The annealing temperature was also adjusted to the primer set: 55ºC for 27F-1391R and 51ºC for 27F-1492R. To assess possible reagent contamination, each PCR reaction included a no template control sample, which did not amplify.

In the second round, PCR mixture (100 µl) contained 0.5 nM of the first-round PCR product, 20 µl of 5X Phusion Buffer HF, 0.2 mM of dNTPs, 0.02 U/µl of Phusion High Fidelity Taq Polymerase (Thermo Scientific), and 2 µl of each specific barcode (EXP-PBC001) as recommended in the Oxford Nanopore protocol 1D PCR barcoding amplicons (SQK-LSK108). The PCR thermal profile consisted of an initial denaturation step for 30s at 98°C, followed by 15 cycles for 15s at 98°C, 15s at 62ºC for annealing, 45s at 72°C for extension, and a final extension step for 7 min at 72°C.

Following each PCR round, a clean-up step using AMPure XP beads at 0.5X concentration was used to discard short fragments as recommended by the manufacturer. DNA quantity was assessed using Qubit fluorimeter.

A final equimolar pool containing 1ug of the barcoded DNA samples in 45 uL of DNAse and RNAse free water will be used to prepare the sequencing library.

Library preparation

The Ligation Sequencing Kit 1D (SQK-LSK108) was used to prepare the amplicon library to load into the MinION™ following the instructions of the 1D PCR barcoding amplicon protocol of ONT. Input DNA samples were1 µg of the barcoded DNA pool in a volume of 45 µL and 5 µL of DNA CS (DNA from lambda phage, used as a sequencing positive control). The DNA was processed for end repair and dA-tailing using the NEBNext End Repair / dA-tailing Module (New England Biolabs). A purification step using Agencourt AMPure XP beads (Beckman Coulter) was performed and approximately the expected 700ng of total DNA were recovered as assessed by Qubit quantification.

For the adapter ligation step, a total of 0.2 pmol of the end-prepped DNA (approximately 200 ng of our 1,500 bp fragment) were added in a mix containing 50µL of Blunt/TA ligase master mix (New England Biolabs) and 20µL of adapter mix, and were incubated at room temperature for 10 min. We performed a purification step using Agencourt AMPure XP beads (Beckman Coulter) and Adapter Bead Binding buffer provided on SQK-LSK108 kit to finally obtain the DNA library also called pre-sequencing mix.

Preparing and loading the flow cell

We used SpotON Flow Cell Mk I (R9.4) (FLO-MIN106), which had been previously stored at 4ºC. We fitted the flow cell to the MinION™ and performed the Quality Control. We continue with the priming of the flowcell with a mixture of Running Buffer with fuel mix (RBF from SQK-LSK108) and Nuclease-free water (500 µL + 500 µL).

We prepared the pre-sequencing mix (12 µL of DNA library) to be loaded by mixing it with Library Loading beads (25.5 µL) and Running Buffer with fuel mix (37.5 µL). Immediately after priming, the nanopore sequencing library was loaded in a dropwise fashion using the spot-on port.

We run a total of two flow cells including several barcoded samples. The first one contained the mock communities amplified using V1-V9 primer set (M1 and M2 are technical replicates) together with other skin microbiota samples not included in this study. The second flow cell included the skin microbiota sample from the inner pinna amplified using V1-V8 and V1-V9 primer sets (biological replicates) described in this study, among others.

Once the library was loaded we initiated a standard 48h sequencing protocol using the MinKNOW™ software.

For the mock communities, basecalling was performed using the Metrichor™ agent 1D barcoding for pre-existing basecalls and demultiplexing using Epi2me debarcoding workflow. Finally, fast5 files were converted to fastq files using poRe (40)(Watson et al., 2015) and adapters were trimmed using Porechop (41).

On the other hand, for the skin microbiota samples basecalling was performed using Albacore v0.8.4 software. Again, fast5 files were converted to fastq files using poRe (40). Afterwards sequences were demultiplexed and adapters trimmed using Porechop (41).

As a final step, we trimmed the universal tags of the sequences using a custom script and filtered out those sequences shorter than 1,100 bp for V1-V8 and 1,200 bp for V1-V9 amplifications respectively, using split_libraries.py from QIIME software (42).

Downstream analyses

For both the mock community and the skin microbiota samples, we performed the analysis using NanoOK (43) with LAST aligner (44) against two databases. We used a subset of Greengenes database (45,46) and the *rrn* database (33).

Greengenes database offers annotated, chimera-checked, and full-length 16S rRNA gene sequences and it is one of the most commonly used databases when performing microbiota analyses(45,46). We used the Greengenes database clustered at 99% of similarity to reduce redundancy and filtered out those sequences that did not reach species level or that did not have a minimum length of 1,400bp. This database contained 20,745 sequences belonging to 3,147 different species.

rrn database is a custom database created by Benitez-Paez and Sanz to analyze the rrn operons. It contains information of the ribosomal RNA operon (16S-ITS-23S genes) retrieved from bacterial genomes of GenBank at NCBI. This database contained 22,351 sequences belonging to 2,384 different species (33).

Results and discussion

Mock community analyses

We amplified full-length 16S rRNA sequences from the staggered community with primers V1-V9 by duplicate (M1 and M2). We processed a total of 11,284 sequences for M1 and 22,995 for M2. The taxonomic results obtained are shown in Table 2, both for Greengenes and *rrn* databases.

Table 2. Taxonomic assignment of the mock community at species level. Results obtained after MinION™ sequencing of two replicates of the staggered mock community (M1 and M2) aligned against Greengenes and rrn databases. Relative abundances that correlated to operon counts are in bold. (*) Is the species annotated in the database?

		Greengenes database			rrn database		
Taxonomy	**nº of operons**	**Tax at sps?***	**% reads M1**	**% reads M2**	**Tax at sps?***	**% reads M1**	**% reads M2**
Escherichia coli	1,000,000	**Yes**	**15.29**	**16.45**	**Yes**	**22.49**	**23.92**
Rhodobacter sphaeroides	1,000,000	Yes	8.54	7.19	Yes	8.18	6.85
Staphylococcus epidermidis	1,000,000	**Yes**	**18.81**	**19.04**	**Yes**	**15.99**	**16.71**
Streptococcus mutans	1,000,000	No	-	-	**Yes**	**17.34**	**17.12**
Bacillus cereus	100,000	**Yes**	**2.26**	**2.1**	Yes	0.93	0.96
Clostridium beijerinckii	100,000	No	-	-	Yes	0.46	0.67
Pseudomonas aeruginosa	100,000	Yes	0.01	0.03	Yes	0.85	0.88
Staphylococcus aureus	100,000	**Yes**	**1.59**	**1.81**	**Yes**	**6.81**	**6.49**
Streptococcus agalactiae	100,000	**Yes**	**1.27**	**1.42**	**Yes**	**2.36**	**2.36**
Acinetobacter baumannii	10,000	No	-	-	**Yes**	**0.11**	**0.09**
Helicobacter pylori	10,000	**Yes**	**0.1**	**0.14**	**Yes**	**0.11**	**0.13**
Lactobacillus gasseri	10,000	No	-	-	**Yes**	**0.17**	**0.12**
Listeria monocytogenes	10,000	**Yes**	**0.21**	**0.29**	**Yes**	**0.12**	**0.23**
Neisseria meningitides	10,000	No	-	-	**Yes**	**0.14**	**0.24**
Propionibacterium acnes	10,000	**Yes**	**0.17**	**0.22**	Yes	0.03	0.05
Actinomyces odontolyticus	1,000	**Yes**	-	-	Yes	0	0
Bacteroides vulgatus	1,000	No	-	-	**Yes**	**0.02**	**0.01**
Deinococcus radiodurans	1,000	No	-	-	No	-	-
Enterococcus faecalis	1,000	No	-	-	Yes	0.11	**0.02**
Streptococcus pneumoniae	1,000	No	-	-	Yes	3.63	3.38

Greengenes is one of the most commonly used databases in microbiota studies because it is curated and checked for chimeras (46). However, half of the bacterial species of the mock community were not annotated in the database up to the species level, so we expected seeing only genus level (marked as "No" Table 2).

rrn database (33) contains information of the ribosomal RNA operon (16S-ITS-23S genes) for 22,351 sequences belonging to 2,384 different species. This database lacks information for only one member of the mock community (*Deinococcus radiodurans*).

On one hand, Greengenes contains taxonomic annotation up to the species level for 10 out of 20 bacterial species included in the mock bacterial community. From these, we were able to detect all of them. Moreover, it's worthy to note that despite *Streptococcus mutans* was not in Greengenes database at species level, we detected the closely related species *Streptococcus sobrinus* in high abundance (M1=14.4% and M2=11.3%); in fact both belong to the *mutans* group (47). Moreover, we saw *Streptococcus infantis* (M1=9.6 and M2= 7.9%) as another abundant species. On the other hand, *rrn* database contained species level information from 19 out of the 20 bacterial species of the mock community and we were able to detect all of them (Table 2).

The overall trend when looking at relative abundances is that operon counts correlated to relative abundances: when operon count changed in one magnitude order also relative abundances did. The exceptions were *Rhodobacter sphaeroides*, *Pseudomonas aeruginosa* and *Actinomyces odontolyticus*, which were detected at lower abundances than expected or not detected at all.

When looking at the results of *rrn* database, we could see that not only *Rhodobacter sphaeroides* and *Pseudomonas aeruginosa* were underrepresented but also *Bacillus cereus* and *Clostridium beijerinckii*. Finally, *Streptococcus pneumoniae* was overrepresented, probably suggesting that the sequencing errors together with the large amount of Streptoccous entries in the *rrn* database (44 *Streptococcus* species in *rrn* vs 9 *Streptococcus* species in Greengenes) produced an incorrect identification of this species. Thus, some of these biases could be explained by sequencing errors, primers used, or low resolution of 16S rRNA gene to distinguish among species from certain genera.

At the genus level, we were able to identify all the members of the mock community except *Actinomyces odontolyticus* even when it was represented on both databases (Table 3). Aproximately 10% of the reads for Greengenes and 2% for *rrn* database belonged to other genera theoretically not present in the mock community. Among those "other genera", the most abundant belonged to *Shigella*, *Enterobacter* and *Salmonella*, which are closely related to *Escherichia coli* (48). If we not consider these taxa probably wrongly assigned because they are closely related to *Escherichia coli*, only ~2.5% and ~0.5 % of the reads of Greengenes and *rrn* database respectively belong to other genera rather than the expected, which could be due to sequencing errors or to cross-contamination from dog skin microbiota samples.

Table 3. Taxonomic assignment of the mock community at genus level. Results obtained at the genus level after MinION™ sequencing of two replicates of the staggered mock community (M1 and M2) aligned against Greengenes and rrn databases.

		Greengenes		rrn	
Genus	**Expected abund.**	**% of reads M1**	**% of reads M2**	**% of reads M1**	**% of reads M2**
Streptococcus	++++	33.99	33.44	35.44	34.82
Staphylococcus	++++	24.00	24.30	24.32	24.73
Escherichia	++++	15.41	16.52	22.91	24.23
Rhodobacter	++++	8.58	7.21	8.35	6.99
Bacillus	+++	3.66	3.69	3.28	3.37
Clostridium	+++	1.31	1.46	1.25	1.3
Pseudomonas	+++	1.04	1.05	1.07	1.01
Listeria	++	0.27	0.30	0.20	0.30
Neisseria	++	0.17	0.27	0.17	0.30
Propionibacterium	++	0.17	0.22	0.13	0.17
Lactobacillus	++	0.27	0.17	0.28	0.18
Helicobacter	++	0.10	0.14	0.12	0.13
Acinetobacter	++	0.13	0.12	0.13	0.11
Enterococcus	+	0.14	0.11	0.12	0.06
Bacteroides	+	0.02	0.01	0.02	0.01
Deinococcus	+	0.02	0	0.02	0.01
Actinomyces	+	0	0	0	0
Other genera*	-	10.72	10.99	2.19	2.28

We can conclude from mock community analyses that full-length 16S rRNA sequencing with Oxford Nanopore is able to detect taxonomy assignments and retrieve diversity information, provided that the target species are in the database. At the genus level, we were able to accurately retrieve the mock community composition. It's also worthy to note the good technical replicates obtained for M1 and M2 samples.

Evaluation of primer sets V1-V8 and V1-V9 in microbial richness

Dog skin microbiota samples were sequenced as a pool with MinION™ after amplification of full-length 16S rRNA gene with primers targeting regions V1-V8 or V1-V9 (see Table 1). These complex microbiota samples were basecalled with Albacore v0.8.4 (Oxford Nanopore Technologies) and fast5 files were converted to fastq files. After demultiplexing, adapters and universal tags were trimmed and sequences analyzed using NanoOK (43) with LAST aligner (44). We finally obtained a total of 79,083 sequences for V1-V9 and 74,243 for V1-V8.

The same samples had been previously sequenced individually with Ion Torrent PGM™ with primers targeting V1-V2 hypervariable regions. In that case sequences were analyzed with QIIME 1.9.1 (42) with operational taxonomic units (OTUs) picking a representative

sequence of a group of sequences with a 97% similarity and taxonomy was assigned with RDP classifier (49) against the whole Greengenes database (45,46)(that contains many entries that do not reach low taxonomic levels). Using RDP classifier, if the taxonomy assignment does not reach a specific threshold, the sequences included in the OTU are set as "Other". We finally obtained a total of 249,572 sequences for V1-V2 region.

We performed the evaluation of the primer sets using exclusively the Greengenes database, because V1-V2 short-reads were analyzed using this same database. We used a subset of the Greengenes database that contained only those taxa that reached species level for the long-reads obtained for V1-V8 and V1-V9 regions with MinION™. As a consequence of the stricter criteria for taxonomy assignment of V1-V2 short-reads, lower taxonomic levels were poorly annotated. Thus, we compared diversity estimates of higher taxa (from kingdom to order).

Both V1-V8 and V1-V9 primer sets for long-reads were able to retrieve more bacterial taxa than V1-V2 short reads. The bacterial richness was higher at different taxonomic levels when assessed with long-reads rather than with short-reads, as seen in Table 4, and this trend increased as we were lowering taxonomic level.

Table 4. Bacterial richness estimates for skin microbiota samples. Bacterial richness retrieved with different primer pairs targeting short (V1-V2) or full-length (V1-V8 and V1-V9) 16S rRNA gene.

	V1-V2 (~350 bp)	**V1-V8 (~1300 bp)**	**V1-V9 (~1400 bp)**
Kingdom	1	2	2
Phylum	16	22	22
Class	33	46	47
Order	53	88	91

At the highest taxonomic level, we were able to detect not only Bacteria but also Archaea kingdom, despite using universal primers specific for Bacteria (25). However, they were present at really low proportions (< 0.01% of total reads), which agrees with results previously reported on human skin microbiota samples (50).

Delving into phylum level, we detected that the most common and better characterized phyla were retrieved by all the primers. These taxa represented >98% of the total skin microbiota composition. Long-read primers were able to detect 8 phyla previously unseen using V1-V2 short-reads (Table 5). It has already been reported the low coverage of this primer set for some specific phyla (21,51). Some phyla were only detected with a specific primer set, such as Lentisphaerae with V1-V8 or Fibrobacteres with V1-V9. On the other hand, GN02, TM7 and Thermi phyla belong to candidate divisions and none of their members have been cultivated (52)(Camanocha et al., 2014), so databases do not have taxonomy information up to species level. Thus, since we used the Greengenes subset database with species-level sequences, no representative of those phyla were included for taxonomy assignment of V1-V8 and V1-V9 long-reads and that is probably the reason why they are only detected with V1-V2 primers.

Table 5. Microbial richness on skin samples. Table containing all the observed phyla per primer subset. Long-reads taxonomy was obtained from a species-level subset of Greengenes database (see materials and methods) that did not contain information of GN02, SR1 and TM7. *phyla that belong to Archaea kingdom. ** phyla with no representative at the spp level in the database used for assigning taxonomy to long-reads (V1-V8 and V1-V9).

	Long-reads		Short-reads
Phylum	**V1-V9**	**V1-V8**	**V1-V2**
[Thermi]	+	+	+
Acidobacteria	+	+	+
Actinobacteria	+	+	+
Bacteroidetes	+	+	+
Chlorobi	+	+	+
Chloroflexi	+	+	+
Cyanobacteria	+	+	+
Deferribacteres	+	+	+
Firmicutes	+	+	+
Fusobacteria	+	+	+
Proteobacteria	+	+	+
Spirochaetes	+	+	+
Tenericutes	+	+	+
Aquificae	+	+	
Chlamydiae	+	+	
Crenarchaeota*	+	+	
Elusimicrobia	+	+	
Euryarchaeota*	+	+	
Planctomycetes	+	+	
Synergistetes	+	+	
Verrucomicrobia	+	+	
Fibrobacteres	+		
Lentisphaerae		+	
GN02**			+
SR1**			+
TM7**			+
Total	**22/26**	**22/26**	**16/26**

So, we were able to detect previously unseen bacteria phyla on dog skin using MinION™ long-amplicons for full-length 16S rRNA sequences, which provided better richness estimates. These previously unseen bacteria phyla presented low relative abundances on canine skin microbiota, but the use of long-amplicons in more uncharacterized environments will provide better diversity estimates. In conclusion, as previously reported, the ability to classify sequences to the genus or species level is a function of read length, but also of sample type and the reference database (23).

Skin microbiota analyses

We could see from the mock community results that species-level resolution was a little bit tricky because sometimes the target species was not present in the database. However, here we aimed to compare the taxonomies at species level obtained from two independent databases containing different taxonomic annotations. We considered a species-level assignment valid when two unrelated databases gave identical annotation.

When comparing the most abundant species (>1% of total relative abundances), we could see 7 bacterial species identified by two independent databases: *Bergeyella zoohelcum*, *Capnocytophaga canimorsus*, *Pasteurella multocida*, *Neisseria shayeganii*, *Lactobacillus reuteri*, *Bibersteinia trehalosi*, and *Neisseria weaver* (Figure 1).

This study included 5 skin microbiota samples together in a pool, which had been previously sequenced individually using V1-V2 short-reads and Ion Torrent PGM™. Figure 2A shows the microbiota profile of the most abundant taxa per individual sample (taxa representing > 0.5%). We can see that all of them have been identified up to family level and some of them also to genus level (Figure 2A). Despite working with an equimolar pool of these samples for nanopore sequencing, we could see in Figure 2B that this experiment is not representing equally all the individual samples and is missing some taxa.

When comparing the results of nanopore sequencing with the short reads, we could detect that *Lactobacillus* species were exclusive to one single sample of the pool (18A), so probably not the most abundant taxa on the inner pinna skin microbiota. *Fusobacteriaceae* (specifically *Fusobacterium* genus) is one of the most abundant families on the inner pinna skin microbiota, and both databases detected that pattern at family level. However, Greengenes database was not able to identify any *Fusobacterium*, since the database lacked entries of the genus at the species level. *Porphyromonadaceae* was also another abundant member of skin microbiota, but species level taxonomy differed in each database giving different results. Long-reads allowed us reaching lower taxonomic levels confirmed by two independent databases for: 1) [*Weeksellaceae*] family, with *Bergeyella zoohelcum*; 2) *Neisseriaceae,* with *Neisseria shayeganii* and *Neisseria weaveri; 3) Pasteurellaceae*, with *Pasteurella multocida* and 4) *Flavobacteriaceae*, with *Capnocitophaga canimorsus.*

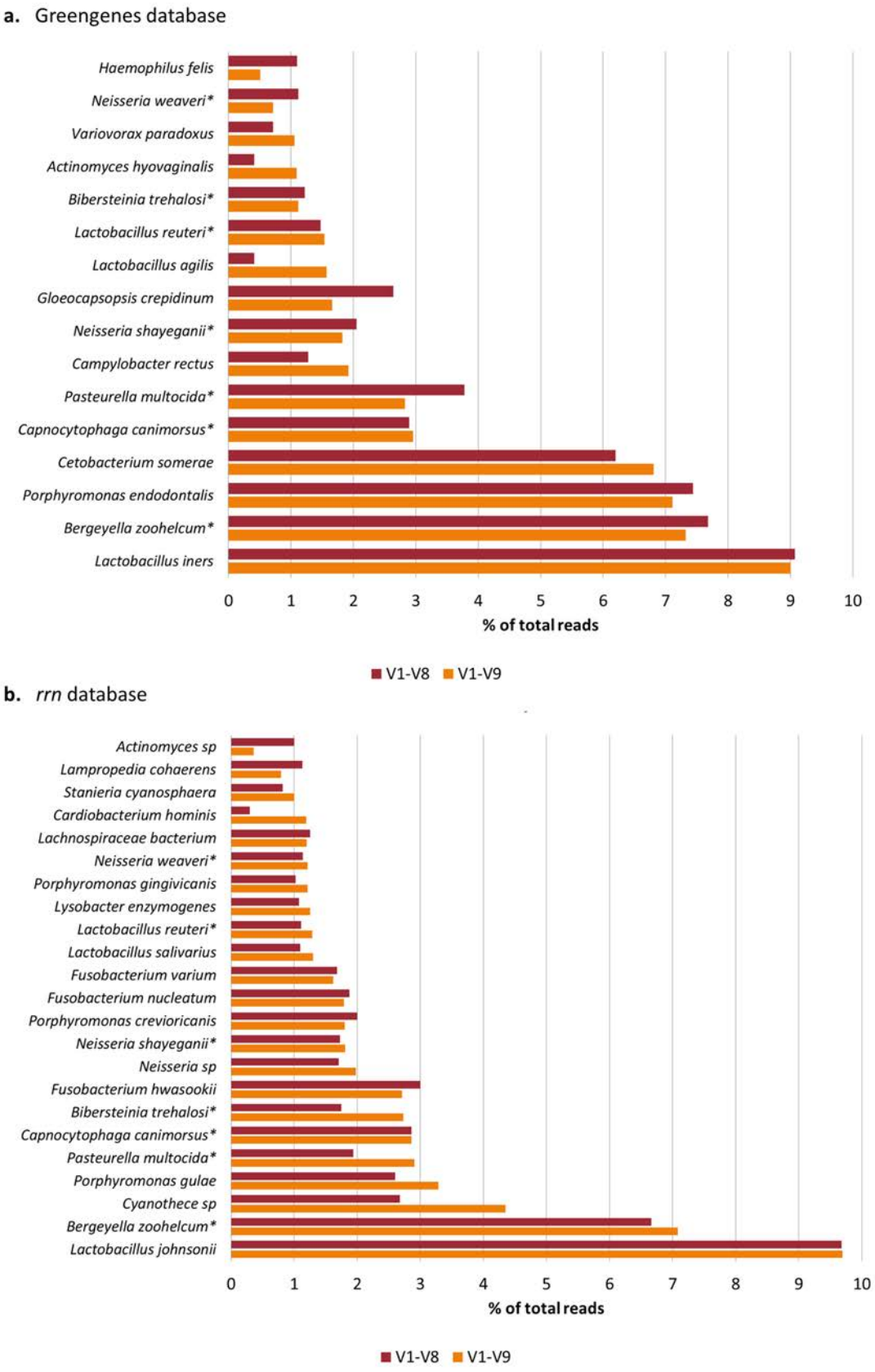


Figure 1. Skin microbiota composition at species level. Comparison of the species detected against the **(a)** Greengenes and **(b)** *rrn* databases in the dog skin microbiota samples amplified with V1-V8 and V1-V9 16S rRNA primers and sequenced with MinION™. (*) coincident taxonomic assignments using independent databases.

