# **Supplementary Materials**

The number of sequence variations among the known phage genomes can provide an estimate of the *P. acnes* phage population diversity within each skin community. We calculated the range of nucleotide variations among a given number of phage genomes. For each given number of phage genomes, subsets of the 43 unique phage strains were selected randomly and analyzed with the core SNP analysis method to calculate the number of nucleotide variations. One hundred resamplings of each subset size were calculated when applicable.

We found that the HMP sample HMP20 had 1,353 SNPs in the assembled phage metagenome. Although the number of SNPs is smaller than the range of nucleotide variation sites between any two unique phage strains (2,099 - 2,760), it is much greater than the maximum number of 14 SNPs within a single phage strain. This suggests that the *P. acnes* phage population in HMP20 may consist of more than one phage strain.



### Figure S1. *P. acnes* phage morphology.

A representative electron micrograph of phage PHL066M04 shows that it has an icosahedral head (~50 nm in diameter) and a ~150 nm flexible tail. The phage morphology is similar to those previously described (Farrar *et al.*, 2007; Marinelli *et al.*, 2012; Lood *et al.*, 2008; Webster & Cummins, 1978).

PHL009M11	<u> </u>	
PHL010M04	<u> </u>	
PHL025M00	<u>♦                                     </u>	
PHL030N00		
PHL037M02	<u>44</u>	
PHL041M10	▲ <u>▲ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓</u>	
PHL055N00		
PHL060L00	<u>+++++++++++++++++++++++++++++++++++++</u>	
PHL064M01		
PHL064M02		
PHL066M04	▶ <b>→ → → → → → → → → → → → → → → → → → →</b>	
PHL067M01		
PHL067M09	<u> </u>	
PHL067M10		
PHL070N00		
PHL071N05	▶ <b>────────────────────────────────────</b>	
PHL073M02	<u> </u>	
PHL082M00		
PHL082M02	▶ <b>→ → → → → → → → → → → → → → → → → → →</b>	
PHL082M03	<u> </u>	
PHL082M04	<u> </u>	
PHL085M01		
PHL085N00	4         4	
PHL092M00	<u>4/4</u> >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>	
Genome position	0 2000 4000 6000 8000 10000 12000 14000 16000 18000 20000 22000 24000 26000 28000	

gp1excinucleasegp2terminasegp3portal proteingp4hypothetical protgp5scaffold proteingp6major head protegp7hypothetical protgp8hypothetical proteingp0bypothetical protein	tein tein tein tein
gp2terminasegp3portal proteingp4hypothetical protgp5scaffold proteingp6major head protegp7hypothetical protgp8hypothetical protgp0bypothetical prote	tein tein tein tein
gp3portal proteingp4hypothetical protgp5scaffold proteingp6major head proteingp7hypothetical protgp8hypothetical proteingp0hypothetical protein	tein tein tein tein
gp4hypothetical protgp5scaffold proteingp6major head protegp7hypothetical protgp8hypothetical protgp0bypothetical prot	tein tein tein tein
gp5scaffold proteingp6major head protegp7hypothetical protgp8hypothetical protgp0hypothetical prot	ein tein tein tein
gp6major head protegp7hypothetical protgp8hypothetical protgp0hypothetical prot	ein tein tein tein
<ul><li>gp7 hypothetical prot</li><li>gp8 hypothetical prot</li><li>gp0 hypothetical prot</li></ul>	tein tein tein
gp8 hypothetical prot	tein tein
and hypothetical proj	tein
gpg hypothetical pro	
gp10 hypothetical prot	tein
gp11 major tail proteir	า
gp12 hypothetical prot	tein
gp13 hypothetical prot	tein
gp14 tape measure pr	otein
gp15 minor tail subuni	t
gp16 protease	
gp17 hypothetical prot	tein
gp18 hypothetical prot	tein
gp19 collagen-like pro	tein
gp20 endolysin	
ap21 holin	
ap22/23 hypothetical prot	tein
<ul> <li>hypothetical prot</li> </ul>	tein
gp24 sigma factor	
gp25 hypothetical prot	tein
<i>ap26</i> hypothetical prot	tein
ap27 hypothetical prot	tein
gp28 hypothetical prot	tein
ap29 hypothetical prot	tein
gp30 hypothetical prot	tein
<i>qp31</i> DNA primase	
ap32 DNA primase	
ap33 hypothetical prot	tein
ap34 DNA helicase	
ap35 hypothetical prot	tein
gp36 hypothetical prot	tein
<i>qp</i> 37 exonuclease	
gp38 hypothetical prot	tein
gp39 hypothetical prot	tein
ap40 hypothetical prot	tein
ap41 hypothetical prot	tein
<i>qp42</i> hypothetical prot	tein
ap43 hypothetical prot	tein
ap44 hypothetical prot	tein
ap45 hypothetical prot	tein
ap46 hypothetical prot	tein
ap47 hypothetical prot	tein
gp48 endonuclease	

PA6



PA6

#### Figure S2. Newly sequenced *P. acnes* phage genomes and annotations.

The genome organization of all 48 phages shows that these phages are highly similar to each other. On average 45 ORFs are encoded in each phage genome. ORF color indicates the PA6 homolog described in the figure legend.



