A unique mode of nucleic acid immunity performed by a single multifunctional enzyme

S. M. Nayeemul Bari[1], Lucy Chou-Zheng[1], Katie Cater[1], Vidya Sree Dandu[1], Alexander Thomas[1], Barbaros Aslan[1], and Asma Hatoum-Aslan[1*]

[1] Department of Biological Sciences, University of Alabama, Tuscaloosa, AL 35487, USA

Organisms spanning all domains of life protect against pathogens using diverse mechanisms of nucleic acid immunity which detect and eliminate foreign genetic material[1]. The perpetual arms race between bacteria and their viruses (phages) has given rise to both innate and adaptive nucleic acid immunity mechanisms, including restriction-modification and CRISPR-Cas, respectively[2]. These sophisticated systems encode multiple components that sense and degrade phage-derived genetic material while leaving the host genome unharmed. Here, we describe a unique mode of innate immunity performed by a single protein, SERP2475, herein renamed to Nhi. We show that this enzyme protects against phages by preventing phage DNA accumulation, and in a purified system it degrades both DNA and RNA substrates. This enzyme also exhibits ATP-dependent helicase activity, and excess ATP abrogates nuclease function, suggesting a possible mechanism for its regulation. Further, using directed evolution, we isolated and characterized a collection of resistant phage mutants and found that a single-stranded DNA binding protein provides a natural means for phages to escape immunity. These observations support a model in which Nhi senses and degrades phage-specific replication intermediates. We also found that this dual-function enzyme protects against diverse phages, and its homologs are distributed across several bacterial phyla. Altogether, our findings reveal a new innate immune system with minimal composition that provides robust defense against diverse bacterial viruses.

Phages are the most abundant entities in the biosphere[3], and as such they impose a tremendous selective pressure upon their bacterial hosts. Phages attach to a specific host, inject their genome, and utilize the host's enzymes and energy stores to replicate exponentially, a process that typically leads to the death of the host (Fig. 1a). In response, bacteria have evolved diverse immune systems that target nearly every step of the phage infection cycle[2]-- such systems may block phage genome entry, digest phage-derived nucleic acids, and/or initiate programmed cell death, a process known as abortive infection (Abi), to prevent the spread of a phage infection. Perhaps the most direct mechanism to block phage replication is through the use of nucleases that sense and destroy foreign genetic material. However, bacteria are known to possess only two modes of nucleic acid immunity which contribute to anti-phage defense: Innate immunity provided by restriction-modification (R-M) and adaptive immunity conferred by CRISPR-Cas (Fig. 1a). Given that bacteria and their phages have been co-evolving for billions of years, it is only natural to expect that many unknown immune systems are yet to be discovered, particularly in non-model organisms[4,5].

Here, we sought to identify and characterize new anti-phage immune mechanisms in the commensal opportunistic pathogen *Staphylococcus epidermidis* RP62a[6]. This organism encodes a Type III-A CRISPR-Cas system[7], an Abi mechanism[8], and a putative Type I R-M system (GenBank Accession CP000029.1). All three systems reside within ~30,000 nucleotides of each other (Extended Data 1a), in agreement with recent evidence that suggests prokaryotic immune systems can be found clustered together within discrete defense islands[5,9]. Importantly, key insights into CRISPR-Cas and Abi in this organism were revealed by studying molecular interactions with the targets of immunity, siphophages ΦNM1 and CNPx, respectively[8,10]. Therefore, we reasoned that the identification of novel immunity mechanisms would necessarily require the expansion of the *S. epidermidis* phage collection to include others unrelated to those commonly studied. Toward that end, we captured and characterized four new *S. epidermidis* podophages: Andhra, JBug18, Pontiff, and Pike[12,13]. These phages share over 95% sequence

identity (Extended Data 2), however, they exhibit distinct host ranges—While Andhra and Pontiff can infect the wild-type RP62a strain, JBug18 and Pike can only infect a mutant variant, LM1680[14], which has lost the defense island and surrounding regions. These observations lead to the hypothesis that JBug18 and Pike are sensitive to genetic element(s) within the defense island.

To test this, we used a set of *S. epidermidis* RP62a mutants that bear deletions of varying extents across the defense island[14] (Extended Data 1a). These strains were plated together with Andhra and JBug18, representative phages with resistant and sensitive phenotypes, respectively. The resulting zones of bacterial growth inhibition (plaques) were enumerated, and the data revealed that only one of the deletion mutants encodes the gene(s) required for complete protection against JBug18. These observations narrowed down the protective genetic element(s) to a stretch of ~12,000 nucleotides containing 12 genes (*SERP2466-SERP2477*) which incidentally encompasses the R-M system (Extended Data 1b). To determine which of the 12 gene(s) are responsible for immunity, they were inserted (individually and in groups) into a derivative of plasmid pC194[15] (herein referred to as p*SERP-*), introduced into *S. epidermidis* LM1680, and the resulting strains were challenged with Andhra and JBug18. Through this analysis, we found that a single gene of unknown function, *SERP2475*, is sufficient to provide robust immunity against both JBug18 (Extended Data 1 c and d) and Pike (Extended Data 3 a and b).

We next sought to characterize the mechanism of *SERP2475*-mediated immunity. An adsorption assay was first used to determine that JBug18 can attach to LM1680/p*SERP-2475* cells just as efficiently as can Andhra (Fig. 1b), thus ruling out an adsorption-blocking mechanism. A cell viability assay was next used to test for abortive infection. The prediction is that if programmed cell death accompanies immunity, then challenge with a high proportion of phages to cells ($\geq$1:1) would lead to significant cell death similar to that observed when there is no immune protection[16,17]. To test this, phages were combined with LM1680/p*SERP-2475* cells in liquid media

at ratios of 1:1, 5:1, or 10:1, and cell viability was measured following 5 hours. We found that while Andhra causes significant death of LM1680/p*SERP-2475* cells, JBug18 elicits only a minor decrease in the viability of this strain, even when phages outnumber bacteria 10:1 (Fig. 1c). These observations indicate that abortive infection is unlikely to be occurring. Finally, quantitative PCR (qPCR) was used to track the accumulation of phage DNA at various time points following infection. The results showed that while Andhra's DNA accumulates to ~20-fold by 20 minutes post-adsorption in LM1680/p*SERP-2475*, JBug18's DNA accumulates to less than 4-fold in the same time period (Fig. 1d). These observations support the hypothesis that *SERP2475* interferes with phage DNA replication and/or expression.

SERP2475 encodes a 606 amino acid protein with predicted ATPase and helicase domains (according to BLASTp, Fig. 2a). Additionally, the HH Pred homology search tool[18] identified similarities between the N-terminus of SERP2475 and the HsdR restriction endonuclease from *Vibrio vulnificus*. Finally, the I-TASSER structure prediction tool[19] showed potential similarities with the eukaryotic DNA2 nuclease-helicase[20]. To test the importance of these predicted activities *in vivo*, two mutant versions of p*SERP-2475* were created, *-mutA* and *-mutB*, which encode multiple amino acid substitutions in the predicted nuclease and helicase active sites, respectively (Fig. 2a). Following challenge with phage, we found that JBug18 could now replicate on both strains (Extended Data 4), indicating that both regions of the protein are required for immune function *in vivo*.

To test for its activities *in vitro*, SERP2475 and both mutants were overexpressed in *E. coli* and purified (Fig. 2b). The proteins were first combined with single-stranded DNA and RNA substrates (Extended Data 5a) and various divalent cations to test nuclease activity, and we observed that SERP2475 can indeed degrade both substrates in the presence of Mg^{2+} or Mn^{2+} (Extended Data 5 b and c). Importantly, this activity is significantly abrogated in both mutants (Fig. 2c), confirming that the activity stems from SERP2475. Similar nuclease assays using circular single-stranded DNA and RNA substrates revealed that a free end is required for nuclease activity

(Extended Data 6 a and b). Finally, we tested DNase activity on various double-stranded substrates (Extended Data 7), and found that while the enzyme is unable to degrade double-stranded DNA with blunt ends, it can cut short stretches of double-stranded DNA (albeit inefficiently), particularly if provided with a 3'- overhang (Fig. 2d). Altogether, these results demonstrate that SERP2475 is primarily a 3'-5' exonuclease with a strong preference for single-stranded nucleic acids.

