Shared signatures and divergence in skin microbiomes of children with atopic dermatitis and their caregivers

Minghao Chia, PhD[1], Ahmad N.M. Naim, BSc[1], Angeline S.L. Tay, PhD[2], Karmun Lim, BSc[1], Chew Kean Lee, FRCPath(UK)[3], See Jie Yow, BSc[4], John Chen, PhD[4], John E.A. Common, PhD[2#], Niranjan Nagarajan, PhD[1,5#], Elizabeth Huiwen Tham, MRCPCH[5,6#]

Affiliations

[1]Genome Institute of Singapore (GIS), Agency for Science, Technology and Research (A*STAR), Singapore

[2]Skin Research Institute of Singapore (SRIS), Agency for Science, Technology and Research (A*STAR), Singapore

[3]Department of Laboratory Medicine, National University Health System, Singapore

[4]Infectious Diseases Programme and Department of Microbiology and Immunology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore.

[5]Yong Loo Lin School of Medicine, National University of Singapore (NUS), Singapore

[6]Khoo Teck Puat-National University Children's Medical Institute, National University Health System (NUHS), Singapore

[#]Corresponding Authors

Elizabeth Huiwen Tham

Address: Department of Paediatrics, Yong Loo Lin School of Medicine

National University of Singapore

1E Kent Ridge Road, Level 12, NUHS Tower Block

Singapore 119228

Email address: elizabeth_tham@nuhs.edu.sg

Phone: (65) 6772-4415

Fax: (65) 6779-7486

Niranjan Nagarajan

Address: Genome Institute of Singapore, A*STAR, 60 Biopolis Street, #02-01 (Genome)

Singapore 138672

Email address: nagarajann@gis.a-star.edu.sg

Phone: (65) 68088071

John E.A. Common

Address: Skin Research Institute of Singapore, A*STAR, 11 Mandalay Road, Clinical Sciences Building, #17-01 Singapore 308232

Email address: john.common@asrl.a-star.edu.sg

Phone: (65) 64070508

Source of funding:

This research was supported by grant funding from the National University Health System Leadership in Academic Medicine (NLAM) Program (PFFR Sept 2016); Master of Clinical Investigation (MCI), Yong Loo Lin School of Medicine, National University of Singapore and

NUS Biomedical Institute for Global Health Research and Technology (BIGHEART) Joint Research Grant 2018 administered by Institute for Health Innovation & Technology, National University of Singapore, and an A*STAR BMRC EDB IAF-PP grant (H18/01/a0/016), titled the Asian Skin Microbiome Program. Minghao Chia and Angeline Tay are research fellows receiving core funding from the A*STAR Graduate Academy (AGA). Elizabeth Tham is supported by the National Medical Research Council (NMRC) Research Training Fellowship [MH 095:003\008-225] and the NMRC Transition Award [MOH-000269]. The computational work for this article was partially performed on resources of the National Supercomputing Centre, Singapore (https://www.nscc.sg).

Conflict of Interest:

The authors have no conflict of interest to declare.

Word count: 7695

Number of Tables: 7 (2 main, 5 supplementary)

Number of Figures: 10 (5 main, 5 supplementary)

Abstract

Background

Atopic dermatitis (AD) is a common chronic skin condition in children (15-20%) that can significantly impair their quality of life. Due to its relapsing nature and enrichment of *Staphylococcus aureus* during flares, clinical management can include eradicating *S. aureus* from the skin of children, however this does not extend to their healthy caregivers who are potential reservoirs.

Objective

Our aim was to understand skin microbiome sharing and microbial features in children with AD and their healthy adult caregivers.

Methods

We utilized whole metagenome profiling at 4 body sites (volar forearm, antecubital fossae, cheeks and lesions) in combination with sequencing of *S. aureus* isolates to characterize a cohort of children with AD and their healthy caregivers (n = 30 families), in comparison to matched pairs from control households (n = 30 families).

Results

Metagenome analysis revealed distinct microbiome configurations in the non-lesional skin of AD children and their healthy caregivers versus controls, which were sufficient to accurately predict case-control status (AUC > 0.8). These differences were accompanied by significant microbiome similarity between children and their caregivers, indicating that microbiome sharing may play a role in recurrent disease flares. Whole genome comparisons with high quality *S. aureus* isolate genomes (n = 55) confirmed significant strain sharing between AD children and their caregivers and AD-specific enrichment of strains expressing enterotoxins Q and K/K2.

Conclusion

Our results highlight the distinctive skin microbiome features of healthy caregivers for children with AD, and support their inclusion in strategies for the treatment of recurrent paediatric AD.

Clinical Implications:

The identification of healthy caregiver skin as a reservoir for the sharing of microbes, especially *S. aureus* strains, supports the inclusion of caregivers in strategies for treating recurrent paediatric AD.

Capsule summary:

Close contact between clinically healthy caregivers and paediatric atopic dermatitis patients could favour the sharing of skin microbiota communities enriched in *S. aureus* and lacking in protective commensals.

Keywords:

Atopic Dermatitis, Caregivers, Household Contact, Skin Microbiome, *Staphylococcus aureus*

Abbreviations:

Atopic dermatitis (AD), Scoring AD Severity (SCORAD), Antecubital fossae (Ac), Volar forearms (Vf), Cheeks (Ch), Methicillin-resistant *Staphylococcus aureus* (MRSA), Methicillin-sensitive Staphylococcus aureus (MSSA), Linear discriminant analysis effect size (LefSe), Cumulative sum scaled (CSS), Enterotoxin Q (SEQ), Enterotoxin K (SEK), Multilocus variable-number tandem repeat fingerprinting (MLVF), Average nucleotide identity (ANI).

Introduction

Atopic dermatitis (AD) is the most common chronic inflammatory skin disease afflicting children, with rapidly rising prevalence.[1-3] Patients with moderate to severe disease suffer from complications such as frequent flares, superimposed bacterial and viral skin infections and significantly impaired quality of life from intense itching and sleep deprivation.[4, 5]

The skin microbiome is the composite of all micro-organisms colonizing the surface of an individual's skin. In AD, it can be polarized during flares towards a predominance of *Staphylococcus aureus* (*S. aureus*), and reduced microbial diversity. Up to 30-40% of the human population is asymptomatically colonized with *S. aureus,*[6] but colonization in children with AD may be as high as 80%[7] to 100%,[8] of which up to 16% are methicillin-resistant, complicating clinical treatment.[7] *S. aureus* colonization is also associated with increased disease severity and morbidity.[8-10] Some of the mechanisms by which *S. aureus* enhances skin inflammation in AD include the production of virulence factors such as superantigens, α-toxin, protein A (SpA), lipotechoic acid (LTA), phenol soluble modulin (PSM)-α, and proteases which induce direct keratinocyte damage, production of proinflammatory cytokines and inactivation of native antimicrobial peptides, to contribute to skin barrier disruption.[11]

Conventional AD treatment is aimed only at disease control and does not effect a permanent cure. Interventions such as the use of emollients and less frequently, bleach baths, only result in transient reduction of S. *aureus* skin burden and recovery in skin microbiome diversity.[12-15] AD patients are often quickly re-colonized with *S. aureus* upon treatment cessation, which could contribute to recurring disease.[16] While a possible source for re-acquisition of *S. aureus* strains could be the skin microbiome of caregivers,[17-19] the extent of skin microbiome sharing between

132 paediatric AD patients and their caregivers, particularly in comparison to healthy households has
133 not been examined.[20-22]

134 In this work, we utilized whole metagenome sequencing of 4 skin sites (volar forearm,
135 antecubital fossae, cheeks and lesions) for in-depth characterization of skin microbiomes of
136 healthy caregivers of children with moderate to severe AD (n = 30 families), as well as age-
137 matched control households (n = 30 families, Total: n = 116 subjects represented by at least one
138 metagenome, 277 metagenomes passing filters). We found that the healthy caregivers of AD
139 children had distinct skin microbiomes compared to control caregivers. These differences were
140 accompanied by microbiome differences in the non-lesional skin of AD children versus controls,
141 and significant microbiome similarity between children and their caregivers, indicating that
142 microbiome sharing is prevalent in paediatric AD. In order to further characterize *S. aureus*
143 reservoirs from children and their caregivers, we analysed the assembled genomes of strains
144 isolated from lesions, nares, axilla and groin (n = 55). Phylogenetic analysis of *S. aureus* isolates
145 confirmed that AD children and their healthy caregivers shared similar strains, with an AD-specific
146 enrichment of strains harbouring enterotoxin genes Q, K and K2. These results demonstrate the
147 utility of whole metagenomic profiling and comparisons of *S. aureus* genomes, for a deeper
148 understanding of skin microbial variations in AD children and their healthy caregivers. Our
149 identification of healthy caregiver skin as a reservoir for the sharing of microbes, especially *S.*
150 *aureus* strains, supports the inclusion of caregivers in strategies for treating recurrent paediatric
151 AD.

Methods

The study was conducted in the outpatient specialist paediatric clinic in a tertiary hospital (National University Hospital) in Singapore. The study received ethics approval by the NHG Domain Specific Review Board, Singapore (Reference number 2016/01080) and informed consent was obtained from all parents and caregivers and assent from children who were capable of providing informed assent. 30 children 0-10 years of age with moderate to severe AD, 30 healthy age-matched children as controls and one healthy primary caregiver for each child were recruited. Diagnosis of AD was made by trained paediatric allergists according to the Hanifin and Rajka diagnostic criteria[23] and its severity assessed using the Scoring AD Severity (SCORAD) grading system.[24] Moderate to severe AD was defined[25] as a SCORAD of more than 25. A primary caregiver was defined as a healthy person living in the same household as the index patient, who spent at least 8 hours a day with the patient for the past 6 months. Healthy caregivers had to be free of a history of allergic diseases such as AD, asthma or allergic rhinitis. Subjects were excluded if they had any other significant co-morbidities in major organ systems, immunodeficiency or were pregnant.

A questionnaire was administered to caregivers to collect clinical data on the child's AD onset, frequency of flares, treatment regimen, childcare setup and house cleaning practices. All subjects were required to avoid application of topical steroids, antibiotics or antiseptics (1 week), moisturizers or showering (12 hours) and systemic antibiotics (1 month) prior to enrolment into the study. Subjects who had been admitted to a hospital in the past 1 year were excluded from participation.

Skin microbiome collection

Skin tape strips were obtained from non-lesional skin at antecubital fossae (Ac), volar forearms (Vf), cheeks (Ch; in children below 3 years of age in whom cheeks are common flare-prone areas) in subjects and their primary caregivers as well as from lesional skin in AD subjects (**Supplementary Figure 1A**). The skin sampling protocol has been described in detail previously.[26] Briefly, 22mm D-Squame® standard sampling discs (CuDerm Corp, Dallas, TX, USA) were applied to the subjects' skin and removed repeatedly for at least 2 minutes and until the adhesive tape discs were fully saturated with skin surface material. Handling control samples were obtained by repeated application and removal of tape strips to alcohol-sterilized gloved hands of the individual performing the tape stripping procedures before proceeding to collection of samples from subjects and caregivers. Gloves were changed, alcohol-sterilized and air-dried before sampling a different individual.

Genomic DNA Extraction

Tape discs were transferred into Lysing Matrix E tubes (MP Biomedicals) and 500 µl of ATL buffer (Qiagen) was added. All samples were subjected to bead-beating with a FastPrep-24 Instrument (MP Biomedicals) at a speed of 6.0 ms^{-1} for 40 seconds. Samples were then centrifuged at 16,000g for 5 min and the supernatants were treated with Proteinase K (Qiagen) and incubated at 56 °C for 15 min. DNA was extracted using an EZ1 Advanced XL Instrument (Qiagen) with the EZ1 DNATissue Kit (Qiagen) and quantified using the Qubit dsDNA HS Assay Kit (Life Technologies) and stored at −20 °C. Blank samples were subjected to the same extraction protocol, except that the tape discs used were only exposed to air, without contacting skin or other laboratory surfaces.

Collection & identification of S. aureus isolates

196 The anterior nares, axilla and groin (inguinal crease) are known reservoirs for *S. aureus*.
197 Sterile cotton tipped swabs (Copan Transystem™ Culture Swab Transport System, Copan, Italy)
198 were rolled in the anterior nares of all subjects for 3 seconds. Skin swabs were obtained by rolling
199 separate sterile cotton tipped swabs over the skin in the axilla and inguinal crease, as well as on
200 lesional skin in AD patients for 3 seconds. Each swab was placed directly into Amies media
201 (Copan Transystem™ Culture Swab Transport System, Copan, Italy) for transportation to the
202 laboratory at room temperature.

203 One swabful of the swab suspension was inoculated onto a Phenylethyl Alcohol Agar plate
204 (RemelTM PEA, ThermoFisher Scientific) for MSSA detection and a chromogenic MRSA Select
205 agar plate (Bio-Rad MRSASelect^{TM} 63747) for MRSA detection, followed by streaking using the
206 3-sector streak method. The plates were incubated in air at 36 °C for 18-20 hours and reviewed for
207 the presence of bacterial colonies. Gram stain (confirming Gram positive cocci in clusters),
208 positive catalase test and positive coagulase test (Staphaurex^{TM} latex agglutination test,
209 ThermoFisher Scientific) were performed as supportive tests. For colonies with atypical or
210 borderline reaction(s), identification of *S. aureus* was confirmed using the Matrix-Assisted Laser
211 Desorption Ionization Time of Flight mass spectrometry (Bruker Biotyper, Germany).

212 Antimicrobial sensitivity testing was performed using VITEK 2 (bioMérieux, Marcy
213 l'Étoile, France), an integrated system that automatically performs antimicrobial susceptibility
214 testing based on kinetic analysis of growth data. Colonies were picked from each agar plate and
215 emulsified in 2ml of 0.45% normal saline, adjusted to the turbidity of a 0.6 McFarland Standard,
216 and used to load the ID-GPC test card for VITEK 2, in accordance with the manufacturer's
217 instructions. Susceptibility results for both MRSA and MSSA were expressed in minimal
218 inhibitory concentrations (MICs), using EUCAST breakpoints as guidance.

For total chromosomal DNA, *S. aureus* isolates were grown in tryptic soy broth at 37°C to mid-log phase and 1ml was pelleted for each strain. Cell pellets were resuspended in P1 buffer (Qiagen) and treated with 0.5 µg/ml lysostaphin (AMBI) at 37°C for 1 hr, then 0.5 µg/ml proteinase K (Sigma-Aldrich) at 37°C for 1 hr, and DNA was extracted with the MasterPure™ DNA Purification Kit (Epicentre).

Library construction for metagenomes

Two different batches of libraries were constructed for this study. For the first batch, an ultrasonicator (Covaris) was used to shear a standard volume of 50 µl of the extracted DNA. DNA libraries were prepared using the GeneRead DNA Library I Core Kit (Qiagen, Cat. No. 180434) according to the manufacturer's protocol with the use of barcode adapters in place of the GeneRead Adapter I Set. Custom 6 bp, 48 index-primers were used for PCR enrichment of the DNA libraries, according to an enrichment protocol adapted from the Multiplexing Sample Preparation Oligonucleotide kit (Illumina, Cat. No. PE-930-1002) (**Supplementary table 1**). Library concentrations and sizes were quantified using an Agilent Bioanalyzer, using the Agilent High Sensitivity DNA Kit (Agilent Technologies, Cat. No. 5067-4626). The DNA was pooled and paired-end sequencing (2×101 bp reads) was performed on the Illumina HiSeq 4000 platform.

For the second batch, 26 µl of extracted DNA was used as an input for the NEBNext ® UltraTM II FS DNA library prep kit (New England Biolabs, Cat. No. E7805). Libraries were made according to manufacturer's instructions. Library concentrations and sizes were quantified using the Agilent D1000 ScreenTape assay (Agilent Technologies, Cat. No. 5067-5582 & 5067-5583). Samples were pooled in an equimolar manner with the following exceptions. DNA extracted from the skin of children were pooled at a 2x higher concentration than that from adults (~ 24M vs 12M reads per sample). In addition, DNA from lesions were pooled at a 3x higher concentration (~36M

reads per sample) than DNA from adult skin samples. Paired-end sequencing (2×151bp reads) was performed on the Illumina HiSeq X platform.

The tally of sequenced libraries was 278 metagenomes from children and caregivers, and 58 laboratory blank and handling controls. After quality control, data from 277 skin metagenomes and all the blank and handling controls were used for subsequent analyses.

Taxonomic profiling

Short reads from Illumina sequencing were processed using a Nextflow[27] pipeline (https://github.com/CSB5/shotgunmetagenomics-nf). In brief, low quality bases and adapter sequences were trimmed using the default settings with fastp[28] (v0.20.0). Next, human reads were removed by mapping to the hg19 human reference genome using BWA-MEM[29] (v0.7.10-r789) and samtools[30] (v1.9), with parameters –f12 –F256. The remaining reads were profiled with Kraken2[31] (v.2.0.8) and Bracken[32] (v 2.5) to obtain taxonomic labels, using the 8 Gb MiniKraken2 database built from Refseq bacteria, archaea and viral genomes and the GrCh38 human genome. To ensure a minimum level of coverage across taxa, we only selected samples with at least 10,000 paired microbial reads and further removed reads which were still taxonomically assigned to *Homo sapiens*.

Removal of contaminant microbial species

To identify microbial contaminants introduced by laboratory reagents and sample handling, we sequenced a set of laboratory blanks and a set of handling controls. The blanks were obtained by constructing libraries from just laboratory reagents with no further inputs and sequencing them (n = 5) (**Supplementary Figure 1C**) The handling controls (n = 19 for batch 1 & n = 34 for batch 2) were obtained by applying tape strips to the gloves of colleagues doing the sampling in hospitals, and sequencing the material on those tape strips. The relative abundances

of major species (median relative abundance ≥ 0.1% within each set) detected in these negative controls are listed in **Supplementary table 6**.

We used a "batch and correlation approach" to identify and remove microbial contaminants[33]. Each metagenomic library was grouped into batches "A" or "B", based on which library preparation kit was used, in order to determine batch specific effects. Very rare taxa were removed on the basis of very low median relative abundances (< 0.1%). Subsequently, microbial species were flagged as contaminants if they were commonly present in one batch (relative abundance > 0.5% in ≥ 25% of libraries), but if they were much less prevalent in the other (≥ 2 fold difference in presence from one batch relative to the other). The identified contaminant taxa were *Achromobacter xylosoxidans*, *Acinetobacter baumannii*, *Janibacter indicus* and *Paracoccus yeei*. We then performed a non-parametric correlation analysis per batch to identify more contaminants based on their high correlation (Spearman's rho ≥ 0.7) with each other, and with *Achromobacter xylosoxidans*, *Acinetobacter baumannii*, *Janibacter indicus* or *Paracoccus yeei*. The additional contaminant species identified were *Brevundimonas vesicularis*, *Brevundimonas sp. GW460-12-10-14-LB2*, *Paracoccus contaminans*, *Pseudomonas pseudoalcaligenes* and *Pseudomonas sp. phDV1*. Reads assigned to these taxa were removed and taxonomic abundances were re-calculated based on the new totals. In addition, no statistically significant correlations (p > 0.05) were found between these contaminants and other skin microbes (e.g. *Cutibacterium acnes* and *Staphylococcus aureus*).

Diversity and differential abundance analysis

Shannon diversity indices (SDI) were computed from relative abundances using the "diversity" function from the R package "vegan" (v2.5-6). The effective species number was calculated as the exponential of the SDI (e^{SDI}). Effective species number was plotted against the

estimated microbial read counts for all libraries to check for robustness to variations in library size. Bray-Curtis dissimilarity indices were computed from community data matrices using the "vegdist" function from the vegan package. These dissimilarity indices were used to perform principle coordinates analysis (PCoA) of species composition. Species richness was determined by first rarefying all metagenomes to the same number of reads (> 12,000), followed by counting the number of species with ≥ 1 read.

