nature chemical biology

ARTICLE

PUBLISHED ONLINE: 1 FEBRUARY 2016 | DOI: 10.1038/NCHEMBIO.2018

Assembly and clustering of natural antibiotics guides target identification

Chad W Johnston[1,2], Michael A Skinnider[1,2], Chris A Dejong[1,2], Philip N Rees[1,2], Gregory M Chen[1,2], Chelsea G Walker[1,2], Shawn French[1], Eric D Brown[1], János Bérdy[3], Dennis Y Liu[1,2] & Nathan A Magarvey[1,2]*

Antibiotics are essential for numerous medical procedures, including the treatment of bacterial infections, but their widespread use has led to the accumulation of resistance, prompting calls for the discovery of antibacterial agents with new targets. A majority of clinically approved antibacterial scaffolds are derived from microbial natural products, but these valuable molecules are not well annotated or organized, limiting the efficacy of modern informatic analyses. Here, we provide a comprehensive resource defining the targets, chemical origins and families of the natural antibacterial collective through a retrobiosynthetic algorithm. From this we also detail the directed mining of biosynthetic scaffolds and resistance determinants to reveal structures with a high likelihood of having previously unknown modes of action. Implementing this pipeline led to investigations of the telomycin family of natural products from *Streptomyces canus*, revealing that these bactericidal molecules possess a new antibacterial mode of action dependent on the bacterial phospholipid cardiolipin.

The increasing prevalence of antibacterial resistance in the clinic has highlighted the need for new antibacterial drugs with divergent mechanisms[1,2]. With few exceptions, clinically relevant antibacterials are derived from microbial natural product antibiotics[3]. Although the scale and size of this natural antibacterial collective has been poorly defined to date, it is composed of valuable chemical structures that have seemingly been honed through evolution to provide microbes with competitive advantages in their native environments[4]. Biosynthetic pathways, including modular assembly line–like enzymes, facilitate the synthesis of families of molecules that can undergo natural selection and yield privileged scaffolds[5]. Significant effort has been devoted to identifying clinically applicable antibacterial targets[6], but it is perhaps not surprising that nearly every known druggable target has a dedicated natural product antibiotic directed toward it[7–9] (**Fig. 1**). As such, new and known natural product antibacterials should provide innovation to derive the next generation of clinical antibiotics if they can be collected and charted adeptly using new chemo- and bioinformatic methods[10].

Natural products produced by modular biosynthetic gene clusters frequently possess bioactive chemical scaffolds that have been honed by natural selection, leading to families of molecules that represent variations on a conserved core directed toward a common molecular target[5]. Using new chemoinformatic techniques that can leverage our knowledge of biosynthesis, these families can be elaborated and charted so as to direct efforts toward exotic natural products with little chemical similarity to families with known modes of action. Bioinformatic analysis can also be used to direct efforts toward molecules that are unlikely to possess known targets. While many natural product biosynthesis genes are useful for inferring natural product structures, others can provide clues to natural product function. Studies of antibacterial resistance genes that are naturally associated with biosynthetic gene clusters[11,12] have revealed that resistance determinants typically have functions related to a specific class of small molecule, corresponding to a given antibiotic scaffold or its molecular target[13–16]. As such, resistance genes may provide a unique opportunity to reveal the target of a given antibiotic, independent of structural similarity to known molecules. Computational pipelines that could incorporate extensive accumulated knowledge of natural product structures and biosynthesis may create a means to identify orphaned antibacterial agents with unique mechanisms.

The golden age of antibiotic discovery (1940s–1970s) produced a massive quantity of data at a time when modern big data organization tools did not exist. Thousands of antibiotics were discovered, but only a small handful were developed and have persisted as drugs until recently. With the advent of widespread antibacterial resistance in the clinic[17], refactoring these old scaffolds is becoming more difficult and less effective, and new molecules will be needed to take their places. To capitalize on the valuable—but disparate—data uncovered from the golden age of discovery through to the present day will require that these data be collected, annotated and processed by modern chemo- and bioinformatic methods[18]. Reconciling this massive quantity of chemical and biological data with modern techniques could provide an efficient, directed means to quickly identify chemical scaffolds with new modes of action to reinvigorate the antibiotic pipeline. Following the collection and organization of the known antibacterials, compelling orphan antibiotics can be identified and revisited, given that many of these valuable structures were discarded for reasons that are not necessarily relevant in the current climate of antibacterial resistance[19]. Revisiting promising orphaned antibiotics at this time makes sense in light of these new clinical realities and of the recent development of chemical and biosynthetic methods that provide means to modify a wider range of scaffolds and improve clinical success[20,21]. Orphaned antibiotics from microbes may also create new infection treatment strategies[22] and antibacterial mechanisms[7,23,24] for development. As flagship antibiotic scaffolds succumb to resistance, new approaches are needed to find molecules with new modes of action. In this work, we compile, for the first time, a comprehensive collection of antibacterial natural products and profile them with a retrobiosynthetic algorithm to define rare classes without known resistance mechanisms or known