We next tested for ATPase and helicase activities. SERP2475 is a predicted superfamily 1 helicase (pfam01443), which functions as a monomer or dimer to unwind double-stranded substrates using energy from ATP[21]. An ATPase assay showed that SERP2475 does in fact hydrolyze ATP (Fig. 2e). A helicase assay was next performed by incubating the enzyme with ATP and various DNA substrates, and the results showed that SERP2475 can indeed unwind double-stranded DNA, particularly when offered a 5'- or 3'- overhang (Fig. 2f). Interestingly, the single-stranded product released upon unwinding remains intact in the helicase assay, suggesting that excess ATP may have an inhibitory effect on nuclease activity. To test this, the nuclease assays were repeated in the presence of ATP, and we found that ATP inhibits nuclease activity in a dose-dependent manner (Fig. 2g). Altogether, these results demonstrate that SERP2475 exhibits both nuclease and helicase activities that likely compete with one another. In addition, given that ATP drives helicase activity while inhibiting nuclease activity, we hypothesize that ATP levels may provide a mechanism for functional regulation of the enzyme.

To refine our understanding of this mechanism, we sought to determine how Andhra escapes immunity. Andhra and JBug18 encode the same 20 protein homologs (Extended Data 8a) and an alignment of their coding regions show that they differ at only 705 positions by either a single-nucleotide polymorphism (SNP) or a gap (Supplementary file 1). To narrow down which SNPs and/or gaps in Andhra are important for resistance to immunity, we first attempted to isolate naturally-evolved JBug18 mutants that can escape immunity by plating concentrated phage preparations with LM1680/p*SERP-2475*. After several failed attempts at recovering plaques, one

attempt yielded resistant phages, which upon further inspection were found to possess hybrid genomes that contain a patchwork of Andhra and JBug18 sequences. These hybrids necessarily arose through the inadvertent mixing of the two phages. Nonetheless, this fortuitous accident proved invaluable in helping to pinpoint the region required for immune resistance--since all hybrids can escape immunity, they must share Andhra-derived sequences in the region required for resistance. Therefore, we purified and sequenced eight such hybrids, and determined the fraction of hybrids that possess Andhra identity at each of the 705 differing positions across their coding regions. We found that all eight hybrids share Andhra identity at positions 891-2117 in the alignment (Fig. 3a and Supplementary file 1). This region overlaps gene products (*gp*) *03-06* in the phage genomes and encompasses 69 SNPs and gaps, of which, 64 are concentrated within *gp03* and *gp04* (Fig. 3b). Accordingly, we speculated that one or both of the latter are responsible for resistance.

To narrow down the protective region even further, a second set of resistant JBug18 hybrids were generated which bear Andhra-derived sequences in *gp03* and *gp04*. This was accomplished by introducing Andhra's *gp03* and/or *gp04* coding regions into *S. epidermidis* LM1680 on plasmids, and then propagating JBug18 on these strains to allow the phage to recombine with the Andhra-derived sequences (Extended Data 8b). The resulting phages were then plated on LM1680/p*SERP-2475* to select for resistant phage recombinants. Ten such hybrids (9-18) were purified, and sequencing across *gp03* and *gp04* revealed that they had all acquired a 60-nucleotide stretch spanning positions 1302-1362 in Andhra's genome (Fig. 3c and Supplementary file 2). This region overlaps *gp03*, which encodes a single-stranded DNA binding protein (SSB, Fig 3d, and Supplementary file 2). Importantly, JBug18 harbors a 5-nucleotide insertion in this region and consequently harbors a truncated SSB (Fig. 3d and Extended Data 8c). However, by acquiring the 60-nucleotide stretch from Andhra, all ten hybrids had restored the reading frame and hence encode a full-length SSB, suggesting that the SSB C-terminus is essential for escape from immunity. In agreement with these observations, Pontiff and Pike

possess the expected *gp03* genotypes: While the resistant phage Pontiff encodes a full-length SSB, the sensitive phage Pike encodes a truncated version, this time due to a single nucleotide deletion (Extended Data 8d). Since SSBs are known to bind and protect DNA and regulate the replication machinery[23], our observations support a preliminary model for immunity in which SERP2475 senses phage-specific replication intermediates and uses its helicase and nuclease activities to unwind and degrade them (Fig. 4). Future work will investigate how these activities are coordinated to achieve anti-phage defense.

Altogether, our findings describe a new mechanism of innate nucleic acid immunity performed by *S. epidermidis* RP62a SERP2475. Taking into account its demonstrated functions, we propose to name this enzyme Nhi (Nuclease-helicase mediated immunity). Importantly, further testing has revealed that Nhi provides robust immunity against diverse phages that infect both *S. epidermidis* and *S. aureus* strains (Extended Data 9). Additionally, Nhi homologs can be found in several bacterial phyla (Extended Data 10). Thus, this unique mode of innate immunity likely represents a common defense strategy used in the battle between bacteria and their viruses.

Methods

Bacterial strains and growth conditions. S. *epidermidis* RP62a and mutant variants were a generous gift from Luciano Marraffini. *S. epidermidis* strains were grown in Brain Heart Infusion (BHI, BD Diagnostics), *S. aureus* strains were grown in Tryptic Soy Broth (TSB, BD Diagnostics), *E. coli* DH5α was grown in Luria Bertani (LB) broth (VWR), and *E. coli* Rossetta2 (DE3) was grown in Terrific broth (VWR) for protein purification. Growth media was supplemented with the following: 10 µg/ml chloramphenicol (to select for pC194-based plasmids), 10 µg/ml tetracycline (to select pT181-based plasmids), 15 µg/ml neomycin (to select for *S. epidermidis* cells), 30 µg/ml chloramphenicol (to select for *E. coli* Rossetta2 plasmids) and 50 µg/ml kanamycin (to select for pET28b-His_{10}Smt3-based plasmids).

182 ***Phage propagation and plate infection assays.*** *S. epidermidis* phages (Andhra, JBug18,
183 Pontiff, Pike, Quidividi, and Twillingate) and *S. aureus* phages (ISP, Lorac, and Pabna) were
184 propagated on their respective hosts to create stocks. Stock concentrations in plaque-forming
185 units per mL (pfu/mL) were determined as previously described[12,13,24,25]. For plate infection
186 assays, 10-fold dilutions of phage stocks were spotted atop a lawn of indicated strains using the
187 double-agar overlay method[12]. Following overnight incubation, phage pfu/mL were enumerated.
188 ***Construction of pC194 and pT181-based plasmids and introduction into staphylococcal***
189 ***strains.*** All pC194- and pT181- based plasmids were constructed using either inverse PCR or
190 Gibson Assembly[26] with the primers listed in Table S1 as described previously[27]. pAH011[28], a
191 derivative of pC194[15], was used as the backbone for plasmids designated as p*SERP*- in this
192 study. Plasmid pT181[22] was used as backbone for pT181-*gp03* and pT181-*gp0304*. All
193 assembled plasmids were first introduced into *S. aureus* RN4220 (pC194-based plasmids) or
194 OS2 (pT181-based plasmids) via electroporation and inserted sequences were confirmed by
195 PCR amplification and Sanger sequencing (performed by Eurofins MWG Operon) using primers
196 shown in Table S1. Confirmed plasmids were purified using EZNA Plasmid Mini Kit (Omega
197 Bio-tek), and where indicated, introduced into *S. epidermidis* LM1680 via electroporation as
198 previously described[27].
199 ***Phage adsorption and cell viability assays.*** Overnight cultures of *S. epidermidis*
200 LM1680/p*SERP-NC* and LM1680/p*SERP-2475* were diluted 1:100 in fresh BHI supplemented
201 with antibiotics and 5 mM $CaCl_2$, and incubated at 37°C with agitation for one hour. For the
202 adsorption assay, Andhra or JBug18 were added to cultures (0.01:1 phage:cell ratio) and
203 incubated at 37°C for 10 min. Cells along with adsorbed phages were pelleted at 8000 *x g* for 5
204 min at 4°C and resulting supernatants were passed through 0.45 μm syringe filter. The number
205 of free phages in the supernatants were enumerated by the double-agar overlay method[12]. The
206 number of adsorbed phages were determined by subtracting the number of phages in

suspension from the number that was initially added. Triplicate samples were prepared for each treatment, and two independent trials were conducted. For cell viability assays, 200 μl of the bacterial cultures were distributed into a 96-well microtiter plate (into triplicate wells for each treatment), and phages were added to cells at ratios of 1:1, 5:1, or 10:1. These bacteria-phage mixtures were incubated at 37°C with agitation for an additional five hours. 25 μl of 0.1% (w/v) 2,3,5 Triphenyltetrazolium chloride (TTC) was added into each well and the microtiter plate was incubated at 37°C for an additional 30 mins to allow the colorless TTC to become enzymatically reduced to the red 1,3,5-triphenylformazan product by actively growing bacterial cells. The relative numbers of viable cells were then determined by measuring the absorbance at 540 nm. Triplicate measurements were averaged, and two independent trials were conducted.