To identify differentially abundant species between any two groups of samples (e.g. AD child vs. control child volar forearm), all counts were first normalized using the cumulative sum scaling (CSS) method from the R package "metagenomeSeq"[34] (v3.11). To reduce the number of statistical comparisons, very rare taxa were filtered out on the basis of very low median relative abundance (< 0.1% by total sum scaling) in the metagenomic libraries. Subsequently, two-sided Wilcoxon rank sum tests were conducted for each species, between cases and controls. *P*-values were adjusted to control for false discovery rates (FDR < 0.05) using the p.adjust function in R.

Feature selection and logistic regression

Two different logistic regression models were trained and tested using the "caret" (v6.0-86) package in R to predict case-control status. In the first model, CSS normalized counts of taxa were filtered to remove very rare species (median relative abundance < 0.1%). Samples were then grouped by anatomical region (volar forearm or antecubital fossae) and recursive feature elimination (10-fold repeated cross-validation, three repeats) was used to select the best fitting subset of features. These features were then used to train a logistic regression model (10-fold repeated cross-validation, three repeats). In the second model, a modified version of the selection of balances procedure[35] was adopted. Taxa were first filtered to remove very rare species (median relative abundance < 0.1%). Raw counts for remaining taxa were kept and a compositional balance

311 defined as the normalized log ratio of *S. aureus* counts over the geometric mean of the counts of
312 subsets of taxa with k components (where k ∈ ℤ+) was computed. A logistic regression with
313 forward selection was then performed on this feature space and a minimal set of compositional
314 features were used to predict case-control status.

315 Accuracy and area under the receiver operating characteristic curve (AUC) metrics were
316 computed for these models. A corrected resampled paired t-test[36] was used to test for significant
317 differences in these metrics from different models using the same training sets.

318 ***Linear regression***

319 Linear regression of the balance of *S. aureus/S. hominis* or normalized *S. aureus*
320 abundances against SCORAD was done using the lm() function in R.

321 ***Microbial source tracking***

322 Source tracking analysis was done using the FEAST package in R.[37] This is an expectation-
323 maximization algorithm which estimates the origins or degree of contribution from different
324 reservoirs or "sources", to the skin microbiomes of different individuals ("sinks"). Each caregiver
325 and child skin sample was defined in turn as a "sink", and all the other skin samples were
326 distributed into different source categories. We modified additional procedures described in
327 Shenhav et al[37] to reduce noise caused by similarities in skin microbiomes within caregivers and
328 within children. For each sink *j*, we calculated the Bray-Curtis dissimilarities between (1)
329 metagenomes from caregiver Vf (*j*) and metagenomes from all other caregiver Vfs, (2)
330 metagenomes from children Vf (*j*) and metagenomes from all other children Vfs, (3) metagenomes
331 from caregiver Ac (*j*) and metagenomes from all other caregiver Acs and (4) metagenomes from
332 children Ac (*j*) and metagenomes from all other children Acs. We removed sources within groups
333 1-4, whose Bray-Curtis dissimilarities were lower than the corresponding median in each

respective group. Data was then processed for each sink *j* according to the documentation in the FEAST package: https://github.com/cozygene/FEAST.

Genome assembly for S. aureus isolates and quality control

10–20 ng DNA extracted from a total of 58 *S. aureus* isolates were used as inputs for the NEBNext ® UltraTM II FS DNA library prep kit (New England Biolabs, Cat. No. E7805). Libraries were made according to manufacturer's instructions. Library concentrations and sizes were quantified using the Agilent D1000 ScreenTape assay (Agilent Technologies, Cat. No. 5067-5582 & 5067-5583). Samples were pooled in an equimolar manner. Paired-end sequencing (2×151bp reads) was performed on the Illumina HiSeq X platform.

Illumina sequencing data for each *S. aureus* isolate was downsampled to at least 1 M read pairs and assembled using Velvet (v1.2.10). Parameters were optimized with VelvetOptimizer (v2.24) with a k-mer length ranging from 19 to 115 and a minimum contig length of 500. Assemblies were assessed for contamination using CheckM[38] (v1.0.7), and high quality genomes were selected for substantial completeness (≥ 95%) and low contamination (≤ 5%). The average nucleotide identity (ANI = #SNPs/alignment size) between each assembly and the *S. aureus* reference genome (strain NCTC 8325) was computed using fastANI[39] (v1.3). Assemblies were excluded from further analysis if the ANI was < 95% compared to the reference. After quality control, 55 genome assemblies were used for subsequent analysis. Quality control metrics for these isolates are described in **Supplementary table 4**.

Hierarchical clustering and phylogenetic analysis of isolates

Pairwise alignments between assembled genomes of isolates was performed using nucmer (v3.23). The calling of SNPs between pairs of genomes was done using MUMmer (v3.23) 'show-snps' function and regions were compared to determine their average nucleotide identity (ANI =

#SNPs/alignment size). Regions containing more than 1 SNP within 20bp were filtered out to account for artefacts resulting from horizontal gene transfer, recombination or repetitive sequences.[40, 41]

A maximum-likelihood phylogenetic tree of *S. aureus* isolates was constructed using Parsnp[42] (v1.2; -c –x) and visualized using the "ggtree" (v3.11) R package.

Virulence factor analysis

The SRST2 software was used with default parameters (v0.20, --min_coverage 90, --max_divergence 10, --min_depth 5) to detect if virulence factor sequences were present in the genomes of the *S. aureus* isolates[43] (https://github.com/katholt/srst2). The list of virulence factors on mobile genetic elements was obtained from Malachowa and DeLeo, 2010.[44] Representative sequences for each virulence factor were downloaded from the NCBI nucleotide database and saved in a custom-built reference fasta file. Virulence factor genes variably present in isolate genomes at frequencies of between 15% and 85% were shortlisted and tested for enrichment in AD households (one-sided Fisher's exact test).

Nucleotide BLAST (BlastN) was performed by querying each virulence factor against each high quality isolate genome. BlastN hits were evaluated based on very stringent criteria[45] (evalue < 1E-10 & percentage identity ≥ 98% & query coverage ≥ 60%).

Reverse transcription-quantitative PCR (RT-qPCR)

RNA samples were quantified using Invitrogen Qubit RNA HS Assay Kit (Catalog no. Q32852) prior to reverse transcription with Invitrogen SuperScript IV VILO Master Mix with ezDNase Enzyme (Catalog no. 11766050). To remove trace DNA, equal amounts of RNA were subjected to ezDNase treatment as per manufacturer's protocol. Samples were incubated at 37°C for 2 min and placed on ice. After DNase treatment, 10 µl nuclease-free water (1st Base Catalog

no. BUF-1180) was added to each sample to total volume of 20 µl. Each sample was split into two reverse transcription reactions (10 µl each); one with reverse transcriptase (RT sample) and the other without (no-RT control sample). Reverse transcription reaction was set up as per manufacturer's protocol and incubated in the thermal cycler at 25°C for 10 min, 50°C for 10 min, and 85°C for 5 min. 2 µl of 10× diluted cDNA was taken from each sample (RT sample, and no-RT control sample) for qPCR.

Three pairs of primers were used to amplify the cDNA (refer to table below for primer sequences). Reactions were prepared on a 384-well plate, in duplicates, using 5 µl of Applied Biosystems PowerUp SYBR Green Master Mix (Catalog no. A25742), 0.5 µl of 5 µM primers (forward and reverse combined) and 2 µl of 10× diluted cDNA, in a total volume of 10 µl for each reaction. The ViiA 7 Real-Time PCR System (Applied Biosystems) was used for qPCR with the following amplification parameters: 1 cycle of 95 °C for 2 min, 40 cycles of 95 °C for 15 s, 60 °C for 15 s, and 72 °C for 1 min. A 16S rRNA standard curve was created using serial dilution of synthesized double-stranded DNA oligomers (16S rRNA gBLOCK) to convert 16S rRNA Ct values to copy numbers. A pool of cDNA (1 uL of cDNA from each sample) was used as a mock standard curve for enterotoxin Q and enterotoxin K Ct values. Primer and gene block sequences are listed in **Supplementary table 1.**

Statistical testing

All statistical tests were conducted using R. Two-sided Wilcoxon rank sum tests were used for differential abundance analyses, unless stated otherwise. All p-values < 0.05 were considered statistically significant, after FDR adjustment where appropriate. Linear discriminant analysis on total sum scaled abundances was done using the LefSe package (v.1.09). The parameters used for

format_input.py were -c 1 -o 1000000. The parameters used for run_lefse.py were –a 1 –w 1 to output all raw p-values for further adjustment for multiple comparisons (FDR < 0.05) in R.

Data and source code availability

Intermediate files and scripts used to generate the figures in this manuscript can be found at https://github.com/CSB5/AD-caregiver-manuscript. Raw sequencing data has been uploaded to the European Nucleotide Archive, under study accession number PRJEB42399.

Results

Clinical characteristics of cohort

We recruited a cohort of children with AD and their healthy caregivers, as well as matched pairs from control (non-AD) households (n = 30 families). The healthy controls were well children on routine follow up for well-child visits, such as vaccinations and growth checks, or who attended the general clinic for other one-time medical issues such as a head injury or limb trauma. These patients are thus unlikely to have spent significantly more time in a tertiary hospital setting than a typical child in the community, as is standard practice within the Singapore medical care structure. Of 120 recruited individuals in total, 17 metagenomes were excluded due to library preparation failure, leaving 277 analyzed metagenomes from 116 subjects represented by at least 1 metagenome (**Supplementary Figure 1A**). Principal coordinates analysis (PCoA) showed that there were no obvious sequencing batch specific effects in the metagenomic libraries (**Supplementary Figure 1B**).

The clinical characteristics of the subjects are summarized in **Table 1**. The mean SCORAD in the AD subjects was 51.3 ± 13.4 and the mean age at diagnosis was 16.5 ± 19.5 months. The majority of patients reported experiencing AD flares at least once a week (43.3%) and 21 subjects (70%) were using topical steroids. There were no significant differences in demographic characteristics, household cleaning and bathing practices between control and AD households. Mothers were most commonly the primary caregivers for both groups.

S. aureus was cultured from nasal, axilla and groin swabs in 57% (17/30) of AD children and 40% (12/30) of healthy children (**Table 2**). These frequencies are also broadly consistent with *S. aureus* carriage frequencies both in other local studies (35%) and in international cohorts.[7-9] Six AD child-caregiver pairs (6/30 pairs = 20%) and five healthy child-caregiver pairs (5/30 pairs=

~17%) were co-colonized with *S. aureus* in these sites and yielded high quality assemblies (**Supplementary figure 1A**). The *S. aureus* strains cultured from the majority of subjects were methicillin-sensitive *S. aureus* (MSSA). Of the AD lesional samples, 55.6% (10/18) were positive for *S. aureus*, all of which were MSSA. Two AD children, 1 AD caregiver, 1 healthy child and 1 healthy caregiver were colonized with methicillin-resistant *S. aureus* (MRSA). None of the MRSA strains were shared between child and caregiver pairs.

AD children and their healthy caregivers have distinct skin microbiomes compared to individuals from control households

Although differences in the relative abundances of various species in the skin microbiomes of AD flare sites have been widely reported, there are no published observations about whether the non-lesional or unaffected skin at the time of active AD in children, together with the skin of their healthy caregivers, have distinct skin microbiomes relative to controls. This prompted us to investigate species level differences in case-control caregiver-child pairs, leveraging information from shotgun metagenomic sequencing (**Supplementary table 2**).

We used Cumulative Sum Scaling (CSS)[46] to obtain normalized species-level abundances, and tested for statistically significant differences between case and controls using two-sided Wilcoxon rank sum tests while adjusting p-values for multiple comparisons. Very rare taxa were filtered out on the basis of very low median relative abundance (< 0.1% by total sum scaling) so that differences in case-control groups were not driven by multiple scarce microbes. Among differentially abundant species most of them were less abundant (adjusted p-value < 0.05) on the non-lesional skin of AD children compared to that of healthy children (**Figure 1A**). In the AD children's Ac samples, there was a reduction in the median normalized abundances of key skin commensal species known to be antagonistic to *S. aureus* such as *Staphylococcus hominis* and

454 *Staphylococcus capitis,*[47] or associated with dampened inflammatory responses, like *Dermacoccus*
455 *nishinomiyaensis.*[26] In particular, the median normalized abundances of *Staphylococcus hominis*
456 was also reduced in volar forearm samples of AD children. Notably, although the microbiomes
457 were sampled from non-lesional skin, *S. aureus* was consistently enriched in the Ac samples of
458 AD children (**Figure 1A**, left panel and **Supplementary Figure 2A**).

459 We also checked if the same species were differentially abundant on caregivers' skin (AD
460 houses vs. control houses) using the same normalization method and statistical tests as before.
461 Although no significant differences were detected in the abundance of *S. aureus* after correcting
462 for multiple comparisons, a subset of caregivers in AD houses had lower abundances of the skin
463 commensals *Staphylococcus hominis* and *Dermacoccus nishinomiyaensis* compared to controls
464 (**Figure 1B**, Kolmogorov-Smirnov test, p-value < 0.05 and **Supplementary Figure 2B**).

465 As expected, AD lesions had the highest burden of *S. aureus* compared to all other sites
466 (**Supplementary Figure 2C**). We further observed a trend of increasing *S. aureus* relative
467 abundances from caregivers in AD houses, to paediatric AD non-lesional skin in both Vf and Ac
468 and to paediatric AD lesions, although not all comparisons were statistically significant with our
469 relatively small sample sizes (**Supplementary Figure 2C**). In contrast, this pattern was not
470 observed for the skin commensal *S. hominis*.

471 We also analyzed differences in the diversity of skin metagenomes between individuals in
472 AD and control households, using a measurement robust to library size differences[48] (Shannon
473 index, SDI; **Supplementary Figure 2D**). The non-lesional Ac skin of AD children had lower
474 diversity than the skin of control children (Wilcoxon rank sum test, $p < 0.05$; **Figure 1C**), which
475 varies from what has previously been shown in non-lesional skin of AD and control adults.[26] In
476 contrast, diversity was not significantly different for non-lesional Vf and Ch skin microbiomes

between AD and control children (Wilcoxon rank sum test, $p > 0.05$). There were also no differences in the diversity of the skin microbiomes of healthy caregivers from AD and control households (**Supplementary Figure 2E**). The richness of non-lesional skin microbiomes at both the Ac and Vf of AD children were reduced compared to that of healthy children (Wilcoxon rank sum test, p-value < 0.05), consistent with the prior observation of a reduction in diversity (**Figure 1C** and **Supplementary Figure 2F**). In contrast, there was no statistically significant difference in species richness in the skin microbiomes of caregivers from AD and healthy households. Taken together these results highlight that the non-lesional skin of AD children and their healthy caregivers have distinct microbiomes compared to individuals from control households.

Individuals in AD houses can be distinguished from those in control houses by the ratio of *Staphylococcus aureus* to *Staphylococcus hominis*

Despite the presence of weak individual signals differentiating clinically healthy caregivers in AD versus control households, we investigated if other statistical models could differentiate them, given a set of normalized microbial abundances. In one approach, we used the linear discriminant analysis effect size (LefSe) method[49] on total sum scaled species-level abundances. We found that *S. aureus* and *S. hominis* were key discriminators of AD skin (lesion and non-lesional) vs control skin in both the Vf and Ac of children even after adjustment for multiple comparisons (adjusted p-value < 0.05, **Supplementary figure 3** and **Supplementary table 5** for raw LefSe results).

In another approach, we separated the metagenomes for caregivers and children by body site (Vf or Ac) and conducted recursive feature elimination. The reduced set of CSS normalized microbial abundances were used as inputs in logistic regression to classify individuals as belonging to either AD or control houses (see **Methods** for more details). This approach yielded a model

with eight statistically significant (adjusted p-value < 0.05) features for the volar forearm dataset (**Supplementary table 3**). Notably, increased abundances of *S. aureus* and *S. epidermidis* were associated with an increased likelihood of an individual belonging to an AD household. This is in line with previous reports showing that these species have a pro-inflammatory role in AD.[26] In contrast, higher *S. hominis* abundance was associated with a decreased likelihood of an individual belonging to an AD household, consistent with previous results. Interestingly, we also found associations not previously reported in other studies. Three novel taxa: *Stenotrophomonas maltophilia*, *Propionibacterium sp. Oral taxon 193* and *Streptococcus pneumoniae,* were negatively associated with AD households. Further experimental validation is required to determine if and how these commensals protect against paediatric AD.

For the antecubital fossae dataset, eight statistically significant (adjusted p-value < 0.05) features were selected in the final model (**Supplementary table 3**). Similar to the volar forearm dataset, higher *S. aureus* abundances and higher *S. hominis* abundances were associated with an increased or decreased likelihood of an individual belonging to an AD household respectively. *Rothia dentocariosa* was also positively associated with AD households, consistent with another study.[50] Other features not previously known to be associated with AD were also significant in this model. Of these, *Propionibacterium sp. Oral taxon 193*, *Stenotrophomonas maltophilia*, *Streptococcus gordonii* and *Streptococcus oralis* were negatively associated, while *Pseudomonas putida* was positively associated with AD households. These novel associations are a starting point for further investigations into the specific roles these species may play in health and disease.

In our final approach using a compositionally robust model similar to the "Selbal" algorithm,[35] we identified the normalized log ratio of *S. aureus* over *S. hominis* (subsequently abbreviated as the A/H ratio) as the single most important feature able to classify individuals as

belonging to an AD or control household, regardless of whether they were otherwise healthy adults or children (**Figures 2A** and **2B**). This model performed well in terms of training accuracy (> 0.74) and AUC (> 0.80) for both the Vf and Ac datasets. Furthermore, it classified individuals with similar accuracy and AUC (repeated paired resampled t-test, $p > 0.05$) as the multiple regression model with more features **(Figure 2A**). These findings support the notion that the skin of clinically healthy caregivers in AD households, has a distinct microbial signature compared to that of controls with *S. aureus* and *S. hominis* abundances being particularly important. Given that it is a distinguishing feature across many caregiver-child pairs in AD houses, an increased A/H ratio may be useful for stratifying families into those at risk of recurrent AD exacerbated by the skin microbiome, and to guide therapeutic intervention in combination with other biomarkers.

Given that the A/H ratio could classify individuals as belonging to AD or control houses, we tested if this feature was also correlated with clinical features such as AD severity (SCORAD). We found that the A/H ratio on the skin of AD children was positively correlated with SCORAD on both the Vf and Ac samples (**Figure 2C**). The strength of this correlation was stronger than that of normalized *S. aureus* abundances against SCORAD (**Supplementary figure 4**). In summary, these results highlight the distinctive skin microbiome features of children with AD and their healthy caregivers. Our observations agree with previous findings showing that *S. aureus* abundances are inversely correlated with the abundances of other *Staphylococcus* species, AD severity and inflammatory parameters.[51] This also provides a basis for further investigation of the importance of skin microbiome sharing between caregiver and child, in the development of recurrent paediatric AD.

Significant microbiome similarities between individuals living in the same household

We hypothesized that microbiome sharing between caregiver and child could in part, explain why healthy caregivers of AD children have a distinct microbial signature from caregivers in control houses. Bray-Curtis dissimilarity indices indicated that the Vf and Ac skin microbiomes of caregiver-child pairs from the same household were more significantly similar than that of caregiver-child pairs from different households for both AD and control groups (Wilcoxon rank sum test, p-value < 0.001; **Figure 3A**).

We also used an alternative source-tracking approach to estimate the contribution of different reservoirs (sources) to the host skin microbiomes (sinks).[37] In this analysis, each caregiver skin sample was designated as a sink and for each of these sinks, all other skin samples were classified into different source categories: matched children from the same house; other caregivers; other AD children; other control children and unknown sources. Source-tracking analysis confirmed that sources from the same household contributed more to a target caregiver's skin microbiome than sources from outside the house (**Figure 3B**). Further analysis using children's skin samples as sinks showed similar findings (**Supplementary figure 5**). This supports the notion of shared skin microbiomes in individuals living in the same household.