[1]Department of Biochemistry & Biomedical Sciences, M.G. DeGroote Institute for Infectious Disease Research, McMaster University, Hamilton, Ontario, Canada. [2]Department of Chemistry & Chemical Biology, McMaster University, Hamilton, Ontario, Canada. [3]Eötvös Loránd University, IVAX Drug Research Institute Ltd., Budapest, Hungary. *e-mail: magarv@mcmaster.ca

modes of action. This analysis led to investigation of the telomycin family of natural products (compound **1**), which are shown to possess a new antibacterial target—the phospholipid cardiolipin.

RESULTS

Charting natural antibiotics

In contrast to the burgeoning of genetic databases detailing the biosynthetic origins of natural products[25–27], there has been little progress in the development of a comprehensive annotated chemical database of extant microbial antibiotic natural products. Our systematic analysis of antibacterial natural products was initiated by incorporating an up-to-date version of the *Handbook of Antibiotic Compounds*[28,29], and extensive reviews of published literature and patents, into a single database. Chemical structures for each known antibiotic were generated and rendered in SMILES format (10,343 compounds). Manual annotation of the specificity of each compound indicated that 7,184 molecules demonstrated nonspecific antibacterial activity, while 3,159 displayed specific antibacterial activity. This specificity analysis involved a comprehensive review of primary literature for each natural product described, assessing bioactivity toward bacteria as well as fungi, plants, human cell lines and animals. Nonspecific antibacterials were found from fungi (35%) as well as bacteria (66%), while specific antibacterial agents were significantly enriched in bacterial producers (96%) as opposed to fungi (4%), where specific antibacterials were most often fusidic acids, mutilins and penicillins. Metadata related to the known or suspected mechanisms of action of these specific antibacterials were also incorporated into our analysis, revealing 54 described natural antibacterial targets (**Fig. 1**). Antibiotic classes that make use of modular biosynthetic pathways were segregated for analysis, including ribosomal and nonribosomal peptides, polyketides, β-lactams, and carbohydrate or aminoglycoside superfamilies. To generate high-fidelity subgroupings of these molecules, we developed a retrobiosynthetic algorithm that could utilize biosynthetic and chemoinformatic logic to assign likeness, as opposed to classical chemoinformatic metrics such as Tanimoto similarity scoring. The limitations of chemoinformatic relationship scoring using Tanimoto have been reported previously, particularly with molecules that are as diverse as natural products[30,31]. To facilitate high-fidelity clustering, our new algorithm was designed to integrate the distinct biosynthetic origins of the aforementioned modular natural product superfamilies, identifying substrates contained within these antibiotics and the monomers used in their construction. We reasoned that this retrobiosynthetic algorithm could retrace the evolutionary histories of these selective antibacterial agents and provide a means of allocating them into defined subgroups. This algorithm was incorporated as a publicly available Java-based web application, Antibioticome (**Supplementary Results**, **Supplementary Fig. 1**), that selectively deconstructs molecules from modular natural product superfamilies using a series of custom-tailored rules to provide a proto-scaffold consisting of the originating building blocks. Each series of individual components are used as a fingerprint for their respective antibiotic, facilitating antibacterial fingerprint alignment via a Needleman-Wunsch algorithm (C.A.D., G.M.C., H. Li, C.W.J., P.N.R., M.A.S. *et al.*, unpublished observations; **Fig. 2a**). By processing these data through