Phage infection time course followed by quantitative PCR. Phage infection time course assays were conducted in liquid media as previously described[29]. Briefly, *S. epidermidis* LM1680 mid-log cells bearing p*SERP-NC* or *pSERP-2475* were infected with Andhra or JBug18 (phage:cell ratio of 0.5:1), cells were harvested at 0, 10, or 20 minutes post-infection, and their total DNA was extracted. Each qPCR reaction (25 μL) contained 500 ng of total DNA as template, 0.4 nM of phage-specific primers (N233 and N234) or host-specific primers (S001 and S002) (Table S1), and 1X PerfeCTa SYBR Green SuperMix (Quanta Biosciences). Separate standard reactions containing 10^2–10^9 DNA molecules were also prepared using purified Andhra phage DNA extract, JBug phage DNA extract, or bacterial genomic DNA extract. A CFX Connect Real-Time PCR Detection System (Bio-Rad) was used to amplify the DNA templates. Phage DNA copy number was normalized against host values, and the normalized value for the 0 min time point was set to one to obtain the relative DNA abundance for the rest of the time points as described previously[29]. Triplicate measurements were taken for each of two independent trials.

Construction of pET28b-His$_{10}$Smt3-based plasmids. pET28b-His$_{10}$Smt3-*SERP2475* was constructed using restriction cloning with primers N284 and N266 (Table S1) as previously described[27]. The ligated construct was introduced into *E. coli* DH5α by heat shock, and transformants were confirmed to have the desired plasmids using PCR and DNA sequencing with primers T7P and T7T (Table S1). Confirmed plasmids were purified using the EZNA Plasmid Mini Kit and introduced into *E. coli* Rosetta2 (DE3) cells for protein purification. Plasmids pET28b-H$_{10}$Smt3-*SERP2475-mutA* and pET28b-H$_{10}$Smt3-*SERP2475-mutB* were constructed by inverse PCR as described previously[27] using pET28b-His$_{10}$Smt3-*SERP2475* as the backbone and primers listed in Table S1.

Purification of recombinant SERP2475, SERP2475-mutA, and SERP2475-mutB. Recombinant proteins encoded in pET28b-His$_{10}$Smt3-based plasmids were overexpressed and purified as described previously[29].

Nuclease assays. Single stranded RNA or DNA substrates were labeled on their 5'-ends by incubating with T4 polynucleotide kinase and γ-[^{32}P]-ATP and purified over a G25 column (IBI Scientific). Double-stranded DNA duplexes were prepared by combining 5′-radiolabeled ssDNA oligonucleotides and unlabeled complementary ssDNA oligonucleotides (as shown in Fig. S6) in a 1:1.5 molar ratio. The mixtures were heated to 95°C for 5 min and then slowly cooled down to room temperature over a period of 3 hours. Circular ssRNA was prepared by incubating the radiolabeled substrate with T4 RNA ligase 1 (NEB) and the supplied buffer according to the manufacturer's instructions. For nuclease assays, radiolabeled substrates were combined with 25 pmol of enzyme in nuclease buffer (25 mM Tris-HCl pH 7.5, 2 mM DTT) supplemented with 10 mM of divalent metal ($MgCl_2$, $MnCl_2$, $NiCl_2$ or $ZnCl_2$), EDTA, or ATP as indicated in figures and legends. Reactions were allowed to proceed at 37°C for indicated time periods. Reactions were stopped by adding an equal volume of 95% formamide loading buffer and resolved using denaturing PAGE. Gels were exposed to a Storage Phosphor screen and visualized using a Typhoon FLA 9500 biomolecular imager.

257 ***Helicase Assay***. Double-stranded DNA substrates with the top strand radiolabeled were
258 prepared as described above. The helicase assay was performed by first mixing 10 pmol of
259 radiolabeled DNA duplex with a 5-fold molar excess of unlabeled top-strand DNA. This mixture
260 was combined with 25 pmol SERP2475 in nuclease buffer supplemented with 10 mM $MgCl_2$ and
261 20 mM ATP. Reaction mixtures were incubated at 37°C for 1h. As a positive control, DNA
262 substrates were heated to 95°C for 15 min in the absence of the enzyme. Samples were
263 resolved on an 8% (v/v) non-denaturing polyacrylamide gel. The gel was dried under vacuum at
264 80°C, exposed to a Storage Phosphor screen and visualized by a Typhoon FLA 9500
265 biomolecular imager. Three independent trials were conducted.
266 ***ATPase assay.*** A Quanticrom™ ATPase/GTPase Assay kit (BioAssay Systems) was used
267 according to the manufacturer's protocol with some minor modifications. Briefly, 40 µl reactions
268 were set up in triplicate in a 96-well plate. Reaction wells contained 0.5 X assay buffer, 50 pmol
269 enzyme, and 50 µM ATP. Control wells without enzyme were supplemented with protein dialysis
270 (IMAC) buffer[29]. Reaction wells containing known concentrations of inorganic phosphate were
271 also prepared to generate a standard curve. The microtiter plate was incubated at 37°C for 1 h,
272 and 200 µl of stop reagent was added to each well. The plate was incubated for an additional 30
273 min at room temperature, and absorbance at 620 nm was measured with a Spectramax M2
274 microplate reader (Molecular Devices). The concentration of free phosphate was calculated by
275 plotting the OD_{620} value against the standard curve. The fraction of ATP consumed in the
276 reaction was calculated by dividing the molar amount of free phosphate generated by the molar
277 amount of ATP in the initial reaction. Triplicate wells were prepared for each treatment, and
278 three independent trials were conducted.
279 ***Phage hybrid generation and sequencing.*** JBug18-Andhra Hybrids 1-8 were isolated as
280 immune resistant mutants following challenge of LM1680/p*SERP-2475* with a high titer lysate of
281 JBug18 (~1 x 10^{10} pfu/mL). To generate JBug18-Andhra Hybrids 9-18, overnight cultures of *S.*

epidermidis LM1680 harboring pT181-*gp03* or pT181-*gp0304* were diluted 1:100 in fresh TSB supplemented with antibiotics and 5 mM $CaCl_2$. The mixture was incubated at 37°C for an hour with agitation, then JBug18 was added to the cells in a 1:1 ratio, and the incubation continued with agitation overnight. The next day, cells were pelleted by centrifugation at 8000 *x g* for 5 min and supernatant was filtered through 0.45 μm filter. Filtered lysates were mixed with LM1680-p*SERP2475* overnight culture (1:1) and the mixture was plated on TSA containing 5 mM $CaCl_2$ using the double-agar overlay method[12]. For all phage hybrids, individual plaques were isolated and re-plated three times on LM1680/p*SERP-2475* to purify. Phages were propagated and their DNA was extracted as previously described[30]. Phage genomes were PCR amplified across the entire coding region for Hybrids 1-8 or *gp03-gp04* for Hybrids 9-18, and the PCR products were sequenced by the Sanger method (at Eurofins MWG Operon) using the primers listed in Table S1.

Hybrid phage genome sequence analysis. For JBug18-Andhra Hybrids 1-8, Sanger sequencing reads covering their coding regions were manually assembled using SnapGene software. For JBug18-Andhra Hybrids 9-18, a single read covered the region of interest, therefore no assembly was required. Sequences for each set of hybrids (1-8 and 9-18) were aligned with corresponding genomic regions in Andhra and JBug18 using the Clustal Omega Multiple Sequence Alignment tool (https://www.ebi.ac.uk/Tools/msa/clustalo/). The sequence alignments (Supplementary files 1 and 2) were analyzed by a Python script developed in-house which first scans the alignment of JBug18 and Andhra, identifies each position of non-similarity, and then determines at those positions the fraction of hybrids that possess Andhra identity. The output data was exported into an Excel file, and the graphs showing the fraction of hybrids with Andhra identity at each position were generated using Microsoft Excel.

Data availability. The raw phage sequence reads that support the findings of Figure 4 are publicly available through figshare (10.6084/m9.figshare.9598040). The Accession codes for Andhra, JBug18, Pontiff, and Pike genomes are KY442063, MH972263, MH972262, and

MH972261, respectively. Phages, mutant derivatives, and constructs can be made available upon written request.