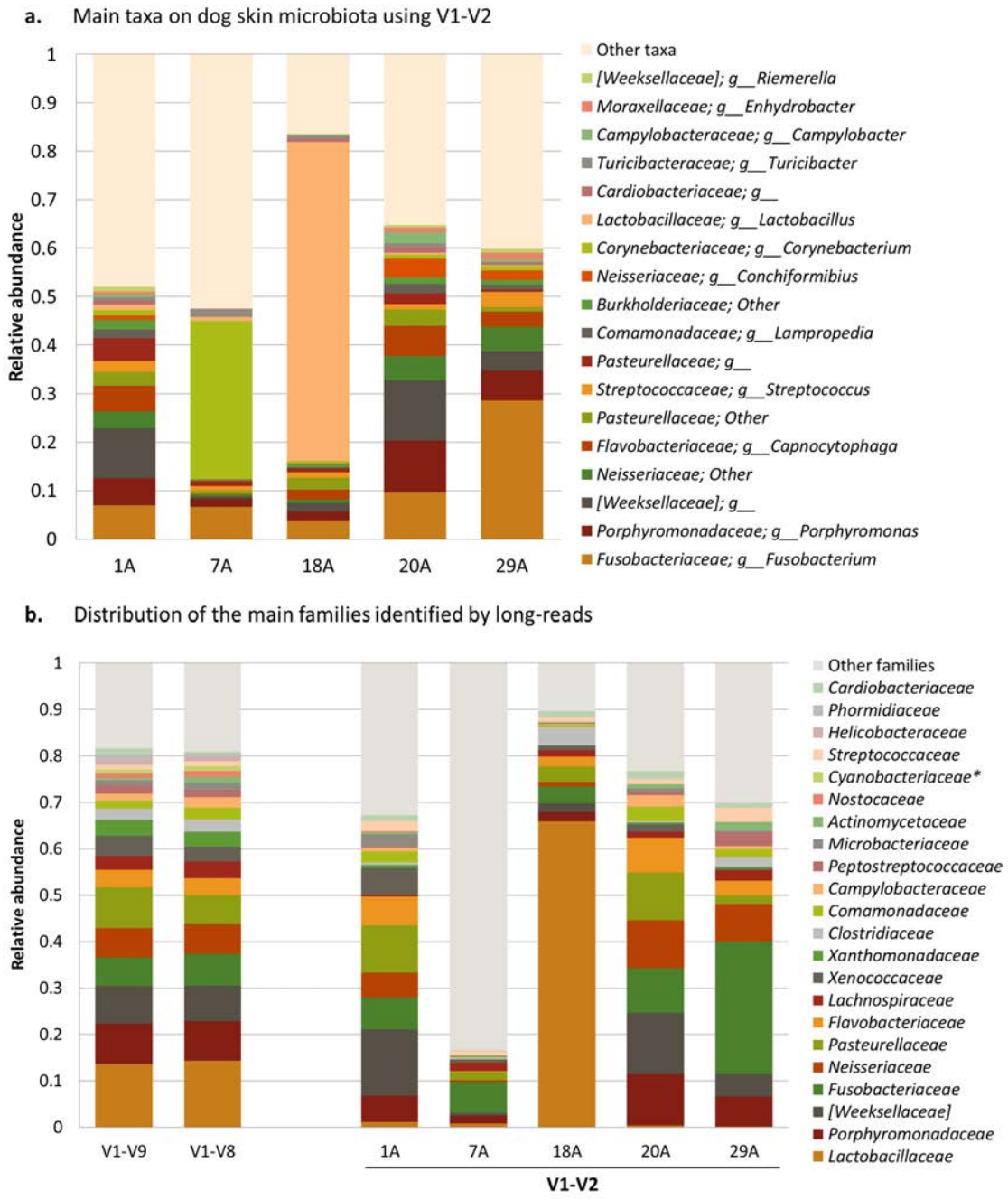


Figure 2. Skin microbiota taxonomic profile using long and short reads. Bar plots of (a) the most abundant taxa on individual skin samples sequenced with V1-V2 primers in short-reads**.** "g__" means there is no information at genus level; and b) main families on dog skin microbiota when sequencing the full-length 16S rRNA gene (V1-V9 and V1-V8) on a pool of samples and representation of the same families when sequencing V1-V2 region on the individual samples that constituted the pool.

Conclusions

Nanopore sequencing of the full-length 16S rRNA gene allowed us inferring microbiota composition from both simple and complex microbial communities. Moreover, long-reads were able to retrieve increased richness estimates, which show us previously unseen phyla on dog skin, despite being at really low abundances.

Taxonomy assignment down to species level was not always feasible because of both the high error rate of 1D reads and the absence of some bacterial taxa on the databases. With the nanopore reads, we assigned taxonomy through alignment strategies executing all-vs-all comparisons that need many computational resources, so we needed a small database to obtain results. When working with a 16S database subset, we should be sure to include the most relevant taxa even if they do not have representative members at species level. Using only species-level taxonomy, we lost insights of candidate division phyla such as SR1, GN02 or TM7. Oxford nanopore offers other bioinformatics tools, such as the cloud-based EPI2ME platform. However, we run out of memory on our hard disk when trying to perform these analyses.

Finally, we should keep in mind that delving into species level with only 16S is sometimes difficult because there are bacterial species that share almost all their 16S rRNA sequences (60). Recently, Benitez-Paez and colleagues proposed to sequence the whole rrn operon constituted by 16S-ITS-23S to obtain better species-level resolution (33).

In conclusion, nanopore sequencing of the full-length 16S rRNA gene allowed identifying bacterial species even from a complex community, obtaining a microbiota profile and improving richness estimates. So, nanopore sequencing has the potential to be used to assess microbial communities. Future studies should be relying on the new $1D^2$ kit that presents higher accuracy and, since sequencing length is not an impediment, other experimental strategies could be assessed.

Bibliography

1. Marchesi JR, Ravel J. The vocabulary of microbiome research: a proposal. *Microbiome* (2015) **3**:31. doi:10.1186/s40168-015-0094-5

2. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. Bacterial community variation in human body habitats across space and time. *Science* (2009) **326**:1694–1697. doi:10.1126/science.1177486

3. The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* (2012) **486**:207–214. doi:10.1038/nature11234

4. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, et al. A human gut microbial gene catalogue established by

metagenomic sequencing. *Nature* (2010) **464**:59–65. doi:10.1038/nature08821

5. Belkaid Y, Segre JA. Dialogue between skin microbiota and immunity. *Science* (2014) **346**:954–9. doi:10.1126/science.1260144

6. Thaiss CA, Zmora N, Levy M, Elinav E. The microbiome and innate immunity. *Nature* (2016) **535**:65–74. doi:10.1038/nature18847

7. Belkaid Y, Hand TW. Role of the microbiota in immunity and inflammation. *Cell* (2014) **157**:121–141. doi:10.1016/j.cell.2014.03.011

8. Fitz-Gibbon S, Tomida S, Chiu B-H, Nguyen L, Du C, Liu M, Elashoff D, Erfe MC, Loncaric A, Kim J, et al. *Propionibacterium acnes* strain populations in the human skin microbiome associated with acne. *J Invest Dermatol* (2013) **133**:2152–60. doi:10.1038/jid.2013.21

9. Dreno B, Martin R, Moyal D, Henley JB, Khammari A, Seité S. Skin microbiome and acne vulgaris: *staphylococcus*, a new actor in acne. *Exp Dermatol* (2017) doi:10.1111/exd.13296

10. Gao Z, Tseng C, Strober BE, Pei Z, Blaser MJ. Substantial alterations of the cutaneous bacterial biota in psoriatic lesions. *PLoS One* (2008) **3**:e2719. doi:10.1371/journal.pone.0002719

11. Fahlén A, Engstrand L, Baker BS, Powles A, Fry L. Comparison of bacterial microbiota in skin biopsies from normal and psoriatic skin. *Arch Dermatol Res* (2012) **304**:15–22. doi:10.1007/s00403-011-1189-x

12. Alekseyenko A V, Perez-Perez GI, De Souza A, Strober B, Gao Z, Bihan M, Li K, Methé BA, Blaser MJ. Community differentiation of the cutaneous microbiota in psoriasis. *Microbiome* (2013) **1**:31. doi:10.1186/2049-2618-1-31

13. Kong HH, Oh J, Deming C, Conlan S, Grice EA, Beatson MA, Nomicos E, Polley EC, Komarow HD, Murray PR, et al. Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis. *Genome Res* (2012). **22**:850-9. doi:10.1101/gr.131029.111

14. Oh J, Freeman AF, Park M, Sokolic R, Candotti F, Holland SM, Segre JA, Kong HH. The altered landscape of the human skin microbiome in patients with primary immunodeficiencies. *Genome Res* (2013) **23**:2103–2114. doi:10.1101/gr.159467.113

15. Seite S, Flores GE, Henley JB, Martin R, Zelenkova H, Aguilar L, Fierer N. Microbiome of affected and unaffected skin of patients with atopic dermatitis before and after emollient treatment. *J Drugs Dermatol* (2014) **13**:1365–72.

16. Chng KR, Tay ASL, Li C, Ng AHQ, Wang J, Suri BK, Matta SA, McGovern N, Janela B, Wong XFCC, et al. Whole metagenome profiling reveals skin microbiome-dependent susceptibility to atopic dermatitis flare. *Nat Microbiol* (2016) **1**:16106. doi:10.1038/nmicrobiol.2016.106

17. Kennedy EA, Connolly J, Hourihane JO, Fallon PG, McLean WHI, Murray D, Jo J-H, Segre JA, Kong HH, Irvine AD. Skin microbiome before development of atopic dermatitis: Early colonization with commensal staphylococci at 2 months is associated with a lower risk of atopic dermatitis at 1 year. *J Allergy Clin Immunol* (2017) **139**:166–172. doi:10.1016/j.jaci.2016.07.029

18. Rodrigues Hoffmann A, Patterson AP, Diesel A, Lawhon SD, Ly HJ, Elkins Stephenson C, Mansell J, Steiner JM, Dowd SE, Olivry T, et al. The skin microbiome in healthy and allergic dogs. *PLoS One* (2014) **9**:e83197. doi:10.1371/journal.pone.0083197

19. Bradley CW, Morris DO, Rankin SC, Cain CL, Misic AM, Houser T, Mauldin EA, Grice EA. Longitudinal evaluation of the skin microbiome and association with microenvironment and treatment in canine atopic dermatitis. *J Invest Dermatol* (2016) 136:1182-90. doi:10.1016/j.jid.2016.01.023

20. Pierezan F, Olivry T, Paps JS, Lawhon SD, Wu J, Steiner JM, Suchodolski JS, Hoffmann AR. The skin microbiome in allergen-induced canine atopic dermatitis. *Vet Dermatol* (2016) **27**:332-e82. doi:10.1111/vde.12366

21. Kuczynski J, Lauber CL, Walters WA., Parfrey LW, Clemente JC, Gevers D, Knight R. Experimental and analytical tools for studying the human microbiome. *Nat Rev Genet* (2011) **13**:47–58. doi:10.1038/nrg3129

22. Clarridge JE. Impact of 16S rRNA Gene Sequence Analysis for Identification of Bacteria on Clinical Microbiology and Infectious Diseases. *Clin Microbiol Rev* (2004) **17**:840–862. doi:10.1128/CMR.17.4.840-862.2004

23. Schloss PD, Jenior ML, Koumpouras CC, Westcott SL, Highlander SK. Sequencing 16S rRNA gene fragments using the PacBio SMRT DNA sequencing system. *PeerJ* (2016) **4**:e1869. doi:10.7717/peerj.1869

24. Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer K-H, Whitman WB, Euzéby J, Amann R, Rosselló-Móra R. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* (2014) **12**:635–645. doi:10.1038/nrmicro3330

25. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glockner FO. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* (2013) **41**:e1. doi:10.1093/nar/gks808

26. Fichot EB, Norman RS. Microbial phylogenetic profiling with the Pacific Biosciences sequencing platform. *Microbiome* (2013) **1**:10. doi:10.1186/2049-2618-1-10

27. Mosher JJ, Bernberg EL, Shevchenko O, Kan J, Kaplan LA. Efficacy of a 3rd generation high-throughput sequencing platform for analyses of 16S rRNA. *J*

Microbiol Methods (2013) **95**:175-81. doi: 10.1016/j.mimet.2013.08.009

28. Mosher JJ, Bowman B, Bernberg EL, Shevchenko O, Kan J, Korlach J, Kaplan LA. Improved performance of the PacBio SMRT technology for 16S rDNA sequencing. *J Microbiol Methods* (2014) **104**:59–60. doi:10.1016/j.mimet.2014.06.012

29. Wagner J, Coupland P, Browne HP, Lawley TD, Francis SC, Parkhill J. Evaluation of PacBio sequencing for full-length bacterial 16S rRNA gene classification. *BMC Microbiol* (2016) **16**:274. doi:10.1186/s12866-016-0891-4

30. Singer E, Bushnell B, Coleman-Derr D, Bowman B, Bowers RM, Levy A, Gies EA, Cheng J-F, Copeland A, Klenk H-P, et al. High-resolution phylogenetic microbial community profiling. *ISME J* (2016) **10**:2020–2032. doi:10.1038/ismej.2015.249

31. Ma X, Stachler E, Bibby K. Evaluation of Oxford Nanopore MinION Sequencing for 16S rRNA Microbiome Characterization. *bioRxiv* (2017). doi:10.1101/099960

32. Benítez-Páez A, Portune KJ, Sanz Y. Species-level resolution of 16S rRNA gene amplicons sequenced through the MinION™ portable nanopore sequencer. *Gigascience* (2016) **5**:4. doi:10.1186/s13742-016-0111-z

33. Benitez-Paez A, Sanz Y. Multi-locus and ultra-long amplicon sequencing approach to study microbial diversity at species level using the MinION TM portable nanopore sequencer. *bioRxiv* (2017) doi:10.1101/117143

34. Mitsuhashi S, Kryukov K, Nakagawa S, Takeuchi JS, Shiraishi Y, Asano K, Imanishi T. A portable system for metagenomic analyses using nanopore-based sequencer and laptop computers can realize rapid on-site determination of bacterial compositions. *bioRxiv* (2017) https://doi.org/10.1101/101865

35. Shin J, Lee S, Go M-J, Lee SY, Kim SC, Lee C-H, Cho B-K. Analysis of the mouse gut microbiome using full-length 16S rRNA amplicon sequencing. *Sci Rep* (2016) **6**:29681. doi:10.1038/srep29681

36. Cuscó A, Belanger JM, Gershony L, Islas-Trejo A, Levy K, Medrano JF, Sánchez A, Oberbauer AM, Francino O. Individual signatures and environmental factors shape skin microbiota on healthy dogs. *Microbiome* (submitted).

37. Cuscó A, Sánchez A, Altet L, Ferrer L, Francino O. Individual Signatures Define Canine Skin Microbiota Composition and Variability. *Front Vet Sci* (2017) **4**:6. doi:10.3389/fvets.2017.00006

38. Alm E, Oerther D, Larsen N, Stahl D. The oligonucleotide probe database. *Appl Environ Microbiol.* (1996) **62**:3557–3559

39. Mao D-P, Zhou Q, Chen C-Y, Quan Z-X. Coverage evaluation of universal bacterial primers using the metagenomic datasets. *BMC Microbiol* (2012) **12**:66. doi:10.1186/1471-2180-12-66

40. Watson M, Thomson M, Risse J, Talbot R, Santoyo-Lopez J, Gharbi K, Blaxter M. PoRe: An R package for the visualization and analysis of nanopore sequencing data. *Bioinformatics* (2015) **31**:114–115. doi:10.1093/bioinformatics/btu590

41. Wick R. Porechop. Available at: https://github.com/rrwick/Porechop

42. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* (2010) **7**:335–336. doi:10.1038/nmeth.f.303

43. Leggett RM, Heavens D, Caccamo M, Clark MD, Davey RP. NanoOK: multi-reference alignment analysis of nanopore sequencing data, quality and error profiles. *Bioinformatics* (2015) **32**:btv540. doi:10.1093/bioinformatics/btv540

44. Kielbasa SM, Wan R, Sato K, Horton P, Frith MC. Adaptive seeds tame genomic sequence comparison. *Genome Res* (2011) **21**:487–493. doi:10.1101/gr.113985.110

45. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* (2006) **72**:5069–72. doi:10.1128/AEM.03006-05

46. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J* (2012) **6**:610–8. doi:10.1038/ismej.2011.139

47. Lal D, Verma M, Lal R. Exploring internal features of 16S rRNA gene for identification of clinically relevant species of the genus *Streptococcus*. *Ann Clin Microbiol Antimicrob* (2011) **10**:28. doi:10.1186/1476-0711-10-28

48. Fukushima M, Kakinuma K, Kawaguchi R. Phylogenetic analysis of *Salmonella*, *Shigella*, and *Escherichia coli* strains on the basis of the gyrB gene sequence. *J Clin Microbiol* (2002) **40**:2779–85. doi:10.1128/JCM.40.8.2779-2785.2002

49. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* (2007) **73**:5261–7. doi:10.1128/AEM.00062-07

50. Oh J, Byrd AL, Deming C, Conlan S, Barnabas B, Blakesley R, Bouffard G, Brooks S, Coleman H, Dekhtyar M, et al. Biogeography and individuality shape function in the human skin metagenome. *Nature* (2014) **514**:59–64. doi:10.1038/nature13786

51. Walters WA, Caporaso JG, Lauber CL, Berg-Lyons D, Fierer N, Knight R. PrimerProspector: de novo design and taxonomic analysis of barcoded polymerase chain reaction primers. *Bioinformatics* (2011) **27**:1159–61. doi:10.1093/bioinformatics/btr087

52. Camanocha A, Dewhirst FE. Host-associated bacterial taxa from Chlorobi, Chloroflexi, GN02, Synergistetes, SR1, TM7, and WPS-2 Phyla/candidate divisions. *J Oral Microbiol* (2014) **6**:25468 doi:10.3402/jom.v6.25468

53. Wallis C, Marshall M, Colyer A, O'Flynn C, Deusch O, Harris S. A longitudinal assessment of changes in bacterial community composition associated with the development of periodontal disease in dogs. *Vet Microbiol* (2015) **181**:271–282. doi:10.1016/j.vetmic.2015.09.003

54. Lin WR, Chen YS, Liu YC. Cellulitis and Bacteremia Caused by *Bergeyella zoohelcum*. *J Formos Med Assoc* (2007) **106**:573–576. doi:10.1016/S0929-6646(07)60008-4

55. Oehler RL, Velez AP, Mizrachi M, Lamarche J, Gompf S. Bite-related and septic syndromes caused by cats and dogs. *Lancet Infect Dis* (2009) **9**:439–447. doi:10.1016/S1473-3099(09)70110-0

56. Sturgeon A, Stull JW, Costa MC, Weese JS. Metagenomic analysis of the canine oral cavity as revealed by high-throughput pyrosequencing of the 16S rRNA gene. *Vet Microbiol* (2013) **162**:891–898. doi:10.1016/j.vetmic.2012.11.018

57. Andersen BM, Steigerwalt AG, O'Connor SP, Hollis DG, Weyant RS, Weaver RE, Brenner DJ. *Neisseria weaveri* sp. nov., formerly CDC group M-5, a gram-negative bacterium associated with dog bite wounds. *J Clin Microbiol* (1993) **31**:2456–66.

58. Kim SY, Adachi Y. Biological and genetic classification of canine intestinal lactic acid bacteria and bifidobacteria. *Microbiol Immunol* (2007) **51**:919–28.

59. Silva BC, Jung LRC, Sandes SHC, Alvim LB, Bomfim MRQ, Nicoli JR, Neumann E, Nunes AC. In vitro assessment of functional properties of lactic acid bacteria isolated from faecal microbiota of healthy dogs for potential use as probiotics. *Benef Microbes* (2013) **4**:267–75. doi:10.3920/BM2012.0048

60. Janda JM, Abbott SL. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol* (2007) **45**:2761–4. doi:10.1128/JCM.01228-07

4. Discussion

This thesis intends to shed light onto the skin microbiota on healthy dogs as well as providing a tool to assess genetic variability in innate immunity.