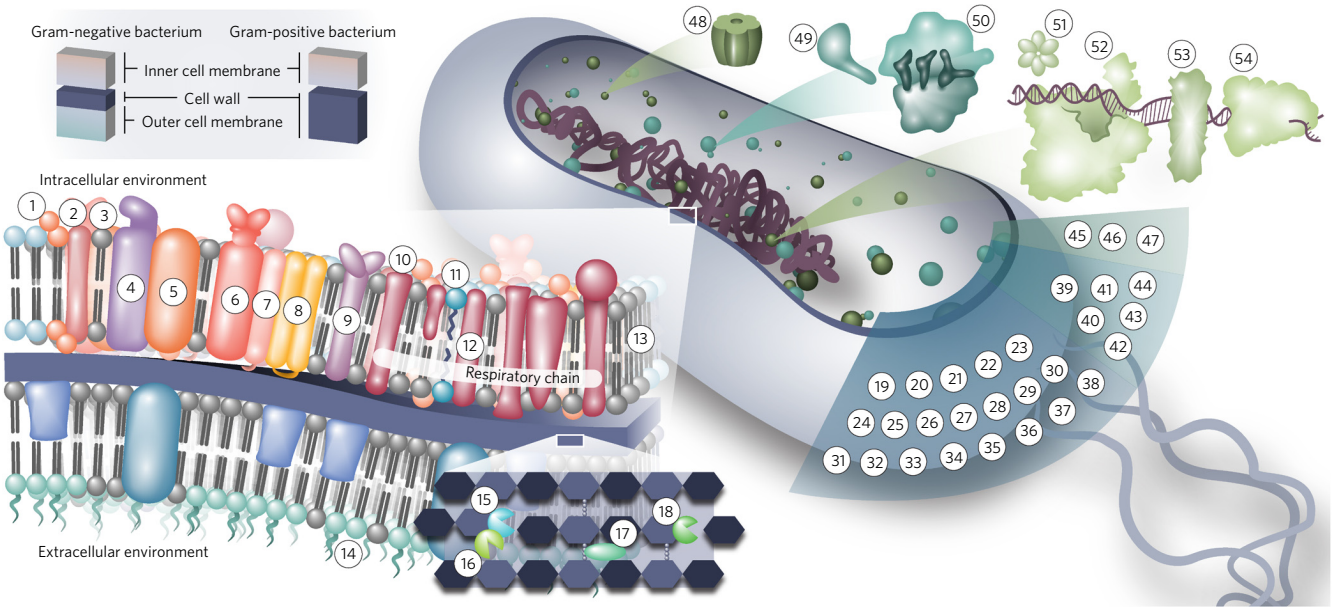

Figure 1 | Microbial natural products with specific antibacterial activity define a diverse range of antibacterial targets. A comprehensive review of the antibacterial natural products identified 54 distinct molecular targets of known antibacterials, including membrane-associated targets (1, lipid II; 2, undecaprenyl phosphate; 3, phosphatidylethanolamine; 4, Na-dependent NADH-quinone reductase; 5, translocase 1; 6, SecA–YEG complex; 7, type 1 signal peptidase; 8, type 2 signal peptidase; 9, WalK-WalR two-component system; 10, NADH oxidase; 11, menaquinone; 12, cytochrome *bd*; 13, membrane stability; 14, LPS), cell wall–associated enzymes (15, peptidoglycan transglycosylase; 16, penicillin binding proteins; 17, D-Ala-D-Ala dipeptide; 18, UDP-N-acetylglucosamine-3-enolpyruvyl transferase), targets associated with amino acid biosynthesis and metabolism (19, L-histidinol phosphate aminotransferase; 20, N-acetylornithine transaminase; 21, isoleucine-tRNA synthetase; 22, leucine-tRNA synthetase; 23, tryptophan-tRNA synthetase; 24, proline-tRNA synthetase; 25, tyrosyl-tRNA synthetase; 26, threonine-tRNA synthetase; 27, phenylalanine-tRNA synthetase; 28, aspartate-tRNA synthetase; 29, serine-tRNA synthetase; 30, aspartate semialdehyde dehydrogenase; 31, alanine racemase; 32, glutamine synthase; 33, pyruvate dehydrogenase; 34, threonine synthase; 35, biotin aminotransferase; 36, homoserine-O-succinyltransferase; 37, pyruvate carboxylase; 38, peptide deformylase), fatty acid biosynthesis (39, FabF; 40, FabB; 41, acyl-CoA carboxylase; 42, FabD; 43, FabG; 44, FabI), individual metabolic enzymes (45, glucosamine-6-phosphate synthase; 46, deoxyxylulose phosphate reductoisomerase; 47, 4-amino-4-deoxychorismate (ADC) synthase) and components of macromolecular machinery (48, ClpP protease; 49, elongation factors; 50, ribosome; 51, transcription termination factor Rho; 52, RNA polymerase; 53, DNA gyrase; 54, DNA polymerase). Image courtesy of Sam Holmes.