Code availability. The Python code written to analyze the data for Figure 4 is freely available at GitHub (https://github.com/ahatoum/Hybrid-phage-genome-sequence-analysis).

References

1. Hartmann, G. *Nucleic Acid Immunity*. *Advances in Immunology* **133,** (Elsevier Inc., 2017).
2. Rostøl, J. T. & Marraffini, L. (Ph)ighting Phages : How Bacteria Resist Their Parasites. *Cell Host Microbe* **25,** 184–194 (2019).
3. Bergh, O., Børsheim, K. Y., Bratbak, G. & Heldal, M. High abundance of viruses found in aquatic environments. *Nature* **340,** 467–468 (1989).
4. Stern, A. & Sorek, R. The phage-host arms race: Shaping the evolution of microbes. *BioEssays* **33,** 43–51 (2011).
5. Doron, S. *et al.* Systematic discovery of antiphage defense systems in the microbial pangenome. *Science.* **359,** eaar4120 (2018).
6. Christensen, G. D., Baddour, L. M. & Simpson, W. A. Phenotypic variation of *Staphylococcus epidermidis* slime production in vitro and in vivo. *Infect. Immun.* **55,** 2870–2877 (1987).
7. Marraffini, L. A. & Sontheimer, E. J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science.* **322,** 1843–1845 (2008).
8. Depardieu, F. *et al.* A Eukaryotic-like Serine/Threonine Kinase Protects Staphylococci against Phages. *Cell Host Microbe* **20,** 471–481 (2016).
9. Makarova, K. S., Wolf, Y. I., Snir, S. & Koonin, E. V. Defense Islands in Bacterial and Archaeal Genomes and Prediction of Novel Defense Systems. *J. Bacteriol.* **193,** 6039–6056 (2011).
10. Goldberg, G. W., Jiang, W., Bikard, D. & Marraffini, L. A. Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting. *Nature* **514,** 633–637 (2014).
11. Deghorain, M. & Van Melderen, L. The staphylococci phages family: An overview. *Viruses* **4,** 3316–3335 (2012).
12. Cater, K. *et al.* A Novel Staphylococcus Podophage Encodes a Unique Lysin with

Unusual Modular Design. *mSphere* **2,** 1–9 (2017).

13. Culbertson, E. K. *et al.* Draft Genome Sequences of Staphylococcus Podophages JBug18, Pike, Pontiff, and Pabna. *Microbiol. Resour. Announc.* **8,** e00054-19 (2019).

14. Jiang, W. *et al.* Dealing with the Evolutionary Downside of CRISPR Immunity: Bacteria and Beneficial Plasmids. *PLoS Genet.* **9,** e1003844 (2013).

15. Ehrlich, S. D. Replication and expression of plasmids from *Staphylococcus aureus* in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U. S. A.* **74,** 1680–1682 (1977).

16. Goldfarb, T. *et al.* BREX is a novel phage resistance system widespread in microbial genomes. *Embo J* **34,** 169–183 (2015).

17. Ofir, G. *et al.* DISARM is a widespread bacterial defence system with broad anti-phage activities. *Nat. Microbiol.* **3,** 90–98 (2017).

18. Zimmermann, L. *et al.* A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J. Mol. Biol.* **430,** 2237–2243 (2018).

19. Yang, J. *et al.* The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* **12,** 7–8 (2015).

20. Zhou, C., Pourmal, S. & Pavletich, N. P. Dna2 nuclease-helicase structure, mechanism and regulation by Rpa. *Elife* **4,** e09832 (2015).

21. Fairman-Williams, M. E., Guenther, U.-P. & Jankowsky, E. SF1 and SF2 helicases: family matters. *Curr. Opin. Struct. Biol.* **20,** 313–324 (2010).

22. Khan, S. A., Carleton, S. M. & Novick, R. P. Replication of plasmid pT181 DNA in vitro : Requirement for a plasmid-encoded product. *Proc. Natl. Acad. Sci.* **78,** 4902–4906 (1981).

23. Shereda, R. D. *et al.* SSB as an Organizer / Mobilizer of Genome Maintenance Complexes. *Crit. Rev. Biochem. Mol. Biol.* **43,** 289–318 (2008).

24. Freeman, M. E. *et al.* Complete Genome Sequences of *Staphylococcus epidermidis* Myophages Quidividi, Terranova, and Twillingate. *Microbiol. Resour. Announc.* **8,**

e00598-19 (2019).

25. Marc, A. *et al.* Complete Genome Sequence of *Staphylococcus aureus* Siphophage Lorac. *Microbiol. Resour. Announc.* **8,** e00586-19 (2019).

26. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Meth* **6,** 343–345 (2009).

27. Walker, F. C., Chou-Zheng, L., Dunkle, J. A. & Hatoum-Aslan, A. Molecular determinants for CRISPR RNA maturation in the Cas10 – Csm complex and roles for non-Cas nucleases. *Nucleic Acids Res.* **45,** 2112–2123 (2017).

28. Hatoum-Aslan, A., Maniv, I. & Marraffini, L. A. Mature clustered, regularly interspaced, short palindromic repeats RNA (crRNA) length is measured by a ruler mechanism anchored at the precursor processing site. *Proc. Natl. Acad. Sci.* **108,** 21218–21222 (2011).

29. Chou-Zheng, L. & Hatoum-Aslan, A. A type III-A CRISPR-Cas system employs degradosome nucleases to ensure robust immunity. *Elife* 8:e45393 (2019).

30. Bari, S. M. N., Walker, F. C., Cater, K., Aslan, B. & Hatoum-Aslan, A. Strategies for Editing Virulent Staphylococcal Phages Using CRISPR-Cas10. *ACS Synth. Biol.* **6,** (2017).

Supplementary Information

The following supplementary files are also available:

Supplementary Table 1. Oligonucleotides used for cloning and PCR in this study.

Supplementary file 1. Multiple sequence alignment for JBug18-Andhra Hybrids 1-8

Supplementary file 2. Multiple sequence alignment for JBug18-Andhra Hybrids 9-18

Acknowledgements

AH-A would like to acknowledge funding for this project from the National Science Foundation CAREER award [MCB/1749886].

Author Contributions

S.M.N.B., L.C-Z, K.C., V.S.D., A.T., and B.A. performed experiments, S.M.N.B., L.C-G, B.A., and A.H.-A. designed experiments, A.H.-A. conceived the study and wrote the manuscript, and all authors have read and approved the manuscript.

Author Information

The authors declare no conflict of interest. Correspondence and requests for materials should be addressed to A. H-A. (ahatoum@ua.edu).

Figures and Legends

Figure 1. *SERP2475* impairs phage DNA accumulation. (**a**) Illustration of the lytic phage replication cycle. Known defense systems that interfere with each step are indicated on the right. (**b** and **c**) Results of an adsorption assay (**b**) and a cell viability assay (**c**) following challenge of *S. epidermidis* LM1680 cells bearing indicated plasmids with Andhra and JBug18. (**d**) Relative abundance of phage DNA at various time points following phage infection as measured by qPCR. For all experiments, the mean ± S.D. of triplicate measurements are shown as a representative of at least two independent trials.