#### Figure S3. Phylogenetic tree of *P. acnes* phages based on non-synonymous SNPs in the core regions.

A phylogenetic tree was constructed from the set of non-synonymous SNPs in the core regions of the 62 sequenced phage genomes. Consistent with the tree in Figure 2, no significant phylogenetic lineages were observed. Branches with bootstrap values less than 80 (based on 1,000 resamplings) were collapsed.



**Figure S4. The genomes of the** *P. acnes* **phages assembled from the HMP metagenomic data are highly similar to the known phage genomes.** High-quality drafts of *P. acnes* phage genomes assembled from the metagenomes of three HMP retroauricular samples (in bold) were included in the phylogenetic analysis of the phages. 7,270 SNPs in the core regions of the 65 genomes were used to generate the phylogenetic tree. The phages in samples HMP04 and HMP09, which were collected from the same individual, are highly similar to each other. Branches with bootstrap values less than 80 (based on 1,000 resamplings) were collapsed. Bootstrap values are displayed for the branches.





#### Figure S5. Phage genes under diversification.

(A) Sites of nucleotide variation (red) between PHL037M02 and Group VI phages (PHL115M02, PHL085M01, PHL085N00) mapped to the PHL037M02 genome. All but one of the nucleotide variations are located in the region encoding Gp16, Gp17, and Gp18. (B) Sites of nucleotide variation between PHL073M02 and Group VIII phages (PHL010M04, PHL066M04, PHL151M00, PHL151N00) mapped to the PHL073M02 genome. The nucleotide variations are primarily located in the endolysin and holin genes. Green arrows represent ORFs in each genome.



Phages used for challenge experiments										L	lony	ony	lony	lony					
								G	irou	up V	/I	G	rou	p VI		S	S	S	S
Phage used to infect ATCC6919 previously	Passage Number	PHL071N05	PHL113M01	PHL111M01	PHL060L00	PHL067M10	PHL112N00	PHL037Z02	PHL115M02	PHL085M01	PHL114L00	PHL073M02	PHL010M04	PHL066M04	12,216 3,054 2.036		•••	<b>, ,</b> .	-
PHL060L00	4	R	R	R	R	R	R	R	R	R	R	R	R	R	1,636	E			
PHL112N00	4	R	R	R	R	R	R	R	R	R	R	R	R	R	1,018 🛌				
PHL037M02	2	R	R	R	R	R	R	R	R	R	R	R	R	R					
PHL073M02	2	R	R	R	R	R	R	R	R	R	R	R	R	R	506 396		-		
PHL114L00	2	R	R	R	R	R	R	R	R	R	R	R	R	R	344				
uninfected	-	s	S	S	S	S	s	S	S	S	S	S	S	s	230				1

Figure S6. Pseudolysogenic properties of *P. acnes* phages. (A) *P. acnes* phage DNA can circularize via the 11-nt overhangs at the ends of the phage genome. Primers targeting a region spanning the two ends of the phage genome were used. A ~735 bp PCR product was amplified from all 15 phages tested. (B) P. acnes infected with phages gained superinfection immunity. Strain ATCC6919 infected with five different phages and passaged 2-4 times was challenged with 13 phages. 'R' indicates resistance/no lysis, and 'S' indicates susceptibility. (C) P. acnes phages exist in a subpopulation of re-grown P. acnes cultures after phage lysis. PCR amplification of the phage gp11 gene produced the expected 437 bp product in two of four colonies from the fourth passage of an ATCC6919 culture infected with PHL060L00.



#### Figure S7. Protospacers identified in the phage genomes are mainly located in the left arm region.

Each of the 15 sequenced and tested phages was aligned to all CRISPR spacers identified in the *P. acnes* strains. Each protospacer sequence in a phage genome is shown in red if it is an exact match or in orange if it contains up to two mismatches. Plus- and minus-strand protospacers are shown above and below the genome lines, respectively.