Caregivers and children in AD households share *S. aureus* strains from common sources

Given the importance of *S. aureus* in AD pathology as well as our results showing that microbiomes are more similar between individuals in matched houses, we looked for direct evidence of *S. aureus* sharing between caregivers and children.

While 30 AD and 30 control caregiver-child pairs were swabbed for culture experiments, most did not yield *S. aureus* positive cultures (**Supplementary Figure 1A**). Of the sequenced *S. aureus* isolates, 55 of them could be assembled into high quality genomes (see **Methods**). Only 6 distinct caregiver-child pairs, each pair from a different AD house, yielded high quality *S. aureus*

568 isolate genomes. There were only 5 control families for which both the caregiver and the child
569 yielded high quality *S. aureus* isolate genomes.

570 We analyzed the high quality *S. aureus* isolate genomes (n=55) from different body sites
571 (nares, axilla and groin combined and lesions) and constructed a phylogenetic tree based on
572 comparing the core genome alignment between all assemblies. The average nucleotide identity
573 (%ANI) between pairs of genomes was 99.518% (s = 0.243%). Strikingly, we observed very high
574 similarity isolates found on matched caregivers and children in 5 out of 6 AD households (99.998-
575 99.999% ANI) (**Figure 4A**). The % ANI between isolates from caregivers and children showed a
576 statistically significant increase in strain similarity (one-sided Wilcoxon rank sum test, p-value <
577 0.05) in caregiver-child pairs from AD households compared to controls, indicative of *S. aureus*
578 co-colonization. In addition, we tested the hypothesis that some known *S. aureus* virulence
579 factors[44] were more abundant in the genomes of isolates from AD households relative to controls.
580 Among the virulence factor genes that were variably present, enterotoxins Q (*seq*), K2 and K (*sek*)
581 were enriched in isolates from AD households but not in control households (one-sided Fisher's
582 exact test, adjusted p-value < 0.05) (**Figure 4B**). We also checked for expression of *seq* and *sek*
583 mRNAs from cultures of the relevant isolates by RT-qPCR. Isolates whose genomes harboured
584 *seq* and *sek* also expressed their mRNAs in vitro, whereas the negative control isolate did not
585 (**Figure 5**). Altogether, our data indicate that the same *S. aureus* strains are more frequently found
586 on both paediatric AD patients and their healthy caregivers, highlighting that microbial risk factors
587 for AD could be transferred by close physical contact in the house.

588 **Discussion**

589 Using a case-control study, we demonstrated for the first time that there are shared features
590 in the skin metagenomes and *S. aureus* strains between AD children and their healthy caregivers

591 living in the same household. We found that sites of disease predilection (non-lesional Ac) were
592 enriched with *S. aureus* but had less diverse communities in AD children relative to controls.[52]
593 Their similarities to the microbiome signatures found on AD lesions suggests that shifts in the
594 community can precede clinical disease.[13]

595 Although microbial community diversity and richness were similar between all adults, a
596 key finding is that healthy caregivers of AD children exhibited some distinctive skin microbiome
597 features compared to controls as well. Both healthy caregivers and their children who suffer from
598 AD were deficient in skin commensals such as *D. nishinomiyaensis* and *S. hominis,* compared to
599 healthy caregiver-child pairs. The relative deficiency of both these skin microbes could be due to
600 their weaker attachment to sensitized or dermatitic skin, or due to out-competition by pathobionts.
601 More studies are required to determine if these commensals have a protective role against AD,
602 either by moderating immune responses or by secreting anti-microbial peptides.[26, 47, 53] In addition,
603 the normalized log ratio of *S. aureus/S. hominis* (the A/H ratio) was found to be a sensitive and
604 specific marker for affiliation to an AD household regardless of the individual's own disease status.
605 This biomarker might be useful for identifying clinically healthy caregivers, who are at risk of
606 exacerbating AD in their children. AD caregiver-child pairs also shared the same, or highly similar
607 *S. aureus* strains with a greater prevalence of certain virulence factors compared to control
608 caregiver-child dyads. We propose that constant close contact between caregivers and children in
609 AD houses could favour the sharing of microbes that exacerbate AD and disfavor the sharing of
610 protective commensals.

611 The notion of shared skin microbiome characteristics implies that the caregiver skin
612 microbiome is an important influence on the skin microbiome of their children, and vice versa.
613 Despite their lack of clinical skin disease, healthy adults living in the same household as AD

children shared similar skin microbial features. Although prior studies have determined that the sharing of skin microbiota is present within occupants of the same household, there is significant variation depending on the nature of the relationships between occupants.[54] Therefore, we confined our sampling only to primary caregivers as determined by a strict definition, instead of all adults in the same household. This allowed us to standardize the nature of the relationship between the two individuals and assess the highest degree of close contact in our study. Such a relationship would be characterized by daily physical contact between the caregiver and child, allowing the greatest exchange of skin microbiota to interrogate our hypothesis. Determining directionality of transmission was beyond the scope of this study; this would require a longitudinal study tracking the presence of specific microbial strains present initially only on one individual and becoming detectable on another individual over time after commencement of close contact at a clearly defined time-point. In a culture-based study, Gaitanis et al. found a significant concordance of bacterial genera cultured from skin swabs of mothers and their infants at almost all time-points in the first year of life, indicative of a bidirectional mode of microbial transmission between the mother and her infant.[55]

The children in our cohort bathe more frequently on average compared to those in other populations, due to the tropical climate which promotes sweating – a known trigger of AD flares. We cannot rule out the possibility that these differences in hygiene habits may contribute to region specific microbiome signatures. To reduce this impact on our findings, we collected samples from all subjects only after $\geq$ 12 hours post-bathing, in line with practices defined by the Human Microbiome Project Consortium.[56] Despite differences in washing regimes in our cohort as compared to other populations, previous work has shown that the skin microbiome is temporally stable even at dry skin sites which are expected to be regularly cleaned.[57] Likewise, another study

found that skin site was the main factor explaining variations in skin microbiome composition as opposed to washing practices.[58] These observations suggest that an individual's skin microbiome can recover robustly from transient disturbances caused by bathing. Therefore, we are confident that our findings are broadly generalizable to other populations.

Published studies have shown that *S. aureus* carriage is significantly higher in close contacts of AD patients compared to controls.[17] Different non-sequencing based approaches such as multilocus variable-number tandem repeat fingerprinting (MLVF) and pulsed field gel electrophoresis have shown a high degree of concordance between *S. aureus* isolates in AD patients and their close contacts, supporting significant intra-familial transmission.[18, 19] *S. aureus* colonization of mothers prior to conception was also a risk factor for infant colonization at birth.[59] These studies, however, neither allowed for in-depth analysis of nucleotide identity, nor of genetic elements such as virulence factors present within the *S. aureus* strains. Our analysis of assembled *S. aureus* genomes indicated that healthy caregiver-child pairs did not share the same strains, likely due to the lower abundances of *S. aureus* on their bodies. Strikingly, strains found on multiple caregiver-child pairs in AD houses were extremely similar in their core genome alignments (% ANI $\geq$ 99.998%). Such high nucleotide identities imply that very little time has elapsed since the separation of these strains on the skin surfaces of caregiver and children. We caveat our strain-related findings by noting our limited sample sizes (low culture positivity) prevents us from making stronger conclusions about trends in the wider population, while the lack of temporal data prevents us from concluding whether these strains are transients or stable residents driving disease. Nonetheless, our findings provide a basis for larger scale investigations into intra-familial transmission of pathogenic *S. aureus* strains in atopic dermatitis, as already seen in the context of community associated MRSA.

660 The sharing of the same *S. aureus* strains could be mediated by different modes of
661 transmission which are not mutually exclusive. Direct transmission and persistence of the same
662 strains between caregiver-child pairs could be promoted by the relatively higher abundances of *S.*
663 *aureus* on the skin of AD children and frequent physical contact. In particular, there could be ample
664 opportunity for transmission or sharing of skin microbes during skin to skin contact while the
665 primary caregiver is applying medication or skin care products to their children. Alternatively, the
666 same strains could be indirectly shared via exposure to common sources within the same house
667 e.g. furniture or beddings, or via behaviorial characteristics peculiar to members of AD households
668 e.g. cleaning habits. In our study, the *S. aureus* cultures from caregivers were obtained by
669 combining swabs of the nares, axilla, and groin. We are therefore unable to determine the specific
670 anatomical reservoirs of these *S. aureus* strains on caregivers. A follow-up study will be important
671 to identify reservoirs or routes of spread that are of most concern. Taken together, we suggest that
672 caregiver skin surfaces are a possible source of *S. aureus* re-acquisition in recurrent paediatric AD.
673 *S. aureus* enterotoxins function as superantigens which are able to induce robust
674 proinflammatory responses in the skin, contributing to AD persistence and severity. Interestingly,
675 we found that *S. aureus* strains isolated from AD child-caregiver pairs harboured genes encoding
676 for enterotoxin Q (*seq*), K2 and K (*sek*) to a greater extent than strains from control houses. These
677 are clinically important toxins whose contribution to AD pathology is still unknown. USA300
678 MSSA and MRSA strains produce these enterotoxins, which have also been implicated in *S.*
679 *aureus* skin and soft tissue infections.[60, 61] The SEQ protein has also been described to be
680 remarkably heat-resistant, making it a potent pathogen in food poisoning.[62] Direct measurement
681 of virulence factor protein levels on the skin of several AD patients showed persistence of SEQ
682 on both lesions and non-lesional skin after bleach bath intervention.[63] *In vivo* studies in mice and

683 *in vitro* studies in human cells have shown that SEQ and SEK are superantigens capable of
684 excessive stimulation of immune cells.[62, 64] In addition, both *seq* and *sek* are amongst the most
685 prevalent toxin genes harboured by Asian *S. aureus* clinical isolates in Thai[65] and Chinese[66]
686 epidemiological studies. The functional role of these enterotoxins in AD pathogenesis and their
687 implications in strain sharing within households are yet to be elucidated. These are crucial research
688 questions to be incorporated into future studies.

689 The clinical significance of these findings translates into implications on antimicrobial
690 therapy and treatment approaches for *S. aureus* eradication. Skin decolonization for *S. aureus*
691 eradication is already part of routine clinical practice in other specialties such as surgery, renal
692 dialysis, intensive care and long term chronic care settings as a means of reducing *S. aureus*-
693 associated infection risks.[67] The two most common methods of decolonization are nasal
694 applications of antibiotic ointment (usually mupirocin) to the nose and antimicrobial skin washes,
695 both of which are non-invasive and cost-effective. Chlorhexidine gluconate (CHG)[68] and sodium
696 hypochlorite (bleach)[69] bath washes have been shown to significantly reduce infections in selected
697 patient groups such as surgical, dialysis, ICU patients as well as long term care.[70-73] Simultaneous
698 decolonization of household contacts together with the index patient has also demonstrated more
699 sustained long-term control in infants with MRSA skin infections than treating the index patient
700 alone.[19] The evidence base for *S. aureus* eradication in AD, such as bleach baths, however is weak
701 and is focused only on the patient. Outcomes have been conflicting and there is still little evidence
702 for sustained long-term effects,[74-79] which may be due to rapid re-colonization.[16] Our findings
703 suggest that one factor contributing to failure of therapy may be microbial sharing between close
704 contacts, rendering decolonization of the index patient alone ineffective. We thus speculate that
705 eradication of *S. aureus* colonization may not be effectively achieved by treatment of the AD

patient alone, but may need to be extended to all close household contacts to break the chain of transmission. Environmental control may also be an important factor, although this was beyond the scope of this study. Our data also did not show any differences in household cleaning practices between AD and control households which could confound bacterial colonization findings.

Long term AD control requires a multi-pronged approach encompassing adequate skin hydration, aggressive treatment of ongoing inflammation and eradication of pathogenic bacterial load. Among these treatment approaches, the latter is most difficult to achieve despite good compliance to prescribed therapy. Future therapeutics to eradicate or control pathogenic *S. aureus* strains may include not only transplanting of bioengineered probiotic strains that are strictly non-pathogenic, but also prebiotics (selectively fermented ingredients) and postbiotics (byproducts of microbial metabolism) as alternatives to antibiotic regimes which are known to promote antibiotic resistance. Future randomized controlled trials are needed to examine the efficacy of extending bacterial eradication beyond the index AD patient to his/her close household contacts and environment, as a step towards better long-term AD control and reduction of complications.

Acknowledgments

We would like to thank the doctors for their assistance in patient recruitment: Nisha Suyien Chandran, Lynette Pei-Chi Shek, Hugo van Bever, Soh Jian Yi, Pauline Poh-Lin Chan Ng, Lydia Su Yin Wong, Alison Joanne Lee, David Ng, Mohana Rajakulendran as well as our clinical research coordinators Ms Poon Chuying, Ms Rabi'atul'adawiyah Binte Abdullah and Ms Liu Chenjing for their help with patient recruitment and study procedures. We would also like to thank Ms Alicia Kang for her help in formatting and manuscript submission. The graphical abstract was created with BioRender.com.

References

1. Asher MI, Stewart AW, Wong G, Strachan DP, García-Marcos L, Anderson HR. Changes over time in the relationship between symptoms of asthma, rhinoconjunctivitis and eczema: a global perspective from the International Study of Asthma and Allergies in Childhood (ISAAC). Allergol Immunopathol (Madr) 2012; 40:267-74.
2. Weidinger S, Beck LA, Bieber T, Kabashima K, Irvine AD. Atopic dermatitis. Nat Rev Dis Primers 2018; 4:1.
3. Langan SM, Irvine AD, Weidinger S. Atopic dermatitis. Lancet 2020; 396:345-60.
4. Mancini AJ, Kaulback K, Chamlin SL. The socioeconomic impact of atopic dermatitis in the United States: a systematic review. Pediatr Dermatol 2008; 25:1-6.
5. Lee BW, Detzel PR. Treatment of childhood atopic dermatitis and economic burden of illness in Asia Pacific countries. Ann Nutr Metab 2015; 66 Suppl 1:18-24.
6. Lowy FD. Staphylococcus aureus infections. N Engl J Med 1998; 339:520-32.
7. Suh L, Coffin S, Leckerman KH, Gelfand JM, Honig PJ, Yan AC. Methicillin-resistant Staphylococcus aureus colonization in children with atopic dermatitis. Pediatr Dermatol 2008; 25:528-34.
8. Goh CL, Wong JS, Giam YC. Skin colonization of Staphylococcus aureus in atopic dermatitis patients seen at the National Skin Centre, Singapore. Int J Dermatol 1997; 36:653-7.
9. Totte JE, van der Feltz WT, Hennekam M, van Belkum A, van Zuuren EJ, Pasmans SG. Prevalence and odds of Staphylococcus aureus carriage in atopic dermatitis: a systematic review and meta-analysis. Br J Dermatol 2016.
10. Spaulding AR, Salgado-Pabon W, Kohler PL, Horswill AR, Leung DY, Schlievert PM. Staphylococcal and streptococcal superantigen exotoxins. Clin Microbiol Rev 2013; 26:422-47.
11. Geoghegan JA, Irvine AD, Foster TJ. Staphylococcus aureus and Atopic Dermatitis: A Complex and Evolving Relationship. Trends Microbiol 2018; 26:484-97.
12. Seite S, Flores GE, Henley JB, Martin R, Zelenkova H, Aguilar L, et al. Microbiome of affected and unaffected skin of patients with atopic dermatitis before and after emollient treatment. J Drugs Dermatol 2014; 13:1365-72.
13. 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.
14. George SM, Karanovic S, Harrison DA, Rani A, Birnie AJ, Bath-Hextall FJ, et al. Interventions to reduce Staphylococcus aureus in the management of eczema. Cochrane Database Syst Rev 2019; 2019.
15. Chopra R, Vakharia PP, Sacotte R, Silverberg JI. Efficacy of bleach baths in reducing severity of atopic dermatitis: A systematic review and meta-analysis. Ann Allergy Asthma Immunol 2017; 119:435-40.
16. Gilani SJ, Gonzalez M, Hussain I, Finlay AY, Patel GK. Staphylococcus aureus re-colonization in atopic dermatitis: beyond the skin. Clin Exp Dermatol 2005; 30:10-3.
17. Chiu LS, Chow VC, Ling JM, Hon KL. Staphylococcus aureus carriage in the anterior nares of close contacts of patients with atopic dermatitis. Arch Dermatol 2010; 146:748-52.

18. Chiu LS, Ho MS, Hsu LY, Tang MB. Prevalence and molecular characteristics of Staphylococcus aureus isolates colonizing patients with atopic dermatitis and their close contacts in Singapore. Br J Dermatol 2009; 160:965-71.
19. Meyer Sauteur PM, Leimer N, Chan PK, Relly C, Seidl K, Zinkernagel AS, et al. Methicillin-resistant Staphylococcus aureus eradication in young infants should include the diaper area and treat housing contacts. Acta Paediatr 2015; 104:e581-4.
20. Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, et al. 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.
21. 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.
22. Ross AA, Doxey AC, Neufeld JD. The Skin Microbiome of Cohabiting Couples. mSystems 2017; 2.
23. Hanifin JM, Rajka G. Diagnostic features od atopic dermatitis. Acta Derm Venereol (Stockh) 1980; Suppl 92: 44-47
24. Kunz B, Oranje AP, Labreze L, Stalder JF, Ring J, Taieb A. Clinical validation and guidelines for the SCORAD index: consensus report of the European Task Force on Atopic Dermatitis. Dermatology 1997; 195:10-9.
25. Severity scoring of atopic dermatitis: the SCORAD index. Consensus Report of the European Task Force on Atopic Dermatitis. Dermatology 1993; 186:23-31.
26. Chng KR, Tay AS, Li C, Ng AH, Wang J, Suri BK, et al. Whole metagenome profiling reveals skin microbiome-dependent susceptibility to atopic dermatitis flare. Nat Microbiol 2016; 1:16106.
27. Di Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E, Notredame C. Nextflow enables reproducible computational workflows. Nature Biotechnology 2017; 35:316-9.
28. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 2018; 34:i884-i90.
29. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 2010; 26:589-95.
30. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009; 25:2078-9.
31. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biology 2019; 20:257.
32. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: estimating species abundance in metagenomics data. PeerJ Computer Science 2017; 3:e104.
33. de Goffau MC, Lager S, Salter SJ, Wagner J, Kronbichler A, Charnock-Jones DS, et al. Recognizing the reagent microbiome. Nature Microbiology 2018; 3:851-3.
34. Berger B, Heinrich J, Niederstätter H, Hecht W, Morf N, Hellmann A, et al. Forensic characterization and statistical considerations of the CaDNAP 13-STR panel in 1,184 domestic dogs from Germany, Austria, and Switzerland. Forensic Sci Int Genet 2019; 42:90-8.
35. Rivera-Pinto J, Egozcue JJ, Pawlowsky-Glahn V, Paredes R, Noguera-Julian M, Calle ML. Balances: a New Perspective for Microbiome Analysis. mSystems 2018; 3:e00053-18.
36. Nadeau C, Bengio Y. Inference for the Generalization Error. Machine Learning 2003; 52:239-81.