In the research article *"Non-synonymous genetic variation in exonic regions of canine Toll-like receptors"* we characterized by massive sequencing the genetic polymorphisms of Toll-like receptors on dogs from seven different breeds and on wolves from two populations. Finally we developed a TaqMan OpenArray® panel with 64 non-synonymous SNPs with a likely functional effect on the TLR proteins and we validated it using a subset of the previously sequenced animals.

In the research article *"Individual signatures define canine skin microbiota composition and variability"* we characterized the variability of the skin microbiota by analyzing eight different skin sites in a cohort of nine healthy dogs from three different breeds: German Shepherd, French Bulldog and West Highland White Terrier. We assessed factors such as breed, skin site and individual as potential microbiota drivers. We found that the individual was the main factor driving skin microbiota structure and composition, followed by the skin site. We clarified that this individual effect should be understood as the dog, its lifestyle, and its environment. Finally, we detected no effect of the breed, however only three animals per breed were used so this effect should be better assessed.

To resolve if that individual effect shaping skin microbiota on healthy dogs came from the dog itself or from the environment we performed a third research article *"Individual signatures and environmental factors shape skin microbiota on healthy dogs"*. Here we characterized the variability of the skin microbiota in healthy dogs cohabiting together in a shared environment, by analyzing eight different skin sites in a cohort of thirty-five Golden-Labrador Retriever crossbred dogs. Even for this cohort with a shared environment, we found that microbiota composition was driven by the individual, confirming the results of

the second research article. When considering abundances, the microbiota structure was driven both by the individual and by the skin site. Considering that the cohort was environmentally uniform, we were able to detect the effect of host-specific factors: the season when the animal was born or the time spent in the kennel was a factor shaping skin microbiota in all skin sites.

Finally in the chapter *"Using MinION™ to characterize dog skin microbiota through full-length 16S rRNA gene sequencing approach"* we assessed the potential of Nanopore sequencing for microbiota analyses using the full-length 16S rRNA gene. We first sequenced a simple microbial mock community containing 20 bacterial species (HM-783D) and then a complex microbial community represented by a pool of different skin microbiota samples from the inner pinna. Despite using 1D chemistry, which presents low accuracy, we were able to obtain better richness estimates and some information at species level. However, we should consider using the new $1D^2$ chemistry from now on, and exploiting the potential of this technology performing other approaches.

In each of these four chapters the obtained results are thoroughly discussed. Thus, the purpose of the present section is to review and unify the four independent studies, as well as to discuss the pitfalls and to propose future directions.

4.1. Dual assessment of innate immunity and skin microbiota

In the research article "*Non-synonymous genetic variation in exonic regions of canine Toll-like receptors*" (Chapter 3.1), we characterized by massive sequencing the genetic variability in the coding regions of the 10 Toll-like Receptor (TLRs) genes in 355 dogs from 7 breeds and in 100 wolves from 2 populations. We functionally annotated the different variants and predicted the likely effect of non-synonymous SNPs (nsSNPs) on the TLR proteins that act as the first sensors of microbes. We found that the frequencies of nsSNPs differed among breeds and that some of them were breed- or species-specific.

Finally, we developed a TaqMan OpenArray genotyping plate with 60 non-synonymous SNPs and 4 frameshift mutations. This tool allows obtaining an innate immunity profile on single dogs by genotyping their TLR polymorphisms. It can be used in many areas, such as characterizing genetic variability on different dog breeds (Chapter 3.1) and even performing evolution studies when including other canid species; screening case-control cohorts to identify variants associated to a certain disease; and finally, characterizing innate immune profile of individual dogs that are also analyzed in a microbiota study.

When we screened the genetic variability of TLRs on several dogs from different breeds, we found that Labrador and German Shepherd dogs clustered separately (Figure 28a). Thus, some breeds present different innate immune profiles that could be influencing their predisposition and outcome to certain diseases, as well as their microbiota structure and composition.

Three of the nsSNPs from TLR5 gene identified by massive sequencing had been previously associated with Inflammatory Bowel diseases in German Shepherd (Kathrani et al., 2010) and in other breeds (Kathrani et al., 2011). In our cohort of healthy dogs, we found that the risk allele for IBD in German Shepherd was absent in this breed although it presented different frequencies in other breeds (Figure 28b). TLR5 is the innate immune receptor for flagellin that is the principal protein component of bacterial flagella (Leifer et al., 2014). German Shepherd dogs carrying this risk allele for canine IBD showed hyper-responsiveness to flagellin and Kathrani and colleagues suggested that this altered innate immunity-microbiota cross-talk could be responsible for the inappropriate inflammation observed in this disease (Kathrani et al., 2012). More recently, Vazquez-Baeza and collaborators confirmed that dogs with IBD presented an altered fecal microbiota that in fact could be used as a diagnosis tool to distinguish them from the healthy individuals (Vázquez-Baeza et al., 2016). Similarly on human IBD, Knights and collaborators reported a significant association between risk alleles of the innate immune receptor NOD2 and an increased relative abundance of *Enterobacteriaceae* in gastrointestinal microbiota (Knights et al., 2014).

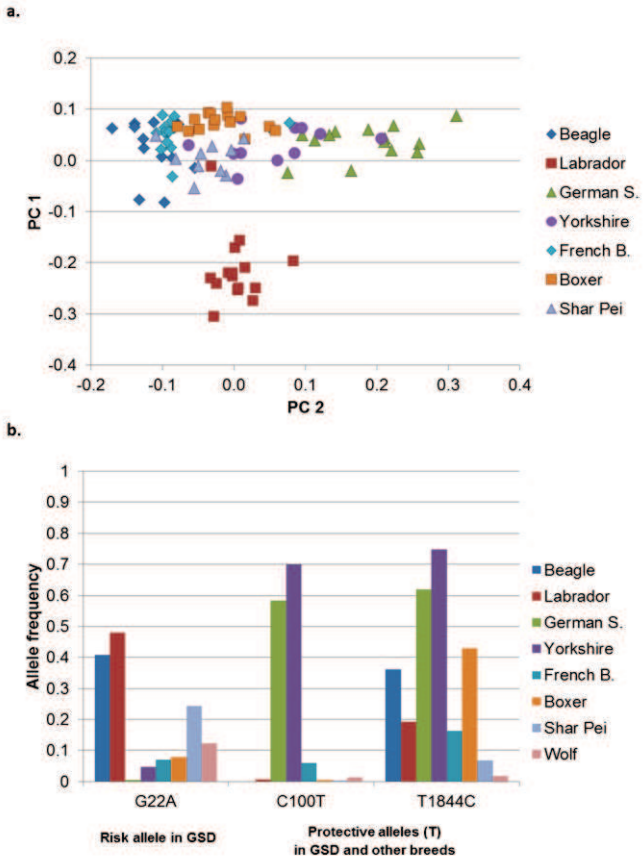


Figure 28. Genetic variation on canine Toll-like Receptors. In a) Principal Component Analysis of dogs individually genotyped using TaqMan Open Array plate. In b) allele frequencies in our cohort of non-synonymous SNPs obtained through massive sequencing and associated to IBD (From Chapter 3.1, Cuscó et al., 2014).

Regarding the aim of this dissertation, this tool will allow us to perform a dual assessment of the dog in health and disease: innate immune profile and microbiota composition.

In healthy skin, microbiota and immunity interact to maintain the homeostasis in front of disruptions. In fact, mice studies have demonstrated impaired and weakened skin immune responses in Germ free mice when compared to Pathogen-Specific Free mice (Naik et al., 2012b). In healthy human skin, some of these cross-talks between specific commensals and innate immunity have already been characterized (Barnard and Li, 2017; Holmes et al., 2015; Nakamizo et al., 2015). *Staphylococcus epidermidis* is one of the most studied commensals and interacts with host cell's TLRs to modulate innate immunity against pathogens (Table 3, Figure 24). Other commensals such as *Propionibacterium acnes*, *Pseudomonas aeruginosa* or *Staphylococcus aureus* contribute to skin immunity by producing antimicrobial substances (Holmes et al., 2015).

In dermatological diseases, the equilibrium and cross-talks between skin immunity and microbiota are disrupted. As we have extensively reviewed in the introduction, most of the cutaneous diseases are associated with an altered microbiota or dysbiosis (Table 4), and/or an altered innate immunity, either through genetic polymorphisms or altered expression of TLRs (Table 2). When comparing the two tables, some common cutaneous diseases appear

to present both an altered microbiota and an altered skin immunity profile. As an example, some links between disease and TLR-microbiota have been merged in Table 5.

Table 5. TLR-microbiota associations in main cutaneous diseases.

Disease	TLR-microbiota link
Atopic dermatitis	- During AD flare *S. aureus* increases, but also *S. epidermidis* which might be a compensatory mechanism (see Table 3 for specific mechanisms) - Topical probiotic of *Vitreoscilla filiformis* reduced AD score, increasing IL-10 production via TLR2 (Volz et al., 2014)
Psoriasis	- Polymorphisms on innate immune system receptor (e.g. TLRs) facilitate the initiation of an inflammatory response to commensal microbiota (Fry et al., 2015)
Acne vulgaris	- *P. acnes* strains are individual-specific, some virulent and some not (Fitz-Gibbon et al., 2013) - Health-associated phylotypes produce thiopeptides, which are antimicrobial compounds that inhibit the growth of gram-positive bacteria (Christensen and Brüggemann, 2014)

Some correlations between host genetic variation and certain microbiota composition have already been described in health and disease. Knights and colleagues, as we have previously explained, found a link between innate immune receptor NOD2 and increased *Enterobacteriaceae* on gastrointestinal microbiota. Moreover, they suggested that genetically altered host functional pathways can shape microbiome structure (Knights et al., 2014). Blekham and colleagues approach was retrieving the host sequences of the shotgun metagenomics dataset of the Human Microbiome Project. First they correlated host functional pathways (groups of aggregated SNPs and genes) with overall microbiome composition and found significant enrichment in several genes involved in (i) complex diseases that have already been linked to the microbiome (such as obesity and Inflammatory Bowel disease); and (ii) other immunity-related pathways, including the Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses. They also correlated host variation to specific microbiome members and found key host genes related to immunity involved. For example, genetic variants on HLA-DRA and TLR1 were correlated to *Selenomonas* presence on throat microbiota and *Lautropia* in the tongue dorsum, respectively (Blekhman et al., 2015).

To sum up, TLRs serve the dual function of sensing pathogens and symbionts and this sensing leads to very different outcomes for both microbes (clearance *vs.* symbiosis) and host (inflammation *vs.* immune homeostasis) (Chu and Mazmanian, 2013). Genetic variants on these genes can affect these responses. Thus, future uses of the canine TLR chip with the non-synonymous SNPs could be in both health and disease: in health, to elucidate if some specific TLR variants are associated to specific microbiota compositions; and in disease, to elucidate if some specific TLR variants are associated to some specific diseases and maybe in turn with an altered microbiota composition.

Finally, we should not forget the third piece of this puzzle: the environment. An excellent example of interaction between host genome, microbiome and environment is observed in humans by the interaction among LCT gene, *Bifidobacterium* and diet (dairy consumption). Lactose is the main sugar of dairy products and can be metabolized in the gastrointestinal tract either by *Bifidobacterium* of the microbiota or the lactase enzyme of the host, which is codified by LCT gene. Lactose intolerance in adults is associated with the SNP rs4988235 with genotype GG in LCT gene. Only in GG carriers, *Bifidobacterium* abundance is correlated with dairy consumption suggesting that microbiota is replacing and/or complementing the host genetics function (Figure 29) (Bonder et al., 2016). Thus, microbiota can adapt to different host genetics backgrounds to complement functions using environmental resources.

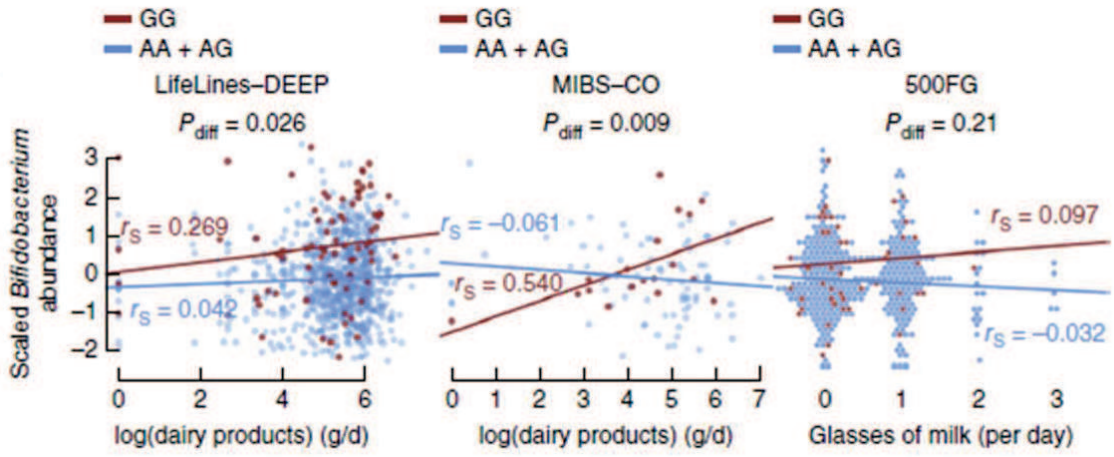


Figure 29. Example of an interaction between host genome, microbiota and environment: LCT gene, Bifidobacterium and dairy consumption. In individuals that present the GG genotype in LCT gene, the dairy consumption is positively correlated with the presence of Bifidobacterium in the GI tract (excerpted from Bonder et al., 2016).

Taken together, these findings motivate the need for larger association studies to characterize host genetic variation linked to the microbiome in the context of various health conditions, environmental effects, and genetic backgrounds.

4.2. Skin site signatures on canine skin microbiota

Human skin is mainly divided in three microhabitats depending on its physicochemical properties that harbor specific microbiota: sebaceous sites, with Propionibacterium spp; moist sites with *Staphylococcus* and *Corynebacterium* spp; and dry sites, with gram-negative microorganisms (Costello et al., 2009; Grice and Segre, 2011; Grice et al., 2009b).

Dog skin is more uniform and almost totally covered by a dense fur, and despite the anatomical and physicochemical differences among certain skin sites (reviewed at Chapter 1.2) these are not as large as in humans. Moreover, environment has probably a homogenizing effect hiding some skin-site bacterial signatures. Despite the more homogeneous microhabitat, skin diseases present different prevalence depending on skin sites (Miller et al., 2013).

Even with this uniformity, the first cross-sectional study of skin microbiota on healthy dogs identified that hairy skin presented higher diversity values when compared to mucosal surfaces or muco-cutaneous junctions, as well as skin-site specific taxonomic profiles (Rodrigues Hoffmann et al., 2014). We have expanded this knowledge, performing two more cross-sectional studies on healthy dogs and obtaining similar results for diversity, with low diversity values in perianal and nasal skin microbiota samples (Chapter 3.2 and 3.3, Cuscó et al., 2017a, 2017b).

Identifying skin-site bacterial signatures is biologically meaningful, and reflects the anatomical and physicochemical differences present on dog skin (Figure 30). We found influences of gastrointestinal and oral microbiota on the perianal region and the chin, reflecting their physical proximity (Chapter 3.2 and 3.3, Cuscó et al., 2017a, 2017b). These other body site microbiotas present higher bacterial concentrations than the skin (Belkaid and Segre, 2014a; Sender et al., 2016). The inner pinna presented the most uniform skin microbiota composition among individuals despite being the most diverse skin site when compared to the other skin sites, probably because it was the most isolated part of the skin of the dogs included in our studies (Chapter 3.2 and 3.3, Cuscó et al., 2017a, 2017b).

To obtain first insights on the functional potential of a microbial community, predictive tools that work with 16S rRNA gene data, such as PICRUSt (Langille et al., 2013), PanFP (Jun et al., 2015) or Tax4fun (Aßhauer et al., 2015) have proven to be useful. At the predicted functional level, we were able to detect an increase on lipid metabolism on back skin, where sebaceous glands are larger and more abundant (Figure 30c). This last finding suggested that microbiota was adapting to the specific microhabitat or resource.

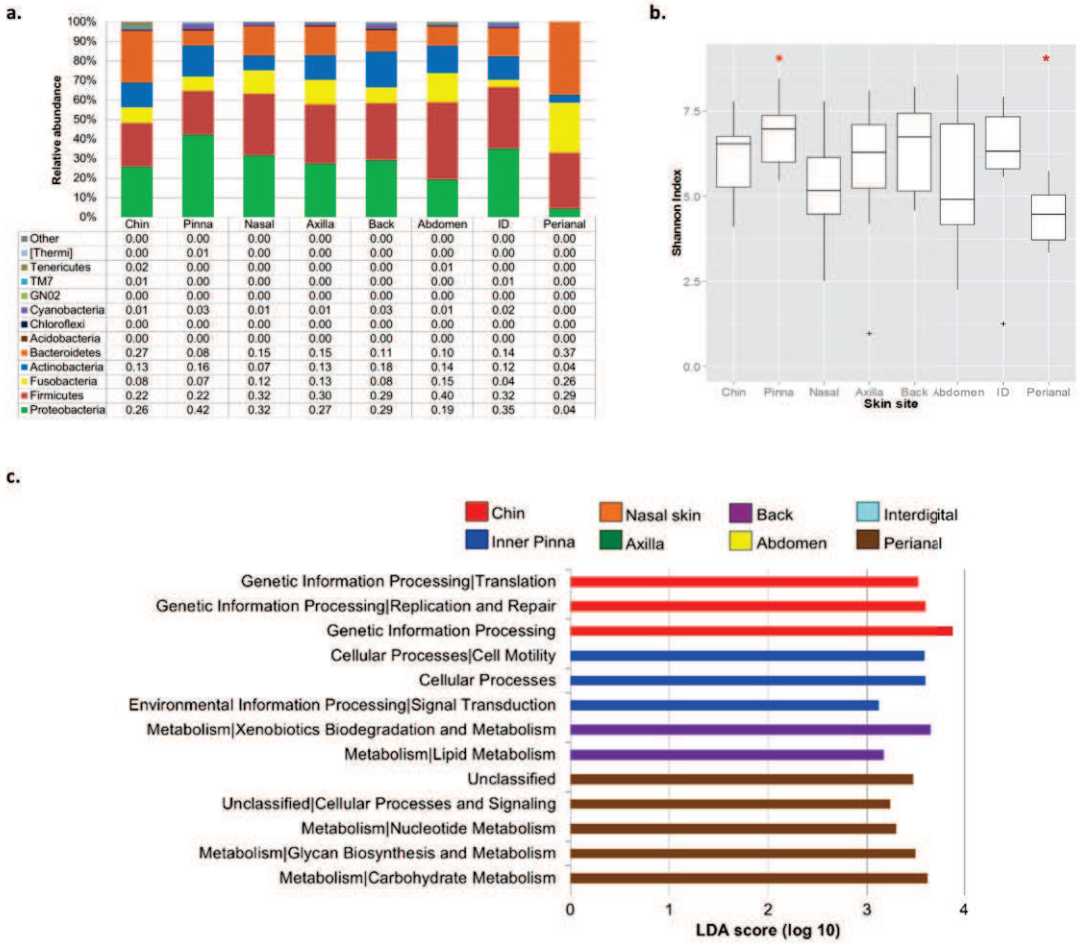


Figure 30. Microbiota of healthy dogs depending on the skin site. In a) microbiota profile at phylum level of the 9 dogs cohort. In b) boxplots of the diversity values depending of each skin site, marked with an asterisk those significantly different skin sites. In c) differentially distributed predicted functions on different skin sites (Figures from Cuscó et al., 2017a).

Further studies assessing the metagenomes, using shotgun whole genome sequencing, will allow detecting rather than predicting the functional potential of the microbial community. However, metagenomics studies on the skin are especially challenging due to the low biomass obtained from this tissue (Kong et al., 2017).

To conclude, we have detected different skin site microbial signatures that elucidate the need of cross-sectional studies and the avoidance of general results or even therapies focused on "the skin" as a unique microhabitat. Specifically, we found few examples on how skin site properties shape its specific microbiota, even one at the functional level. The next step will be finding out how these specific microbial signatures contribute to skin functions to promote health or disease.

4.3. Individual signatures on skin microbiota

Inter-individual diversity is high in human skin microbiota (Grice and Segre, 2011; Human Microbiome Project Consortium, 2012) and can even be used to identify individuals by matching them to objects that the person came in contact with (Fierer et al., 2010; Lax et al., 2015; Meadow et al., 2014). Generally low-abundant microorganisms are the ones with the highest identification power, tend to be spread over all the skin sites, and their abundances remain relatively constant through time (Oh et al., 2016).

Despite prior studies suggested individuality was affecting canine skin microbiota (Rodrigues Hoffmann et al., 2014), they did not assess this factor directly. Here, we specifically found that the individual was the main factor shaping dog skin microbiota composition: samples resembled more each other within the same dog, even when including different skin sites (Figure 31) (Chapter 3.2 and 3.3, Cuscó et al., 2017a, 2017b).

In Chapter 3.2, we could not distinguish if that effect was purely the individual or maybe it was due to their unique associated environment (the 9 dogs came from different households). However, in Chapter 3.3 we worked with a large cohort of same-breed dogs cohabiting together in the same environment and our results confirmed that the individual was the main factor shaping skin microbiota.