# Table S1. Phage genome information and sequencing statistics

	Genome					Assembled						
	Donor	length	N	umber	Sequencing	Sequencing	Sequencing	bases	Assembly	Assembly		
Phage	skin	(bp)	GC% of	ORFs	method	bases (Mb)	coverage	(Mb)	coverage	polymorphisms		
PHL009M11	healthy	29,503	53.9	45	MiSeq	1.5	50x	1.1	39x	3		
PHL010M04	healthy	29,511	54.0	46	454	1.7	57x	1.7	57x	0		
PHL025M00	healthy	29,496	54.0	43	MiSeq	5.1	174x	5.1	174x	2		
PHL030N00	healthy	29,424	53.9	45	MiSeq	4.7	158x	4.7	158x	9		
PHL037M02	healthy	29,443	53.8	44	454	3.1	105x	3.1	105x	2		
PHL041M10	healthy	29,412	54.3	45	MiSeq	1.6	55x	1.6	55x	2		
PHL055N00	healthy	29,264	54.2	45	MiSeq	1.7	58x	1.7	58x	0		
PHL060L00	acne	29,514	54.0	46	454/MiSeq	29.6	1,002x	12.0	407x	7		
PHL064M01	healthy	29,424	53.9	46	MiSeq	2.6	88x	2.6	88x	0		
PHL064M02	healthy	29,407	53.9	46	MiSeq	3.3	112x	3.3	112x	0		
PHL066M04	healthy	29,512	54.0	46	454	2.4	80x	2.4	80x	0		
PHL067M01	acne	29,377	54.3	46	MiSeq	1.8	61x	1.8	61x	0		
PHL067M09	acne	29,386	54.3	46	MiSeq	3.7	128x	3.7	128x	0		
PHL067M10	acne	29,377	54.3	46	454	2.3	79x	2.3	79x	1		
PHL070N00	healthy	29,421	54.0	44	MiSeq	4.8	163x	4.8	163x	0		
PHL071N05	healthy	29,467	53.9	44	454	3.1	104x	3.1	104x	7		
PHL073M02	healthy	29,503	54.0	46	454	2.5	86x	2.5	86x	1		
PHL082M00	acne	29,491	54.4	46	MiSeq	121.9	4,132x	3.3	112x	10		
PHL082M02	acne	29,492	54.4	46	MiSeq	12.3	417x	12.3	417x	3		
PHL082M03	acne	29,491	54.4	46	MiSeq	3.2	107x	3.2	107x	2		
PHL082M04	acne	29,502	54.4	46	MiSeq	9.2	312x	9.2	312x	1		
PHL085M01	healthy	29,451	53.8	44	454	26.1	887x	12.6	426x	4		
PHL085N00	healthy	29,454	53.8	44	MiSeq	74.4	2,525x	3.9	133x	11		
PHL092M00	acne	29,262	53.9	45	MiSeq	3.3	114x	3.3	114x	1		
PHL095N00	acne	29,752	54.0	45	MiSeq	3.6	121x	3.6	121x	0		
PHL111M01	healthy	29,140	54.3	45	454	2.9	98x	2.9	98x	0		
PHL112N00	acne	29,266	54.5	45	454/MiSeq	26.7	911x	12.3	419x	11		
PHL113M01	healthy	29,200	54.1	44	454	2.2	77x	2.2	77x	4		
PHL114L00	acne	29,464	54.2	45	454/MiSeq	13.0	440x	12.3	416x	6		
PHL114N00	acne	29,464	54.2	45	MiSeq	2.1	71x	2.1	71x	0		
PHL115M02	healthy	29,453	53.8	44	454	4.7	159x	4.7	159x	1		
PHL116M00	healthy	29,394	54.0	46	MiSeq	6.0	203x	1.2	42x	1		
PHL116M10	healthy	29,396	54.0	46	MiSeq	5.4	184x	5.4	184x	6		
PHL117M00	acne	29,255	54.2	45	MiSeq	1.7	58x	1.7	58x	0		
PHL117M01	acne	29,422	53.9	46	MiSeq	54.3	1,845x	6.8	231x	1		
PHL132N00	acne	29,003	54.3	42	MiSeq	1.9	66x	1.9	66x	2		
PHL141N00	acne	29,494	53.9	45	MiSeq	1.9	64x	1.9	64x	2		
PHL150M00	healthy	29,439	54.2	45	MiSeq	27.4	932x	27.4	932x	0		
PHL151M00	healthy	29,511	54.0	46	MiSeq	6.0	203x	6.0	203x	0		
PHL151N00	healthy	29,511	54.0	46	MiSeq	4.2	141x	4.2	141x	0		
PHL152M00	acne	29,247	54.2	45	MiSeq	266.7	9,120x	266.7	9,120x	0		
PHL163M00	healthy	29,264	54.2	46	MiSeq	2.0	69x	2.0	69x	0		
PHL171M01	healthy	29,328	53.7	45	MiSeq	132.0	4,500x	132.0	4,500x	2		
PHL179M00	acne	29.428	53.9	45	MiSeq	6.1	207x	6.1	207x	1		
PHL194M00	acne	29,264	54.2	45	MiSeq	1.6	53x	1.6	53x	0		
PHL199M00	healthv	29.806	54.0	45	MiSea	4.8	161x	4.8	161x	4		
PHL301M00	acne	29.323	54.4	45	MiSea	5.6	191x	5.6	191x	0		
PHL308M00	healthv	29,442	54.2	45	MiSea	7.3	247x	7.3	247x	1		
Median	v	29,434	54.0	45	4	4.0	135x	3.5	118x	1		