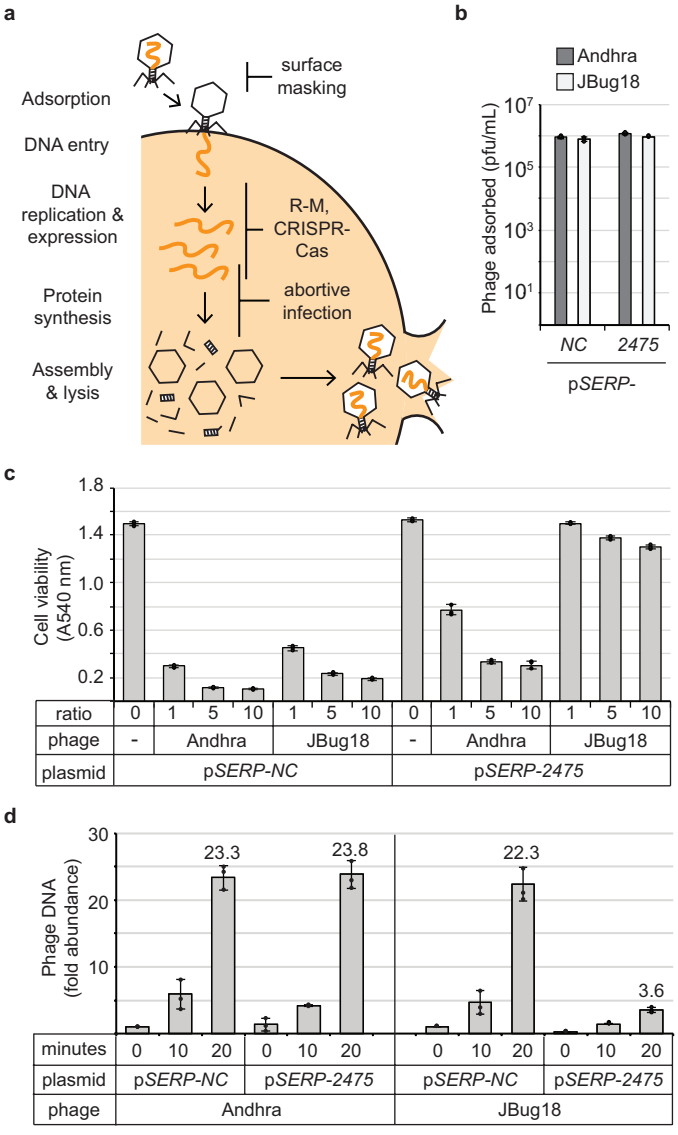
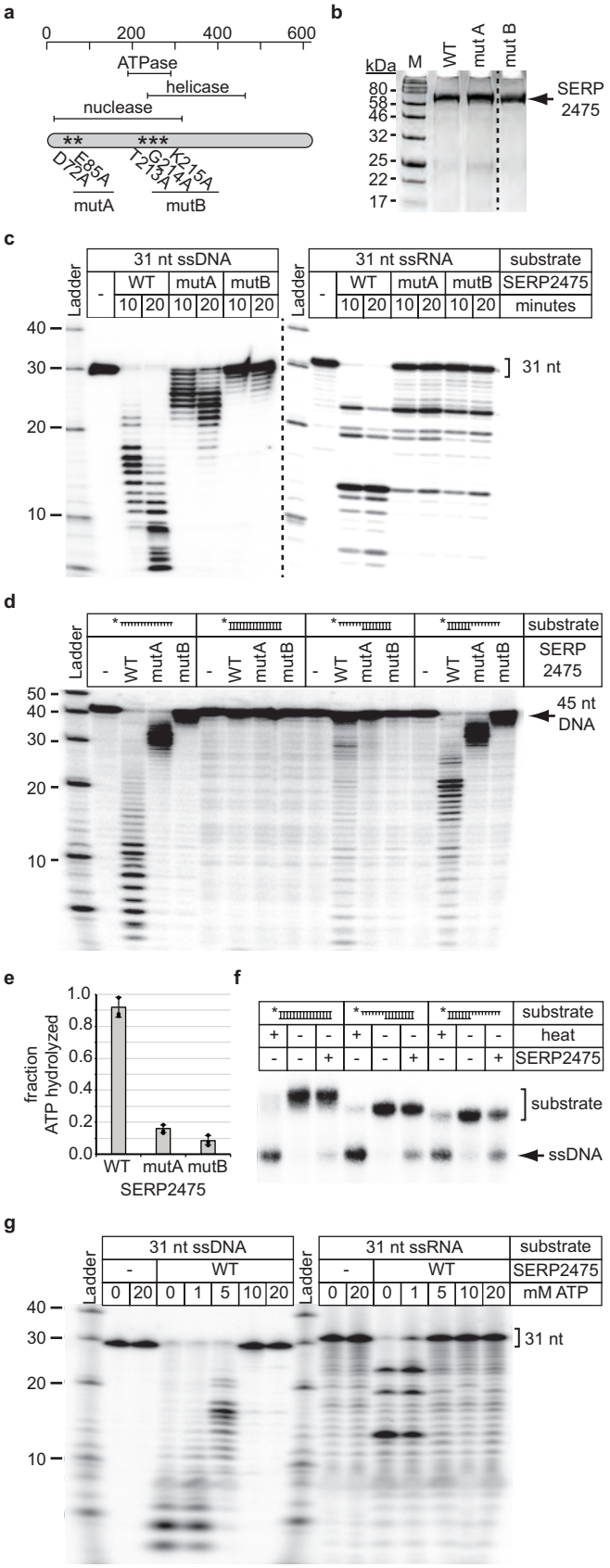


Figure 2. SERP2475 exhibits nuclease, ATPase, and helicase activities. (**a**) Predicted active sites of SERP2475 and corresponding mutations introduced in this study. (**b**) Purified recombinant SERP2475 and mutant variants. (**c**, **d**, and **g**) Nuclease assays in which SERP2475 and mutant variants were combined with radiolabeled substrates and ATP where indicated. (**e**) An ATPase assay with indicated enzymes. Shown is an average of triplicate measurements (±S.D.). (**f**) A helicase assay in which SERP2475 was combined with radiolabeled substrates and 20 mM ATP. Dotted lines separate data derived from different gels or non-contiguous regions of the same gel. For all assays, a representative of at least three independent trials is shown.



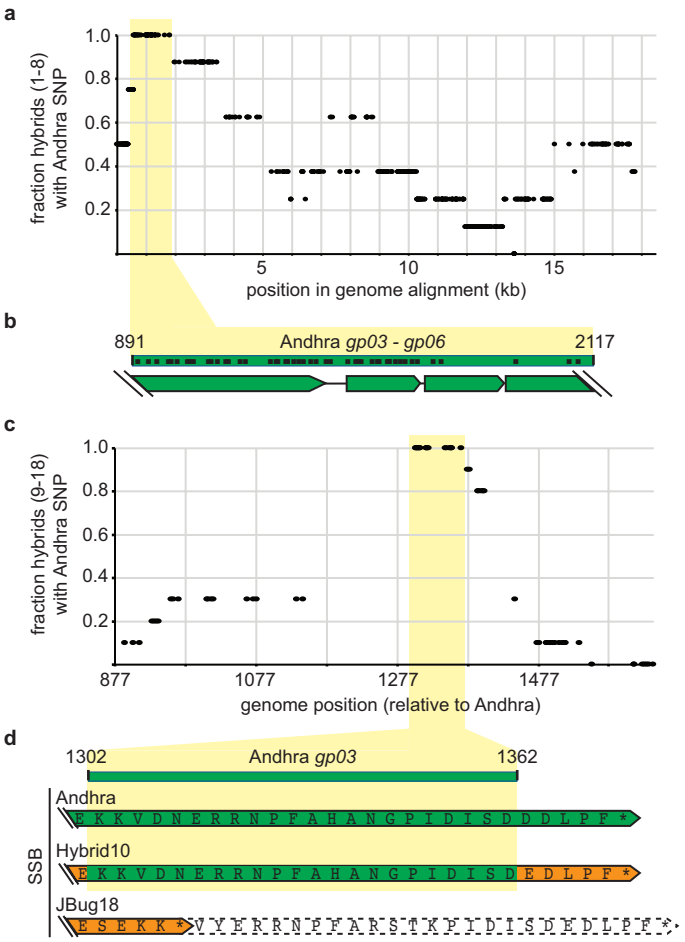


Figure 3. A phage-encoded single-stranded DNA binding protein protects against immunity. (**a** and **c**) Fractions of JBug18-Andhra Hybrids that harbor Andhra identity at positions where the parental phage genomes differ. Segments common to all hybrids are highlighted in yellow. (**b** and **d**) Expanded views of the highlighted common regions. Black dots in (**b**) indicate relative positions where SNPs and/or gaps occur in an alignment between the parental phages. In (**d**), green and orange arrows with solid borders delimit phage-encoded SSBs, and the white arrow with dotted border delimits a truncated portion.