37. Shenhav L, Thompson M, Joseph TA, Briscoe L, Furman O, Bogumil D, et al. FEAST: fast expectation-maximization for microbial source tracking. Nature Methods 2019; 16:627-32.
38. Koutoukidis DA, Jebb SA, Ordóñez-Mena JM, Noreik M, Tsiountsioura M, Kennedy S, et al. Prominent positioning and food swaps are effective interventions to reduce the saturated fat content of the shopping basket in an experimental online supermarket: a randomized controlled trial. Int J Behav Nutr Phys Act 2019; 16:50.
39. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nature Communications 2018; 9:5114.
40. Chng KR, Li C, Bertrand D, Ng AHQ, Kwah JS, Low HM, et al. Cartography of opportunistic pathogens and antibiotic resistance genes in a tertiary hospital environment. Nat Med 2020; 26:941-51.
41. Brooks B, Olm MR, Firek BA, Baker R, Thomas BC, Morowitz MJ, et al. Strain-resolved analysis of hospital rooms and infants reveals overlap between the human and room microbiome. Nat Commun 2017; 8:1814.
42. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome Biol 2014; 15:524.
43. Inouye M, Dashnow H, Raven LA, Schultz MB, Pope BJ, Tomita T, et al. SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. Genome Med 2014; 6:90.
44. Malachowa N, DeLeo FR. Mobile genetic elements of Staphylococcus aureus. Cell Mol Life Sci 2010; 67:3057-71.
45. Clausen PT, Zankari E, Aarestrup FM, Lund O. Benchmarking of methods for identification of antimicrobial resistance genes in bacterial whole genome data. J Antimicrob Chemother 2016; 71:2484-8.
46. Paulson JN, Stine OC, Bravo HC, Pop M. Differential abundance analysis for microbial marker-gene surveys. Nat Methods 2013; 10:1200-2.
47. Williams MR, Costa SK, Zaramela LS, Khalil S, Todd DA, Winter HL, et al. Quorum sensing between bacterial species on the skin protects against epidermal injury in atopic dermatitis. Science Translational Medicine 2019; 11:eaat8329.
48. Haegeman B, Hamelin J, Moriarty J, Neal P, Dushoff J, Weitz JS. Robust estimation of microbial diversity in theory and in practice. Isme j 2013; 7:1092-101.
49. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. Genome Biol 2011; 12:R60.
50. Fyhrquist N, Muirhead G, Prast-Nielsen S, Jeanmougin M, Olah P, Skoog T, et al. Microbe-host interplay in atopic dermatitis and psoriasis. Nat Commun 2019; 10:4703.
51. Ottman N, Barrientos-Somarribas M, Fyhrquist N, Alexander H, Wisgrill L, Olah P, et al. Microbial and transcriptional differences elucidate atopic dermatitis heterogeneity across skin sites. Allergy 2021; 76:1173-87.
52. Sun Z, Huang S, Zhu P, Yue F, Zhao H, Yang M, et al. A Microbiome-Based Index for Assessing Skin Health and Treatment Effects for Atopic Dermatitis in Children. mSystems 2019; 4.

53. Nakatsuji T, Chen TH, Narala S, Chun KA, Two AM, Yun T, et al. Antimicrobials from human skin commensal bacteria protect against Staphylococcus aureus and are deficient in atopic dermatitis. Sci Transl Med 2017; 9.
54. Leung MHY, Tong X, Wilkins D, Cheung HHL, Lee PKH. Individual and household attributes influence the dynamics of the personal skin microbiota and its association network. Microbiome 2018; 6:26.
55. Gaitanis G, Tsiouri G, Spyridonos P, Stefos T, Stamatas GN, Velegraki A, et al. Variation of cultured skin microbiota in mothers and their infants during the first year postpartum. Pediatr Dermatol 2019; 36:460-5.
56. A framework for human microbiome research. Nature 2012; 486:215-21.
57. Oh J, Byrd AL, Park M, Kong HH, Segre JA. Temporal Stability of the Human Skin Microbiome. Cell 2016; 165:854-66.
58. 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.
59. Jimenez-Truque N, Tedeschi S, Saye EJ, McKenna BD, Langdon W, Wright JP, et al. Relationship between maternal and neonatal Staphylococcus aureus colonization. Pediatrics 2012; 129:e1252-9.
60. Diep BA, Gill SR, Chang RF, Phan TH, Chen JH, Davidson MG, et al. Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant Staphylococcus aureus. Lancet 2006; 367:731-9.
61. Aguilar JL, Varshney AK, Wang X, Stanford L, Scharff M, Fries BC. Detection and measurement of staphylococcal enterotoxin-like K (SEl-K) secretion by Staphylococcus aureus clinical isolates. J Clin Microbiol 2014; 52:2536-43.
62. Hu DL, Ono HK, Isayama S, Okada R, Okamura M, Lei LC, et al. Biological characteristics of staphylococcal enterotoxin Q and its potential risk for food poisoning. J Appl Microbiol 2017; 122:1672-9.
63. Moran MC, Cahill MP, Brewer MG, Yoshida T, Knowlden S, Perez-Nazario N, et al. Staphylococcal Virulence Factors on the Skin of Atopic Dermatitis Patients. mSphere 2019; 4.
64. Orwin PM, Leung DY, Donahue HL, Novick RP, Schlievert PM. Biochemical and biological properties of Staphylococcal enterotoxin K. Infect Immun 2001; 69:360-6.
65. Indrawattana N, Sungkhachat O, Sookrung N, Chongsa-nguan M, Tungtrongchitr A, Voravuthikunchai SP, et al. Staphylococcus aureus clinical isolates: antibiotic susceptibility, molecular characteristics, and ability to form biofilm. Biomed Res Int 2013; 2013:314654.
66. Xie Y, He Y, Gehring A, Hu Y, Li Q, Tu SI, et al. Genotypes and toxin gene profiles of Staphylococcus aureus clinical isolates from China. PLoS One 2011; 6:e28276.
67. Septimus EJ, Schweizer ML. Decolonization in Prevention of Health Care-Associated Infections. Clin Microbiol Rev 2016; 29:201-22.
68. Denny J. Chlorhexidine Bathing Effects on Health-Care-Associated Infections. Biol Res Nurs 2016.
69. Liu C, Bayer A, Cosgrove SE, Daum RS, Fridkin SK, Gorwitz RJ, et al. Clinical practice guidelines by the infectious diseases society of america for the treatment of methicillin-resistant Staphylococcus aureus infections in adults and children: executive summary. Clin Infect Dis 2011; 52:285-92.

70. George S, Leasure AR, Horstmanshof D. Effectiveness of Decolonization With Chlorhexidine and Mupirocin in Reducing Surgical Site Infections: A Systematic Review. Dimens Crit Care Nurs 2016; 35:204-22.
71. Schora DM, Boehm S, Das S, Patel PA, O'Brien J, Hines C, et al. Impact of Detection, Education, Research and Decolonization without Isolation in Long-term care (DERAIL) on methicillin-resistant Staphylococcus aureus colonization and transmission at 3 long-term care facilities. Am J Infect Control 2014; 42:S269-73.
72. Schmid H, Romanos A, Schiffl H, Lederer SR. Persistent nasal methicillin-resistant staphylococcus aureus carriage in hemodialysis outpatients: a predictor of worse outcome. BMC Nephrol 2013; 14:93.
73. Gray D, Foster K, Cruz A, Kane G, Toomey M, Bay C, et al. Universal decolonization with hypochlorous solution in a burn intensive care unit in a tertiary care community hospital. Am J Infect Control 2016.
74. Wachs GN, Maibach HI. Co-operative double-blind trial of an antibiotic/corticoid combination in impetiginized atopic dermatitis. Br J Dermatol 1976; 95:323-8.
75. Sugimoto K, Ishikawa N, Terano T, Kitukawa Y, Kubosawa H, Ito S, et al. The importance of bacterial superantigens produced by Staphylococcus aureus in the treatment of atopic dermatitis using povidone-iodine. Dermatology 2006; 212 Suppl 1:26-34.
76. Birnie AJ, Bath-Hextall FJ, Ravenscroft JC, Williams HC. Interventions to reduce Staphylococcus aureus in the management of atopic eczema. Cochrane Database Syst Rev 2008:Cd003871.
77. Bath-Hextall FJ, Birnie AJ, Ravenscroft JC, Williams HC. Interventions to reduce Staphylococcus aureus in the management of atopic eczema: an updated Cochrane review. Br J Dermatol 2011; 164:228.
78. Huang JT, Abrams M, Tlougan B, Rademaker A, Paller AS. Treatment of Staphylococcus aureus colonization in atopic dermatitis decreases disease severity. Pediatrics 2009; 123:e808-14.
79. Gong JQ, Lin L, Lin T, Hao F, Zeng FQ, Bi ZG, et al. Skin colonization by Staphylococcus aureus in patients with eczema and atopic dermatitis and relevant combined topical therapy: a double-blind multicentre randomized controlled trial. Br J Dermatol 2006; 155:680-7.

954 **Table 1. Clinical characteristics of study participants**

	Control (n=30 caregiver-child pairs)	Atopic Dermatitis (n=30 caregiver-child pairs)
Demographics		
Age (months) at recruitment Mean ± SD (Range)	58.9 ± 38.8 (4 – 125)	58.5 ± 39.0 (4 – 128)
Gender, n(%) Male	12 (40.0%)	13 (43.3%)
Ethnicity, n(%)		
Chinese	27 (90.0%)	20 (66.7%)
Malay	2 (6.7%)	10 (33.3%)
Indian	1 (3.3%)	0 (0%)
Age at AD diagnosis (months) Mean ± SD (Range)	**NA**	16.5 ±19.5, (2 – 72)
SCORAD Mean ± SD (Range)	**NA**	51.3 ± 16.4, (27.1 - 82.3)
Housing Type, n(%)		
Subsidized housing	22 (73.3%)	24 (80.0%)
Private housing	7 (23.3%)	6 (20%)
Others	1 (3.3%)	0 (0%)
Primary Caregiver, n(%)		
Mother	25 (83.3%)	24 (80.0%)
Father	0 (0%)	1 (3.3%)
Grandparent	3 (10.0%)	4 (13.3%)
Other Relative/Nanny	2 (6.6%)	1 (3.3%)
Household cleaning practices		
Frequency of floor cleaning per week, n(mean)	4.3 (2.6)	4.9 (2.4)
Frequency of sofa cleaning, n(%)		
Never	4 (13.3%)	9 (30.0%)
Almost daily	6 (20.0%)	8 (26.7%)
At least once a week	5 (16.7%)	8 (26.7%)
At least once a month	14 (46.7%)	4 (13.3%)
Frequency of child's bedsheet change, n(%)		
Daily	0 (0%)	2 (6.7%)
At least once a week	19 (63.3%)	21 (70.0%)
At least once a month	11 (36.7%)	7 (23.3%)
Less than once a month	0 (0%)	0 (0%)
Frequency of caregiver's bedsheet change, n(%)		
Daily	0 (0%)	2 (6.7%)
At least once a week	17 (56.7%)	18 (60.0%)
At least once a month	13 (43.3%)	10 (33.3%)
Child's bathing frequency, n(%)		
Twice a day	17 (56.7%)	21 (70.0%)
Once a day	11 (36.7%)	8 (26.7%)
2-4× a week	2 (6.7%)	1 (3.3%)

Table 2. Microbial isolation from Control and Atopic Dermatitis child-caregiver pairs

	Controls (n=30 caregiver-child pairs)		**Atopic Dermatitis (n=30 caregiver-child pairs)**	
	Child	Caregiver	Child	Caregiver
Nasal, Axilla & Groin cultures (n=30)				
S. aureus isolated	12 (40%)	12 (40.0%)	17 (56.7%)	7 (23.3%)
Both child and caregiver positive for *S. aureus*	4 (13.3%)		6 (20.0%)	
MSSA*	11 (91.7%)	11 (91.7%)	15 (88.2%)	6 (85.7%)
MRSA#	1 (8.3%)	1 (8.3%)	2 (11.8%)	1 (14.3%)
Lesional cultures (n=18)				
S. aureus isolated	NA	NA	10 (55.6%)	NA

*MSSA: Methicillin-sensitive *S. aureus*
#MRSA: Methicillin-resistant *S. aureus*

Figure 1. **Distinctive skin microbiome signatures on the non-lesional skin of individuals in AD houses.**

(A) Scatter plot showing general depletion of microbial species on the non-lesional skin of AD children compared to control children for the antecubital fossae (left, n = 57) and volar forearm (right, n = 56). Median cumulative sum scaled (CSS) values for selected taxa were plotted here. These taxa were represented only if their normalized abundances were significantly different (Wilcoxon rank sum test, adjusted p-value < 0.05) between AD and control children. For visual clarity, only selected taxa of known biological relevance to AD were labelled for the antecubital fossae dataset (left, red dots), while all species labels were added for the volar forearm dataset (right, red dots).

(B) The distribution of normalized abundances of *Staphylococcus aureus* (left), *Staphylococcus hominis* (middle), and *Dermacoccus nishinomiyaensis* (right) in the metagenomes of caregivers from the antecubital fossae (top, n = 57) and volar forearm (bottom, n = 57), in AD and control houses. The p-values for two sided Kolmogorov-Smirnov tests are shown.

(C) Shannon diversities for different metagenomes of children from the volar forearm (n = 56), antecubital fossae (n = 57) and cheeks (n = 24). The p-values for Wilcoxon rank sum tests are shown.

Figure 2. Individuals in AD houses can be distinguished from those in control houses by the ratio of *Staphylococcus aureus* to *Staphylococcus hominis*

(A) Accuracy and the area under the receiver operating characteristics curve (AUC) for two different types of logistic regression models. The first model uses CSS normalized read counts of selected species to classify individuals as belonging to control or AD households (see Methods for species selection criteria). The second model uses the normalized log ratio of *S. aureus*/*S. hominis*

read counts (A/H ratio) for classification instead. There were no statistically significant differences in accuracy or AUC between the two models (repeated re-sampled paired t-test, $p > 0.05$). For this analysis, n = 114 and n = 113 for the antecubital fossae and volar forearm datasets respectively.

(B) Box plots showing how the A/H ratios are distributed between different individuals from control and AD houses. The sample sizes (n) and p-values for Wilcoxon rank sum tests are shown.

(C) Linear regression of A/H ratios from the metagenomes of non-lesioned skin of AD children, against their SCORAD values. The sample sizes (n), Pearson correlation coefficient *R*, and the associated p-values are shown.

Figure 3. The skin microbiomes of individuals in the same house are more similar than that of individuals in different houses

(A) Bray Curtis dissimilarity indices comparing metagenomes (same skin site) of matched caregiver-child pairs or comparing that of unmatched caregiver-child pairs in AD or control houses. The p-values for Wilcoxon rank sum tests are shown. Sample sizes are: n = 26 pairs of comparisons from matched AD houses and n = 30 pairs of comparisons from matched control houses. For pairwise comparisons of individuals from different AD houses, there were n = 676 and n = 703 pairs of comparisons between unmatched AD houses for Vf and Ac metagenomes respectively. Unmatched control houses had n = 870 pairs of comparisons.

(B) Estimated contributions of different sources (horizontal axis) to skin metagenomes of caregivers. Each metagenome from a caregiver was defined as a sink and other metagenomes in this study were split into distinct groups: matched child from same house, other caregivers from different houses, other children with AD from different houses, and other healthy children from different houses. The contribution of unknown sources to each sink was also included in these

models. The p-values for Wilcoxon rank sum tests are shown. For this analysis, n = 54 sinks from AD houses and n = 60 sinks from control houses.

Figure 4. Some caregivers and children in AD households share *S. aureus* strains from common sources

(A) Approximate maximum likelihood phylogenetic trees constructed using the Parsnp package, showing relationships between different *S. aureus* isolates (n = 55) in this study. The numbers in bold are the % average nucleotide identities (% ANI) between isolates of interest. "Matched isolates" refer to *S. aureus* cultured from the skin of caregivers and children from the same house. There is a statistically significant increase in % ANI between isolates from matched caregiver-child pairs in AD houses versus control houses (one-sided Wilcoxon rank sum test, p-value < 0.05). The scale represents the number of substitutions per site in the core alignment.

(B) Heatmap showing presence or absence of different virulence genes on mobile genetic elements in DNA samples from *S. aureus* cultures (n = 55) isolated from different individuals and different body sites. Boxes are color coded by whether the virulence factor was detected by SRST2, by BlastN, by both or by none of these approaches. Red boxes highlight the enterotoxin genes which were statistically enriched in isolates from individuals in AD houses versus control houses (one-sided Fisher's exact test, adjusted p-value < 0.05).

Figure 5. Enterotoxins Q and K/K2 mRNAs are expressed by a set of shared *S. aureus* strains

Relative expression of enterotoxin Q (top) or K/K2 (bottom) in a panel of *S. aureus* isolates whose genomes were predicted to harbour these genes, measured by RT-qPCR. The negative control isolate genome does not contain these genes. Enterotoxin expression was normalized to 16S rRNA levels.

Figure 1

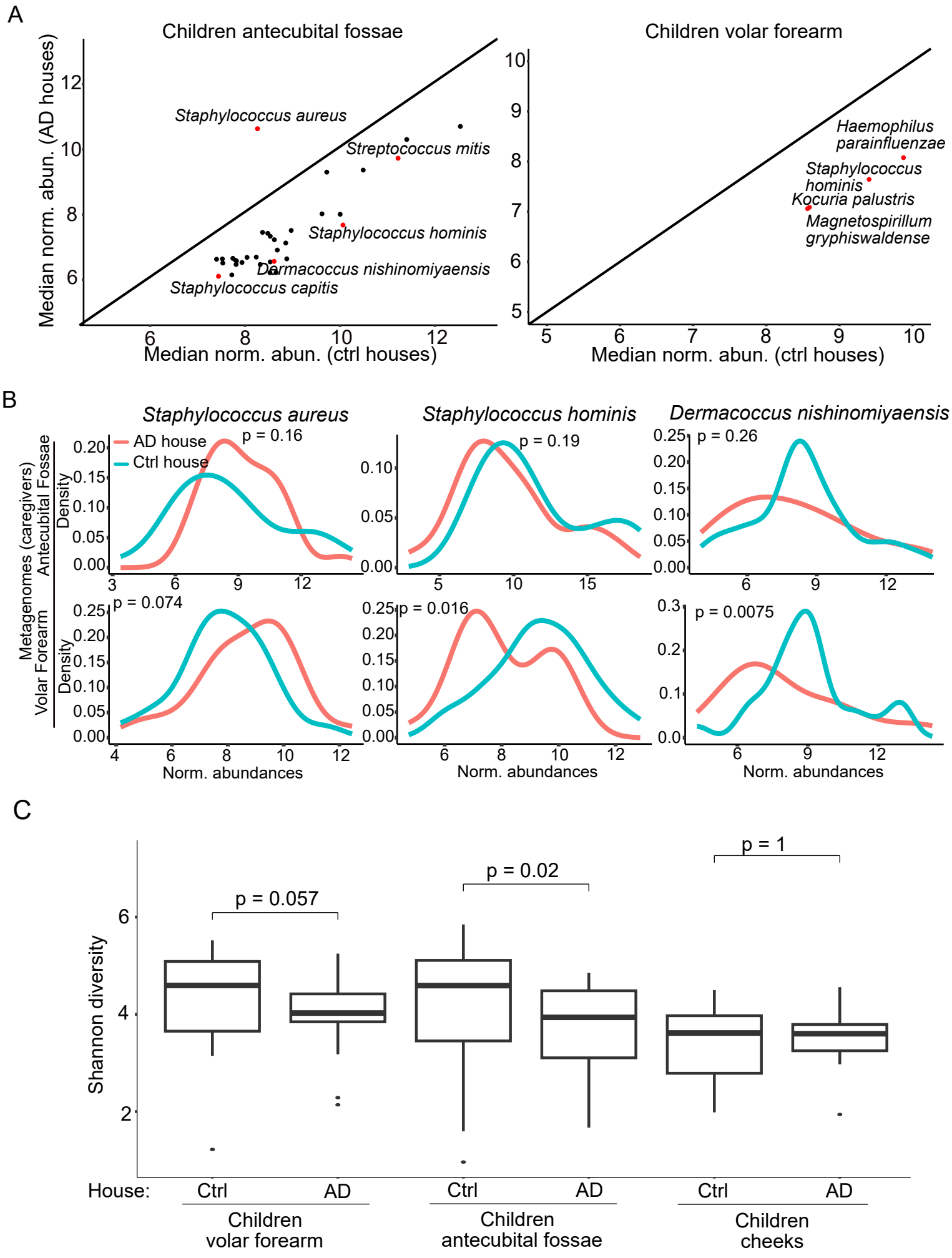
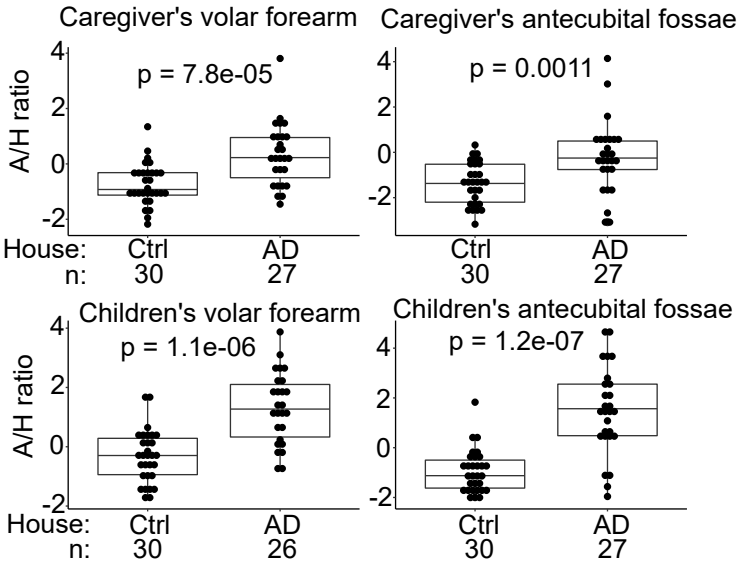