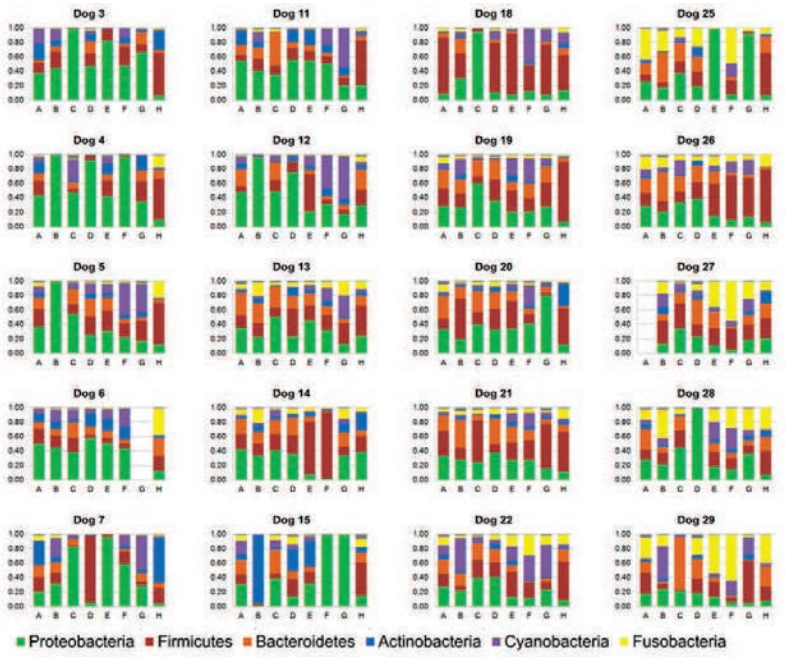


Figure 31. Skin microbiota profile of healthy dogs at phylum level. A, B, C, D, E, F, G and H represent inner pinna, chin, nasal skin, dorsal back, axilla, abdomen, interdigital region and perianal region. (Figure modified from Cuscó et al., 2017a).

This individuality of the skin microbiota should be taken into account on disease treatments. Skin microbiota is altered during several cutaneous diseases on humans and on atopic dermatitis on dogs (reviewed at Section 1.3.4). Considering that fact together with host genetics, generic therapies are probably not the best option. These individual effects could explain the differences on efficacy of the different treatments for skin diseases, such as those seen for canine atopic dermatitis (Olivry and Bizikova, 2013).

Microbiota manipulation can be potentially used to treat diseases as reviewed previously on Section 1.3.5. Using a skin microbiota survey, Bradley and colleagues found *Staphylococcus* and *Corynebacterium* increased in canine atopic dermatitis (Figure 32a). Despite these general characteristics, they treated the affected dogs with antimicrobials and see a highly personalized trend on both the degree of improvement (Figure 32b, Visit 2) and the outcome after antibiotics removal (Figure 32b, Visit 3) (Bradley et al., 2016b).

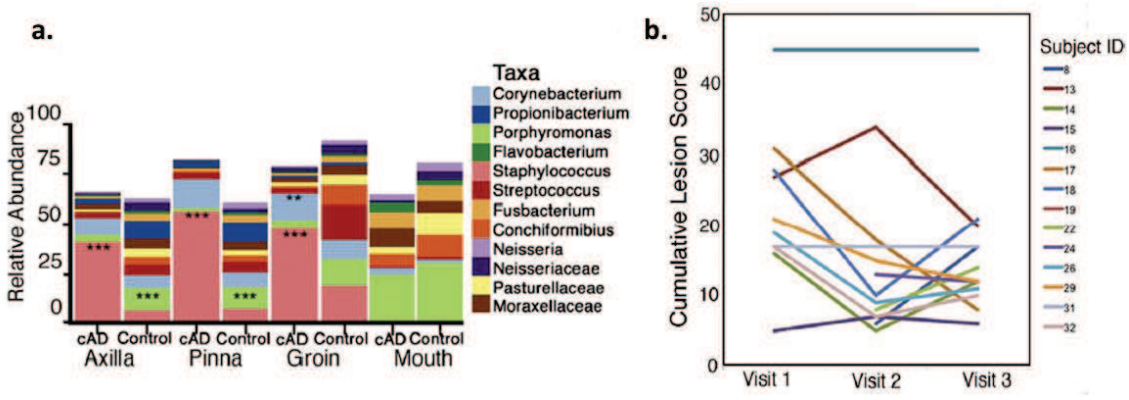


Figure 32. Dog skin microbiota on atopic dermatitis. a) Microbiota profile at genus level on healthy and atopic dermatitis groups. b) Atopic dermatitis score through time: visit 1, is the initial point; visit 2, during antimicrobial treatment; visit 3, after removing antimicrobials (Figures from Bradley et al., 2016).

Besides classical antimicrobial administration, other therapies such as administration of pre- and probiotics could be used to treat or prevent certain diseases. In the light of our results, where the individual is the main factor shaping dog skin microbiota, assessing individual microbiota profile for each dog would help to understand the different outcomes to a common treatment between individuals. The study from Bradley and colleagues demonstrates the necessity to assess the individual microbial signatures and to use this information to develop personalized therapies (Bradley et al., 2016b).

These individual signatures are often seen at low taxonomic levels, such as species level (Oh et al., 2014). To achieve that, we assessed 3rd generation sequencing technologies using MinION™ and found that species-level assignment was possible in some cases when sequencing the full-length 16S rRNA gene (Chapter 3.4).

These individual-specific characteristics of skin microbiota have been seen not only at compositional level, but also in the dynamics of microbial populations using network studies. Bashan and colleagues found that among the skin sites tested (left and right antecubital fossae and retro-auricular creases), only retro-auricular creases skin microbiota presented universal dynamics in humans (Bashan et al., 2016).

In conclusion, individuality of dog skin microbiota is a fact that veterinary clinicians and researchers should account for, especially when developing new treatments to modify or alter skin microbiota. It is worthy to note that individuality is not synonymous of host genetics neither host environment, but it is probably a combination of many factors and it remains to be clarified to which extent one affect to the other. Despite this individuality, it will also be interesting to assess if any universal interactions exist in specific skin sites, as those seen in retro-auricular creases in human (Bashan et al., 2016). If microbial dynamics of dog skin are not universal, generic microbiota manipulations may result ineffective or even detrimental.

4.4. Environmental bacteria on skin: transient or resident microbiota?

A recent study on caterpillars has demonstrated that they lack a resident gut microbiota, and they only have bacteria that come from the environment and the ingested food, with more than 80% of the 16S rRNA reads belonging to chloroplasts. Treating the animals with antibiotics did not affect their fitness or survival, demonstrating that the transient microbiota was not contributing on their overall health (Hammer et al., 2017).

In a relatively isolated environment such as the gut, it is easier to control whether the microbiota comes from the environment or not. However that is much more difficult in the skin, which is totally exposed to the outer environment. For example, tribal populations that live mostly outdoors and have close contact with nature presented an incredible rich microbiota with many environmental-derived bacteria, when compared to westernized civilizations (Clemente et al., 2015; Hospodsky et al., 2014).

In Chapter 3.2 and 3.3 we assessed skin microbiota of healthy dogs and detected a clear influence of the environment, both by the chloroplast amplification (ranging from 2 to 77% of the total sequences per sample in Chapter 3.2) and by the environmental-associated bacteria such as *Xenococcaceae, Sphingomonadaceae* or *Pseudomonadaceae* (Cuscó et al., 2017a)(Cuscó et al., 2017b). Previous studies also detected environmental-derived bacteria on canine skin microbiota (Rodrigues Hoffmann et al., 2014; Song et al., 2013a; Torres et al., 2017).

Song and colleagues suggested that dogs not only harbor a resident microbiota but also shed a transient one (Song et al., 2013a). Even some built environment microbiota studies considered dogs as vectors of the outer environment to the household, increasing indoors-bacterial diversity (Kettleson et al., 2015; Miletto and Lindow, 2015).

The large proportion of chloroplasts sequences seen on some dog skin microbiota samples probably showed their close and recent contact with vegetation and could be indicating that some of the bacteria that we detected on these samples come from plants (Figure 33). Comparing low-chloroplast with high-chloroplast skin microbiota samples after chloroplasts removal will give powerful insights in identifying environment-associated microbiota. Moreover in future studies it will be interesting to sample not only the dog skin but also the environment (the soil where it walks, the plants which it interacts with, the water where it plays, etc.).

Figure 33. Dogs interacting with the environment.

In conclusion, we detected that a variable proportion of the skin microbiota on dogs came from the outer environment. Temporal studies through seasons can give some clues of environmental-derived taxa (Torres et al., 2017), but if they interact or not with the host microbial dynamics and the resident microbiota remains to be elucidated and it will probably be harder to determine. Using some of the approaches proposed here could help to give light to this question.

4.5. Conducting microbiota studies: 16S rRNA gene and beyond

Depending on the research question to answer, alternative experimental scenarios can be used besides targeting specific 16S rRNA gene hypervariable regions and sequencing them using 2nd generation sequencers. In Figure 34, we provide an overview of the main experimental approaches to perform a microbiota study and their main benefits and pitfalls.

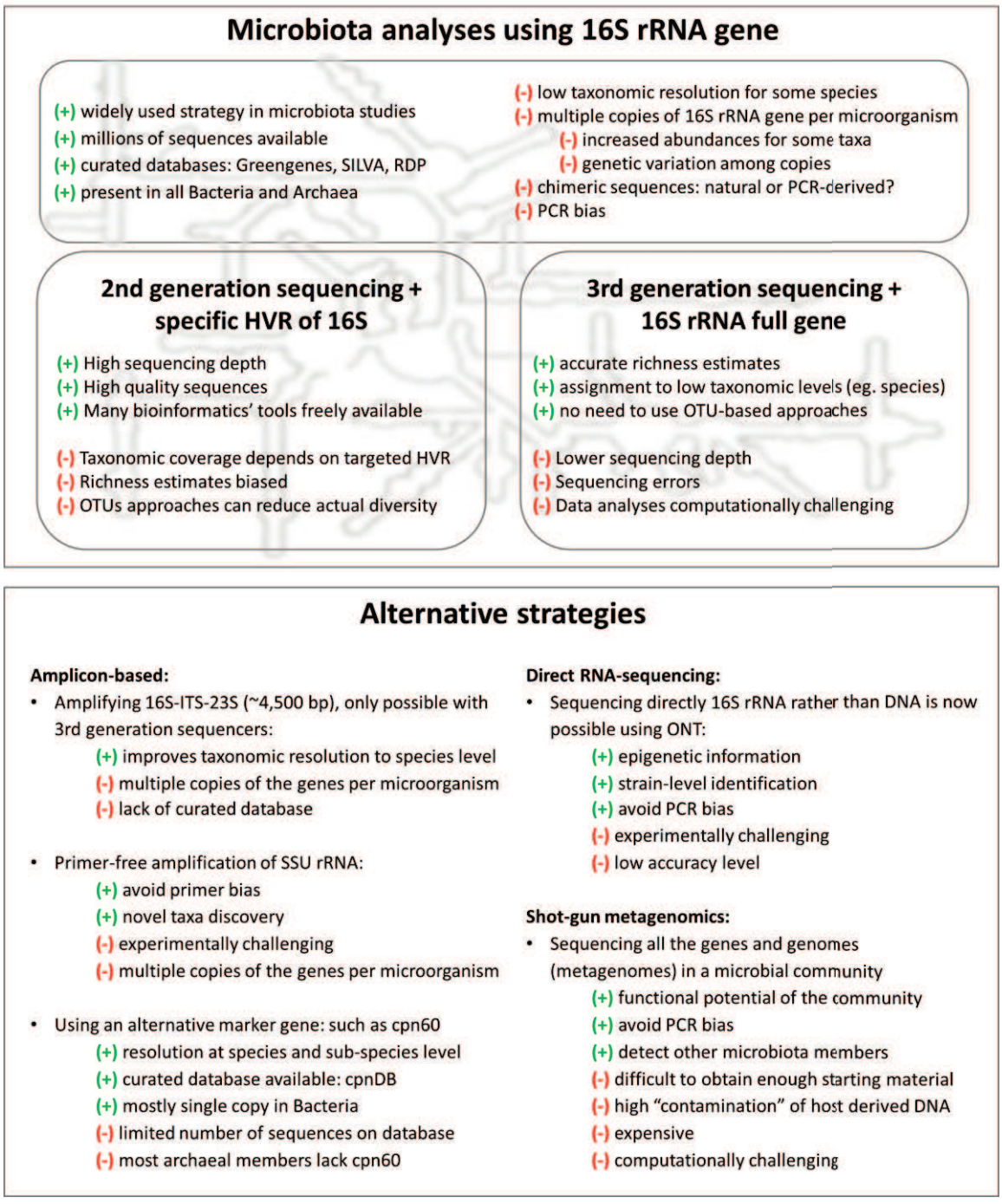


Figure 34. Experimental approaches to perform a microbiota study. (Source: own preparation)

Amplicon-based approach targeting specific hypervariable regions of 16S rRNA gene using second generation sequencing is the preferred choice when conducting microbiota studies. This approach, despite working with short fragments, allows high sequencing depth, low error rates and affordable prices (depending on the technology chosen). Moreover, a broad range of bioinformatics' tools are freely available to analyze the data. Thus, the main approach used on this dissertation will be valid for future analyses.

The most commonly used primers for microbiota studies are F27–R338 that amplify V1-V2 hypervariable regions and F515–R806 that amplify V4 (Kuczynski et al., 2011). In this thesis we choose the primer set F27-R338 (V1-V2), because despite F515–R806 (V4) primer set is more universal (Walters et al., 2011), it lacks sensitivity for *Propionibacterium acnes*, which is a common commensal of human skin microbiota (Figure 35)(Grice et al., 2009).

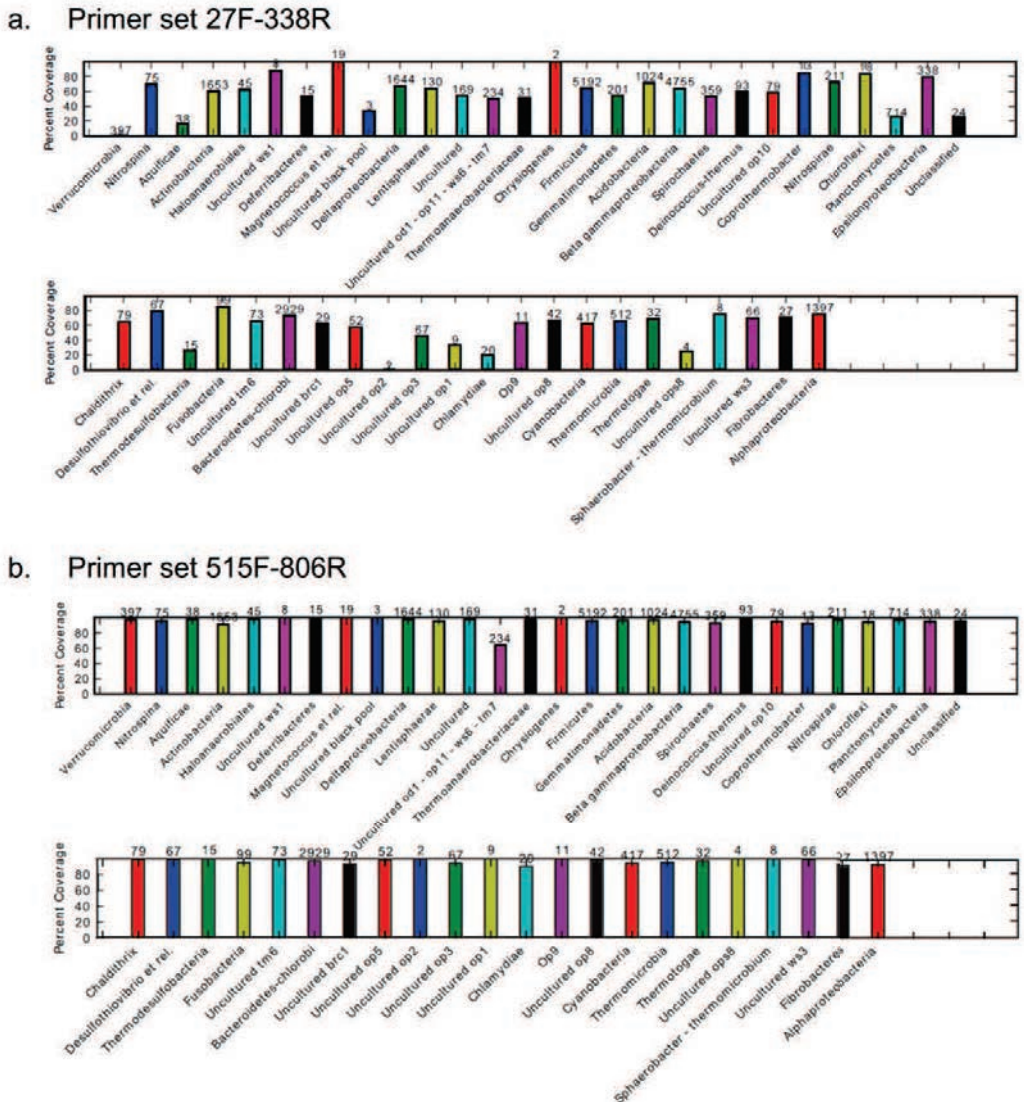


Figure 35. Taxonomic coverage at phylum level for bacteria of two universal 16S rRNA gene primer sets. In a) the 27F-338R primer set (used in this study) and in b) the 515F/806R primer set. The y-axes represent percent coverage and the value on top of each bar is the total number of reference sequences in each taxon (adapted from Walters et al., 2011).

On the other hand, when working with other biological samples rather than skin, researchers should choose the most appropriate primer set. In GI microbiota the most suitable primer set is F515–R806 (V4), considering that F27–R338 (V1-V2) lacks sensitivity for *Bifidobacterium* or the whole phylum *Verrumicrobia*, both expected to be in the GI tract (Kuczynski et al., 2011). For example, the primer set used in this thesis is not able to identify *Akkermansia muciniphila* (from *Verrumicrobia* phylum), one of the species with the highest anti-inflammatory effects that is being studied as a potential probiotic. Moreover, it is associated with weight-loss and lean phenotypes and also its relative abundance on GI microbiota is lower in several metabolic diseases, such as obesity or inflammatory bowel diseases, when compared to health status (Belzer and de Vos, 2012; Derrien et al., 2017; Gómez-Gallego et al., 2016).

Therefore, in the light of our results (Chapter 3.2 and Chapter 3.3), where *Propionibacterium* species represented really low abundances of the total skin microbiota on healthy dogs, it would also be interesting to replicate the analyses with F515–R806 (V4) primer set, especially in cases where an anti-inflammatory effect is studied.

When analyzing short-reads, most bioinformatics approaches rely on operational taxonomic units (OTUs). Working with OTUs helps to reduce data volume and facilitates the posterior analyses by clustering together all the sequences above a specific threshold (usually 97%). As it has to process all the sequences, this step is usually slow and is considered as the bottleneck on microbiota data analyses. This methodology can present several disadvantages, such as 1) irreproducibility, OTU clustering algorithms lead to divergent result from a same dataset (Schmidt et al., 2015); 2) instability, sequences forming an OTUs vary depending on the number of sequences in the analysis, unless clustering against a database (He et al., 2015); 3) inaccuracy, OTU diversity estimates can be inflated due to sequencing errors or PCR artifacts (Patin et al., 2013). Recently, new algorithms have been developed aimed at improving velocity and accuracy of OTU picking, such as OptiClust (Westcott and Schloss, 2017), Deblur (Amir et al., 2017) or DADA2 (Callahan et al., 2016) among many others.

To sum up, if using second generation sequencing and 16S rRNA gene, the hypervariable region targeted and the universal primer sets used should be carefully chosen depending on the sample of interest. Moreover, using the most recently released software and applying strict quality filters will allow obtaining more accurate and reliable results.

As we have seen in chapters 3.2 and 3.3, taxonomy assignments of short fragments of 16S rRNA obtained through second generation massive sequencing do not reach species level. The main reasons are that: i) fragment length analyzed cannot discriminate among several bacterial species and full 16S rRNA gene is required (Drancourt et al., 2000; Janda and Abbott, 2007; Schloss et al., 2016a; Yarza et al., 2014); ii) analyses of microbiota data rely on OTU-based clustering approaches, where all the sequences that are similar above a specific threshold (usually 97%) are considered the same species altering diversity estimates (He et al., 2015; Patin et al., 2013; Rossi-Tamisier et al., 2015); and iii) RDP taxonomy

classifier only computes reliable assignments up to genus level (Schloss et al., 2016a; Wang et al., 2007).

Third generation sequencing generates longer reads that allow targeting the full 16S rRNA gene and obtaining better richness estimates and taxonomic resolution (Yarza et al., 2014). In fact, the ability to classify a sequence to the genus or species level is a function of read length, sample type and the reference database (Schloss et al., 2016a). Thus, to overcome some of the short-read associated problems, we tested 3rd generation MinION™ platform to sequence the full 16S rRNA gene. On one hand, we were able to obtain information down to species level, as well as we did see increased bacterial richness. On the other hand, this was not possible in all the cases because of the databases and data analyses used were not optimum, and the chemistry used (1D reads) still presented low accuracy values. Moreover, the single 16S rRNA gene has high sequence similarity and not enough taxonomic resolution for some species.

Benitez-Paez and Sanz suggested another approach to improve taxonomic resolution using nanopore sequencing by targeting not only 16S rRNA gene, but the whole ribosomal RNA operon (rrn region) constituted by 16S-ITS-23S (~4,500 bp). This approach allowed obtaining a better taxonomic resolution at species level as well as seeing 2-fold more diversity than when using only 16S rRNA gene (Benitez-Paez and Sanz, 2017).

Some of the pitfalls associated with the sequencing technology are being improved by Oxford Nanopore Technologies. They have recently launched the $1D^2$ sequencing chemistry (May 2017), which improves the accuracy of the reads by sequencing both template and complement strands (Figure 36). To improve accuracy, Intramolecular-ligated Nanopore Consensus Sequencing (INC-Seq) approach could be applied to microbiota studies. INC-Seq uses rolling circle amplification of circularized templates to generate linear products (with tandem copies of the template) that can be sequenced on the nanopore platform (Li et al., 2016). Besides that, several bioinformatics tools and strategies have been suggested to further correct the raw sequences, such as NanoCORR (Goodwin et al., 2015) that corrects with an hybrid approach combining the long nanopore reads with short high-quality reads or Nanocorrect and Nanopolish (Loman et al., 2015) that use self-correction algorithms. Some genome assemblers, such as canu (Koren et al., 2017) also perform previous steps of correcting and consensus sequences that can improve accuracy. These tools are mostly developed for *de novo* genome assembly, and take advantage of the fact that you are sequencing one genome and some regions overlap, which is not so useful for amplicon-based microbiota studies. Further studies of microbiota should be using $1D^2$ sequencing kit together with suitable correction methods.