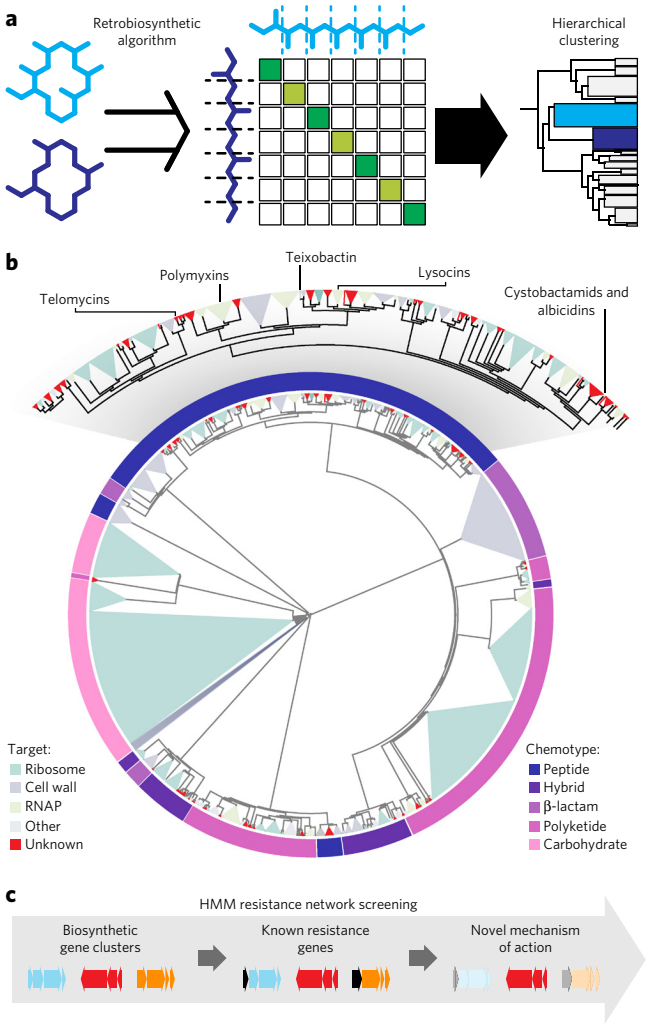

Figure 2 | A retrobiosynthetic strategy for charting antibacterial natural products and identifying rare scaffolds with new molecular targets. (**a**) A retrobiosynthetic algorithm was devised to deconstruct antibiotics according to their modular biosynthetic origins and facilitate high-fidelity clustering of related families. Products of modular assembly architectures are broken down with class-specific rules, providing biosynthetic monomers that can be aligned and scored using a similarity matrix that can be used for hierarchical molecular clustering. (**b**) A sampled collection of antibacterials, including 1,908 modular natural products with specific antibacterial activity. Families are colored by known or unknown mechanism of action, and natural product chemotypes are highlighted by color (outer ring). A section of the peptide antibiotics is inset, demonstrating distinct branches for a number of antibacterial peptide subfamilies. (**c**) Schematic for PRISM-based HMM resistance network screening, outlining the automated detection of biosynthetic gene clusters and detection of known resistance determinants guiding focus toward gene clusters with potentially novel mechanisms of action.

hierarchical clustering, we reasoned that we would create a protocol (see Online Methods) sufficient to reveal the chemoinformatic relationships of all members of these classes and define groups of specific antibiotics with no known mode of action. Results of this analysis were readily organized into subfamilies comprising 1,908 distinct molecules, with an exceptionally low rate of incorrect localization during clustering (<0.5%; **Fig. 2b**).