A
Children antecubital fossae
Median norm. abun. (AD houses)
Median norm. abun. (ctrl houses)
Staphylococcus aureus
Streptococcus mitis
Staphylococcus hominis
Dermacoccus nishinomiyaensis
Staphylococcus capitis
Children volar forearm
Haemophilus parainfluenzae
Staphylococcus hominis
Kocuria palustris
Magnetospirillum gryphiswaldense
B
Staphylococcus aureus
Staphylococcus hominis
Dermacoccus nishinomiyaensis
AD house
Ctrl house
p = 0.16
p = 0.19
p = 0.26
p = 0.074
p = 0.016
p = 0.0075
Metagenomes (caregivers)
Antecubital Fossae
Volar Forearm
Density
Norm. abundances
C
Shannon diversity
p = 0.057
p = 0.02
p = 1
House:
Ctrl
AD
Children volar forearm
Children antecubital fossae
Children cheeks

Figure 2

A

Dataset # samples	Antecubital fossae 114	Volar forearm 113
Feature	CSS normalized values (species)	
Accuracy	0.790	0.788
AUC	0.870	0.843
Feature	log(*S.aureus*/*S.hominis*)	
Accuracy	0.769	0.745
AUC	0.813	0.815

B



C

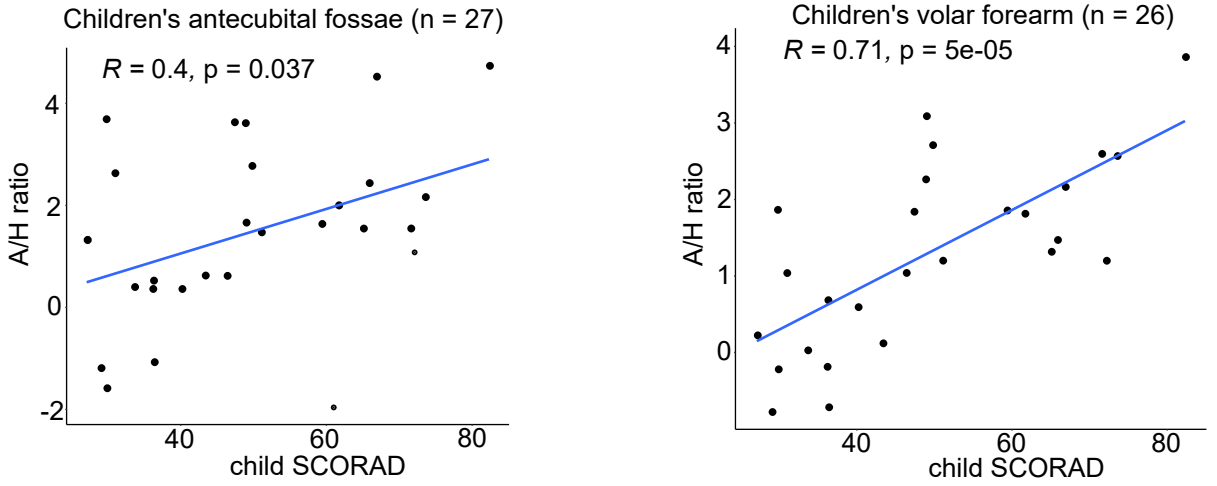


Figure 3

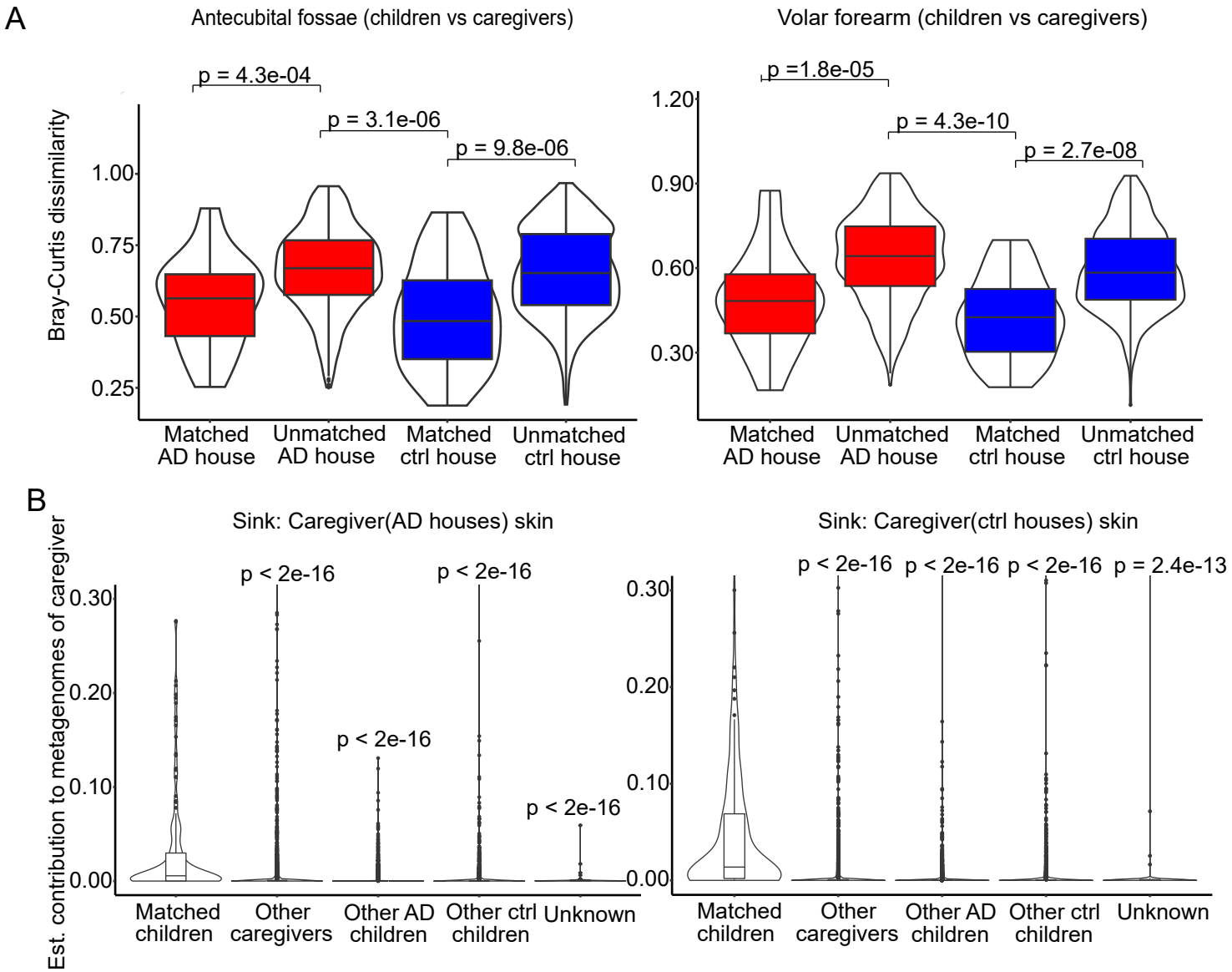

A
Antecubital fossae (children vs caregivers)
Volar forearm (children vs caregivers)
Bray-Curtis dissimilarity
p = 4.3e-04
p = 3.1e-06
p = 9.8e-06
p =1.8e-05
p = 4.3e-10
p = 2.7e-08
Matched AD house
Unmatched AD house
Matched ctrl house
Unmatched ctrl house
B
Sink: Caregiver(AD houses) skin
Sink: Caregiver(ctrl houses) skin
Est. contribution to metagenomes of caregiver
p < 2e-16
p = 2.4e-13
Matched children
Other caregivers
Other AD children
Other ctrl children
Unknown

Figure 4

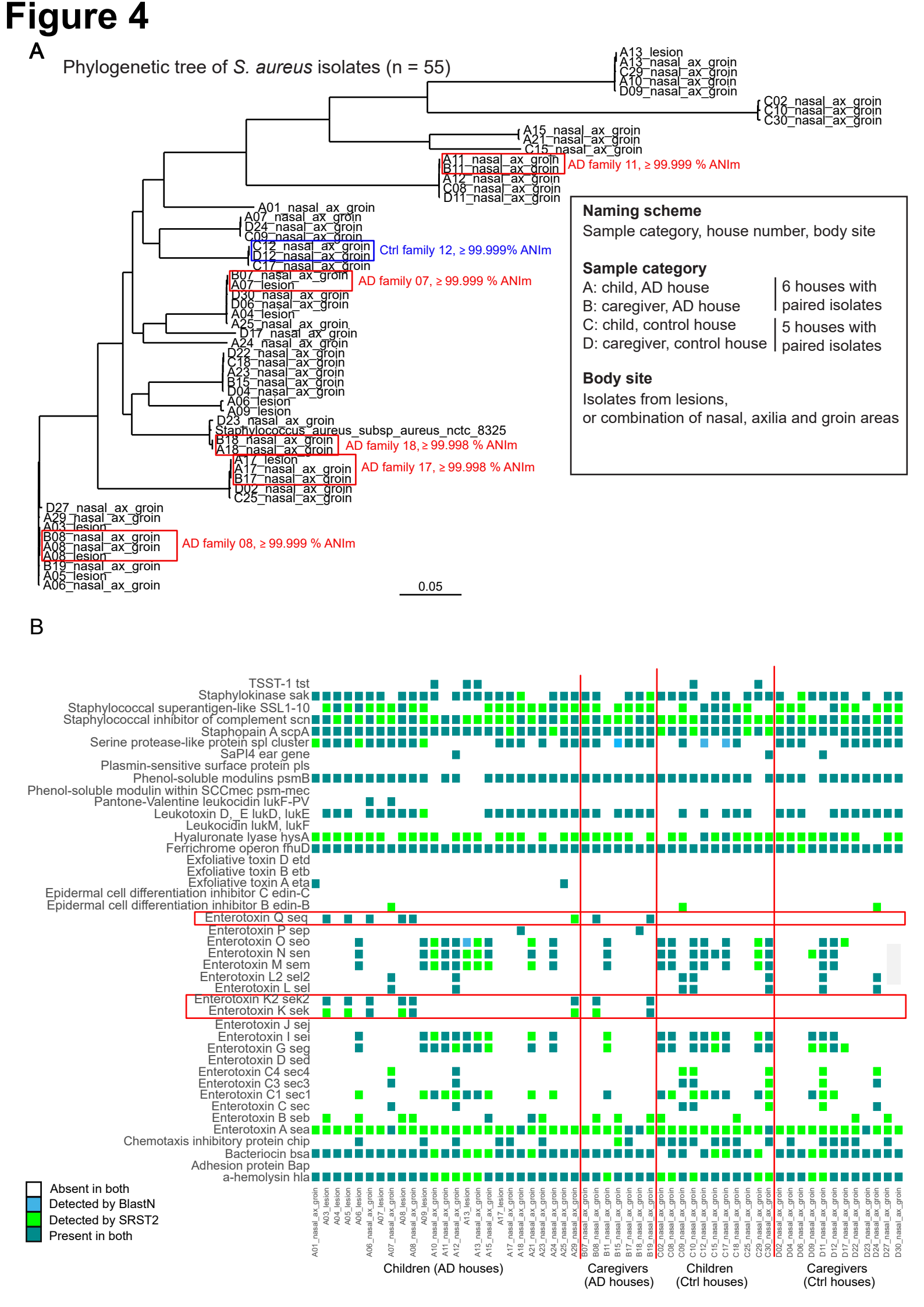
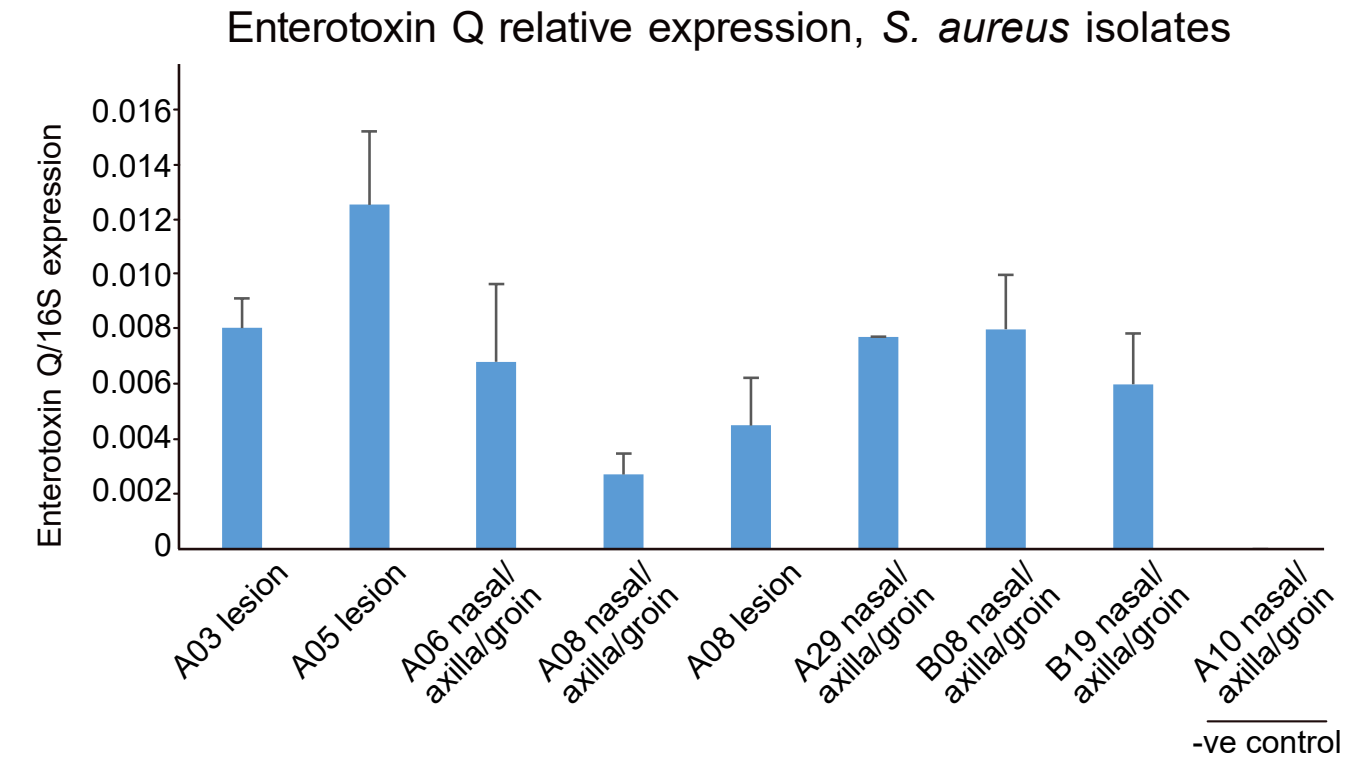

A
Phylogenetic tree of S. aureus isolates (n = 55)
A13_lesion
A13_nasal_ax_groin
C29_nasal_ax_groin
A10_nasal_ax_groin
D09_nasal_ax_groin
C02_nasal_ax_groin
C10_nasal_ax_groin
C30_nasal_ax_groin
A15_nasal_ax_groin
A21_nasal_ax_groin
C15_nasal_ax_groin
A11_nasal_ax_groin
B11_nasal_ax_groin
AD family 11, ≥ 99.999 % ANIm
A12_nasal_ax_groin
C08_nasal_ax_groin
D11_nasal_ax_groin
A01_nasal_ax_groin
A07_nasal_ax_groin
D24_nasal_ax_groin
C09_nasal_ax_groin
C12_nasal_ax_groin
D12_nasal_ax_groin
Ctrl family 12, ≥ 99.999% ANIm
C17_nasal_ax_groin
B07_nasal_ax_groin
A07_lesion
AD family 07, ≥ 99.999 % ANIm
D30_nasal_ax_groin
D06_nasal_ax_groin
A04_lesion
A25_nasal_ax_groin
D17_nasal_ax_groin
A24_nasal_ax_groin
D22_nasal_ax_groin
C18_nasal_ax_groin
A23_nasal_ax_groin
B15_nasal_ax_groin
D04_nasal_ax_groin
A06_lesion
A09_lesion
D23_nasal_ax_groin
Staphylococcus_aureus_subsp_aureus_nctc_8325
B18_nasal_ax_groin
A18_nasal_ax_groin
AD family 18, ≥ 99.998 % ANIm
A17_lesion
A17_nasal_ax_groin
B17_nasal_ax_groin
AD family 17, ≥ 99.998 % ANIm
D02_nasal_ax_groin
C25_nasal_ax_groin
D27_nasal_ax_groin
A29_nasal_ax_groin
A03_lesion
B08_nasal_ax_groin
A08_nasal_ax_groin
A08_lesion
AD family 08, ≥ 99.999 % ANIm
B19_nasal_ax_groin
A05_lesion
A06_nasal_ax_groin
0.05
Naming scheme
Sample category, house number, body site
Sample category
A: child, AD house
B: caregiver, AD house
C: child, control house
D: caregiver, control house
6 houses with paired isolates
5 houses with paired isolates
Body site
Isolates from lesions,
or combination of nasal, axilia and groin areas
B
TSST-1 tst
Staphylokinase sak
Staphylococcal superantigen-like SSL1-10
Staphylococcal inhibitor of complement scn
Staphopain A scpA
Serine protease-like protein spl cluster
SaPI4 ear gene
Plasmin-sensitive surface protein pls
Phenol-soluble modulins psmB
Phenol-soluble modulin within SCCmec psm-mec
Pantone-Valentine leukocidin lukF-PV
Leukotoxin D, E lukD, lukE
Leukocidin lukM, lukF
Hyaluronate lyase hysA
Ferrichrome operon fhuD
Exfoliative toxin D etd
Exfoliative toxin B etb
Exfoliative toxin A eta
Epidermal cell differentiation inhibitor C edin-C
Epidermal cell differentiation inhibitor B edin-B
Enterotoxin Q seq
Enterotoxin P sep
Enterotoxin O seo
Enterotoxin N sen
Enterotoxin M sem
Enterotoxin L2 sel2
Enterotoxin L sel
Enterotoxin K2 sek2
Enterotoxin K sek
Enterotoxin J sej
Enterotoxin I sei
Enterotoxin G seg
Enterotoxin D sed
Enterotoxin C4 sec4
Enterotoxin C3 sec3
Enterotoxin C1 sec1
Enterotoxin C sec
Enterotoxin B seb
Enterotoxin A sea
Chemotaxis inhibitory protein chip
Bacteriocin bsa
Adhesion protein Bap
a-hemolysin hla
Absent in both
Detected by BlastN
Detected by SRST2
Present in both
Children (AD houses)
Caregivers (AD houses)
Children (Ctrl houses)
Caregivers (Ctrl houses)

Figure 5


Enterotoxin Q relative expression, S. aureus isolates
Enterotoxin Q/16S expression
0.016
0.014
0.012
0.010
0.008
0.006
0.004
0.002
0
A03 lesion
A05 lesion
A06 nasal/ axilla/groin
A08 nasal/ axilla/groin
A08 lesion
A29 nasal/ axilla/groin
B08 nasal/ axilla/groin
B19 nasal/ axilla/groin
A10 nasal/ axilla/groin
-ve control