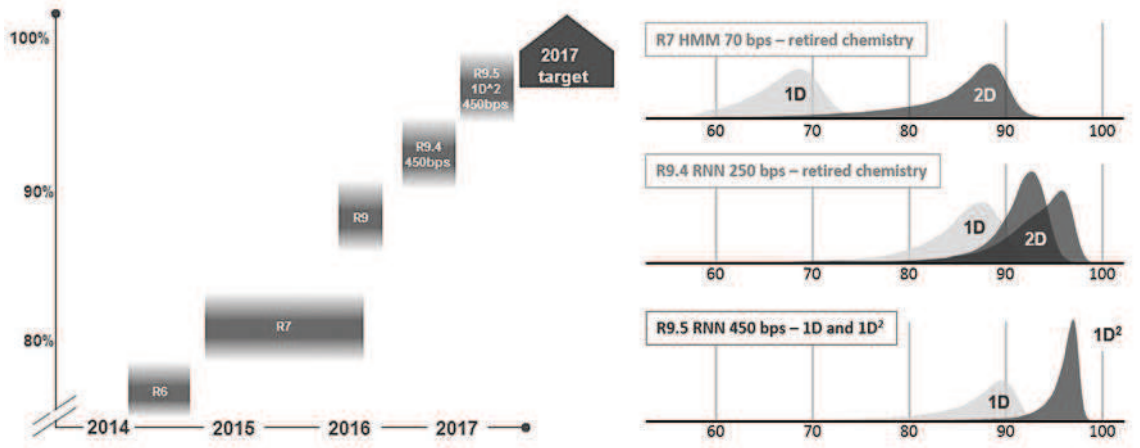


Figure 36. Evolution on the accuracy of Oxford Nanopore Technologies sequencing kits (Figure from ONT, Brown, 2017).

Sometimes 16S rRNA marker gene is not enough. It has several disadvantages such as: i) "universal primers" used to amplify this gene were designed using sequences obtained from cultured microorganisms, which could be biasing the actual diversity (Rajendhran and Gunasekaran, 2011; Schloss and Handelsman, 2004); ii) databases are ecosystem-skewed (Schloss et al., 2016b); 3) the copy number of rRNA operons per bacterial genome varies from 1 to 15 (Klappenbach et al., 2001), which can lead to inflated bacterial diversity estimates for some species; and 4) some species present a high similarity in this gene, as we have corroborated in Chapter 3.4 for *Streptococcus mutans* or for *Escherichia coli*.

To overcome the first issue, Karst and colleagues developed a primer-free amplification of SSU rRNA based on tagging single cDNA molecules obtained from size-selected rRNA (aimed to enrich SSU rRNA). When comparing to SILVA database, they found that 30% of all bacterial OTUs were novel and that the degree of novelty was highly ecosystem specific, ranging from 36% in soil sample to 5% in human gut sample (Karst et al., 2016). This approach elucidated that databases are environmental-skewed (Figure 37), and the authors suggested that creating, maintaining and using an environment-specific 16S rRNA gene database would be more feasible and appropriate to analyze specific microhabitats(Karst et al., 2016), such as The Human Oral Microbiome Database (Chen et al., 2010).

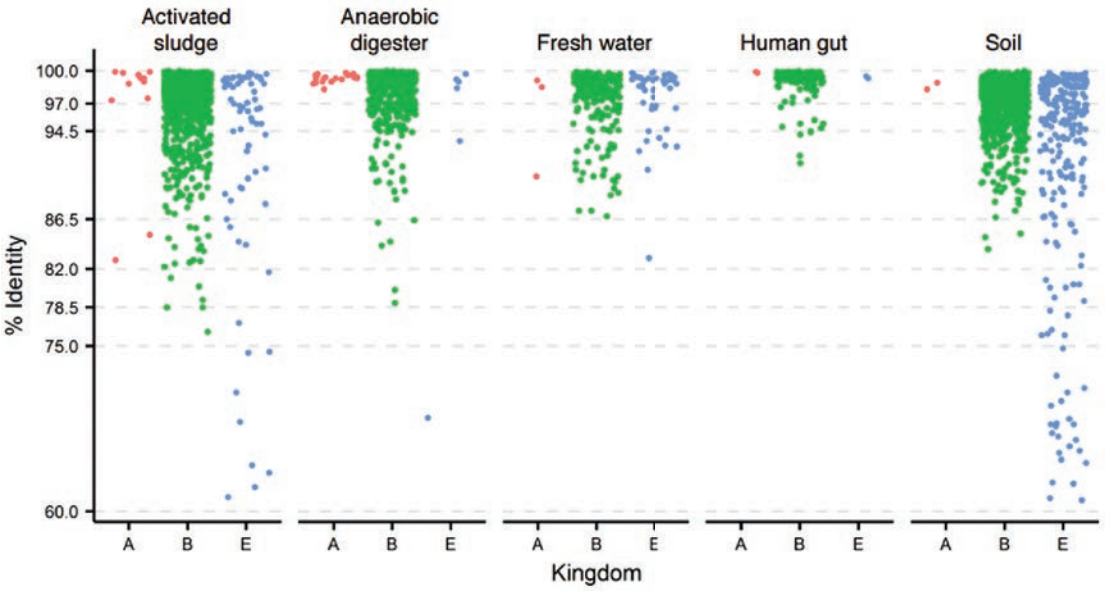


Figure 37. Coverage of the tree of life. The percent identity of SSU rRNA gene sequences in the samples compared to their closest relatives in the SILVA database. Red, green, and blue represent archaea, bacteria, and eukarya, respectively (Figure from Karst et al., 2016).

To overcome the third and fourth problems, which are more linked to the biological properties of 16S rRNA gene, some researchers rely on another universal barcode for bacteria: cpn60 protein coding gene. This gene is usually present as a single copy on bacteria, which would improve relative abundance profiles. As exceptions to this general rule *Chlamydia*, and some members of *Rhizobia* and *Actinobacteria,* present multiple copies of the gene (Lund, 2009) (Figure 38a), whereas several species of *Mycoplasma* and *Ureaplasma parvum* lack this gene (Hill et al., 2004). Links and colleagues compared the two bacterial barcodes under the framework of the International Barcode of Life project and found that the most informative regions of the 16S rRNA gene (V1-V3 region) are less taxonomically informative than the most conserved segments of the cpn60 UT, for which average sequence identity does not exceed 92% (Figure 38b). Thus, cpn60 UT has more taxonomic resolution and is able to reach species and even subspecies level (Links et al., 2012). Moreover, this marker gene also has a curated reference database: cpnDB (Hill et al., 2004).

Another approach would be sequencing RNA directly, which is now possible with nanopore sequencing (Garalde et al., 2016) and even has been optimized for 16S rRNA specifically (Smith et al., 2017). Some divergences at strain level can be consequence of base modifications on the 16S rRNA. For example, 16S ribosomal RNA methylation has been associated to resistance against aminoglycosides on several bacterial species (Doi and Arakawa, 2007). Using direct RNA sequencing on MinION, Smith and colleagues were able to detect the modified bases of *Escherichia coli*, which can identify a pathogenic strain (Smith et al., 2017). Thus, this approach has potential to be used on microbiota studies to assess taxonomy without PCR bias as well as to detect epigenetic divergences among bacterial strains.

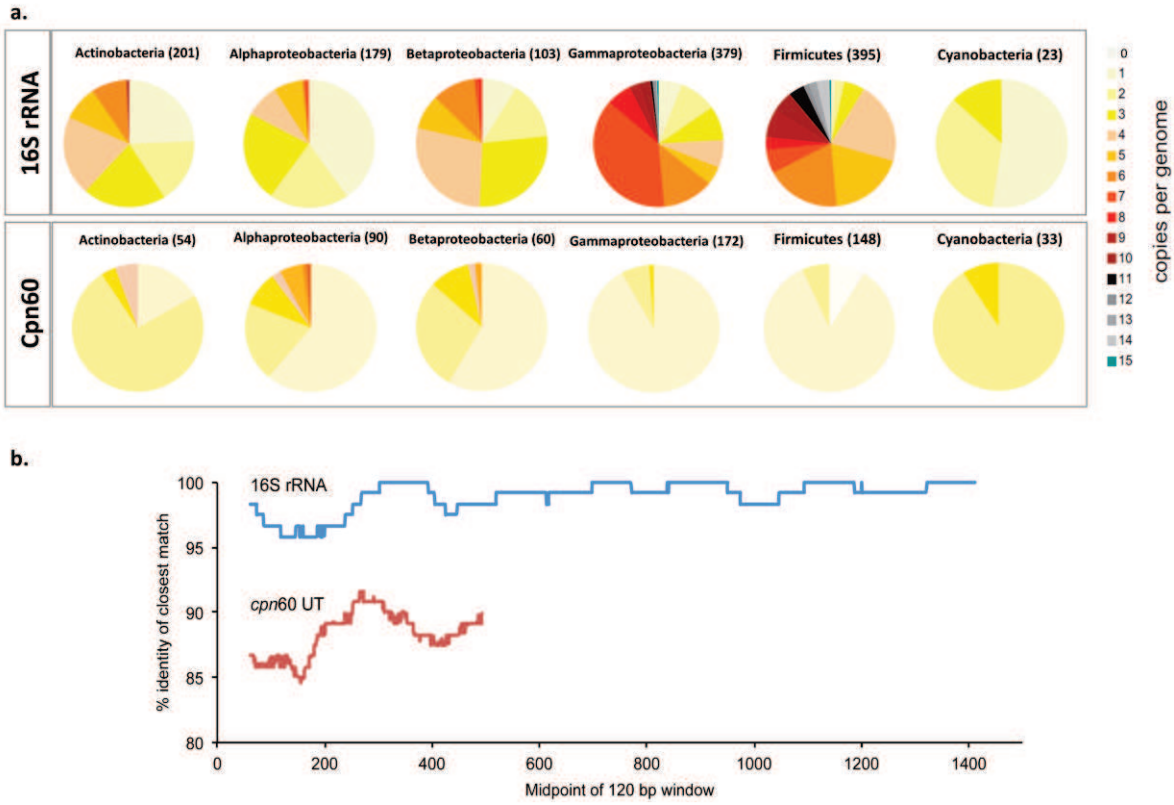


Figure 38. Comparison of universal bacterial barcodes: 16S rRNA and Cpn60. In a) pie charts representing the number of bacterial barcode gene copies per bacterial genome; in parenthesis number of genomes included per taxa. In b) sequence diversity across the 16S rRNA gene and cpn60 UT. Figure a) data from Větrovský and Baldrian, 2013 for 16S and Lund, 2009 for cpn60. Figure b) Excerpted from Links et al., 2012.

The broadest approach to study the microbiota is whole-genome sequencing that allows getting the complete picture by sequencing all microbial members, including bacteria, archaea, fungi, viruses or even eukaryotes. This approach aims to sequence all the genomes in a specific environment and not only a marker gene, which means getting information of the functional potential of the community as well as strain level identification (Kong et al., 2017; Meisel et al., 2016). Few studies on human skin have already used it (Chng et al., 2016; Oh et al., 2014, 2016). The main challenge of this technique is data analysis complexity and host-associated DNA contamination. Moreover, skin samples present low DNA biomass, which could add more contamination problems associated with laboratory reagents or laboratory environment (Kong et al., 2017; Salter et al., 2014). Moreover, on dog skin we would also expect to see contamination from chloroplasts DNA, as we have already detected with 16S rRNA (Chapter 3.2 and 3.3, Cuscó et al., 2017a, 2017b).

Thus, plenty of options have already been proposed and assessed to characterize microbiota structure and composition (Figure 34) and in this thesis we focused on two of them. In the near future, new technologic and bioinformatics advances will allow creating and developing new strategies and some of the new approaches stated here will become routine. It is always important to keep in mind that depending on the research question to answer, some scenarios are more suitable than others.

5. Conclusions

Regarding the innate immunity on healthy dogs, we can conclude that:

1. Genetic variation on TLR genes is greater than previously reported, we have identified 105 novel coding variants by massive sequencing on different dog breeds and wolf populations.
2. Frequencies of TLR genetic variants differ among dog breeds and wolf populations, even some of them are breed-specific or species-specific.
3. The extracellular TLR5 is the most polymorphic among all the canine TLR genes.
4. We designed and validated a TaqMan OpenArray plate containing 60 non-synonymous variants and 4 frameshift mutations spread in the 10 canine TLR genes.
5. Individual genotyping of some dogs representing different breeds allowed us detecting that Labrador and to lesser extent German Shepherd breeds presented different innate immune profiles, when screening TLR genes.

Regarding the skin microbiota on healthy dogs, we can conclude that:

6. The individual is the main factor shaping skin microbiota: different skin sites within the same dog resemble more than the same skin site between different dogs. That is true even in a homogeneous environment.
7. The skin site is defining skin microbiota structure and composition. The inner pinna (hairy skin) is the skin site with the highest diversity, whereas nasal skin and the perianal region (muco-cutaneous regions) are the skin sites with the lowest.
8. The season of birth or the time spent in the kennel was affecting dog skin microbiota in all the skin sites of dogs cohabiting together for at least 3 months. *Sphingomonas* was the most abundant bacterium driving this difference.

9. Bacteria from the environment are highly present on dog skin microbiota, although whether they are transient or permanent members of the microbiota remains to be elucidated.
10. Other factors such as breed, sex and surgery seem to be affecting canine skin microbiota, although larger studies should be performed to obtain stronger evidence.

Regarding the assessment of single molecule sequencing experimental approach to obtain microbiota data, we can conclude that:

11. Nanopore sequencing of full-length 16S rRNA gene using MinION™ allowed obtaining better richness estimates and detecting previously unseen taxa on dog skin microbiota.
12. Nanopore sequencing of full-length 16S rRNA gene using MinION™ allowed obtaining taxonomy assignment at lower levels, even at species level in some cases.

6. Bibliography

Abida, O., Bouzid, D., Krichen-Makni, S., Kharrat, N., Masmoudi, A., Abdelmoula, M., Ben Ayed, M., Turki, H., Sellami-Boudawara, T., and Masmoudi, H. (2013). Potential Role of TLR Ligand in Aethiopathogenesis of Tunisian Endemic Pemphigus Foliaceus. Biochem. Physiol. Open Access *2*, 1–5.

Agarwal, S.K., Wu, M., Livingston, C.K., Parks, D.H., Mayes, M.D., Arnett, F.C., and Tan, F.K. (2011). Toll-like receptor 3 upregulation by type I interferon in healthy and scleroderma dermal fibroblasts. Arthritis Res. Ther. *13*, R3.

Al-Bagdadi, F. (2013). The Integument. In Miller's Anatomy of the Dog, H.E. Evans, and A. de Lahunta, eds. (Elsevier Inc.), 61–79.

Albertsen, M., Karst, S.M., Ziegler, A.S., Kirkegaard, R.H., Nielsen, P.H., and Stokholm-Bjerregaard, M. (2015). Back to Basics – The Influence of DNA Extraction and Primer Choice on Phylogenetic Analysis of Activated Sludge Communities. PLoS One *10*, e0132783.

Alekseyenko, A. V, Perez-Perez, G.I., De Souza, A., Strober, B., Gao, Z., Bihan, M., Li, K., Methé, B.A., and Blaser, M.J. (2013). Community differentiation of the cutaneous microbiota in psoriasis. Microbiome *1*, 31.

Allenspach, K., House, A., Smith, K., McNeill, F.M., Hendricks, A., Elson-Riggins, J., Riddle, A., Steiner, J.M., Werling, D., Garden, O.A., et al. (2010). Evaluation of mucosal bacteria and histopathology, clinical disease activity and expression of Toll-like receptors in German shepherd dogs with chronic enteropathies. Vet. Microbiol. *146*, 326–335.

Amir, A., McDonald, D., Navas-Molina, J.A., Kopylova, E., Morton, J.T., Zech Xu, Z., Kightley, E.P., Thompson, L.R., Hyde, E.R., Gonzalez, A., et al. (2017). Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. mSystems *2*, e00191-16

Arumugam, M., Raes, J., Pelletier, E., Paslier, D., Yamada, T., Mende, D.R., Fernandes, G.R., Tap, J., Bruls, T., Batto, J., et al. (2011). Enterotypes of the human gut microbiome. Nature *473*, 174–180.

Ashelford, K.E., Chuzhanova, N.A., Fry, J.C., Jones, A.J., and Weightman, A.J. (2005). At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies. Appl. Environ. Microbiol. *71*, 7724–7736.

Aßhauer, K.P., Wemheuer, B., Daniel, R., and Meinicke, P. (2015). Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data. Bioinformatics *31*, 2882–2884.

Bakry, O.A., Samaka, R.M., Sebika, H., and Seleit, I. (2014). Toll-like receptor 2 and P. acnes: do they trigger initial acne vulgaris lesions? Anal. Quant. Cytopathol. Histopathol. *36*, 100–110.

Bangert, C., Brunner, P.M., and Stingl, G. (2011). Immune functions of the skin. Clin. Dermatol. *29*, 360–376.

Baquerizo Nole, K.L., Yim, E., and Keri, J.E. (2014). Probiotics and prebiotics in dermatology. J. Am. Acad. Dermatol. *71*, 814–821.

Barnard, E., and Li, H. (2017). Shaping of cutaneous function by encounters with commensals. J. Physiol. *595*, 437–450.

Barreiro, L.B., Ben-Ali, M., Quach, H., Laval, G., Patin, E., Pickrell, J.K., Bouchier, C., Tichit, M., Neyrolles, O., Gicquel, B., et al. (2009). Evolutionary dynamics of human Toll-like receptors and their different contributions to host defense. PLoS Genet. *5*, e1000562.

Bashan, A., Gibson, T.E., Friedman, J., Carey, V.J., Weiss, S.T., Hohmann, E.L., and Liu, Y.-Y. (2016). Universality of human microbial dynamics. Nature *534*, 259–262.

Bastos, M.C.F., Ceotto, H., Coelho, M.L. V, and Nascimento, J.S. (2009). Staphylococcal antimicrobial peptides: relevant properties and potential biotechnological applications. Curr. Pharm. Biotechnol. *10*, 38–61.

Begon, E., Michel, L., Flageul, B., Beaudoin, I., Jean-Louis, F., Bachelez, H., Dubertret, L., and Musette, P. (2007). Expression, subcellular localization and cytokinic modulation of Toll-like receptors (TLRs) in normal human keratinocytes: TLR2 up-regulation in psoriatic skin. Eur. J. Dermatol. *17*, 497–506.

Belkaid, Y., and Segre, J.A. (2014). Dialogue between skin microbiota and immunity. Science *346*, 954–959.

Belzer, C., and de Vos, W.M. (2012). Microbes inside—from diversity to function: the case of Akkermansia. ISME J. *6*, 1449–1458.

Benitez-Paez, A., and Sanz, Y. (2017). Multi-locus and ultra-long amplicon sequencing approach to study microbial diversity at species level using the MinION TM portable nanopore sequencer. bioRxiv.

Beutler, B. (2004). Inferences, questions and possibilities in Toll-like receptor signalling. Nature *430*, 257–263.

Beutler, B. (2009). Microbe sensing, positive feedback loops, and the pathogenesis of inflammatory diseases. Immunol. Rev. *227*, 248–263.

Blekhman, R., Goodrich, J.K., Huang, K., Sun, Q., Bukowski, R., Bell, J.T., Spector, T.D., Keinan, A., Ley, R.E., Gevers, D., et al. (2015). Host genetic variation impacts microbiome composition across human body sites. Genome Biol. *16*, 1–12.

Bochud, P.Y., Sinsimer, D., Aderem, A., Siddiqui, M.R., Saunderson, P., Britton, S., Abraham, I., Tadesse Argaw, A., Janer, M., Hawn, T.R., et al. (2009). Polymorphisms in Toll-like receptor 4 (TLR4) are associated with protection against leprosy. Eur. J. Clin. Microbiol. Infect. Dis. *28*, 1055–1065.

Bonder, M.J., Kurilshikov, A., Tigchelaar, E.F., Mujagic, Z., Imhann, F., Vila, A.V., Deelen, P., Vatanen, T., Schirmer, M., Smeekens, S.P., et al. (2016). The effect of host genetics on the gut microbiome. Nat. Genet. *48*, 1407–1412.

Bradley, C.W., Morris, D.O., Rankin, S.C., Cain, C.L., Misic, A.M., Houser, T., Mauldin, E.A., and Grice, E.A. (2016). Longitudinal evaluation of the skin microbiome and association with microenvironment and treatment in canine atopic dermatitis. J. Invest. Dermatol. *136*, 1182-90.

Brown, C. (2017). Nanopore technology update, London Calling conference 2017.

Burgener, I.A., König, A., Allenspach, K., Sauter, S.N., Boisclair, J., Doherr, M.G., and Jungi, T.W. (2008). Upregulation of Toll-Like Receptors in Chronic Enteropathies in Dogs. J. Vet. Intern. Med. *22*, 553–560.

Cadieu, E., Neff, M.W., Quignon, P., Walsh, K., Chase, K., Parker, H.G., VonHoldt, B.M., Rhue, A., Boyko, A., Byers, A., et al. (2009). Coat Variation in the Domestic Dog Is Governed by Variants in Three Genes. Science. *326*, 150–153.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods *13*, 581–583.

Callewaert, C., Kerckhof, F., Granitsiotis, M.S., Gele, M. Van, Wiele, T. Van De, and Boon, N. (2013). Characterization of Staphylococcus and Corynebacterium Clusters in the Human Axillary Region. PLoS One. *8*, e70538.