Parsing the accumulated antibiotic data revealed the distribution of molecules with specific antibacterial activity originating from modular biosynthetic assemblies. Within these groups, 668 molecules were of peptidic origin, followed in abundance by polyketides (571), carbohydrates (354), β-lactams (196) and hybrid molecules (119). The majority of the molecules in our analysis belong to subfamilies with known mechanisms of action (88.3%), spanning 32 established antibacterial mechanisms (**Fig. 1**), and molecules without established molecular targets are rare (11.7%; **Fig. 2b**). The discrepancy between all known targets and the targets identified in our analyzed subset of antibacterials largely reflects the presence in the former group of small antimetabolites that interfere with amino acid metabolism and that did not yield usable alignment data following deconstruction via our retrobiosynthetic algorithm. Nonribosomal peptides showed the most variation in mode of action, including 22 distinct molecular targets. Polyketides and hybrid assembly structures also demonstrated considerable diversity, with 10 and 7 antibacterial targets, respectively. Collectively, these products were shown to affect 30 distinct targets. For the β-lactams (3) and carbohydrates (2) the list is much smaller: despite the fact that these groupings have a large number of individual members, the target diversity is relatively narrow. Inhibition of the ribosome was the most frequently observed antibacterial mechanism, applicable to 45% (851 molecules) of the specific antibacterials from our modular antibacterial natural product sample, which corresponds to 51% of those with known mechanisms (**Fig. 2b**). Inhibition of cell wall biosynthesis through various molecular targets was the second most frequently observed mechanism among the sampled antibacterials with known mechanisms (26%; 441 molecules), followed by inhibition of RNA polymerase (4.3%; 73 molecules). Analysis of higher-order families (such as thiazolyl peptides, glycopeptides or aminoglycosides) demonstrated the frequency of distinct, evolved scaffolds inhibiting a given target. The ribosome was again the most frequent hit and was the target of 19 distinct families from our sample of antibacterials with modular biosynthetic origins. RNA polymerase was also a frequent target, with 9 distinct families of inhibitors, as was DNA gyrase, which was the target of 5 distinct classes. Notably, of the 137 distinct families observed in our sample of 1,908 modular, specific antibacterials, 54 did not possess known mechanisms of action. This also means that the 83 molecular classes with known mechanisms in our sample covered 32 molecular targets, indicating a frequency for new scaffolds inhibiting new targets of nearly 40%. Closer inspection of well-established targets such as the ribosome or RNA polymerase shows that individual natural product scaffolds bind a wide variety of sites on these macromolecular assemblies[32,33], demonstrating that assembly-line antibiotics will evolve to create chemical and biological diversity. Analysis regarding the specificity of these diverse natural antibiotics for their established targets is well documented in primary literature and reviews (**Supplementary Data Set 1**).

This sampling of natural antibacterials demonstrates that certain structural classes are more effective than others at generating chemical and biological diversity. Peptide antibiotics were found to have the highest degree of chemical and mechanistic diversity, as well as the largest number of subfamilies without known mechanisms of action. Given the nature of their modular biosynthetic machinery and the degree to which monomers and scaffolds can be tailored, nonribosomal peptides offer the greatest number of combinations and permutations for natural selection. These peptides appear to be uniquely capable of interacting specifically with structural small molecules or membrane components and can act as potent bactericidal agents[22,24].

Creating a resistance-determinant library

A salient feature in the hierarchical clustering data obtained through our retrobiosynthetic analysis is that subfamilies share a chemical scaffold and a common mechanism of action. With this unique analytical capacity, we could rapidly define all subfamilies from our sampling of the natural antibacterial collective, on the basis of their