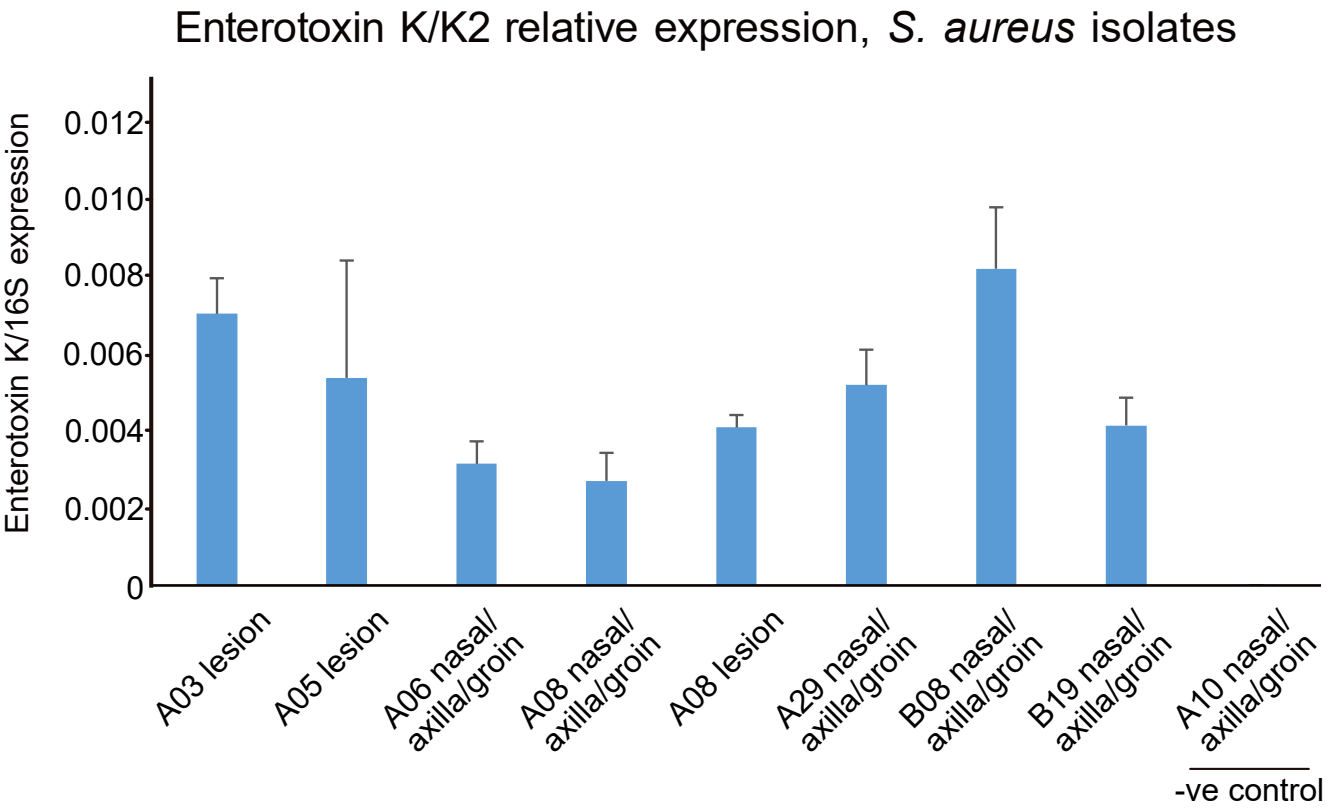
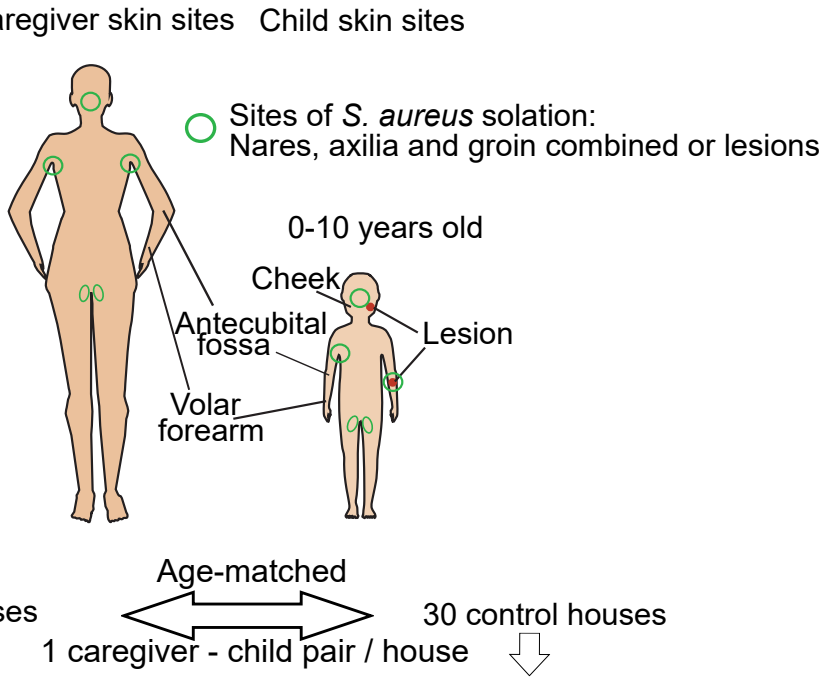

Enterotoxin K/K2 relative expression, S. aureus isolates
Enterotoxin K/16S expression
0.012
0.010
0.008
0.006
0.004
0.002
0
A03 lesion
A05 lesion
A06 nasal/ axilla/groin
A08 nasal/ axilla/groin
A08 lesion
A29 nasal/ axilla/groin
B08 nasal/ axilla/groin
B19 nasal/ axilla/groin
A10 nasal/ axilla/groin
-ve control

Figure S1

A



Age-matched

30 AD houses ⟷ 30 control houses

1 caregiver - child pair / house

Metagenomic sampling experiments

Antecubital fossae:	54/60 non-lesion swabs (27 children + 27 caregivers)	60/60 control swabs (30 children + 30 caregivers)
Volar forearms:	53/60 non-lesion swabs (26 children + 27 caregivers)	60/60 control swabs (30 children + 30 caregivers)
Lesions: (Children only)	26/30 swabs	N/A
Cheeks: (Children < 3 YO)	12/12 swabs	12/12 swabs

Total: 277/294 metagenomes passing QC

***S. aureus* isolation experiments**

Swabbing for strains (lesions):	10/18 swabs	N/A
Swabbing for strains (nasal/axilla/groin):	7/30 swabs from caregivers 17/30 swabs from children	12/30 swabs from caregivers 12/30 swabs from children

Total: 58 *S. aureus* positive cultures/138 swabs

⇩ Assembly and QC

Total: 55 high quality *S. aureus* genomes

⇩ Presence of isolates from different individuals in same house

6 AD caregiver - child pairs with isolates from lesions or nasal/axilla/groin

5 control caregiver - child pairs with isolates from nasal/axilla/groin

Figure S1

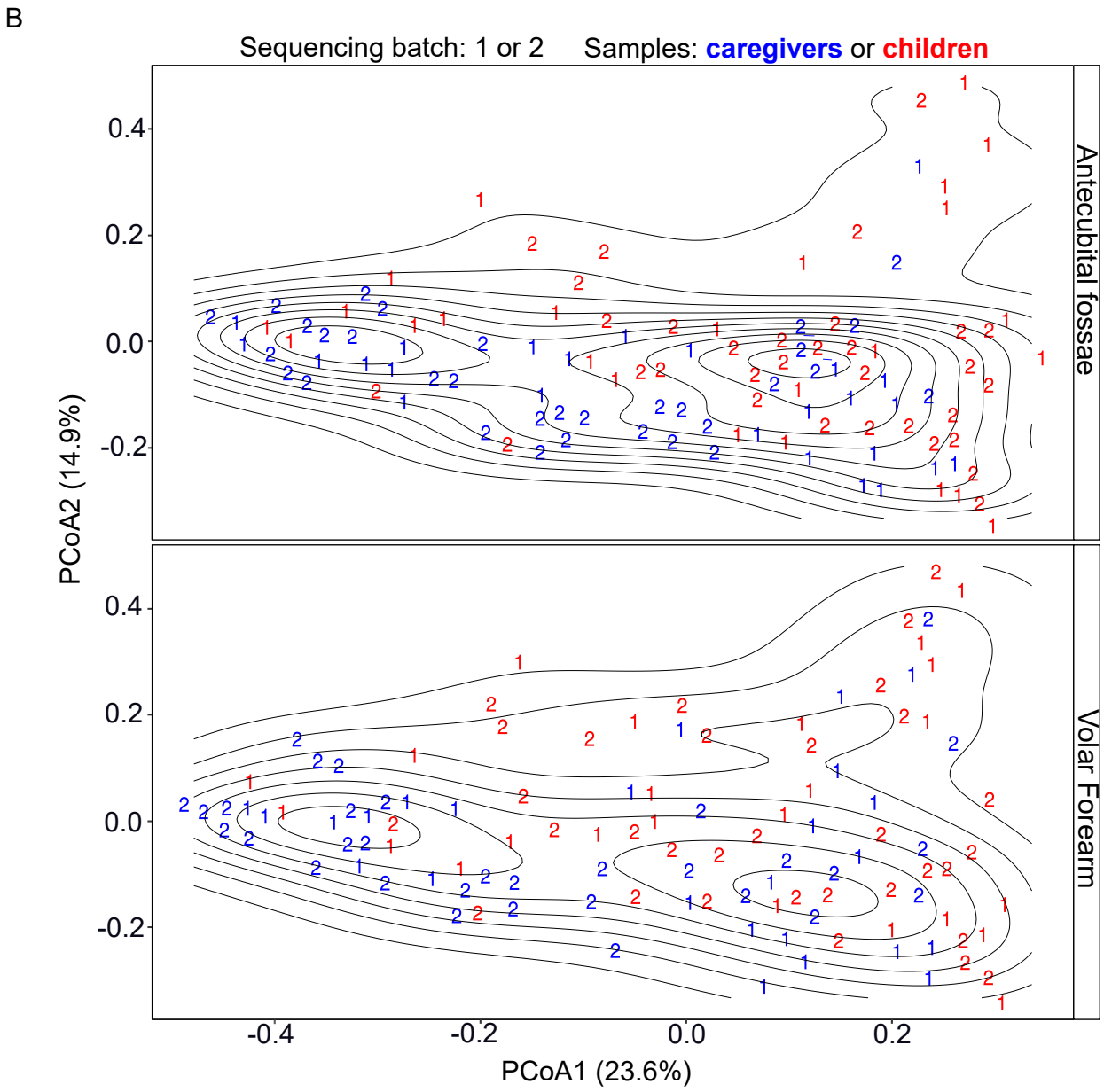

B
Sequencing batch: 1 or 2
Samples: caregivers or children
Antecubital fossae
Volar Forearm
PCoA2 (14.9%)
PCoA1 (23.6%)
0.4
0.2
0.0
-0.2
-0.4

Figure S2

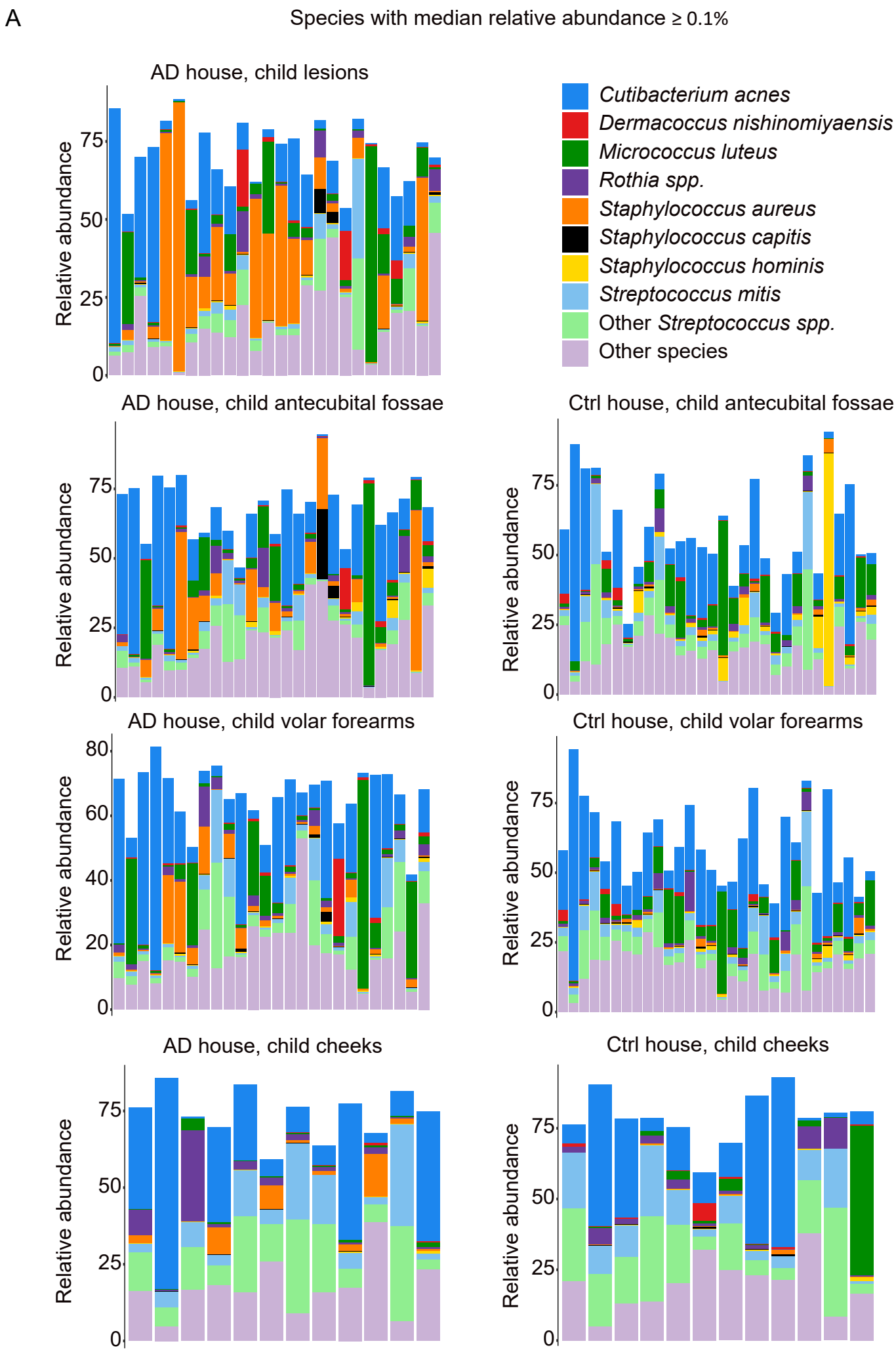

A
Species with median relative abundance ≥ 0.1%
AD house, child lesions
AD house, child antecubital fossae
Ctrl house, child antecubital fossae
AD house, child volar forearms
Ctrl house, child volar forearms
AD house, child cheeks
Ctrl house, child cheeks
Relative abundance
Cutibacterium acnes
Dermacoccus nishinomiyaensis
Micrococcus luteus
Rothia spp.
Staphylococcus aureus
Staphylococcus capitis
Staphylococcus hominis
Streptococcus mitis
Other Streptococcus spp.
Other species

Figure S2

B

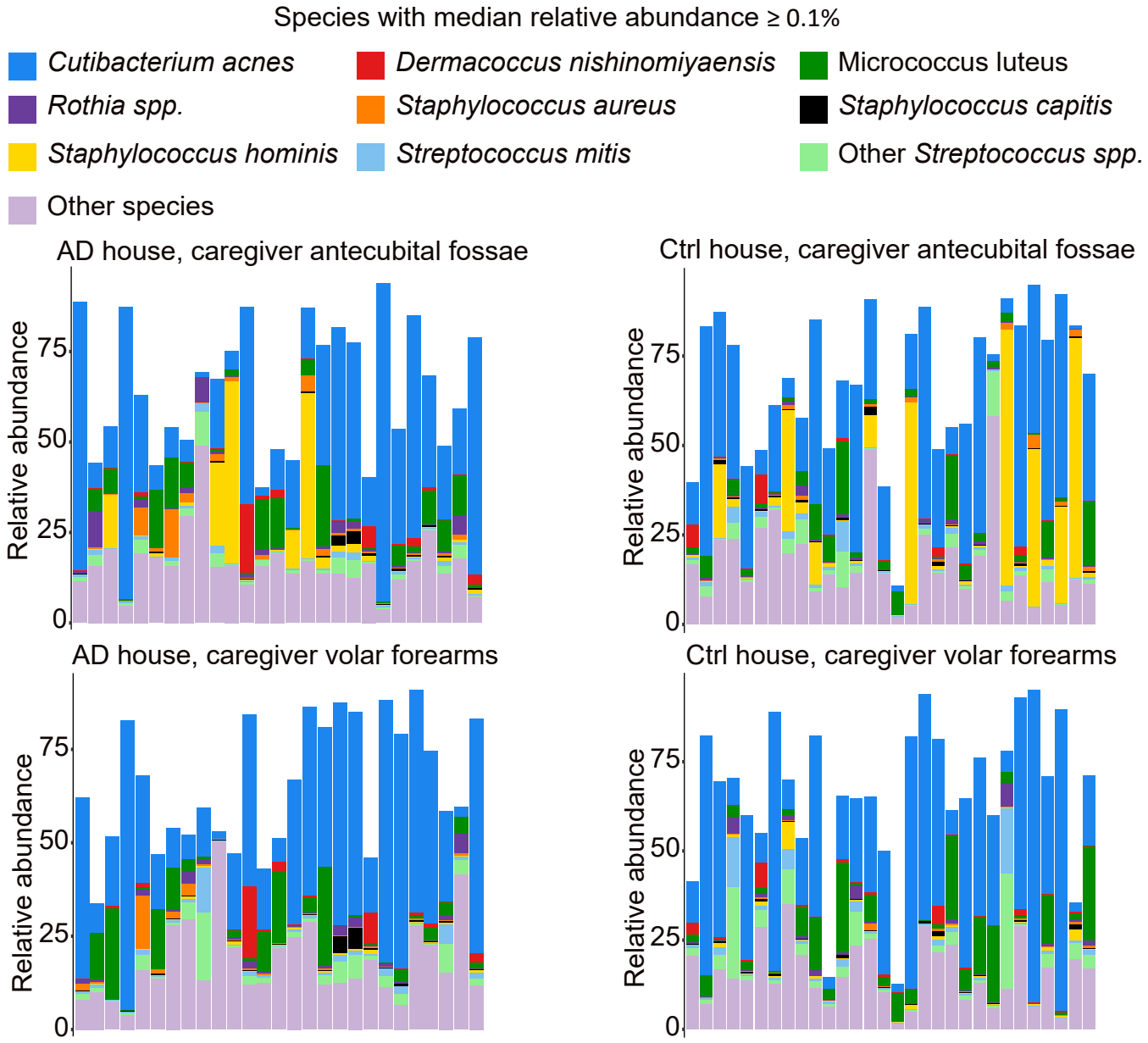

Species with median relative abundance ≥ 0.1%
Cutibacterium acnes
Dermacoccus nishinomiyaensis
Micrococcus luteus
Rothia spp.
Staphylococcus aureus
Staphylococcus capitis
Staphylococcus hominis
Streptococcus mitis
Other Streptococcus spp.
Other species
AD house, caregiver antecubital fossae
Ctrl house, caregiver antecubital fossae
AD house, caregiver volar forearms
Ctrl house, caregiver volar forearms
Relative abundance
0
25
50
75