Group	Phage	Isolated from the same sample of the same individual	Isolated from the same individual at different times	Isolated from different individuals	Isolated from related individuals	Maximum nucleotide variations within group	
	PHL067M01	X					
Ι	PHL067M09	X				1	
	PHL067M10	X					
	PHL082M00	X					
п	PHL082M02	X					
11	PHL082M03	X				5	
	PHL082M04	X					
	PHL055N00						
TTT	PHL117M00			v			
111	PHL163M00			Λ			
	PHL194M00						
TV.	PHL150M00			V	X	14	
1 V	PHL308M00			Λ	X	14	
	PHL064M01	X					
V	PHL064M02	X		v		12	
v	PHL117M01			Λ		12	
	PHL030N00						
	PHL115M02						
VI	PHL085M01		Х	Х		8	
	PHL085N00		X (4 months later)				
VII	PHL114L00		X			0	
V II	PHL114N00		X (3 months later)			0	
	PHL151M00		X				
VIII	PHL151N00		X (5 months later)				
	PHL010M04			Х	X	10	
	PHL066M04				X		
	PHL073M021				X		
IV	PHL116M00	X				4	
	PHL116M10	X				] 4	

Table S2. Relationships among the phages in groups I–IX

<sup>1</sup>This phage is not a member of Group VIII based on our criterion but shows a high genetic similarity to Group VIII members.

Table S3. P. acnes phage reads in the	skin microbiome data from HMP
---------------------------------------	-------------------------------

Sample	Retroauricular crease	Total number of sequenced reads	Mappable P. acnes phage	Percentage of phage reads in	Estimated coverage of	Percentage of phage reads
HMP09	right	61 912 421	612 512	0.9890%	2.060x	99.6%
HMP04	left	76 307 200	435.067	0.5700%	1 463x	99.6%
HMP15	right	68,504,708	299.707	0.4370%	1,008x	99.7%
HMP20	right	9.005.077	160.073	1.7800%	538x	-
HMP08	left	15.217.789	6.430	0.0423%	22x	-
HMP03	right	55.078.145	6.371	0.0116%	21x	-
HMP24	right	26,584,816	5,166	0.0194%	17x	_
HMP19	right	24,662,865	2,920	0.0118%	10x	-
HMP14	right	14,343,937	2,278	0.0159%	8x	-
HMP05	left	6,977,560	2,086	0.0299%	7x	-
HMP12	right	11,463,286	1,778	0.0155%	6x	-
HMP01	right	54,571,597	1,124	0.0021%	4x	-
HMP11	right	18,789,759	1,105	0.0059%	4x	-
HMP23	left	4,970,921	924	0.0186%	3x	-
HMP22	right	9,248,315	903	0.0098%	3x	-
HMP10	right	47,468,302	444	0.0009%	1x	-
HMP07	left	40,649,268	403	0.0010%	1x	-
HMP17	left	23,548,107	350	0.0015%	1x	-
HMP26	left	43,795,837	188	0.0004%	1x	-
HMP06	right	17,284,054	172	0.0010%	1x	-
HMP18	right	20,952,929	131	0.0006%	0x	-
HMP27	right	11,501,845	118	0.0010%	0x	-
HMP25	right	17,024,554	111	0.0007%	0x	-
HMP21	right	7,432,395	72	0.0010%	0x	-
HMP02	left	5,742,482	34	0.0006%	0x	-
HMP16	left	3,485,645	32	0.0009%	0x	-
HMP13	right	3,000,119	24	0.0008%	0x	-

Skin sample	Number of	Total phage genome coverage	Number of SNPs detected at various phage genome coverages										
	reads		20X	50X	100X	200X	500X	1,000X	1,500X	All reads			
HMP09	612,512	2,060X	0	1	1	1	1	1	1	1			
HMP04	435,067	1,463X	0	0	0	0	0	0	-	0			
HMP15	299,707	1,008X	0	0	0	0	0	0	-	0			
HMP20	160,073	538X	37	581	1,000	1,239	1,333	-	-	1,353			
HMP08	6,430	22X	0	-	-	-	-	-	-	0			
HMP03	6,371	21X	16	-	-	-	-	-	-	19			
HMP24	5,166	17X	-	_	-	-	-	-	-	0			

## Table S4. SNPs detected in the assembled phage genomes