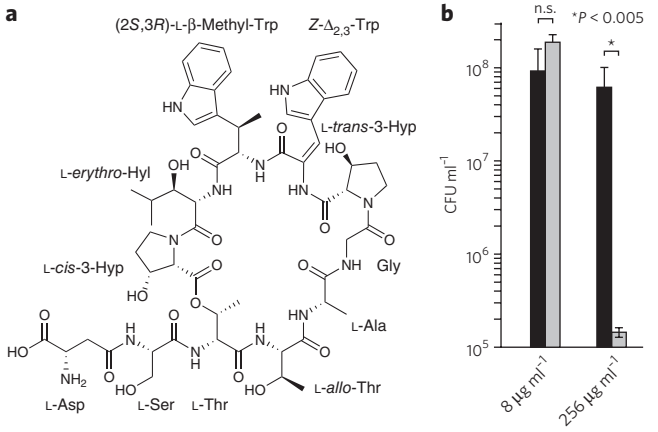

Figure 3 | Telomycin is a nonribosomal peptide that lyses bacteria through an unknown mechanism. (**a**) Structure of telomycin. (**b**) Telomycin lyses *S. aureus* at high concentrations. Wild-type *S. aureus* (OD_{600} 0.2) was exposed to 8 (1× MIC) or 256 μg mL^{-1} (32× MIC) telomycin. Colony-forming units (CFU) were counted before treatment (black) and after 90 min (white). Results are shown as mean ± s.d.; $n = 4$; two-tailed Student's *t*-test; $P = 0.0023$. Data are representative of three independent experiments.

chemical structures and shared biosynthetic components. To interrogate the genetic origins of promising subfamilies, we devised a bioinformatic web application for detecting and displaying biosynthetic gene clusters, PRISM[34] (**Supplementary Fig. 2**). By using a series of hidden Markov models (HMMs), BLAST databases and chemoinformatic algorithms, PRISM detects known and unknown gene clusters, provides similarity scores with known biosynthetic clusters, produces high-fidelity predictions of chemical scaffolds and identifies all known self-resistance determinants (Online Methods). Several key works have underscored that resistance is often disseminated from antibiotic-producing organisms and is associated with biosynthetic gene clusters for antibacterial agents[12,13,16]. Classically, this collection of known resistance determinants has been used to expedite the tracking and spread of antibacterial resistance in a clinical setting. Now, by associating self-resistance determinants with known antibacterial mechanisms, PRISM can scan biosynthetic gene clusters for resistance genes that can predict antibacterial targets. By combining known cluster scoring, chemical predictions and resistance screens, this pipeline can assist in validating antibiotics and prioritizing those with unique mechanisms that lack cross-resistance (**Fig. 2c**).

To facilitate optimal detection of resistance genes, we collected self-resistance determinants from all known antibiotic pathways and derived 108 HMMs that span mechanisms involved in drug efflux, drug modification and target alteration as well as decoy targets (**Fig. 2c**; **Supplementary Data Set 2**). This database was subsequently paired with previously collected HMMs from clinically observed resistance genes to provide a comprehensive means of detecting known resistance genes[35]. To assess the use of our pipeline for profiling self-resistance determinants that indicate mechanism, we loaded a series of DNA sequences from biosynthetic gene clusters into PRISM. Biosynthetic genes were automatically detected and depicted, and known molecule identity was determined. PRISM generated predictions of the natural product structure, assessed similarity to gene clusters with known products, and identified a diverse series of scaffold- and target-specific resistance genes (**Supplementary Figs. 3–11**). Using the integrated HMMs, PRISM detected resistance determinants in each analyzed cluster that were revealing of the activity for the chemical class produced. As unique structures are necessary, but not sufficient, for a unique mechanism of action, this pipeline provides a crucial second measure to ensure that targeted rare scaffolds possess similarly uncommon modes of action without cross-resistance.

Investigation of the mode of action of telomycin

On the basis of our defined retrobiosynthetic clustering of antibacterial natural products, we sought to investigate under-represented chemical scaffolds with the potential to possess uncharacterized molecular targets. Given their structural and mechanistic diversity, we chose to investigate nonribosomal peptides in particular. From a set of promising candidates including, among others, the pyloricidins[36] and griselimycins[37], we chose to investigate the telomycins[38–41] (**Fig. 3a**), which possess potent bactericidal activity against a wide range of Gram-positive organisms[39,42]. The telomycin producer *Streptomyces canus* ATCC 12647 was purchased from the ATCC and submitted for genomic sequencing. Using PRISM, we were able to quickly identify a candidate telomycin biosynthetic gene cluster from the assembled contigs. The cluster was found to encode three nonribosomal peptide synthetases and a series of accessory enzymes consistent with formation of the telomycin undecapeptide core, precursor supply and tailoring (**Supplementary Table 1**). When this biosynthetic gene cluster was passed through our resistance screening pipeline, no genes were identified with similarity to known specific self-resistance determinants (**Supplementary Fig. 12**), indicating a potentially unique mechanism of action. These results are consistent with previous observations that the telomycin-related molecule LL-AO341β does not demonstrate cross-resistance with other antibacterial agents[42]. This automated assessment of potential resistance profiles supported additional efforts to be directed toward describing the mechanism of action of telomycin.