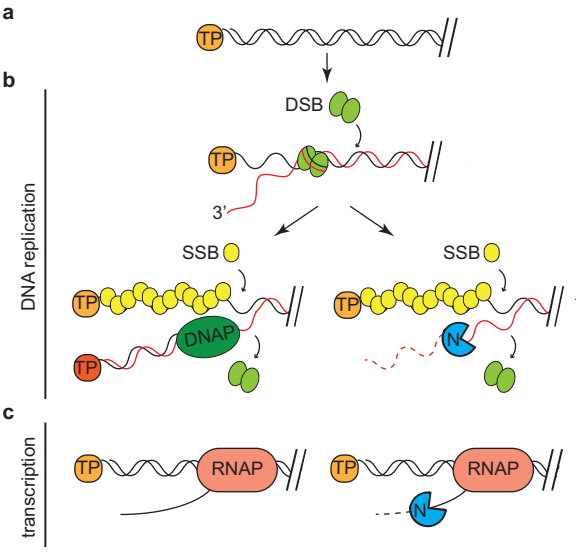
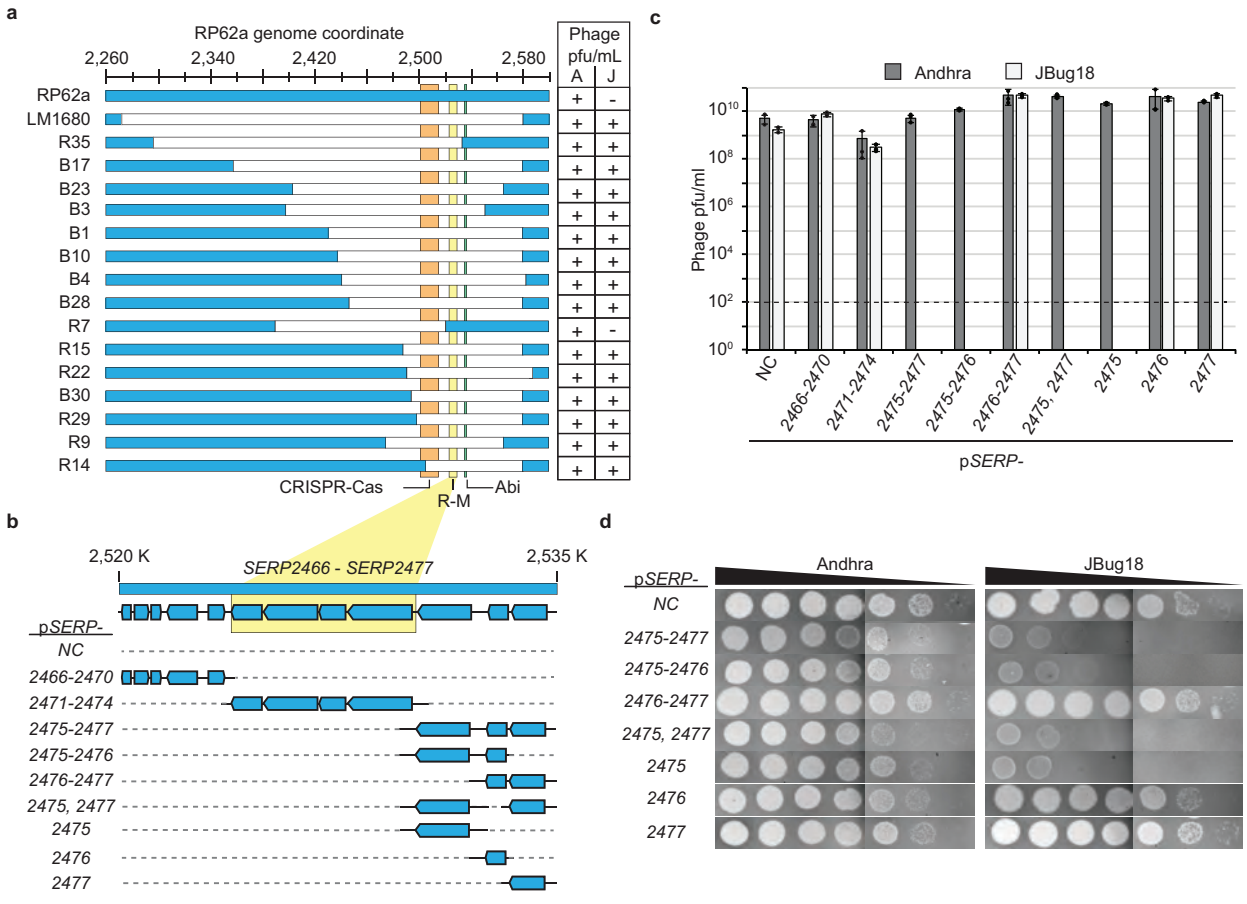
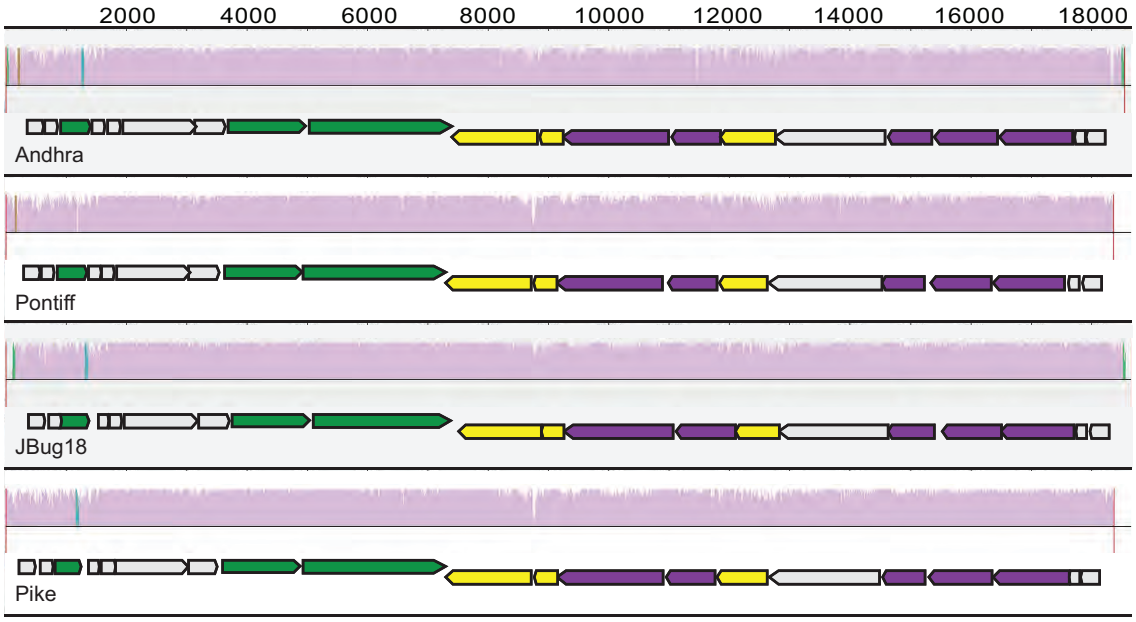


Figure 4. Proposed mechanism for SERP2475 (Nhi) immunity. (**a**) Image of the linear double-stranded genomic DNA of staphylococcal *Podoviridae* phages, in which each 5'-end is covalently linked to a terminal protein (TP). (**b**) During DNA replication, double-stranded DNA binding proteins (DSB) bind and unwind the DNA ends. Under normal circumstances (left) a DNA polymerase (DNAP)-TP complex binds to the free 3'-end and initiates DNA replication. However, in the presence of Nhi (N, right), the 3'-end is susceptible to unwinding/degradation by this enzyme. (**c**) Phage transcripts are also targeted by this enzyme. RNAP, RNA Polymerase.



Extended Data 1. *SERP2475* provides robust immunity against phage JBug18. (**a**) A segment of the S. *epidermidis* RP62a genome and deletion mutants. Regions encoding CRISPR-Cas, R-M, and Abi systems are highlighted. Strains were challenged with Andhra (A) and JBug18 (J), and resulting plaque-forming units per milliliter (pfu/mL) are indicated: +, ~1 x 10^9 pfu/mL; -, 0 pfu/mL. (**b**) Magnified view of the genomic region responsible for immunity and corresponding plasmids that were created. (**c**) *S. epidermidis* LM1680 strains harboring indicated plasmids were challenged with phage, and the resulting pfu/mL are shown as an average of triplicate measurements (± S.D.). The dotted line indicates the limit of detection. (d) Representative plaque images following the application of ten-fold dilutions of Andhra and JBug18 ($1x10^{-1}$ – $1x10^{-7}$) atop *S. epidermidis* LM1680 strains bearing indicated plasmids.