Capone, K.A., Dowd, S.E., Stamatas, G.N., and Nikolovski, J. (2011). Diversity of the human skin microbiome early in life. J. Invest. Dermatol. *131*, 2026–2032.

Caporaso, J.G., Bittinger, K., Bushman, F.D., Desantis, T.Z., Andersen, G.L., and Knight, R. (2010). PyNAST: A flexible tool for aligning sequences to a template alignment. Bioinformatics *26*, 266–267.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat. Methods *7*, 335–336.

Castelino, M., Eyre, S., Moat, J., Fox, G., Martin, P., Ho, P., Upton, M., and Barton, A. (2017). Optimisation of methods for bacterial skin microbiome investigation: primer selection and comparison of the 454 versus MiSeq platform. BMC Microbiol. *17*, 23.

Celhar, T., Magalhães, R., and Fairhurst, A.M. (2012). TLR7 and TLR9 in SLE: when sensing self goes wrong. Immunol. Res. *53*, 58–77.

Chakravorty, S., Helb, D., Burday, M., and Connell, N. (2007). A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. J Microbiol Methods *69*, 330–339.

Chen, T., Yu, W.-H., Izard, J., Baranova, O. V., Lakshmanan, A., and Dewhirst, F.E. (2010). The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information. Database *2010*, baq013.

Chng, K.R., Tay, A.S.L., Li, C., Ng, A.H.Q., Wang, J., Suri, B.K., Matta, S.A., McGovern, N., Janela, B., Wong, X.F.C.C., et al. (2016). Whole metagenome profiling reveals skin microbiome-dependent susceptibility to atopic dermatitis flare. Nat. Microbiol. *1*, 16106.

Chotimanukul, S., and Sirivaidyapong, S. (2011). Differential expression of Toll-like receptor 4 (TLR4) in healthy and infected canine endometrium. Theriogenology *76*, 1152–1161.

Christensen, G.J.M., and Brüggemann, H. (2014). Bacterial skin commensals and their role as host guardians. Benef. Microbes *5*, 201–215.

Chu, H., and Mazmanian, S.K. (2013). Innate immune recognition of the microbiota promotes host-microbial symbiosis. Nat Immunol *14*, 668–675.

Clarke, J., Wu, H.-C., Jayasinghe, L., Patel, A., Reid, S., and Bayley, H. (2009). Continuous base identification for single-molecule nanopore DNA sequencing. Nat Nano *4*, 265–270.

Clarke, K.R. (1993). Non-parametric multivariate analyses of changes in community structure. Aust. J. Ecol. *18*, 117–143.

Clarridge, J.E. (2004). Impact of 16S rRNA Gene Sequence Analysis for Identification of Bacteria on Clinical Microbiology and Infectious Diseases. Clin. Microbiol. Rev. *17*, 840–862.

Clarys, P., and Barel, A. (1995). Quantitative evaluation of skin surface lipids. Clin. Dermatol. 13, 307–321.

Clemente, J.C., Pehrsson, E.C., Blaser, M.J., Sandhu, K., Gao, Z., Wang, B., Magris, M., Hidalgo, G., Contreras, M., Noya-Alarcon, O., et al. (2015). The microbiome of uncontacted Amerindians. Sci. Adv. *1*, e1500183.

Clooney, A.G., Fouhy, F., Sleator, R.D., and Driscoll, A.O. (2016). Comparing Apples and Oranges?: Next Generation Sequencing and Its Impact on Microbiome Analysis. PLoS One. *11*, e0148028.

Cogen, A.L., Yamasaki, K., Muto, J., Sanchez, K.M., Crotty Alexander, L., Tanios, J., Lai, Y., Kim, J.E., Nizet, V., and Gallo, R.L. (2010). Staphylococcus epidermidis Antimicrobial δ-Toxin (Phenol-Soluble Modulin-γ) Cooperates with Host Antimicrobial Peptides to Kill Group A Streptococcus. PLoS One *5*, e8557.

Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., Brown, C.T., Porras-Alfaro, A., Kuske, C.R., and Tiedje, J.M. (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. *42*, D633–D642.

Costello, E.K., Lauber, C.L., Hamady, M., Fierer, N., Gordon, J.I., and Knight, R. (2009). Bacterial community variation in human body habitats across space and time. Science *326*, 1694–1697.

Cotton, D., van Hasselt, P., and Bergers, A. (1975). Nature of the sweat glands in the hairy skin of the beagle. Dermatologica *150*, 75–85.

Cuscó, A., Belanger, J.M., Gershony, L., Islas-Trejo, A., Levy, K., Medrano, J.F., and Sánchez, Armand Oberbauer, Anita M. Francino, O. (2017b). Individual signatures and environmental factors shape skin microbiota on healthy dogs. Microbiome (submitted).

Cuscó, A., Sánchez, A., Altet, L., Ferrer, L., and Francino, O. (2017a). Individual Signatures Define Canine Skin Microbiota Composition and Variability. Front. Vet. Sci. *4*, 6.

Derrien, M., Belzer, C., and de Vos, W.M. (2017). Akkermansia muciniphila and its role in regulating host functions. Microb. Pathog. *106*, 171–181.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., and Andersen, G.L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl. Environ. Microbiol. *72*, 5069–5072.

Dimarzio, L., Cinque, B., Cupelli, F., De Simone, C., Cifone, M.G., and Giuliani, M. (2008). Increase of Skin-Ceramide Levels in Aged Subjects following a Short-Term Topical Application of Bacterial Sphingomyelinase from Streptococcus thermophilus. Int. J. Immunopathol. Pharmacol. *21*, 137–143.

Doi, Y., and Arakawa, Y. (2007). 16S Ribosomal RNA Methylation: Emerging Resistance Mechanism against Aminoglycosides. Clin. Infect. Dis. *45*, 88–94.

Dominianni, C., Wu, J., Hayes, R.B., and Ahn, J. (2014). Comparison of methods for fecal microbiome biospecimen collection. BMC Microbiol. *14*, 103.

Drancourt, M., Bollet, C., Carlioz, A., Martelin, R., Gayral, J.P., and Raoult, D. (2000). 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. J. Clin. Microbiol. *38*, 3623–3630.

Dreno, B., Martin, R., Moyal, D., Henley, J.B., Khammari, A., and Seité, S. (2017). Skin microbiome and acne vulgaris: Staphylococcus, a new actor in acne. Exp. Dermatol. 1-6.

Dressman, D., Yan, H., Traverso, G., Kinzler, K.W., and Vogelstein, B. (2003). Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. Proc. Natl. Acad. Sci. U.S.A. *100*, 8817–8822.

Eckburg, P.B., Lepp, P.W., and Relman, D.A. (2003). Archaea and their potential role in human disease. Infect. Immun. *71*, 591–596.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. Bioinformatics *27*, 2194–2200.

Egert, M., and Simmering, R. (2016). The Microbiota of the Human Skin. In Advances in Experimental Medicine and Biology, Adv Exp Med Biol. *902,* 61–81.

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., et al. (2009). Real-Time DNA Sequencing from Single Polymerase Molecules. Science *323*, 133–138.

Fahlén, A., Engstrand, L., Baker, B.S., Powles, A., and Fry, L. (2012). Comparison of bacterial microbiota in skin biopsies from normal and psoriatic skin. Arch. Dermatol. Res. *304*, 15–22.

Farina, G.A., York, M.R., Di Marzio, M., Collins, C.A., Meller, S., Homey, B., Rifkin, I.R., Marshak-Rothstein, A., Radstake, T.R, and Lafyatis, R. (2010). Poly(I:C) Drives Type I IFN- and TGFβ-Mediated Inflammation and Dermal Fibrosis Simulating Altered Gene Expression in Systemic Sclerosis. J. Invest. Dermatol. *130*, 2583–2593.

Faust, K., and Raes, J. (2016). CoNet app: inference of biological association networks using Cytoscape. F1000Research *5*, 1519.

Faye, T., Holo, H., Langsrud, T., Nes, I.F., and Brede, D.A. (2011). The unconventional antimicrobial peptides of the classical propionibacteria. Appl. Microbiol. Biotechnol. *89*, 549–554.

Ferrer, M., Méndez-García, C., Rojo, D., Barbas, C., and Moya, A. (2017). Antibiotic use and microbiome function. Biochem. Pharmacol. *134*, 114–126.

Ferrer-Admetlla, A., Bosch, E., Sikora, M., Marquès-Bonet, T., Ramírez-Soriano, A., Muntasell, A., Navarro, A., Lazarus, R., Calafell, F., Bertranpetit, J., et al. (2008). Balancing selection is the main force shaping the evolution of innate immunity genes. J. Immunol. *181*, 1315–1322.

Fierer, N., Hamady, M., Lauber, C.L., and Knight, R. (2008). The influence of sex, handedness, and washing on the diversity of hand surface bacteria. Proc. Natl. Acad. Sci. U.S.A. *105*, 17994–17999.

Fierer, N., Lauber, C.L., Zhou, N., McDonald, D., Costello, E.K., and Knight, R. (2010). Forensic identification using skin bacterial communities. Proc. Natl. Acad. Sci. U.S.A. *107*, 6477–6481.

Figueiredo, M.M., Amorim, I.F.G., Pinto, A.J.W., Barbosa, V.S., Pinheiro, L.D.J., Deoti, B., Faria, A.M.C., and Tafuri, W.L. (2013). Expression of Toll-like receptors 2 and 9 in cells of dog jejunum and colon naturally infected with Leishmania infantum. BMC Immunol. *14*, 22.

Findley, K., Oh, J., Yang, J., Conlan, S., Deming, C., Meyer, J.A., Schoenfeld, D., Nomicos, E., Park, M., Kong, H.H., et al. (2013). Topographic diversity of fungal and bacterial communities in human skin. Nature *498*, 367–370.

Fischer, M., and Ehlers, M. (2008). Toll-like Receptors in Autoimmunity. Ann. N. Y. Acad. Sci. *1143*, 21–34.

Fisher, R.A. (1936). The use of multiple measurements in taxonomic problems. Ann. Eugen. *7*, 179–188.

Fitz-Gibbon, S., Tomida, S., Chiu, B.-H., Nguyen, L., Du, C., Liu, M., Elashoff, D., Erfe, M.C., Loncaric, A., Kim, J., et al. (2013). Propionibacterium acnes strain populations in the human skin microbiome associated with acne. J. Invest. Dermatol. *133*, 2152–2160.

Flores, G.E., Caporaso, J., Henley, J.B., Rideout, J., Domogala, D., Chase, J., Leff, J.W., Vázquez-Baeza, Y., Gonzalez, A., Knight, R., et al. (2014). Temporal variability is a personalized feature of the human microbiome. Genome Biol. *15*, 531.

Fouhy, F., Clooney, A.G., Stanton, C., Claesson, M.J., and Cotter, P.D. (2016). 16S rRNA gene sequencing of mock microbial populations- impact of DNA extraction method, primer choice and sequencing platform. BMC Microbiol. *16*, 123.

Foulongne, V., Sauvage, V., Hebert, C., Dereure, O., Cheval, J., Gouilh, M.A., Pariente, K., Segondy, M., Burguière, A., Manuguerra, J.C., et al. (2012). Human skin microbiota: high diversity of DNA viruses identified on the human skin by high throughput sequencing. PLoS One *7*, e38499.

Fry, L., Baker, B.S., Powles, A. V., and Engstrand, L. (2015). Psoriasis is not an autoimmune disease? Exp. Dermatol. *24*, 241–244.

Fuentes, S., van Nood, E., Tims, S., Heikamp-de Jong, I., ter Braak, C.J., Keller, J.J., Zoetendal, E.G., and de Vos, W.M. (2014). Reset of a critically disturbed microbial ecosystem: faecal transplant in recurrent Clostridium difficile infection. ISME J. *8,* 1621–1633.

Gallo, R.L., and Nakatsuji, T. (2011). Microbial symbiosis with the innate immune defense system of the skin. J. Invest. Dermatol. *131*, 1974–1980.

Ganju, P., Nagpal, S., Mohammed, M., Nishal Kumar, P., Pandey, R., Natarajan, V.T., Mande, S.S., and Gokhale, R.S. (2016). Microbial community profiling shows dysbiosis in the lesional skin of Vitiligo subjects. Sci. Rep. *6*, 18761.

Gao, Z., Tseng, C., Strober, B.E., Pei, Z., and Blaser, M.J. (2008). Substantial alterations of the cutaneous bacterial biota in psoriatic lesions. PLoS One *3*, e2719.

Garalde, D.R., Snell, E.A., Jachimowicz, D., Heron, A.J., Bruce, M., Lloyd, J., Warland, A., Pantic, N., Admassu, T., Ciccone, J., et al. (2016). Highly parallel direct RNA sequencing on an array of nanopores. bioRxiv.

Gay, N.J., and Gangloff, M. (2007). Structure and Function of Toll Receptors and Their Ligands. Annu. Rev. Biochem. *76*, 141–165.

Georgel, P., Macquin, C., and Bahram, S. (2009). The heterogeneous allelic repertoire of human toll-like receptor (TLR) genes. PLoS One *4*, e7803.

Gevers, D., Cohan, F.M., Lawrence, J.G., Spratt, B.G., Coenye, T., Feil, E.J., Stackebrandt, E., de Peer, Y. Van, Vandamme, P., Thompson, F.L., et al. (2005). Opinion: Re-evaluating prokaryotic species. Nat. Rev. Microbiol. *3*, 733–739.

Glenn, T.C. (2011). Field guide to next-generation DNA sequencers. Mol. Ecol. Resour. *11*, 759–769.

Gómez-Gallego, C., Pohl, S., Salminen, S., De Vos, W.M., and Kneifel, W. (2016). Akkermansia muciniphila: a novel functional microbe with probiotic properties. Benef. Microbes *7*, 571–584.

Gontcharova, V., Youn, E., Sun, Y., Wolcott, R.D., and Dowd, S.E. (2010). A comparison of bacterial composition in diabetic ulcers and contralateral intact skin. Open Microbiol. J. *4*, 8–19.

Goodwin, S., Gurtowski, J., Ethe-Sayers, S., Deshpande, P., Schatz, M.C., and McCombie, W.R. (2015). Oxford Nanopore sequencing, hybrid error correction, and de novo assembly of a eukaryotic genome. Genome Res. *25*, 1750–1756.

Goodwin, S., McPherson, J.D., and McCombie, W.R. (2016). Coming of age: ten years of next-generation sequencing technologies. Nat. Rev. Genet. *17*, 333–351.

Gough, E., Shaikh, H., and Manges, A.R. (2011). Systematic Review of Intestinal Microbiota Transplantation (Fecal Bacteriotherapy) for Recurrent Clostridium difficile Infection. Clin. Infect. Dis. *53*, 994–1002.

Grice, E.A. (2014). The skin microbiome: potential for novel diagnostic and therapeutic approaches to cutaneous disease. Semin. Cutan. Med. Surg. *33*, 98–103.

Grice, E.A. (2015). The intersection of microbiome and host at the skin interface : genomic- and metagenomic-based insights. Genome Res. *25,* 1514–1520.

Grice, E.A., and Segre, J.A. (2011). The skin microbiome. Nat. Rev. Microbiol. *9*, 244–253.

Grice, E.A., and Segre, J.A. (2012). The Human Microbiome: Our Second Genome. Annu. Rev. Genomics Hum. Genet. *13*, 151–170.

Grice, E.A., Kong, H.H., Conlan, S., Deming, C.B., Davis, J., Young, A.C., Bouffard, G.G., Blakesley, R.W., Murray, P.R., Green, E.D., et al. (2009a). Topographical and temporal diversity of the human skin microbiome. Science *324*, 1190–1192.

Grice, E.A., Kong, H.H., Conlan, S., Deming, C.B., Davis, J., Young, A.C., NISC Comparative Sequencing Program, Bouffard, G.G., Blakesley, R.W., Murray, P.R., et al. (2009b). Topographical and temporal diversity of the human skin. Science *324*, 1190–1192.

Grice, E.A., Kong, H.H., Renaud, G., Young, A.C., NISC Comparative Sequencing Program, Bouffard, G.G., Blakesley, R.W., Wolfsberg, T.G., Turner, M.L., and Segre, J.A. (2008). A diversity profile of the human skin microbiota. Genome Res. *18*, 1043–1050.

Guéniche, A., Bastien, P., Ovigne, J.M., Kermici, M., Courchay, G., Chevalier, V., Breton, L., and Castiel-Higounenc, I. (2009). Bifidobacterium longum lysate, a new ingredient for reactive skin. Exp. Dermatol. *19*, e1–e8.

Haas, B.J., Gevers, D., Earl, A.M., Feldgarden, M., Ward, D. V, Giannoukos, G., Ciulla, D., Tabbaa, D., Highlander, S.K., Desantis, T.Z., et al. (2011). Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. Genome Res. *21*, 494–504.

Halwachs, B., Madhusudhan, N., Krause, R., Nilsson, R.H., Moissl-Eichinger, C., Högenauer, C., Thallinger, G.G., and Gorkiewicz, G. (2017). Critical Issues in Mycobiota Analysis. Front. Microbiol. *8,* 180.

Hammer, T.J., Janzen, D.H., Hallwachs, W., Jaffe, S.L., and Fierer, N. (2017). Caterpillars lack a resident gut microbiome. bioRxiv.

Hampton-Marcell, J.T., Lopez, J. V., and Gilbert, J.A. (2017). The human microbiome: an emerging tool in forensics. Microb. Biotechnol. *10*, 228–230.

Handley, S.A. (2016). The virome: a missing component of biological interaction networks in health and disease. Genome Med. *8*, 32.

Hannigan, G.D., Hodkinson, B.P., McGinnis, K., Tyldsley, A.S., Anari, J.B., Horan, A.D., Grice, E.A., and Mehta, S. (2014). Culture-independent pilot study of microbiota colonizing open fractures and association with severity, mechanism, location, and complication from presentation to early outpatient follow-up. J. Orthop. Res. *32*, 597–605.

Hannigan, G.D., Meisel, J.S., Tyldsley, A.S., Zheng, Q., Hodkinson, B.P., Sanmiguel, A.J., Minot, S., Bushman, F.D., and Grice, A. (2015). The Human Skin Double-Stranded DNA Virome : Topographical and Temporal Diversity , Genetic Enrichment , and Dynamic Associations with the Host Microbiome. *6*, 1–13.

Hanski, I., von Hertzen, L., Fyhrquist, N., Koskinen, K., Torppa, K., Laatikainen, T., Karisola, P., Auvinen, P., Paulin, L., Makela, M.J., et al. (2012). Environmental biodiversity, human microbiota, and allergy are interrelated. Proc. Natl. Acad. Sci. *109*, 8334–8339.

Harris, E.E., and Meyer, D. (2006). The molecular signature of selection underlying human adaptations. Am. J. Phys. Anthropol. *131*, 89–130.

He, Y., Caporaso, J.G., Jiang, X.-T., Sheng, H.-F., Huse, S.M., Rideout, J.R., Edgar, R.C., Kopylova, E., Walters, W.A., Knight, R., et al. (2015). Stability of operational taxonomic units: an important but neglected property for analyzing microbial diversity. Microbiome *3*, 20.

Heather, J.M., and Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. Genomics *107*, 1–8.

Hill, J.E., Penny, S.L., Crowell, K.G., Goh, S.H., and Hemmingsen, S.M. (2004). cpnDB: a chaperonin sequence database. Genome Res. *14*, 1669–1675.

Hilmi, D., Parcina, M., Stollewerk, D., Ostrop, J., Josten, M., Meilaender, A., Zaehringer, U., Wichelhaus, T.A., Bierbaum, G., Heeg, K., et al. (2014). Heterogeneity of Host TLR2 Stimulation by Staphylocoocus aureus Isolates. PLoS One *9*, e96416.

Hoang, B.X., Shaw, G., Pham, P., and Levine, S.A. (2010). Lactobacillus rhamnosus cell lysate in the management of resistant childhood atopic eczema. Inflamm. Allergy Drug Targets *9*, 192–196.

Holmes, C.J., Plichta, J.K., Gamelli, R.L., and Radek, K.A. (2015). Dynamic Role of Host Stress Responses in Modulating the Cutaneous Microbiome: Implications for Wound Healing and Infection. Adv. Wound Care *4*, 24–37.

Hosein, S., Rodríguez-Cortés, A., Blake, D.P., Allenspach, K., Alberola, J., and Solano-Gallego, L. (2015). Transcription of Toll-Like Receptors 2, 3, 4 and 9, FoxP3 and Th17 Cytokines in a Susceptible Experimental Model of Canine Leishmania infantum Infection. PLoS One *10*, e0140325.

Hospodsky, D., Pickering, A.J., Julian, T.R., Miller, D., Gorthala, S., Boehm, A.B., and Peccia, J. (2014). Hand bacterial communities vary across two different human populations. Microbiology *160*, 1144–1152.

Hughes, J.B., and Hellmann, J.J. (2005). The Application of Rarefaction Techniques to Molecular Inventories of Microbial Diversity. In Methods in Enzymology *397*, 292-308

Huizenga, T., Kado, J., Mehregan, D.R., and Diamond, S. (2015). Identifying Toll-like Receptor Expression in Cutaneous Sarcoidosis. Am. J. Dermatopathol. *37*, 67–72.

The Human Microbiome Project Consortium. (2012a). A framework for human microbiome research. Nature *486*, 215–221.

The Human Microbiome Project Consortium. (2012b). Structure, function and diversity of the healthy human microbiome. Nature *486*, 207–214.

Huseini, H.F., Rahimzadeh, G., Fazeli, M.R., Mehrazma, M., Salehi, M., Hooman, Y., and al., et (2012). Evaluation of wound healing activities of kefir products. Burns *38*, 719–723.

