Supplementary Tables

Tables S1-S4. Benchmark datasets genome information and statistics.

Genomes used for benchmarking downloaded on: *Bifidobacterium*: 2017 Jan., *Escherichia / Shigella*: 2018 Feb., *P. acnes*: 2016 Jun., *Staphylococcus*: 2016 Apr. Selection of genomes was based on all genomes publicly available at NCBI assembled genomes and NCBI Traces WGS databases [1]. For *Staphylococcus* and *Escherichia / Shigella* a maximum of 200 genomes was selected in order to allow for practically feasible computation times. Therefore, careful effort was undertaken to take into account all available biodiversity based on provided strain annotations. High-quality genomes were first selected (i.e. those assigned as *'reference'* or *'representative'*, thereafter those with closed genomes assemblies or with the lowest number of contigs / scaffolds).

File :
S1-S4 Datasets-for-Benchmarking.xlsx (online) http://ederveen.science/Thesis/Chapter3/Table-S1-S4_Datasets-for-Benchmarking.PRESUBMISSION.xlsx

Table S5. 16S and SLST primers used in this study for benchmarking and sequencing.

Method	Region	Primer ID	FW primer (5'-3' sense)	RV primer (5'-3' antisense)	Used for ..	Reference
16S	V1-V2	27Fx338R	AGAGTTTGATCCTGGCTCAG	TGCTGCCTCCCGTAGGAGT	Benchmarking	Baurecht *et al.,* 2018 [2]
16S	V1-V9	8Fx1391R	AGAGTTTGATCCTGGCTCAG	GACGGGCGGTG**W**GT**R**CA	Benchmarking	Kong *et al.,* 2012 [3]
16S	V3-V4	357Fx802RV2	CCTACGGGAGGCAGCAG	TAC**NV**GGGTATCTAA**K**CC	Sequencing	Zeeuwen *et al.,* 2017 [4]
SLST	*30S ribosomal protein S11 (OG #1123)*	not applicable	TGGCACGTAAACAAGTATC	GACGACGTTTTGGTGGAC	Sequencing	Current study (and Smits *et al.,* 2018, manuscript in preparation)

Tables S6-S9. Benchmark datasets TaxPhlAn run output reports.

File :
S6-S9 TaxPhlAn-Reference-Runs.xlsx (online) http://ederveen.science/Thesis/Chapter3/Table-S6-S9_TaxPhlAn-Reference-Runs.PRESUBMISSION.xlsx

Table S10. Study samples overview Ederveen *et al.*

The samples used in this study are from a recently generated larger cohort from our department by Smits *et al.*, (manuscript in preparation), from which we have selected all available baseline samples for which both 16S and SLST sequences were generated. In the study by Smits *et al.,* we studied the effect of coal tar treatment on the skin microbiome; however, the for this study selected samples are from before start of coal tar treatment (hence, baseline) and thus considered the endogenous skin microbiome of these volunteers. “COMB_” samples (from *combination*) are samples with 1:1 pooled amplicons of 16S and SLST of the same volunteer (note that 16S and SLST amplicons were generated from the same DNA sample of each volunteer). The general idea behind the “COMB_” samples is (i) to have biological replicates, and (ii) to test feasibility of combining 16S and SLST amplicons of the same volunteer in one sequencing sample with the same barcodes (i.e. to reduce sequencing costs for application of SLST). * samples codes as reported by Smits *et al.,* (manuscript in preparation), these are the same sample identifiers as have been uploaded to ENA. ** sample names are similar for 16S and for SLST, but preceded by "16S_" or "SLST_", respectively (with exception of “COMB_” samples because these contain both 16S and SLST amplicons). AD = atopic dermatitis patient. EASI = Eczema Area and Severity Index. HC = healthy control volunteer.