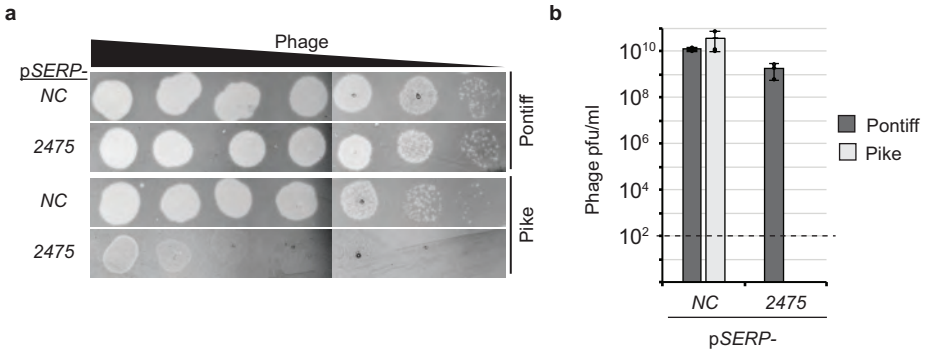


487

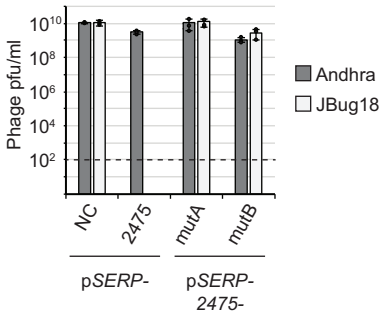
488

489

490 **Extended Data 2. Four related *Podoviridae* phages with different host ranges.** Shown is a
491 multiple genome alignment of *S. epidermidis* podophages Andhra, Pontiff, JBug18, and Pike.
492 Genome coordinates are shown on top, and colored histograms indicate the nucleotide
493 similarity at each position derived from a multiple sequence alignment. The open reading frames
494 for each phage are shown underneath the corresponding histogram. The histograms were
495 generated using the MAUVE open source software (http://darlinglab.org/mauve/mauve.html)
496 and the outlines of open reading frames from the MAUVE output were overlaid using Adobe
497 Illustrator.



Extended Data 3. *SERP2475* is sufficient to provide robust protection against phage Pike. (**a**) 10-fold dilutions of phages Pontiff and Pike ($1x10^{0}$ – $1x10^{-6}$) were plated atop *S. epidermidis* LM1680 cells bearing indicated plasmids. Shown are representative plate images for triplicate experiments. (**b**) The average of triplicate plaque measurements (±S.D.) in pfu/mL. The dotted line indicates the limit of detection for this assay.



505

506

507

508 **Extended Data 4. Accompanies Fig. 2. Predicted nuclease and ATP binding sites in**

509 **SERP2475 are required for immunity *in vivo*.** *S. epidermidis* LM1680 strains bearing the

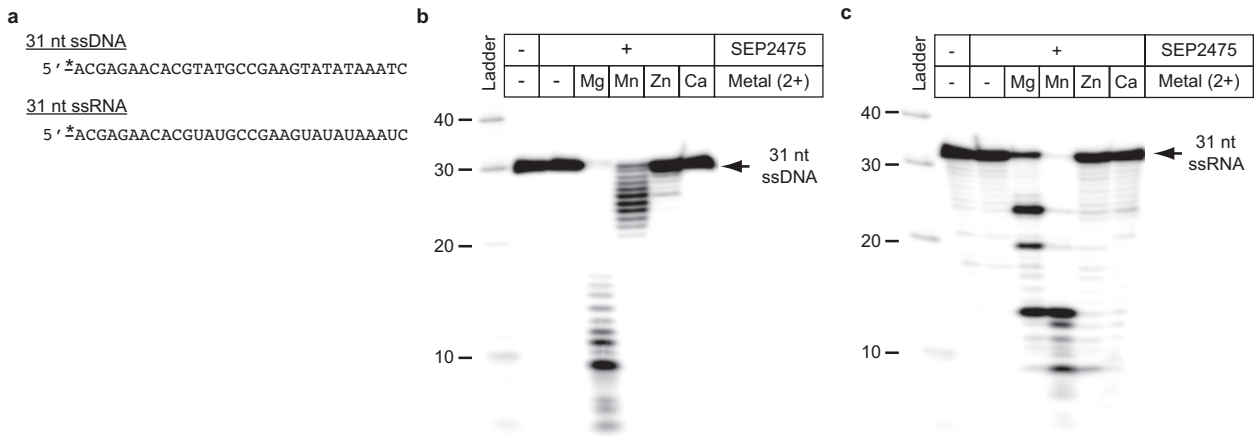
510 indicated plasmids were challenged with phages Andhra and JBug18, and resulting pfu/mL

511 were enumerated. While the p*SERP-2475 -mutA* plasmid encodes mutations in the predicted

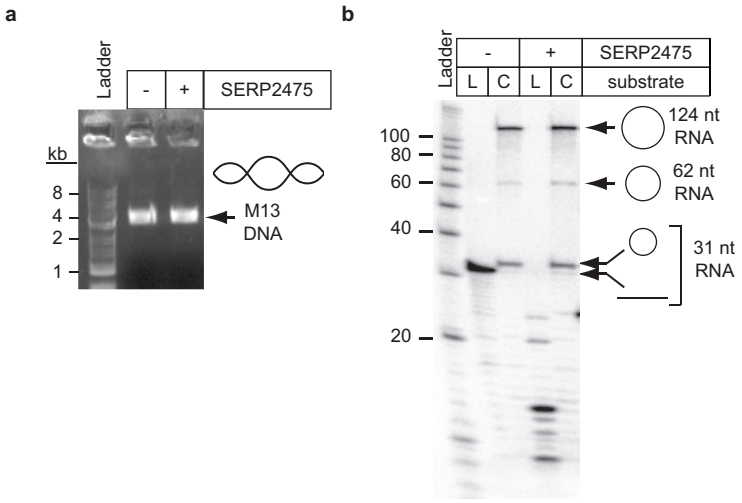
512 nuclease active site, the *-mutB* plasmid encodes mutations in the predicted ATP binding domain

513 (see Fig. 2a for more details). Shown are an average of triplicate measurements (±S.D.). The

514 dotted line indicates the limit of detection for this assay.



Extended Data 5. Accompanies Fig. 2. SERP2475 exhibits metal-dependent nuclease activity. (**a**) The 31-nucleotide single-stranded DNA and RNA substrates used for nuclease assays in this figure and Fig. 2 panels (**c**) and (**g**). Asterisks indicate the position of the radiolabel. (**b** and **c**) DNA and RNA substrates (respectively) were combined with SERP2475 in nuclease buffer supplemented with indicated divalent metals. Shown are representative images for three independent trials.



526

527

528

529

530 **Extended Data 6. Accompanies Fig. 2. SERP2475 cannot cut circular single-stranded**

531 **substrates**. Shown are nuclease assays in which SERP2475 was combined with indicated

532 single-stranded circular substrates. (**a**) For circular DNA, the reaction was supplemented with 10

533 mM $MgCl_2$ and the mixture was incubated at 37°C for 60 min. The reaction products were then

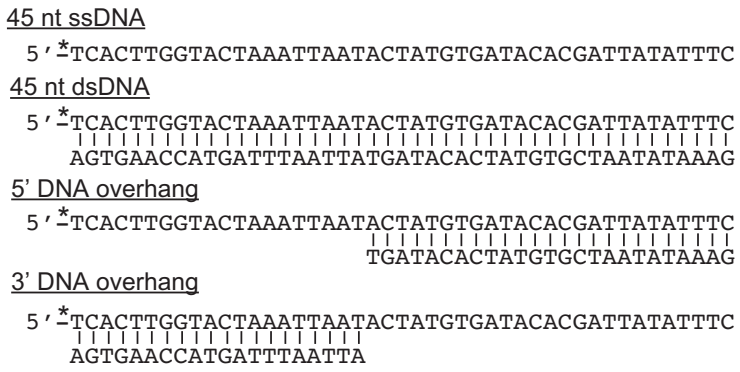
534 resolved on a 1% agarose gel. (**b**) For circular RNA, the reaction was supplemented with 10 mM

535 $MnCl_2$ and the mixture was incubated at 37°C for 20 min. The products were then resolved

536 using denaturing PAGE. L, linear; C, circular. Shown are representative images for two

537 independent trials.