Figure S2

C

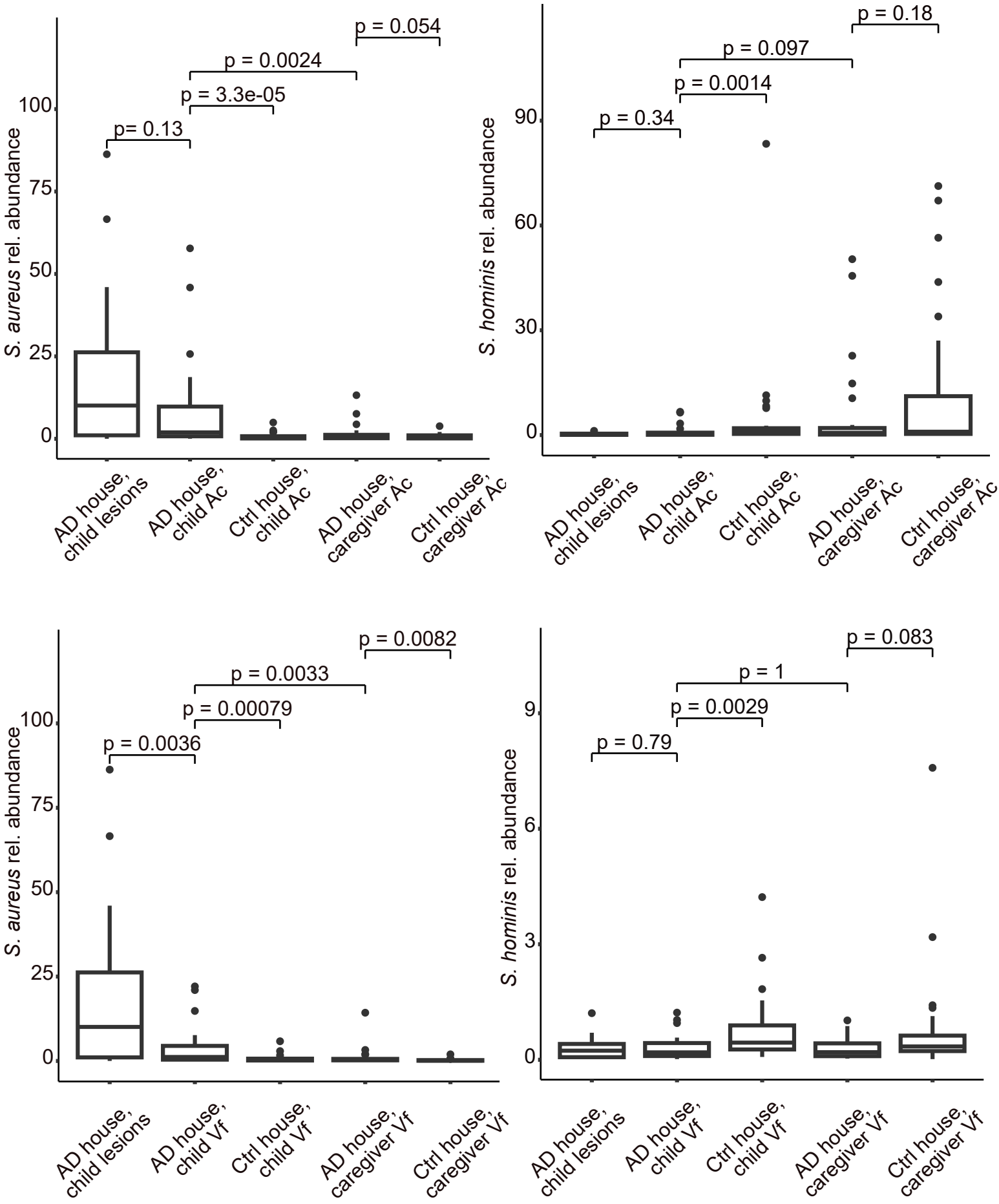

S. aureus rel. abundance
p = 0.054
p = 0.0024
p = 3.3e-05
p= 0.13
100
75
50
25
0
AD house, child lesions
AD house, child Ac
Ctrl house, child Ac
AD house, caregiver Ac
Ctrl house, caregiver Ac
S. hominis rel. abundance
p = 0.18
p = 0.097
p = 0.0014
p = 0.34
90
60
30
0
AD house, child lesions
AD house, child Ac
Ctrl house, child Ac
AD house, caregiver Ac
Ctrl house, caregiver Ac
S. aureus rel. abundance
p = 0.0082
p = 0.0033
p = 0.00079
p = 0.0036
100
75
50
25
0
AD house, child lesions
AD house, child Vf
Ctrl house, child Vf
AD house, caregiver Vf
Ctrl house, caregiver Vf
S. hominis rel. abundance
p = 0.083
p = 1
p = 0.0029
p = 0.79
9
6
3
0
AD house, child lesions
AD house, child Vf
Ctrl house, child Vf
AD house, caregiver Vf
Ctrl house, caregiver Vf

Figure S2

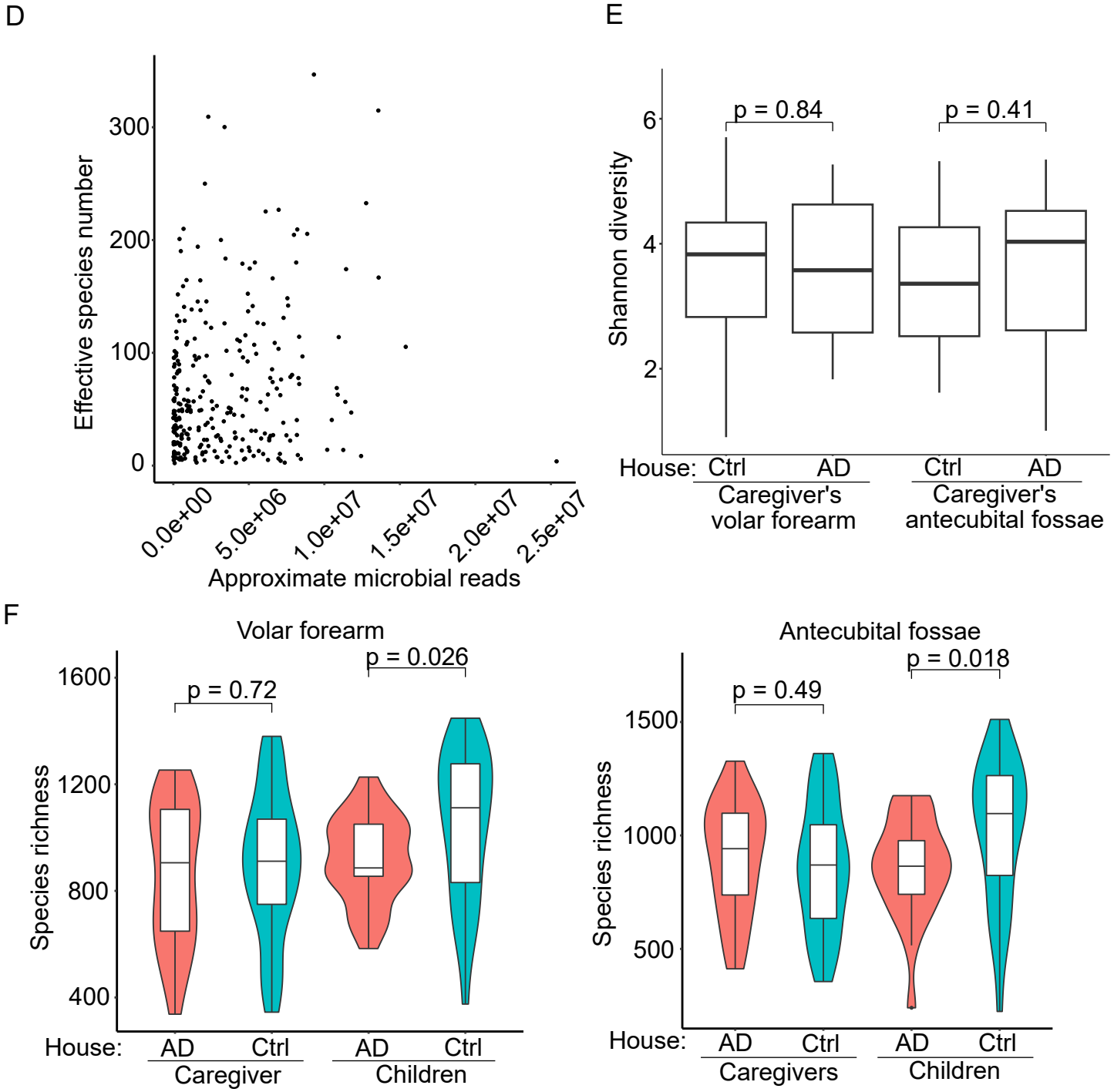
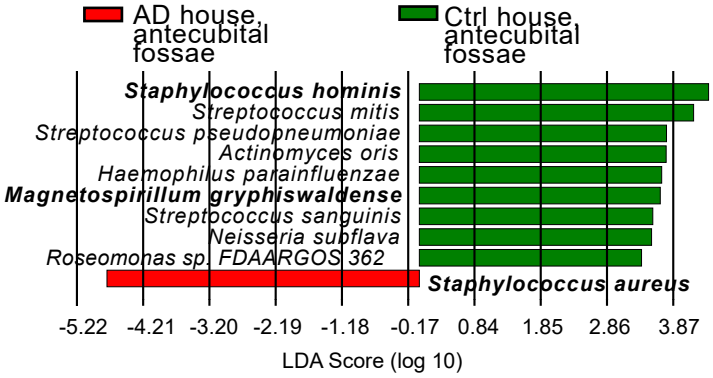
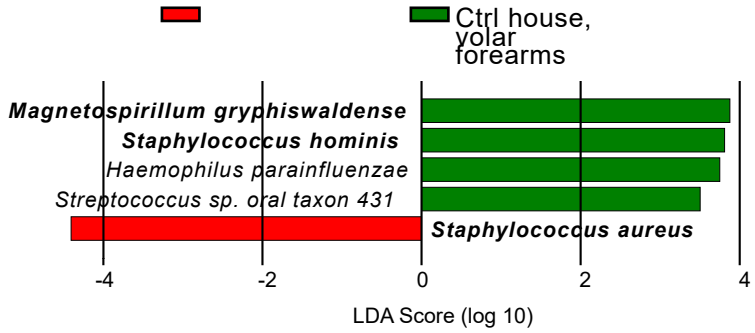

D
Effective species number
300
200
100
0
0.0e+00
5.0e+06
1.0e+07
1.5e+07
2.0e+07
2.5e+07
Approximate microbial reads
E
Shannon diversity
6
4
2
p = 0.84
p = 0.41
House: Ctrl AD Ctrl AD
Caregiver's volar forearm
Caregiver's antecubital fossae
F
Volar forearm
p = 0.72
p = 0.026
Species richness
1600
1200
800
400
House: AD Ctrl AD Ctrl
Caregiver
Children
Antecubital fossae
p = 0.49
p = 0.018
Species richness
1500
1000
500
House: AD Ctrl AD Ctrl
Caregivers
Children

Figure S3

Children metagenomes, AD non-lesioned skin vs control

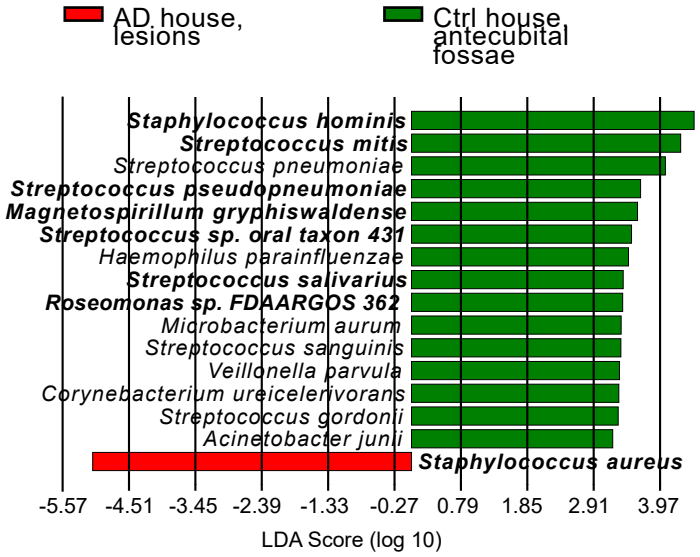




Children metagenomes, lesions vs non-lesioned skin

No statistically significant features
for lesions vs. non-lesioned Ac or Vf by LEfSe
after adjusting for multiple comparisons (FDR)

Children metagenomes, lesions vs control skin



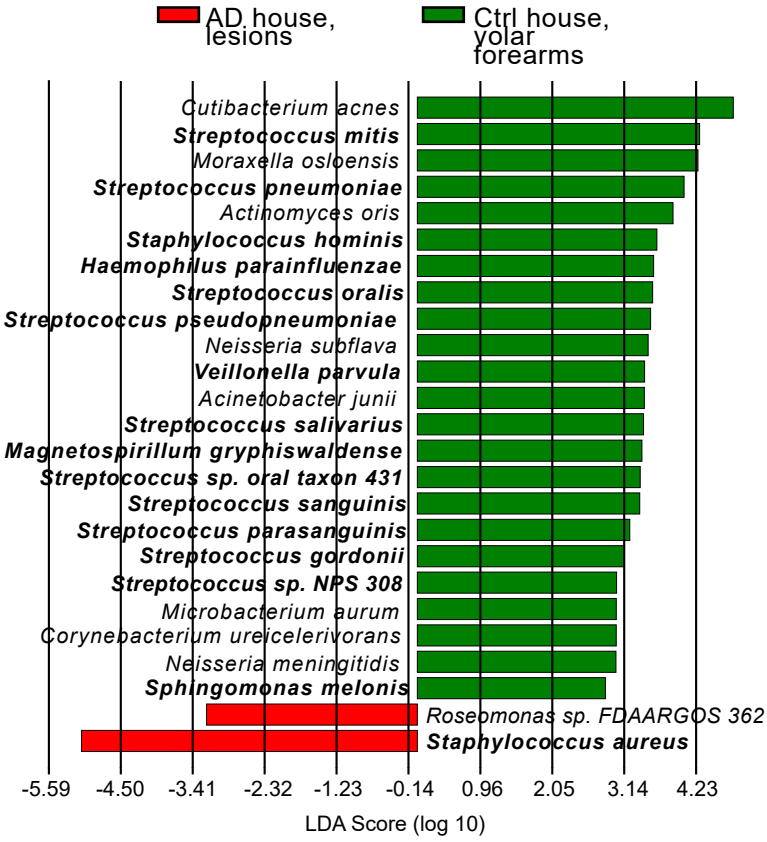


Figure S4

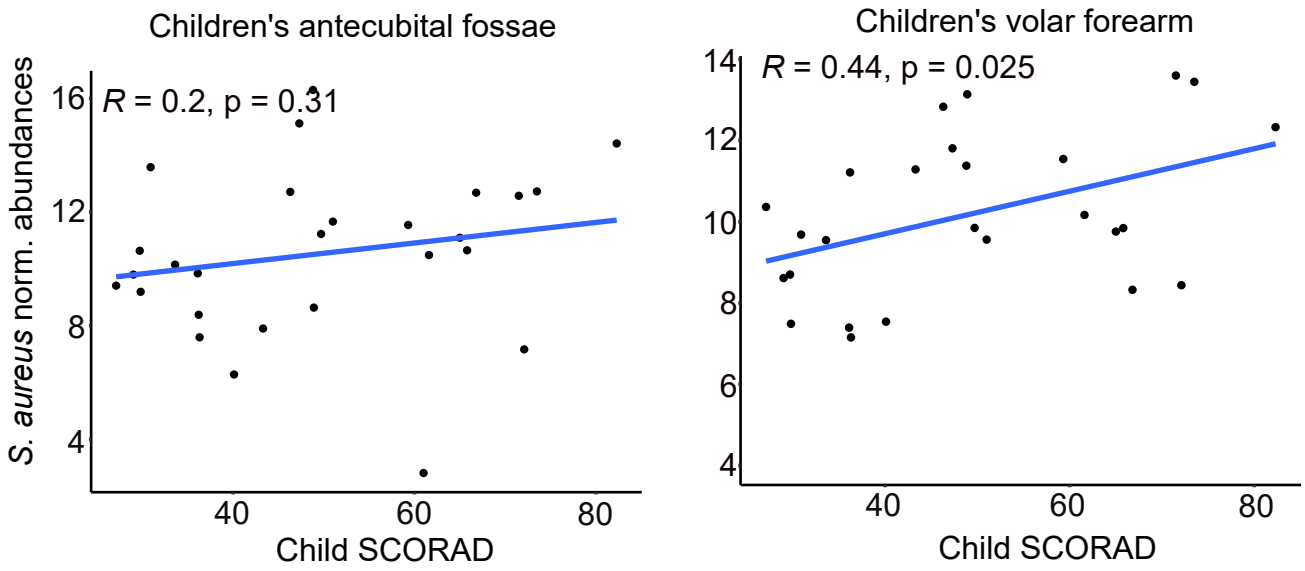
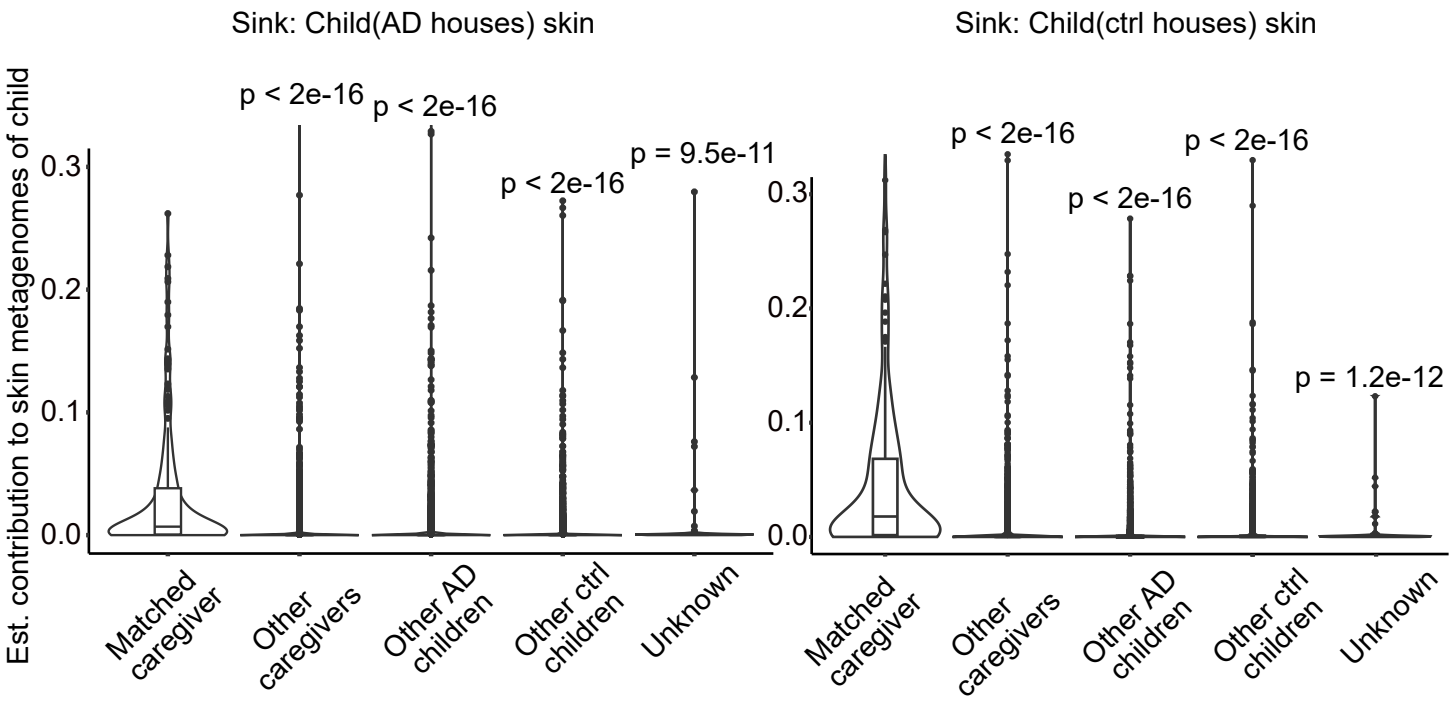