Sample ID Smits *et al.* *	Sample ID Ederveen *et al.* **	Individual	Sample Site	Arm (L = left, R = right)	Condition	Skin Status	AD Severity (EASI)	Age (years)	Gender (M/F)
AD1_T_0	AD1_L	AD1	inner elbow	L	AD	lesional	13.8	27	M
AD1_V_0	AD1_R	AD1	inner elbow	R	AD	lesional	13.8	27	M
AD2_T_0	AD2_L	AD2	inner elbow	L	AD	lesional	44.4	21	M
AD2_V_0	AD2_R	AD2	inner elbow	R	AD	lesional	44.4	21	M
AD3_T_0	AD3_L	AD3	inner elbow	L	AD	lesional	8.1	27	F
AD3_V_0	AD3_R	AD3	inner elbow	R	AD	lesional	8.1	27	F
AD5_T_0	AD4_L	AD4	inner elbow	L	AD	lesional	25.4	25	F
AD5_V_0	AD4_R	AD4	inner elbow	R	AD	lesional	25.4	25	F
AD7_T_0	AD5_L	AD5	inner elbow	L	AD	lesional	11.6	23	M
AD7_V_0	AD5_R	AD5	inner elbow	R	AD	lesional	11.6	23	M
HC10_T_0	HC1_L	HC1	inner elbow	L	HC	normal	0	30	M
HC10_V_0	HC1_R	HC1	inner elbow	R	HC	normal	0	30	M
HC2_T_0	HC2_L	HC2	inner elbow	L	HC	normal	0	30	F
HC2_V_0	HC2_R	HC2	inner elbow	R	HC	normal	0	30	F
HC3_T_0	HC3_L	HC3	inner elbow	L	HC	normal	0	30	F
HC3_V_0	HC3_R	HC3	inner elbow	R	HC	normal	0	30	F
HC4_T_0	HC4_L	HC4	inner elbow	L	HC	normal	0	35	F
HC4_V_0	HC4_R	HC4	inner elbow	R	HC	normal	0	35	F
HC5_T_0	HC5_L	HC5	inner elbow	L	HC	normal	0	22	F
HC5_V_0	HC5_R	HC5	inner elbow	R	HC	normal	0	22	F
HC6_T_0	HC6_L	HC6	inner elbow	L	HC	normal	0	30	F
HC6_V_0	HC6_R	HC6	inner elbow	R	HC	normal	0	30	F
HC7_T_0	HC7_L	HC7	inner elbow	L	HC	normal	0	51	F
HC7_V_0	HC7_R	HC7	inner elbow	R	HC	normal	0	51	F
HC8_T_0	HC8_L	HC8	inner elbow	L	HC	normal	0	27	M
HC8_V_0	HC8_R	HC8	inner elbow	R	HC	normal	0	27	M
HC9_T_0	HC9_L	HC9	inner elbow	L	HC	normal	0	22	F
HC9_V_0	HC9_R	HC9	inner elbow	R	HC	normal	0	22	F
COMB_AD1_V_0	COMB_AD1_R	AD1	inner elbow	R	AD	lesional	13.8	27	M
COMB_AD2_V_0	COMB_AD2_R	AD2	inner elbow	R	AD	lesional	44.4	21	M
COMB_AD3_T_0	COMB_AD3_L	AD3	inner elbow	L	AD	lesional	8.1	27	F
COMB_AD5_T_0	COMB_AD4_L	AD4	inner elbow	L	AD	lesional	25.4	25	F
COMB_AD7_T_0	COMB_AD5_L	AD5	inner elbow	L	AD	lesional	11.6	23	M

Suppl. Table S11. 16S sequencing read statistics and alpha diversity metrics.

File :
S11 16S-Sequencing-Read-Statistics.xlsx (online) http://ederveen.science/Thesis/Chapter3/Table-S11_16S-Sequencing-Read-Statistics.PRESUBMISSION.xlsx

Suppl. Table S12. 16S analysis of *Staphylococcus* by OTU clustering of full-length 16S, V1-2 and V3-4.

File :
S12 16S-Analysis-Staphylococcus(V1-9)(V1-2)(V3-4).xlsx (online) http://ederveen.science/Thesis/Chapter3/Table-S12_16S-Analysis-Staphylococcus(V1-9)(V1-2)(V3-4).PRESUBMISSION.xlsx

Suppl. Table S13. SLST candidates for *Staphylococcus* profiling.

In total, 9 primer sets as listed in this table were initially screened on the lab by PCR with genomic DNA from *S. aureus* strain ATCC 29213 to check for recognition of *Staphylococcus* species. The SLST primer sets marked in green showed positive PCR responses and were further evaluated as reported in Suppl. Fig. S8. The SLST OG #1123 (dark green) is our final SLST target which survived all *in silico* and lab validations, and with which our study was performed. SLST candidates were found by running TaxPhlAn on a set of 200 *Staphylococcus* reference genomes (Table 1, Suppl. Table S9), and their characteristics according to TaxPhlAn default reports and additional *in silico* tests are listed below. The *in silico* PCR were performed on 7085 *Staphylococcus* reference genomes, and on available family Staphylococcaceae ' out-genomes ' and common skin commensals.