Ip, C.L.C., Loose, M., Tyson, J.R., de Cesare, M., Brown, B.L., Jain, M., Leggett, R.M., Eccles, D.A., Zalunin, V., Urban, J.M., et al. (2015). MinION Analysis and Reference Consortium: Phase 1 data release and analysis. F1000Research *4*, 1075.

Janardhanam, S.B., Prakasam, S., Swaminathan, V.T., Kodumudi, K.N., Zunt, S.L., and Srinivasan, M. (2012). Differential expression of TLR-2 and TLR-4 in the epithelial cells in oral lichen planus. Arch. Oral Biol. *57*, 495–502.

Janda, J.M., and Abbott, S.L. (2007). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. J. Clin. Microbiol. *45*, 2761–2764.

Jo, J.H., Deming, C., Kennedy, E.A., Conlan, S., Polley, E.C., Ng, W., Segre, J.A., Kong, H.H., and Kong, H.H. (2016). Diverse Human Skin Fungal Communities in Children Converge in Adulthood. J. Invest. Dermatol. *136*, 2356–2363.

Jumpstart Consortium Human Microbiome Project Data Generation Working Group (2012). Evaluation of 16S rDNA-based community profiling for human microbiome research. PLoS One *7*, e39315.

Jun, S.-R., Robeson, M.S., Hauser, L.J., Schadt, C.W., and Gorin, A.A. (2015). PanFP: pangenome-based functional profiles for microbial communities. BMC Res Notes *8*, 479.

Kang, J.Y., and Lee, J.O. (2011). Structural Biology of the Toll-Like Receptor Family. Annu. Rev. Biochem. *80*, 917–941.

Karst, S.M., Dueholm, M.S., McIlroy, S.J., Kirkegaard, R.H., Nielsen, P.H., and Albertsen, M. (2016). Thousands of primer-free, high-quality, full-length SSU rRNA sequences from all domains of life. bioRxiv.

Kasperkovitz, P.V., Khan, N.S., Tam, J.M., Mansour, M.K., Davids, P.J., and Vyas, J.M. (2011). Toll-Like Receptor 9 Modulates Macrophage Antifungal Effector Function during Innate Recognition of Candida albicans and Saccharomyces cerevisiae. Infect. Immun. *79*, 4858–4867.

Kathrani, A., House, A., Catchpole, B., Murphy, A., Werling, D., and Allenspach, K. (2011). Breed-independent toll-like receptor 5 polymorphisms show association with canine inflammatory bowel disease. Tissue Antigens *78*, 94–101.

Kathrani, A., Holder, A., Catchpole, B., Alvarez, L., Simpson, K., Werling, D., and Allenspach, K. (2012). TLR5 risk-associated haplotype for canine inflammatory bowel disease confers hyper-responsiveness to flagellin. PLoS One *7*, e30117.

Kathrani, A., House, A., Catchpole, B., Murphy, A., German, A., Werling, D., and Allenspach, K. (2010). Polymorphisms in the TLR4 and TLR5 gene are significantly associated with inflammatory bowel disease in German shepherd dogs. PLoS One *5*, e15740.

Kawai, T., and Akira, S. (2009). The roles of TLRs, RLRs and NLRs in pathogen recognition. Int. Immunol. *21*, 317–337.

Kawai, T., and Akira, S. (2010). The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. Nat. Immunol. *11*, 373–384.

Kelly, B.J., Gross, R., Bittinger, K., Sherrill-Mix, S., Lewis, J.D., Collman, R.G., Bushman, F.D., and Li, H. (2015). Power and sample-size estimation for microbiome studies using pairwise distances and PERMANOVA. Bioinformatics *31*, 2461–2468.

Kennedy, E.A., Connolly, J., Hourihane, J.O., Fallon, P.G., McLean, W.H.I., Murray, D., Jo, J.H., Segre, J.A., Kong, H.H., and Irvine, A.D. (2017). Skin microbiome before development of atopic dermatitis: Early colonization with commensal staphylococci at 2 months is associated with a lower risk of atopic dermatitis at 1 year. J. Allergy Clin. Immunol. *139*, 166–172.

Kettleson, E.M., Adhikari, A., Vesper, S., Coombs, K., Indugula, R., and Reponen, T. (2015). Key determinants of the fungal and bacterial microbiomes in homes. Environ. Res. *138*, 130–135.

Kim, J., Ochoa, M.-T., Krutzik, S.R., Takeuchi, O., Uematsu, S., Legaspi, A.J., Brightbill, H.D., Holland, D., Cunliffe, W.J., Akira, S., et al. (2002). Activation of toll-like receptor 2 in acne triggers inflammatory cytokine responses. J. Immunol. *169*, 1535–1541.

Klappenbach, J.A., Saxman, P.R., Cole, J.R., and Schmidt, T.M. (2001). rrndb: the Ribosomal RNA Operon Copy Number Database. Nucleic Acids Res. *29*, 181–184.

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., and Glockner, F.O. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res. *41*, e1–e1.

Knights, D., Silverberg, M.S., Weersma, R.K., Gevers, D., Dijkstra, G., Huang, H., Tyler, A.D., van Sommeren, S., Imhann, F., Stempak, J.M., et al. (2014). Complex host genetics influence the microbiome in inflammatory bowel disease. Genome Med. *6*, 107.

Kong, H.H., Andersson, B., Clavel, T., Common, J.E., Jackson, S.A., Olson, N.D., Segre, J.A., and Traidl-Hoffmann, C. (2017). Performing Skin Microbiome Research: A Method to the Madness. J. Invest. Dermatol. *137*, 561–568.

Kong, H.H., Oh, J., Deming, C., Conlan, S., Grice, E.A., Beatson, M.A., Nomicos, E., Polley, E.C., Komarow, H.D., Murray, P.R., et al. (2012). Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis. Genome Res. *22*, 850–859.

Kopylova, E., Navas-Molina, J.A., Mercier, C., Xu, Z.Z., Mahé, F., He, Y., Zhou, H. W., Rognes, T., Caporaso, J.G., and Knight, R. (2016). Open-Source Sequence Clustering Methods Improve the State Of the Art. mSystems *1*, e00003-15.

Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., and Phillippy, A.M. (2017). Canu: scalable and accurate long-read assembly via adaptive k -mer weighting and repeat separation. Genome Res. *27*, 722–736.

Kruskal, W.H., and Wallis, W.A. (1952). Use of Ranks in One-Criterion Variance Analysis. J. Am. Stat. Assoc. *47*, 583.

Krutmann, J. (2009). Pre- and probiotics for human skin. J. Dermatol. Sci. *54*, 1–5.

Krutzik, S.R., Ochoa, M.T., Sieling, P.A., Uematsu, S., Ng, Y.W., Legaspi, A., Liu, P.T., Cole, S.T., Godowski, P.J., Maeda, Y., et al. (2003). Activation and regulation of Toll-like receptors 2 and 1 in human leprosy. Nat. Med. *9*, 525–532.

Kubinak, J.L., and Round, J.L. (2012). Toll-like receptors promote mutually beneficial commensal-host interactions. PLoS Pathog. *8*, 1.

Kuczynski, J., Lauber, C.L., Walters, W. a., Parfrey, L.W., Clemente, J.C., Gevers, D., and Knight, R. (2011). Experimental and analytical tools for studying the human microbiome. Nat. Rev. Genet. *13*, 47–58.

Kumar, H., Kawai, T., and Akira, S. (2009). Pathogen recognition in the innate immune response. Biochem. J. *420*, 1–16.

Kumar, H., Kawai, T., and Akira, S. (2011). Pathogen recognition by the innate immune system. Int. Rev. Immunol. *30*, 16–34.

Kumar, S., Ingle, H., Prasad, D.V.R., and Kumar, H. (2013). Recognition of bacterial infection by innate immune sensors. Crit. Rev. Microbiol. *39*, 229–246.

Kumar, V., and Sharma, A. (2010). Mast cells: Emerging sentinel innate immune cells with diverse role in immunity. Mol. Immunol. *48*, 14–25.

Kuroki, K., Stoker, A.M., Sims, H.J., and Cook, J.L. (2010). Expression of Toll-like receptors 2 and 4 in stifle joint synovial tissues of dogs with or without osteoarthritis. Am. J. Vet. Res. *71*, 750–754.

La Rosa, P.S., Brooks, J.P., Deych, E., Boone, E.L., Edwards, D.J., Wang, Q., Sodergren, E., Weinstock, G., and Shannon, W.D. (2012). Hypothesis Testing and Power Calculations for Taxonomic-Based Human Microbiome Data. PLoS One *7,* e52078.

Lai, Y., Di Nardo, A., Nakatsuji, T., Leichtle, A., Yang, Y., Cogen, A.L., Wu, Z.-R., Hooper, L. V, Schmidt, R.R., von Aulock, S., et al. (2009). Commensal bacteria regulate Toll-like receptor 3-dependent inflammation after skin injury. Nat. Med. *15*, 1377–1382.

Langille, M.G.I., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J. a, Clemente, J.C., Burkepile, D.E., Vega Thurber, R.L., Knight, R., et al. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat. Biotechnol. *31*, 814–821.

Lauber, C.L., Zhou, N., Gordon, J.I., Knight, R., and Fierer, N. (2010). Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples. FEMS Microbiol. Lett. *307*, 80–86.

Lax, S., Hampton-Marcell, J.T., Gibbons, S.M., Colares, G.B., Smith, D., Eisen, J.A., and Gilbert, J.A. (2015). Forensic analysis of the microbiome of phones and shoes. Microbiome *3*, 21.

Leber, A., Viladomiu, M., Hontecillas, R., Abedi, V., Philipson, C., Hoops, S., Howard, B., and Bassaganya-Riera, J. (2015). Systems Modeling of Interactions between Mucosal Immunity and the Gut Microbiome during Clostridium difficile Infection. PLoS One *10*, e0134849.

Lehtimäki, J., Karkman, A., Laatikainen, T., Paalanen, L., von Hertzen, L., Haahtela, T., Hanski, I., and Ruokolainen, L. (2017). Patterns in the skin microbiota differ in children and teenagers between rural and urban environments. Sci. Rep. *7*, 45651.

Leifer, C.A., McConkey, C., Li, S., Chassaing, B., Gewirtz, A.T., and Ley, R.E. (2014). Linking genetic variation in human Toll-like receptor 5 genes to the gut microbiome's potential to cause inflammation. Immunol. Lett. *162*, 3–9.

Leigh, I.M., and Watt, F.M. (1994). The keratinocyte handbook (New York: Cambridge University Press).

Lesiak, A., Smolewski, P., Sobolewska-Sztychny, D., Sysa-Jedrzejowska, A., and Narbutt, J. (2012). The Role of T-Regulatory Cells and Toll-Like Receptors 2 and 4 In Atopic Dermatitis. Scand. J. Immunol. *76*, 405–410.

Leung, M.H.Y., Wilkins, D., and Lee, P.K.H. (2015). Insights into the pan-microbiome: skin microbial communities of Chinese individuals differ from other racial groups. Sci. Rep. *5*, 11845.

Levene, M.J., Korlach, J., Turner, S.W., Foquet, M., Craighead, H.G., and Webb, W.W. (2003). Zero-Mode Waveguides for Single-Molecule Analysis at High Concentrations. Science *299*, 682–686.

Lew, L.C., and Liong, M.T. (2013). Bioactives from probiotics for dermal health: functions and benefits. J. Appl. Microbiol. *114*, 1241–1253.

Li, C., Chng, K.R., Boey, E.J.H., Ng, A.H.Q., Wilm, A., and Nagarajan, N. (2016). INC-Seq: accurate single molecule reads using nanopore sequencing. Gigascience *5*, 34.

Lindblad-Toh, K., Wade, C.M., Mikkelsen, T.S., Karlsson, E.K., Jaffe, D.B., Kamal, M., Clamp, M., Chang, J.L., Kulbokas, E.J., Zody, M.C., et al. (2005). Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature *438*, 803–819.

Links, M.G., Dumonceaux, T.J., Hemmingsen, S.M., Hill, J.E., and Dalin, E. (2012). The Chaperonin-60 Universal Target Is a Barcode for Bacteria That Enables De Novo Assembly of Metagenomic Sequence Data. PLoS One *7*, e49755.

Loman, N.J., Misra, R. V, Dallman, T.J., Constantinidou, C., Gharbia, S.E., Wain, J., and Pallen, M.J. (2012). Performance comparison of benchtop high-throughput sequencing platforms. Nat. Biotechnol. *30*, 434–439.

Loman, N.J., Quick, J., and Simpson, J.T. (2015). A complete bacterial genome assembled de novo using only nanopore sequencing data. Nat. Methods *12*, 733–735.

Lozupone, C., and Knight, R. (2005). UniFrac : a New Phylogenetic Method for Comparing Microbial Communities. Appl. Environ. Microbiol. *71*, 8228–8235.

Lozupone, C., Lladser, M.E., Knights, D., Stombaugh, J., and Knight, R. (2011). UniFrac: an effective distance metric for microbial community comparison. ISME J. *5*, 169–172.

Lozupone, C.A., and Knight, R. (2007). Global patterns in bacterial diversity. Proc. Natl. Acad. Sci. U. S. A. *104*, 11436–11440.

Lozupone, C.A., and Knight, R. (2008). Species divergence and the measurement of microbial diversity. FEMS Microbiol. Rev. *32*, 557–578.

Lozupone, C.A., Hamady, M., Kelley, S.T., and Knight, R. (2007). Quantitative and Qualitative Diversity Measures Lead to Different Insights into Factors That Structure Microbial Communities. Appl. Environ. Microbiol. *73*, 1576–1585.

Lund, P.A. (2009). Multiple chaperonins in bacteria – why so many? FEMS Microbiol. Rev. *33*, 785–800.

Luo, C., Rodriguez-R, L.M., and Konstantinidis, K.T. (2013). A User's Guide to Quantitative and Comparative Analysis of Metagenomic Datasets. In Methods in Enzymology, pp. 525–547.

Mao, D.P., Zhou, Q., Chen, C.-Y., and Quan, Z.X. (2012). Coverage evaluation of universal bacterial primers using the metagenomic datasets. BMC Microbiol. *12*, 66.

Marchesi, J.R., and Ravel, J. (2015). The vocabulary of microbiome research: a proposal. Microbiome *3*, 31.

Mardis, E.R. (2017). DNA sequencing technologies: 2006-2016. Nat. Protoc. *12*, 213–218.

Matin, N., Tabatabaie, O., Mohammadinejad, P., and Rezaei, N. (2015). Therapeutic targeting of Toll-like receptors in cutaneous disorders. Expert Opin. Ther. Targets *19*, 1651–1663.

McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A., Andersen, G.L., Knight, R., and Hugenholtz, P. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. *6*, 610–618.

McMahon, L.A., House, A.K., Catchpole, B., Elson-Riggins, J., Riddle, A., Smith, K., Werling, D., Burgener, I.A., and Allenspach, K. (2010). Expression of Toll-like receptor 2 in duodenal biopsies from dogs with inflammatory bowel disease is associated with severity of disease. Vet. Immunol. Immunopathol. *135*, 158–163.

Meadow, J.F., Altrichter, A.E., and Green, J.L. (2014). Mobile phones carry the personal microbiome of their owners. PeerJ *2*, e447.

Meason-Smith, C., Diesel, A., Patterson, A.P., Older, C.E., Johnson, T.J., Mansell, J.M., Suchodolski, J.S., and Rodrigues Hoffmann, A. (2017). Characterization of the cutaneous mycobiota in healthy and allergic cats using next generation sequencing. Vet. Dermatol. *28*, 71–e17.

Meason-smith, C., Diesel, A., Patterson, A.P., Older, C.E., Mansell, J.M., Suchodolski, J.S., and Hoffmann, A.R. (2015). What is living on your dog 's skin ? Characterization of the canine cutaneous mycobiota and fungal dysbiosis in canine allergic dermatitis. FEMS Microbiol. Ecol. *91*, 1–12.

Meisel, J.S., Hannigan, G.D., Tyldsley, A.S., Sanmiguel, A.J., Hodkinson, B.P., Zheng, Q., and Grice, E.A. (2016). Skin Microbiome Surveys Are Strongly Influenced by Experimental Design. J. Invest. Dermatol. *136*, 947–956.

Melo, G.D., Silva, J.E.S., Grano, F.G., Homem, C.G., and Machado, G.F. (2014). Compartmentalized gene expression of toll-like receptors 2, 4 and 9 in the brain and peripheral lymphoid organs during canine visceral leishmaniasis. Parasite Immunol. *36*, 726–731.

Mercier, E., Peters, I.R., Day, M.J., Clercx, C., and Peeters, D. (2012). Toll- and NOD-like receptor mRNA expression in canine sino-nasal aspergillosis and idiopathic lymphoplasmacytic rhinitis. Vet. Immunol. Immunopathol. *145*, 618–624.

Miletto, M., and Lindow, S.E. (2015). Relative and contextual contribution of different sources to the composition and abundance of indoor air bacteria in residences. Microbiome *3*, 61.

Miller, W.H., Griffin, C.E., and Campbell, K.L. (2013). Muller & Kirk's Small Animal Dermatology (St Louis: Saunders).

Misic, A.M., Davis, M.F., Tyldsley, A.S., Hodkinson, B.P., Tolomeo, P., Hu, B., Nachamkin, I., Lautenbach, E., Morris, D.O., and Grice, E.A. (2015). The shared microbiota of humans and companion animals as evaluated from Staphylococcus carriage sites. Microbiome *3*, 2.

Mizrahi-Man, O., Davenport, E.R., and Gilad, Y. (2013). Taxonomic Classification of Bacterial 16S rRNA Genes Using Short Sequencing Reads: Evaluation of Effective Study Designs. PLoS One *8*, 18–23.

Morey, M., Fernández-Marmiesse, A., Castiñeiras, D., Fraga, J.M., Couce, M.L., and Cocho, J.A. (2013). A glimpse into past, present, and future DNA sequencing. Mol. Genet. Metab. *110*, 3–24.

Muehleisen, B., Jiang, S.B., Gladsjo, J.A., Gerber, M., Hata, T., and Gallo, R.L. (2012). Distinct Innate Immune Gene Expression Profiles in Non-Melanoma Skin Cancer of Immunocompetent and Immunosuppressed Patients. PLoS One *7*, e40754.

Mukherjee, S., Ganguli, D., and Majumder, P.P. (2014). Global Footprints of Purifying Selection on Toll-Like Receptor Genes Primarily Associated with Response to Bacterial Infections in Humans. Genome Biol. Evol. *6,* 551–558.

Mukherjee, S., Sarkar-Roy, N., Wagener, D.K., and Majumder, P.P. (2009). Signatures of natural selection are not uniform across genes of innate immune system, but purifying selection is the dominant signature. Proc. Natl. Acad. Sci. *106*, 7073–7078.

Myles, I.A., Williams, K.W., Reckhow, J.D., Jammeh, M.L., Pincus, N.B., Sastalla, I., Saleem, D., Stone, K.D., and Datta, S.K. (2016). Transplantation of human skin microbiota in models of atopic dermatitis. JCI Insight *1*, e86955.

Naik, S., Bouladoux, N., Linehan, J.L., Han, S.-J., Harrison, O.J., Wilhelm, C., Conlan, S., Himmelfarb, S., Byrd, A.L., Deming, C., et al. (2015). Commensal-dendritic-cell interaction specifies a unique protective skin immune signature. Nature *520*, 104–108.

Naik, S., Bouladoux, N., Wilhelm, C., Molloy, M.J., Salcedo, R., Kastenmuller, W., Deming, C., Quinones, M., Koo, L., Conlan, S., et al. (2012). Compartmentalized Control of Skin Immunity by Resident Commensals. Science *337*, 1115–1119.

Nakamizo, S., Egawa, G., Honda, T., Nakajima, S., Belkaid, Y., and Kabashima, K. (2015). Commensal bacteria and cutaneous immunity. Semin. Immunopathol. *37*, 73–80.

Nakatsuji, T., Chiang, H.-I., Jiang, S.B., Nagarajan, H., Zengler, K., and Gallo, R.L. (2013). The microbiome extends to subepidermal compartments of normal skin. Nat. Commun. *4*, 1431.

Nestle, F.O., Di Meglio, P., Qin, J.-Z., and Nickoloff, B.J. (2009). Skin immune sentinels in health and disease. Nat. Rev. Immunol. *9*, 679–691.

Netea, M.G., Wijmenga, C., and O'Neill, L. a J. (2012). Genetic variation in Toll-like receptors and disease susceptibility. Nat. Immunol. *13*, 535–542.

O'Neill, L.A.J., Golenbock, D., and Bowie, A.G. (2013). The history of Toll-like receptors — redefining innate immunity. Nat. Rev. Immunol. *13*, 453–460.

Oh, J., Byrd, A.L., Deming, C., Conlan, S., Barnabas, B., Blakesley, R., Bouffard, G., Brooks, S., Coleman, H., Dekhtyar, M., et al. (2014). Biogeography and individuality shape function in the human skin metagenome. Nature *514*, 59–64.

Oh, J., Byrd, A.L., Park, M., Kong, H.H., and Segre, J.A. (2016). Temporal Stability of the Human Skin Microbiome. Cell *165*, 854–866.

Oh, J., Conlan, S., Polley, E., Segre, J. a, and Kong, H.H. (2012). Shifts in human skin and nares microbiota of healthy children and adults. Genome Med. *4*, 77.

Oh, J., Freeman, A.F., Park, M., Sokolic, R., Candotti, F., Holland, S.M., Segre, J.A., and Kong, H.H. (2013). The altered landscape of the human skin microbiome in patients with primary immunodeficiencies. Genome Res. *23*, 2103–2114.

Olasagasti, F., Lieberman, K.R., Benner, S., Cherf, G.M., Dahl, J.M., Deamer, D.W., and Akeson, M. (2010). Replication of individual DNA molecules under electronic control using a protein nanopore. Nat. Nanotechnol. *5*, 798–806.