Children's antecubital fossae
R = 0.2, p = 0.31
S. aureus norm. abundances
16
12
8
4
40
60
80
Child SCORAD
Children's volar forearm
R = 0.44, p = 0.025
14
12
10
8
6
4
40
60
80
Child SCORAD

Figure S5


Sink: Child(AD houses) skin
Sink: Child(ctrl houses) skin
Est. contribution to skin metagenomes of child
0.0
0.1
0.2
0.3
p < 2e-16
p < 2e-16
p < 2e-16
p = 9.5e-11
p < 2e-16
p < 2e-16
p < 2e-16
p = 1.2e-12
Matched caregiver
Other caregivers
Other AD children
Other ctrl children
Unknown

Supplementary table 1. Sequences of oligonucleotides or primers used in this study

Oligo name	Sequence
Barcode_Adapter_1	P-GAT CGG AAG AGC ACA CGT CT
Barcode_Adapter_2	ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT
MP_PCR_HE_1_0_	AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC* T
MP_PCR_Index1	CAA GCA GAA GAC GGC ATA CGA GAT **CGT GAT** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index2	CAA GCA GAA GAC GGC ATA CGA GAT **ACA TCG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index3	CAA GCA GAA GAC GGC ATA CGA GAT **GCC TAA** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index4	CAA GCA GAA GAC GGC ATA CGA GAT **TGG TCA** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index5	CAA GCA GAA GAC GGC ATA CGA GAT **CAC TGT** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index6	CAA GCA GAA GAC GGC ATA CGA GAT **ATT GGC** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index7	CAA GCA GAA GAC GGC ATA CGA GAT **GAT CTG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index8	CAA GCA GAA GAC GGC ATA CGA GAT **TCA AGT** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index9	CAA GCA GAA GAC GGC ATA CGA GAT **CTG ATC** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index10	CAA GCA GAA GAC GGC ATA CGA GAT **AAG CTA** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index11	CAA GCA GAA GAC GGC ATA CGA GAT **GTA GCC** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index12	CAA GCA GAA GAC GGC ATA CGA GAT **TAC AAG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index13	CAA GCA GAA GAC GGC ATA CGA GAT **TTG ACT** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index14	CAA GCA GAA GAC GGC ATA CGA GAT **GGA ACT** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index15	CAA GCA GAA GAC GGC ATA CGA GAT **TGA CAT** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index16	CAA GCA GAA GAC GGC ATA CGA GAT **GGA CGG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index17	CAA GCA GAA GAC GGC ATA CGA GAT **CTC TAC** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index18	CAA GCA GAA GAC GGC ATA CGA GAT **GCG GAC** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T

MP_PCR_Index19	CAA GCA GAA GAC GGC ATA CGA GAT **TTT CAC** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index20	CAA GCA GAA GAC GGC ATA CGA GAT **GGC CAC** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index21	CAA GCA GAA GAC GGC ATA CGA GAT **CGA AAC** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index22	CAA GCA GAA GAC GGC ATA CGA GAT **CGT ACG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index23	CAA GCA GAA GAC GGC ATA CGA GAT **CCA CTC** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index24	CAA GCA GAA GAC GGC ATA CGA GAT **GCT ACC** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index25	CAA GCA GAA GAC GGC ATA CGA GAT **ATC AGT** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index26	CAA GCA GAA GAC GGC ATA CGA GAT **GCT CAT** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index27	CAA GCA GAA GAC GGC ATA CGA GAT **AGG AAT** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index28	CAA GCA GAA GAC GGC ATA CGA GAT **CTT TTG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index29	CAA GCA GAA GAC GGC ATA CGA GAT **TAG TTG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index30	CAA GCA GAA GAC GGC ATA CGA GAT **CCG GTG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index31	CAA GCA GAA GAC GGC ATA CGA GAT **ATC GTG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index32	CAA GCA GAA GAC GGC ATA CGA GAT **TGA GTG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index33	CAA GCA GAA GAC GGC ATA CGA GAT **CGC CTG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index34	CAA GCA GAA GAC GGC ATA CGA GAT **GCC ATG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index35	CAA GCA GAA GAC GGC ATA CGA GAT **AAA ATG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index36	CAA GCA GAA GAC GGC ATA CGA GAT **TGT TGG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index37	CAA GCA GAA GAC GGC ATA CGA GAT **ATT CCG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index38	CAA GCA GAA GAC GGC ATA CGA GAT **AGC TAG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index39	CAA GCA GAA GAC GGC ATA CGA GAT **GTA TAG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index40	CAA GCA GAA GAC GGC ATA CGA GAT **TCT GAG** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index41	CAA GCA GAA GAC GGC ATA CGA GAT **GTC GTC** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T

MP_PCR_Index42	CAA GCA GAA GAC GGC ATA CGA GAT **CGA TTA** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index43	CAA GCA GAA GAC GGC ATA CGA GAT **GCT GTA** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index44	CAA GCA GAA GAC GGC ATA CGA GAT **ATT ATA** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index45	CAA GCA GAA GAC GGC ATA CGA GAT **GAA TGA** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index46	CAA GCA GAA GAC GGC ATA CGA GAT **TCG GGA** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index47	CAA GCA GAA GAC GGC ATA CGA GAT **CTT CGA** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
MP_PCR_Index48	CAA GCA GAA GAC GGC ATA CGA GAT **TGC CGA** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC* T
16S_fwd	ACTCCTACGGGAGGCAGC
16S_rv	TTACCGCGGCTGCTGGCAC
Enterotoxin_Q_fwd	TGGTGGAATTACGTTGGCAA
Enterotoxin_Q_rv	TGCTTACCATTGACCCAGAGA
Enterotoxin_K_fwd	CCCATCATCTCCTGTGTAGAAT
Enterotoxin_K_rv	CCGCTCAAGAGATTGATGTCA
16s gBLOCK	GGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTC GGCAATGGACGGAAGTCTGACCGAGCAACGCCGCGTGA G TGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTAAGAGA A GAACGAGTGTGAGAGTGGAAAGTTCACACTGTGACGGTA T CTTACCAGAAAGGGACGGCTAACTACGTGCCAGCAGCCG C GGTAATACGTAGGTCCCGAG

The barcode adapters and all oligos whose names begin with “MP” were modified at the 3'-terminus with a phosphorothioate internucleotide linkage and were HPLC purified. Index sequences are bolded.

Supplementary table 2. List of differentially abundant species between individuals in AD or control houses

Dataset: Volar forearm metagenomes of children (n = 26 AD children, n = 30 control children). Features are cumulative sum scaled read counts.

Species	**p-value**	**Adjusted p-value**	**Cliff’s delta**	**AD median**	**AD IQR**	**Control median**	**Control IQR**
Haemophilus parainfluenzae	2.41E-03	3.29E-02	-0.47	8.08	2.50	9.88	1.86
Kocuria palustris	3.40E-03	3.49E-02	-0.45	7.09	2.23	8.59	1.81

Magnetospirillum gryphiswaldense	5.33E-04	1.09E-02	-0.53	7.06	1.64	8.59	1.63
Staphylococcus hominis	2.48E-04	1.02E-02	-0.56	7.64	1.98	9.41	2.04

Dataset: Antecubital fossae metagenomes of children (n = 27 AD children, n = 30 control children). Features are cumulative sum scaled read counts.

Species	**p-value**	**Adjusted p-value**	**Cliff's delta**	**AD median**	**AD IQR**	**Control median**	**Control IQR**
Staphylococcus aureus	2.29E-03	6.72E-03	0.46	10.64	3.71	8.26	2.29
Pseudomonas aeruginosa	3.08E-02	3.72E-02	-0.33	6.59	2.44	7.88	1.68
Veillonella parvula	2.72E-02	3.38E-02	-0.34	7.23	4.10	8.62	2.36
Stenotrophomonas maltophilia	2.29E-02	2.94E-02	-0.35	7.45	2.52	8.37	2.86
Sphingomonas melonis	2.10E-02	2.78E-02	-0.36	6.65	2.65	7.76	1.88
Streptococcus parasanguinis	1.84E-02	2.51E-02	-0.36	6.51	3.05	7.53	2.14
Moraxella osloensis	1.61E-02	2.27E-02	-0.37	10.31	2.36	11.41	1.67
Staphylococcus capitis	1.27E-02	1.87E-02	-0.38	6.10	2.91	7.45	2.36
Neisseria mucosa	1.11E-02	1.68E-02	-0.39	6.46	3.61	8.34	2.43
Streptococcus sp. NPS 308	1.00E-02	1.58E-02	-0.40	6.63	3.26	7.40	1.63
Neisseria elongata	7.85E-03	1.29E-02	-0.41	6.64	2.96	7.54	2.02
Streptococcus gordonii	7.47E-03	1.28E-02	-0.41	6.57	3.79	7.88	1.97
Streptococcus oralis	7.47E-03	1.28E-02	-0.41	9.30	3.61	9.75	1.41
Rothia mucilaginosa	6.75E-03	1.26E-02	-0.41	8.02	2.83	9.62	1.76
Streptococcus pneumoniae	6.75E-03	1.26E-02	-0.41	9.37	3.18	10.53	1.25
Neisseria sicca	6.41E-03	1.26E-02	-0.42	7.42	3.08	8.50	2.41
Streptococcus sp. oral taxon 431	4.45E-03	9.60E-03	-0.43	6.68	2.89	8.09	2.28
Dermacoccus nishinomiyaensis	3.79E-03	8.63E-03	-0.44	6.56	2.68	8.61	2.55
Acinetobacter junii	3.59E-03	8.63E-03	-0.44	6.47	2.07	7.81	1.49

Streptococcus salivarius	3.59E-03	8.63E-03	-0.44	6.22	3.45	8.55	2.10
Streptococcus pseudopneumoniae	3.21E-03	8.63E-03	-0.45	7.33	2.67	8.58	1.28
Microbacterium aurum	1.82E-03	6.22E-03	-0.47	6.69	2.33	8.24	2.47
Neisseria meningitidis	2.21E-03	6.72E-03	-0.47	6.22	2.86	8.66	2.14
Micrococcus luteus	1.44E-03	5.36E-03	-0.48	10.71	2.88	12.53	2.87
Streptococcus mitis	1.44E-03	5.36E-03	-0.48	9.73	2.60	11.24	1.12
Corynebacterium ureicelerivorans	1.35E-03	5.36E-03	-0.49	6.15	2.36	7.72	1.87
Neisseria subflava	1.35E-03	5.36E-03	-0.49	6.91	2.90	8.70	2.44
Actinomyces oris	6.84E-04	4.01E-03	-0.51	8.01	3.12	10.00	2.07
Kocuria palustris	6.41E-04	4.01E-03	-0.52	6.58	1.85	7.94	2.05
Streptococcus sanguinis	2.34E-04	1.92E-03	-0.55	7.13	2.82	8.86	1.94
Haemophilus parainfluenzae	4.19E-05	4.30E-04	-0.61	7.51	3.46	9.03	2.08
Roseomonas sp. FDAARGOS 362	3.29E-05	4.30E-04	-0.62	6.64	2.53	8.89	1.65
Magnetospirillum gryphiswaldense	1.65E-05	3.38E-04	-0.67	6.54	6.62	8.53	1.63
Staphylococcus hominis	3.39E-06	1.39E-04	-0.68	7.68	3.36	10.06	2.14

Dataset: Volar forearm metagenomes of caregivers (n = 27 caregivers from AD houses, n = 30 caregivers from control houses). Features are cumulative sum scaled read counts.

Species	**p-value**	**Adjusted p-value**
No statistically significant results after adjustment for multiple comparisons		

Dataset: Antecubital fossae metagenomes of caregivers (n = 27 caregivers from AD houses, n = 30 caregivers from control houses). Features are cumulative sum scaled read counts.

Species	**p-value**	**Adjusted p-value**
No statistically significant results after adjustment for multiple comparisons		

All relative abundances were normalized using cumulative sum scaling (CSS) and very rare taxa were filtered out on the basis of very low median relative abundance (< 0.1%). Only statistically significant results are shown (Wilcoxon rank sum test, adjusted p-value < 0.05). Features are sorted by effect size (Cliff's delta). A Cliff's delta score closer to +1 represents larger effects differences of the given species in AD vs control individuals. Conversely, a Cliff's delta score closer to -1 represents larger effects and differences of the given species in control vs AD individuals. Details about the cumulative sum scaling (CSS) normalization method can be found in the methods section of this manuscript.

Supplementary Table 3. Regression table for metagenomic features classifying individuals (both caregivers AND children) as belonging to AD or control houses

Dataset: Volar forearm metagenomes (n = 53 individuals from AD houses and n = 60 individuals from control houses). Features are species level CSS normalized read counts.					
Feature	**Estimate**	**Std. Error**	**z value**	**p-value**	**Adj. p-value**
(Intercept)	18.61	5.96	3.12	1.79E-03	**8.33E-03**
Streptococcus pseudopneumoniae	4.33	1.55	2.80	5.18E-03	**1.45E-02**
Staphylococcus epidermidis	0.72	0.32	2.25	2.43E-02	**3.78E-02**
Staphylococcus aureus	0.70	0.20	3.54	4.01E-04	**2.80E-03**
Sphingomonas melonis	0.64	0.26	2.46	1.40E-02	**2.79E-02**
Rothia dentocariosa	0.26	0.22	1.20	2.31E-01	2.94E-01
Veillonella parvula	0.05	0.30	0.18	8.55E-01	8.55E-01
Acinetobacter junii	-0.05	0.17	-0.28	7.83E-01	8.43E-01
Streptococcus parasanguinis	-0.14	0.28	-0.51	6.10E-01	7.11E-01
Dermacoccus nishinomiyaensis	-0.29	0.14	-2.05	4.06E-02	5.68E-02
Propionibacterium sp. oral taxon 193	-0.31	0.12	-2.56	1.04E-02	**2.44E-02**
Stenotrophomonas maltophilia	-0.60	0.26	-2.29	2.22E-02	**3.78E-02**
Staphylococcus hominis	-1.42	0.35	-4.07	4.64E-05	**6.49E-04**
Streptococcus pneumoniae	-4.93	1.65	-2.98	2.86E-03	**1.00E-02**
Dataset: Antecubital fossae metagenomes (n = 54 individuals from AD houses and n = 60 individuals from control houses). Features are species level CSS normalized read counts.					
Feature	**Estimate**	**Std. Error**	**z value**	**p-value**	**Adj. p-value**
(Intercept)	11.37	4.58	2.48	1.31E-02	**3.00E-02**

Streptococcus sp. NPS 308	3.20	1.55	2.07	3.83E-02	6.12E-02
Pseudomonas putida	1.30	0.36	3.59	3.34E-04	**1.78E-03**
Rothia dentocariosa	1.16	0.40	2.94	3.33E-03	**9.52E-03**
Staphylococcus aureus	0.92	0.26	3.60	3.23E-04	**1.78E-03**
Veillonella parvula	0.76	0.46	1.64	1.01E-01	1.47E-01
Kocuria palustris	0.22	0.21	1.07	2.87E-01	3.82E-01
Sphingomonas melonis	0.11	0.21	0.52	6.00E-01	6.65E-01
Staphylococcus epidermidis	-0.10	0.30	-0.34	7.37E-01	7.37E-01
Streptococcus pneumoniae	-0.28	0.58	-0.49	6.24E-01	6.65E-01
Streptococcus sp. oral taxon 431	-0.30	0.60	-0.50	6.19E-01	6.65E-01
Propionibacterium sp. oral taxon 193	-0.35	0.16	-2.20	2.76E-02	**4.91E-02**
Staphylococcus hominis	-0.73	0.18	-4.08	4.57E-05	**7.32E-04**
Stenotrophomonas maltophilia	-1.26	0.41	-3.10	1.97E-03	**7.86E-03**
Streptococcus gordonii	-1.86	0.64	-2.91	3.57E-03	**9.52E-03**
Streptococcus oralis	-3.15	1.36	-2.32	2.05E-02	**4.11E-02**

Bolded features denotes statistical significance ($p < 0.05$ after adjustment for multiple comparisons). Rows are arranged in order of the magnitude of the estimated coefficients. Details about the cumulative sum scaling (CSS) normalization method can be found in the methods section of this manuscript.

Figure S1. Community-based measures of skin microbiomes do not show obvious sample preparation and sequencing batch effects

(A) Study design and the body sites from which metagenomic samples and *S. aureus* strains were sampled from.

(B) Principal coordinates analysis (PCoA) diagram using species level Bray-Curtis dissimilarity, showing that skin microbiomes from different sample preparation and sequencing batches do not form distinct clusters. Numbers in blue and red represent skin metagenomes from caregivers and children respectively. Libraries from batch 1 & 2 were constructed with different library preparation kits and sequenced in different months.

Figure S2. Microbial diversity is reduced in AD children but not in their caregivers.

(A) Relative abundance plots of selected skin microbes in different sites for children suffering from atopic dermatitis (left column) and for control children (right column).

Columns do not sum to 100% because very rare taxa (median relative abundance < 0.1% across libraries) were not plotted.
(B) Relative abundance plots of selected skin microbes in different sites for clinically healthy caregivers caring for paediatric AD patients (left column) or caring for control children (right column). Columns do not sum to 100% because very rare taxa were not plotted.
(C) Distribution of relative abundance of *S. aureus* or *S. hominis* between different categories of metagenomes. The p-values for Wilcoxon rank sum tests are shown.
(D) Effective species number ($e^{\text{Shannon diversity index}}$) against estimated microbial read counts in the metagenomic dataset (n = 277).
(E) Shannon diversities for metagenomes of caregivers for the volar forearm (n = 57) and the antecubital fossae (n = 57). The p-values for Wilcoxon rank sum tests are shown.
(F) Violin plots showing species richness (≥ 1 read) detected in all metagenomes (n = 277), rarefied to the same number of reads (> 12000). The p-values for Wilcoxon rank sum tests are shown.

Figure S3. Linear discriminant analysis effect size (LefSe) for species-level relative abundances in metagenomes from children for various comparisons
Very rare taxa (median abundance < 0.1%) were excluded from this analysis.
Top left: Non-lesional antecubital fossae (Ac) (n = 27) vs control Ac (n = 30).
Top right: Non-lesional volar forearms (Vf) (n =26) vs control Vf (n = 30).
Middle: Lesions (n = 26) vs non-lesional Ac (n = 27) or Vf (n = 26) in AD children.
Bottom left: Lesions in AD children (n = 26) vs control Ac in healthy children (n = 30).
Bottom right: Lesions in AD children (n = 26) vs control Vf in healthy children (n = 30).
All features shown here were statistically significant features ($p < 0.05$) reported by the LefSe algorithm, before adjusting for multiple comparisons. Features highlighted in bold were statistically significant features after adjusting for multiple comparisons (FDR < 0.05)

Figure S4. Normalized *S. aureus* abundances are only weakly correlated with SCORAD
Correlation of normalized *S. aureus* abundances from the metagenomes of non-lesional skin of AD children, against their SCORAD values (n = 27 for Ac, n = 26 for Vf) . The Pearson correlation coefficient *R*, and the associated p-values are shown. The blue line indicates the regression line.

Figure S5. Caregivers have a higher contribution to skin microbiomes of children from the same house, relative to sources from other houses
Estimated contributions of different sources (horizontal axis) to the skin metagenomes of children. Each metagenome (Vf, Ac, cheeks and lesions) from a child was defined as a sink and other metagenomes in this study were split into distinct groups: matched caregiver from same house, other caregivers from different houses, other children with AD from different houses, and other healthy children from different houses. The contribution of unknown sources to each sink was also included in these models.

Differences between the “matched caregiver” sources and other sources were tested, and p-values for Wilcoxon rank sum tests are shown. For this analysis, n = 91 sinks from AD houses and n = 72 sinks from control houses.