538



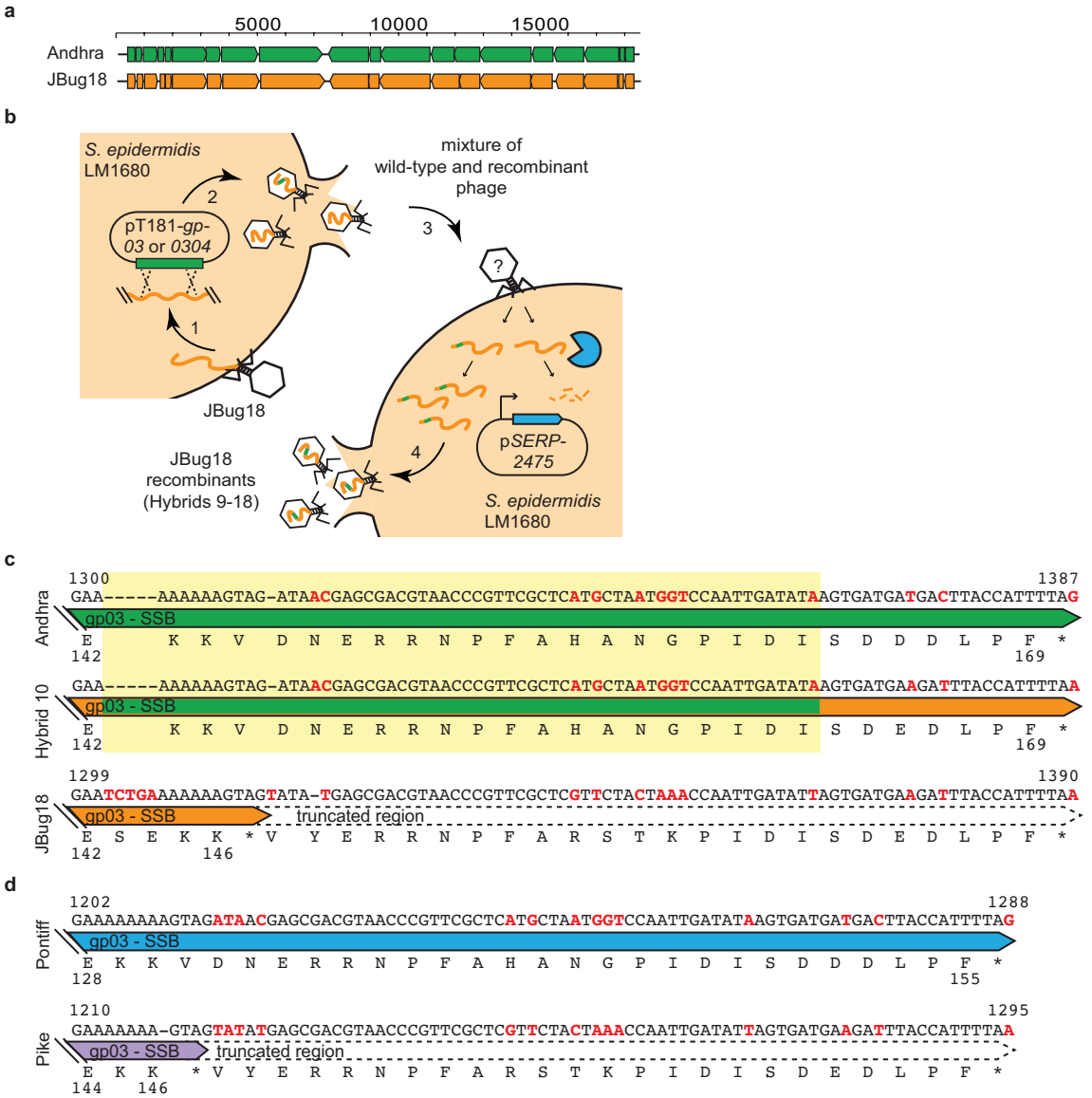
539

540

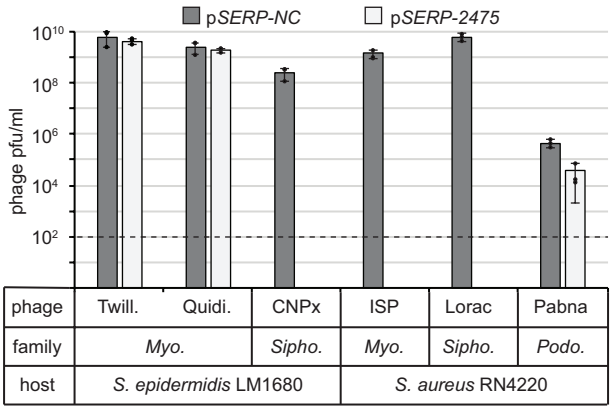
541

542 **Extended Data 7. Accompanies Fig. 2.** Sequences of DNA substrates used in nuclease and

543 helicase assays in Fig. 2, panels (**d**) and (**f**). Asterisks indicate the position of the radiolabel.

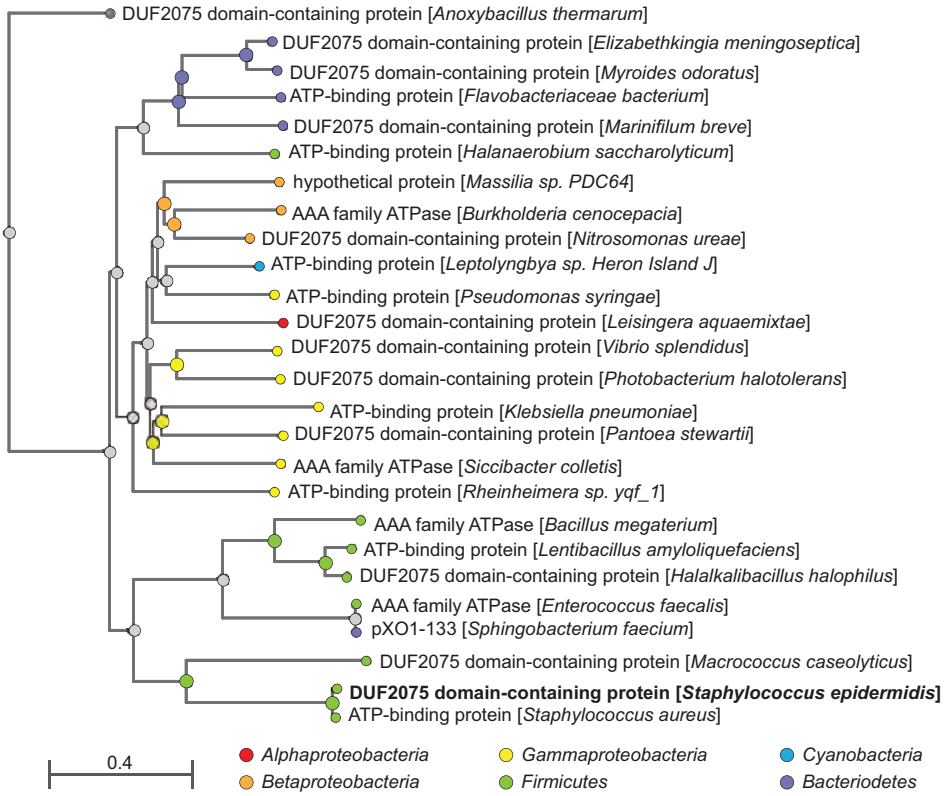


Extended Data 8. Accompanies Fig. 3. Generation of Hybrids 9-18 and comparison of SSB sequences. (**a**) A pairwise comparison of the open reading frames of Andhra and JBug18. (**b**) A diagram of the method used to generate JBug18-Andhra Hybrids 9-18. (**c**) Sequence comparison between the SSBs of Andhra, JBug18, and Hybrid 10, which gained resistance to immunity through the acquisition of only 60 nucleotides of Andhra-derived sequence (highlighted in yellow). (**d**) a similar comparison between the SSBs of phages Pontiff and Pike.



Extended Data 9. SERP2475 protects against diverse staphylococcal phages.

Staphylococcal strains bearing p*SERP-NC* or p*SERP-2475* were plated together with indicated phages. The resulting pfu/mL are shown as an average of triplicate measurements (±S.D.). The dotted line indicates the limit of detection. Twill., Twilingate; Quidi., Quidividi; *Myo.*, *Myoviridae*; *Sipho.*, *Siphoviridae*; *Podo.*, *Podoviridae.*



Extended Data 10. SERP2475 homologs are found in diverse bacterial phyla. A phylogenetic tree constructed from the alignment of SERP2475 homologs found in 25 representative genera. All homologs selected share >30% sequence identity across >70% of their protein sequence. The tree was generated using the NCBI BLAST tool (https://blast.ncbi.nlm.nih.gov/Blast.cg).