Olivry, T., and Bizikova, P. (2013). A systematic review of randomized controlled trials for prevention or treatment of atopic dermatitis in dogs: 2008-2011 update. Vet. Dermatol. *24*, 97–e26.

Pasparakis, M., Haase, I., and Nestle, F.O. (2014). Mechanisms regulating skin immunity and inflammation. Nat Rev Immunol *14*, 289–301.

Patin, N. V., Kunin, V., Lidström, U., and Ashby, M.N. (2013). Effects of OTU Clustering and PCR Artifacts on Microbial Diversity Estimates. Microb. Ecol. *65*, 709–719.

Paulino, L.C., Tseng, C. H., Strober, B.E., and Blaser, M.J. (2006). Molecular Analysis of Fungal Microbiota in Samples from Healthy Human Skin and Psoriatic Lesions. J. Clin. Microbiol. *44*, 2933–2941.

Peleg, A.Y., Hogan, D.A., and Mylonakis, E. (2010). Medically important bacterial–fungal interactions. Nat. Rev. Microbiol. *8*, 340–349.

Perez Perez, G.I., Gao, Z., Jourdain, R., Ramirez, J., Gany, F., Clavaud, C., Demaude, J., Breton, L., and Blaser, M.J. (2016). Body Site Is a More Determinant Factor than Human Population Diversity in the Healthy Skin Microbiome. PLoS One *11*, e0151990.

Pierezan, F., Olivry, T., Paps, J.S., Lawhon, S.D., Wu, J., Steiner, J.M., Suchodolski, J.S., and Hoffmann, A.R. (2016). The skin microbiome in allergen-induced canine atopic dermatitis. Vet Dermatol. 2016 *27*, 332-e82

Potaczek, D., Nastalek, M., Okumura, K., Wojas-Pelc, A., Undas, A., and Nishiyama, C. (2011). An association of TLR2-16934A>T polymorphism and severity/phenotype of atopic dermatitis. J. Eur. Acad. Dermatology Venereol. *25*, 715–721.

Price, L.B., Liu, C.M., Melendez, J.H., Frankel, Y.M., Engelthaler, D., Aziz, M., Bowers, J., Rattray, R., Ravel, J., Kingsley, C., et al. (2009). Community analysis of chronic wound bacteria using 16S rRNA gene-based pyrosequencing: impact of diabetes and antibiotics on chronic wound microbiota. PLoS One *4*, e6462.

Quach, H., Wilson, D., Laval, G., Patin, E., Manry, J., Guibert, J., Barreiro, L.B., Nerrienet, E., Verschoor, E., Gessain, A., et al. (2013). Different selective pressures shape the evolution of Toll-like receptors in human and African great ape populations. Hum. Mol. Genet. *22*, 4829–4840.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. *41*, D590–D596.

Quick, J., and Loman, N.J. Thar she blows! Ultra-long read methods for nanopore sequencing. Oral presentation at London Calling 2017 by Oxford Nanopore Technologies.

Quintana-Murci, L. (2016). Understanding rare and common diseases in the context of human evolution. Genome Biol. *17*, 225.

Rajendhran, J., and Gunasekaran, P. (2011). Microbial phylogeny and diversity: Small subunit ribosomal RNA sequence analysis and beyond. Microbiol. Res. *166*, 99–110.

Reid, G., Younes, J.A., Van der Mei, H.C., Gloor, G.B., Knight, R., and Busscher, H.J. (2011). Microbiota restoration: natural and supplemented recovery of human microbial communities. Nat. Rev. Microbiol. *9*, 27–38.

Rensburg, J.J. Van, Lin, H., Gao, X., Toh, E., Fortney, K.R., Ellinger, S., Zwickl, B., Janowicz, D.M., Katz, B.P., Nelson, D.E., et al. (2015). The Human Skin Microbiome Associates with the Outcome of and Is Influenced by Bacterial Infection. *6,* 1–13.

Reuter, J.A., Spacek, D. V, and Snyder, M.P. (2015). High-throughput sequencing technologies. Mol. Cell *58*, 586–597.

Richter, M., and Rossello-Mora, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. Proc. Natl. Acad. Sci. *106*, 19126–19131.

Rideout, J.R., He, Y., Navas-Molina, J.A., Walters, W.A., Ursell, L.K., Gibbons, S.M., Chase, J., McDonald, D., Gonzalez, A., Robbins-Pianka, A., et al. (2014). Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. PeerJ *2*, e545.

Riggio, M.P., Lappin, D.F., and Bennett, D. (2014). Bacteria and Toll-like receptor and cytokine mRNA expression profiles associated with canine arthritis. Vet. Immunol. Immunopathol. *160*, 158–166.

Roach, J.C., Glusman, G., Rowen, L., Kaur, A., Purcell, M.K., Smith, K.D., Hood, L.E., and Aderem, A. (2005). The evolution of vertebrate Toll-like receptors. Proc. Natl. Acad. Sci. U. S. A. *102*, 9577–9582.

Rodrigues Hoffmann, A., Patterson, A.P., Diesel, A., Lawhon, S.D., Ly, H.J., Elkins Stephenson, C., Mansell, J., Steiner, J.M., Dowd, S.E., Olivry, T., et al. (2014). The skin microbiome in healthy and allergic dogs. PLoS One *9,* e83197.

Rogers, G.B., and Bruce, K.D. (2010). Next-Generation Sequencing in the Analysis of Human Microbiota. Mol. Diagn. Ther. *14*, 343–350.

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. PeerJ *4*, e2584.

Rossi-Tamisier, M., Benamar, S., Raoult, D., and Fournier, P.-E. (2015). Cautionary tale of using 16S rRNA gene sequence similarity values in identification of human-associated bacterial species. Int. J. Syst. Evol. Microbiol. *65*, 1929–1934.

Rothberg, J.M., Hinz, W., Rearick, T.M., Schultz, J., Mileski, W., Davey, M., Leamon, J.H., Johnson, K., Milgrew, M.J., Edwards, M., et al. (2011). An integrated semiconductor device enabling non-optical genome sequencing. Nature *475*, 348–352.

Ruedisueli, F.L., Eastwood, N.J., Gunn, N.K., and Watson, T.D.G. (1998). The measurement of skin pH in normal dogs of different breeds. In Advances in Veterinary Dermatology III, K. Kwochka, T. Willemse, and C. Von Tscharner, eds. (Boston: Butterworth-Heinemann Ltd), p. 596.

Salem, S.A.M., Abu-Zeid, R.M., and Nada, O.H. (2013). Immunohistochemical study of toll-like receptors 1 and 2 expression in cutaneous lichen planus lesions. Arch. Dermatol. Res. *305*, 125–131.

Salter, S.J., Cox, M.J., Turek, E.M., Calus, S.T., Cookson, W.O., Moffatt, M.F., Turner, P., Parkhill, J., Loman, N.J., and Walker, A.W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. *12,* 87

Sanford, J. a., and Gallo, R.L. (2013). Functions of the skin microbiota in health and disease. Semin. Immunol. *25*, 370–377.

Sanmiguel, A., and Grice, E.A. (2015). Interactions between host factors and the skin microbiome. Cell. Mol. Life Sci. *72*, 1499–1515.

Santiago, A., Panda, S., Mengels, G., Martinez, X., Azpiroz, F., Dore, J., Guarner, F., and Manichanh, C. (2014). Processing faecal samples: a step forward for standards in microbial community analysis. BMC Microbiol. *14*, 112.

Schloss, P.D., and Handelsman, J. (2004). Status of the microbial census. Microbiol. Mol. Biol. Rev. *68*, 686–691.

Schloss, P.D., Gevers, D., and Westcott, S.L. (2011). Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PLoS One *6*, e27310.

Schloss, P.D., Girard, R.A., Martin, T., Edwards, J., and Thrash, J.C. (2016b). Status of the Archaeal and Bacterial Census: an Update. MBio *7*, e00201-16

Schloss, P.D., Jenior, M.L., Koumpouras, C.C., Westcott, S.L., and Highlander, S.K. (2016a). Sequencing 16S rRNA gene fragments using the PacBio SMRT DNA sequencing system. PeerJ *4*, e1869.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., et al. (2009). Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Appl. Environ. Microbiol. *75*, 7537–7541.

Schmidt, T.S.B., Matias Rodrigues, J.F., and von Mering, C. (2014). Ecological consistency of SSU rRNA-based operational taxonomic units at a global scale. PLoS Comput. Biol. *10*, e1003594.

Schmidt, T.S.B., Matias Rodrigues, J.F., and von Mering, C. (2015). Limits to Robustness and Reproducibility in the Demarcation of Operational Taxonomic Units. Environ. Microbiol. *17*, 1689–1706.

Schommer, N.N., and Gallo, R.L. (2013). Structure and function of the human skin microbiome. Trends Microbiol. *21*, 660–668.

Schrezenmeir, J., and de Vrese, M. (2001). Probiotics, prebiotics, and synbiotics--approaching a definition. Am. J. Clin. Nutr. *73*, 361S – 364S.

Schwarz, R., Le Roux, J., Schaller, R., and Neurand, K. (1979). Micromorphology of the skin (epidermis, dermis, subcutis) of the dog. Onderstepoort J. Vet. Res. *46*, 105–109.

Scott, D.W., Miller, W.H. (William H., and Griffin, C.E. (1995). Muller and Kirk's Small Animal Dermatology (W.B. Saunder Co).

Seekatz, A.M., Rao, K., Santhosh, K., and Young, V.B. (2016). Dynamics of the fecal microbiome in patients with recurrent and nonrecurrent Clostridium difficile infection. Genome Med. *8*, 47.

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., and Huttenhower, C. (2011). Metagenomic biomarker discovery and explanation. Genome Biol. *12*, R60.

Seite, S., and Bieber, T. (2015). Barrier function and microbiotic dysbiosis in atopic dermatitis. Clin. Cosmet. Investig. Dermatol. *8*, 479–483.

Sender, R., Fuchs, S., and Milo, R. (2016). Revised Estimates for the Number of Human and Bacteria Cells in the Body. PLoS Biol. *14*, e1002533.

Sergeant, M.J., Constantinidou, C., Cogan, T., Penn, C.W., and Pallen, M.J. (2012). High-Throughput Sequencing of 16S rRNA Gene Amplicons: Effects of Extraction Procedure, Primer Length and Annealing Temperature. PLoS One *7*, e38094.

Shannon, C.E., and Weaver, W. (1949). The mathematical theory of information. Urbana University of Illinois Press, 623-656.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. *13*, 2498–2504.

Shu, M., Wang, Y., Yu, J., Kuo, S., Coda, A., Jiang, Y., Gallo, R.L., and Huang, C.M. (2013). Fermentation of Propionibacterium acnes, a commensal bacterium in the human skin microbiome, as skin probiotics against methicillin-resistant Staphylococcus aureus. PLoS One *8*, e55380.

Siponen, M., Kauppila, J.H., Soini, Y., and Salo, T. (2012). TLR4 and TLR9 are induced in oral lichen planus. J. Oral Pathol. Med. *41*, 741–747.

Smeekens, S.P., Huttenhower, C., Riza, A., van de Veerdonk, F.L., Zeeuwen, P.L.J.M., Schalkwijk, J., van der Meer, J.W.M., Xavier, R.J., Netea, M.G., and Gevers, D. (2013). Skin Microbiome Imbalance in Patients with STAT1/STAT3 Defects Impairs Innate Host Defense Responses. J. Innate Immun. *6*, 253–262.

Smith, A.M., Jain, M., Mulroney, L., Garalde, D.R., and Akeson, M. (2017). Reading canonical and modified nucleotides in 16S ribosomal RNA using nanopore direct RNA sequencing. bioRxiv.

Smith, K., and Thiboutot, D. (2008). Thematic review series: skin lipids. Sebaceous gland lipids: friend or foe? J. Lipid Res. *49*, 271–281.

Song, S.J., Lauber, C., Costello, E.K., Lozupone, C.A., Humphrey, G., Berg-Lyons, D., Caporaso, J.G., Knights, D., Clemente, J.C., Nakielny, S., et al. (2013). Cohabiting family members share microbiota with one another and with their dogs. Elife *2*, e00458.

Song, X., Jin, P., Qin, S., Chen, L., and Ma, F. (2012). The evolution and origin of animal Toll-like receptor signaling pathway revealed by network-level molecular evolutionary analyses. PLoS One *7*, e51657.

Song, Y., Garg, S., Girotra, M., Maddox, C., von Rosenvinge, E.C., Dutta, A., Dutta, S., and Fricke, W.F. (2013b). Microbiota Dynamics in Patients Treated with Fecal Microbiota Transplantation for Recurrent Clostridium difficile Infection. PLoS One *8*, e81330.

Stackebrandt, E., and Ebers, J. (2006). Taxonomic parameters revisited: tarnished gold standards. Microbiol Today *33*, 152–155.

Staley, C., Vaughn, B.P., Graiziger, C.T., Singroy, S., Hamilton, M.J., Yao, D., Chen, C., Khoruts, A., and Sadowsky, M.J. (2017). Community dynamics drive punctuated engraftment of the fecal microbiome following transplantation using freeze-dried, encapsulated fecal microbiota. Gut Microbes *28*, 1–13.

Suter, M.M., Schulze, K., Bergman, W., Welle, M., Roosje, P., and Müller, E.J. (2009). The keratinocyte in epidermal renewal and defence. Vet. Dermatol. *20*, 515–532.

Torres, S., Clayton, J.B., Danzeisen, J.L., Ward, T., Huang, H., Knights, D., and Johnson, T.J. (2017). Diverse bacterial communities exist on canine skin and are impacted by cohabitation and time. PeerJ *5*, e3075.

Travers, K.J., Chin, C.-S., Rank, D.R., Eid, J.S., and Turner, S.W. (2010). A flexible and efficient template format for circular consensus sequencing and SNP detection. Nucleic Acids Res. *38*, e159.

Tuon, F.F., Fernandes, E.R., Duarte, M.I.S., and Amato, V.S. (2010). The expression of TLR2, TLR4 and TLR9 in the epidermis of patients with cutaneous leishmaniasis. J. Dermatol. Sci. *59*, 55–57.

Ueta, M., Sotozono, C., Inatomi, T., and Kojima, K. (2007). Toll-like receptor 3 gene polymorphisms in Japanese patients with Stevens–Johnson syndrome. Br J Ophthalmol. *91*, 962-5

Ushijima, T., Takahashi, M., and Ozaki, Y. (1984). Acetic, propionic, and oleic acid as the possible factors influencing the predominant residence of some species of Propionibacterium and coagulase-negative Staphylococcus on normal human skin. Can. J. Microbiol. *30*, 647–652.

van Dijk, E.L., Auger, H., Jaszczyszyn, Y., and Thermes, C. (2014). Ten years of next-generation sequencing technology. Trends Genet. *30*, 418–426.

Vázquez-Baeza, Y., Hyde, E.R., Suchodolski, J.S., and Knight, R. (2016). Dog and human inflammatory bowel disease rely on overlapping yet distinct dysbiosis networks. Nat. Microbiol. *1*, 16177.

Větrovský, T., and Baldrian, P. (2013). The Variability of the 16S rRNA Gene in Bacterial Genomes and Its Consequences for Bacterial Community Analyses. PLoS One. *8*, e57923.

Viljanen, M., Kuitunen, M., Haahtela, T., Juntunen-Backman, K., Korpela, R., and Savilahti, E. (2005). Probiotic effects on faecal inflammatory markers and on faecal IgA in food allergic atopic eczema/dermatitis syndrome infants. Pediatr. Allergy Immunol. *16*, 65–71.

Vincent, A.T., Derome, N., Boyle, B., Culley, A.I., and Charette, S.J. (2016). Next-generation sequencing (NGS) in the microbiological world: How to make the most of your money. J. Microbiol. Methods. *S0167-7012*, 30031-8.

Virgin, H.W. (2014). The virome in mammalian physiology and disease. Cell *157*, 142–150.

Volz, T., Skabytska, Y., Guenova, E., Chen, K.-M., Frick, J.-S., Kirschning, C.J., Kaesler, S., Röcken, M., and Biedermann, T. (2014). Nonpathogenic Bacteria Alleviating Atopic Dermatitis Inflammation Induce IL-10-Producing Dendritic Cells and Regulatory Tr1 Cells. J. Invest. Dermatol. *134*, 96–104.

Wagner Mackenzie, B., Waite, D.W., and Taylor, M.W. (2015). Evaluating variation in human gut microbiota profiles due to DNA extraction method and inter-subject differences. Front. Microbiol. *6*, 130.

Walker, A.W., Martin, J.C., Scott, P., Parkhill, J., Flint, H.J., and Scott, K.P. (2015). 16S rRNA gene-based profiling of the human infant gut microbiota is strongly influenced by sample processing and PCR primer choice. Microbiome *3*, 26.

Walters, W.A., Caporaso, J.G., Lauber, C.L., Berg-Lyons, D., Fierer, N., and Knight, R. (2011). PrimerProspector: de novo design and taxonomic analysis of barcoded polymerase chain reaction primers. Bioinformatics 27, 1159–1161.

Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl. Environ. Microbiol. *73*, 5261–5267.

Wang, Y., Yang, Q., and Wang, Z. (2015). The evolution of nanopore sequencing. Front. Genet. *5*, 449.

Wanke, I., Steffen, H., Christ, C., Krismer, B., Götz, F., Peschel, A., Schaller, M., and Schittek, B. (2011). Skin Commensals Amplify the Innate Immune Response to Pathogens by Activation of Distinct Signaling Pathways. J. Invest. Dermatol. *131*, 382–390.

Werling, D., and Jungi, T.W. (2003). TOLL-like receptors linking innate and adaptive immune response. Vet. Immunol. Immunopathol. *91*, 1–12.

Werling, D., Jann, O.C., Offord, V., Glass, E.J., and Coffey, T.J. (2009). Variation matters: TLR structure and species-specific pathogen recognition. Trends Immunol. *30*, 124–130.

Wesolowska-Andersen, A., Bahl, M., Carvalho, V., Kristiansen, K., Sicheritz-Pontén, T., Gupta, R., and Licht, T. (2014). Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis. Microbiome *2*, 19.

Westcott, S.L., and Schloss, P.D. (2017). OptiClust, an Improved Method for Assigning Amplicon-Based Sequence Data to Operational Taxonomic Units. mSphere *2*, e00073-17.

White, S.D., and Yager, J.A. (1995). Resident dendritic cells in the epidermis: Langerhans cells, Merkel cells and melanocytes. Vet. Dermatol. *6*, 1–8.

Wlasiuk, G., and Nachman, M.W. (2010). Adaptation and constraint at Toll-like receptors in primates. Mol. Biol. Evol. 27, 2172–2186.

Wolcott, R.D., Gontcharova, V., Sun, Y., and Dowd, S.E. (2009). Evaluation of the bacterial diversity among and within individual venous leg ulcers using bacterial tag-encoded FLX and titanium amplicon pyrosequencing and metagenomic approaches. BMC Microbiol. *9*, 226.

Wong, S.H., Gochhait, S., Malhotra, D., Pettersson, F.H., Teo, Y.Y., Khor, C.C., Rautanen, A., Chapman, S.J., Mills, T.C., Srivastava, A., et al. (2010). Leprosy and the Adaptation of Human Toll-Like Receptor 1. PLoS Pathog. *6*, e1000979.

Wylie, K.M., Truty, R.M., Sharpton, T.J., Mihindukulasuriya, K. a, Zhou, Y., Gao, H., Sodergren, E., Weinstock, G.M., and Pollard, K.S. (2012). Novel bacterial taxa in the human microbiome. PLoS One *7*, e35294.

Yamasaki, K., Kanada, K., Macleod, D.T., Borkowski, A.W., Morizane, S., Nakatsuji, T., Cogen, A.L., and Gallo, R.L. (2011). TLR2 Expression Is Increased in Rosacea and Stimulates Enhanced Serine Protease Production by Keratinocytes. J. Invest. Dermatol. *131*, 688–697.

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.-H., Whitman, W.B., Euzéby, J., Amann, R., and Rosselló-Móra, R. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat. Rev. Microbiol. *12*, 635–645.

Ying, S., Zeng, D., Chi, L., Tan, Y., Galzote, C., and Cardona, C. (2015). The Influence of Age and Gender on Skin- Associated Microbial Communities in Urban and Rural Human Populations. PLoS One. *10*, e0141842

Yokoyama, N., Ohta, H., Yamazaki, J., Kagawa, Y., Ichii, O., Khoirun, N., Morita, T., Osuga, T., Lim, S.Y., Sasaki, N., et al. (2017). Localization of Toll-like Receptor (TLR) 2 and TLR4 mRNA in the Colorectal Mucosa of Miniature Dachshunds with Inflammatory Colorectal Polyps. J. Comp. Pathol. *156*, 183–190.

Yuan, S., Cohen, D.B., Ravel, J., Abdo, Z., and Forney, L.J. (2012). Evaluation of Methods for the Extraction and Purification of DNA from the Human Microbiome. PLoS One *7*, e33865.

Zahl, S. (1977). Jackknifing An Index of Diversity. Ecology *58*, 907–913.

Zeeuwen, P.L., Boekhorst, J., van den Bogaard, E.H., de Koning, H.D., van de Kerkhof, P.M., Saulnier, D.M., van Swam, I.I., van Hijum, S.A., Kleerebezem, M., Schalkwijk, J., et al. (2012). Microbiome dynamics of human epidermis following skin barrier disruption. Genome Biol. *13*, R101.

Zhang, E., Tanaka, T., Tajima, M., Tsuboi, R., Nishikawa, A., and Sugita, T. (2011). Characterization of the skin fungal microbiota in patients with atopic dermatitis and in healthy subjects. Microbiol. Immunol. *55*, 625–632.

Zhou, H., Gu, Æ.J., and Lamont, Æ.S.J. (2007). Evolutionary Analysis for Functional Divergence of the Toll-Like Receptor Gene Family and Altered Functional Constraints. *65*, 119–123.