Table S13 *part 1 of 2*

#OG	predicted gene product	amplicon length (based on alignment incl. primers)	# of unique sequences (VR), i.e. SLST types	ML-based correlation to phylogeny (rho)	FW len.	RV len.	FW T_m	RV T_m	delta T_m
810	50S ribosomal protein L16	152	46	0.91	19	19	49.0	53.3	4.3
440	**Elongation factor G**	**482**	**69**	**0.91**	**20**	**20**	**55.2**	**53.7**	**1.5**
426	Elongation factor G	452	66	0.91	20	20	55.2	49.5	5.7
1076	**DNA-directed RNA polymerase subunit beta**	**453**	**80**	**0.89**	**18**	**20**	**46.0**	**51.4**	**5.4**
1123	**30S ribosomal protein S11**	**381**	**48**	**0.87**	**19**	**18**	**50.0**	**53.9**	**3.9**
420	Elongation factor T	205	51	0.83	21	19	50.3	53.5	3.2
1116	**50S ribosomal protein L16**	**241**	**47**	**0.80**	**20**	**20**	**52.1**	**54.6**	**2.5**
952	50S ribosomal protein L16	203	42	0.78	20	19	50.1	49.3	0.8
435	50S ribosomal protein L1	171	49	0.73	19	19	50.6	52.5	1.9
					primer set predicted characteristics				

Table S13 *part 2 of 2*

#OG	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	FW primer (5'-3' sense)	RV primer (5'-3' antisense)	# Staph. genomes recognized (7085 references)	off-targets based on *in silico* PCR * in red distant (undesired) off-targets, in blue off-targets within the family of Staphylococcaceae
810	52	40	32	26	21	9	7	4	GTCAAGGTAAACAAACWCG	ACCGTTACGTCCWACTTCA	7080	
440	80	65	51	36	21	10	6	5	CGTGAACGTGTAGGTCGTTT	ACATCACCRTATTGACCACG	7084	
426	76	61	48	34	21	10	6	5	CGTGAACGTGTAGGTCGTTT	TGACGAGAGAATTTACCTTG	7084	Bacillus
1076	80	65	60	51	37	23	17	12	CWGGTGAAAAAGTTGAAG	CACCTTGCATACGRTAAACT	7063	Macrococcus, Salinicoccus
1123	59	51	44	36	27	12	9	4	TGGCACGTAAACAAGTATC	GACGACGTTTTGGTGGAC	7078	Macrococcus
420	61	49	40	33	26	13	10	5	GCWGGTATGATGGATTGTAAA	TACGAGCAACGAAGTCWGT	7081	Aliicoccus, Jeotgalicoccus, Nosocomiicoccus
1116	46	29	22	15	13	8	3	1	ATATCGTCGTCAACATCGTC	ACCTTTACCAGCACCCATAC	7082	
952	38	25	18	12	10	7	3	1	CGTGTWAAATATCGTCGTCA	GGTGTATGWGGGAAGATTT	7082	Bacillus
435	49	43	34	31	27	13	7	3	GTGCTGAAAAAGCWGGTAT	CAGGACCCATTGTTGTWGT	7076	Bacillus, Macrococcus
	number of SNP per SDI threshold											

Suppl. Table S14-S15. 16S and SLST compositional matrix.

File :
S14-15 16S-and-SLST-Compositional-Tables.xlsx (online) http://ederveen.science/Thesis/Chapter3/Table-S14-S15_16S-and-SLST-Compositional-Tables.PRESUBMISSION.xlsx

References Supplementary Tables

1. Kitts, P.A., et al., *Assembly: a resource for assembled genomes at NCBI.* Nucleic Acids Res, 2016. **44**(D1): p. D73-80.
2. Baurecht, H., et al., *Epidermal lipid composition, barrier integrity, and eczematous inflammation are associated with skin microbiome configuration.* Journal of Allergy and Clinical Immunology, 2018. **141**(5): p. 1668-1676.e16.
3. Kong, H.H., et al., *Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis.* Genome Res, 2012. **22**(5): p. 850-9.
4. Zeeuwen, P.L.J.M., et al., *Reply to Meisel et al.* Journal of Investigative Dermatology, 2017. **137**(4): p. 961-962.