Telomycin is a lytic antibacterial agent, inducing rapid lysis of Gram-positive bacteria including *Staphylococcus aureus* and *Bacillus subtilis* at concentrations above the minimum inhibitory concentration (MIC; **Fig. 3b**). Telomycin also does not have any apparent cytotoxicity, as previous studies[43] with Chinese hamster ovary (CHO) cells did not observe toxicity at concentrations up to 40 μg mL^{-1}, and we did not observe cytotoxicity with human embryonic kidney (HEK293) cells at concentrations up to 128 μg mL^{-1}. Further, telomycin has demonstrated efficacy in animal studies without apparent toxicity[40] and is not hemolytic (**Supplementary Fig. 13**). To determine the mechanism of specific bacterial lysis, *S. aureus* and *B. subtilis* strains were exposed to telomycin to generate and select for spontaneous suppressor mutants, which were sequenced to reveal candidate target loci. Telomycin-resistant bacteria generated with limited passaging were refractory to lysis (**Fig. 4a**) and demonstrated a 16-fold increase in MIC (**Supplementary Table 2**). Comparison of spontaneously resistant mutant genomes with those of their sensitive parent strains revealed a series of inactivating mutations in cardiolipin synthase (*S. aureus cls2*, *B. subtilis clsA*; **Supplementary Fig. 14**). Cardiolipin is a phospholipid dimer, formed from the condensation of two phosphatidylglycerol monomers, that is found in the lipid membranes of bacteria and the inner membranes of mitochondria. As telomycin-resistant mutants possessed truncating or missense mutations affecting the active site of cardiolipin synthase and were refractory to telomycin-mediated lysis, cardiolipin was investigated as a potential target. Resistant mutants possessed levels of cardiolipin ranging from 2% to 10% of wild-type concentrations, with residual cardiolipin presumably produced by the remaining auxiliary cardiolipin synthase gene (*S. aureus cls1*, *B. subtilis clsB*; **Supplementary Table 2**). Cardiolipin levels are known to increase following growth in conditions of high salinity[44], and we found that greater salinity correlated with increased telomycin activity (**Supplementary Table 3**). Having demonstrated that cardiolipin levels were strongly correlated with the antibacterial activity of telomycin, we sought to test whether telomycin specifically interacts with cardiolipin. In contrast to mixtures of telomycin with other lipids, telomycin and cardiolipin rapidly formed an opaque

precipitate in a dose-dependent manner (**Fig. 4b**). This molecule-specific turbidity often occurs with peptide antibiotics that specifically bind membrane lipids, for instance when cinnamycin binds phosphatidylethanolamine[45] or when lysocin binds bacterial menaquinone[24]. Consistent with this observation, and with the hypothesis that telomycin exerts lytic activity through a specific interaction with cardiolipin, excess cardiolipin completely abolished the antibacterial action of telomycin, in contrast to other bacterial phospholipids (**Fig. 4c**,**d**). This effect was dose dependent (**Fig. 4c**) and was largely cardiolipin specific, although the monomeric cardiolipin analog phosphatidylglycerol also demonstrated some protective effect when provided in 10× molar excess.

Cardiolipin possesses significant negative curvature relative to other phospholipids, which causes it to accumulate at poles and septa of bacterial cells[46–48]. Although this has previously been established through microscopy, the only dye available to image cardiolipin is a simple derivative of acridine orange (10-*N*-nonyl acridine orange; NAO)—which was recently found to promiscuously stain anionic lipids in general, with limited selectivity for cardiolipin[48]. Nevertheless, to investigate whether telomycin also localizes to these cardiolipin rich areas, we generated a fluorescein-labeled telomycin conjugate. Telomycin possesses a single free amine from the N-terminal aspartate residue, which is notably absent from the related molecule LL-AO341β and does not appear to affect activity. This amine was reacted with *N*-hydroxysuccinimide carboxyfluorescein to generate a fluorescent conjugate that showed a slightly lower antimicrobial activity than telomycin. Imaging studies with live *B. subtilis* revealed that telomycin localizes to poles and septa (**Fig. 4e**) in a manner similar to that of the established cardiolipin dye NAO[46–48], albeit with less background signal from seemingly nonspecific staining (**Supplementary Fig. 15**).

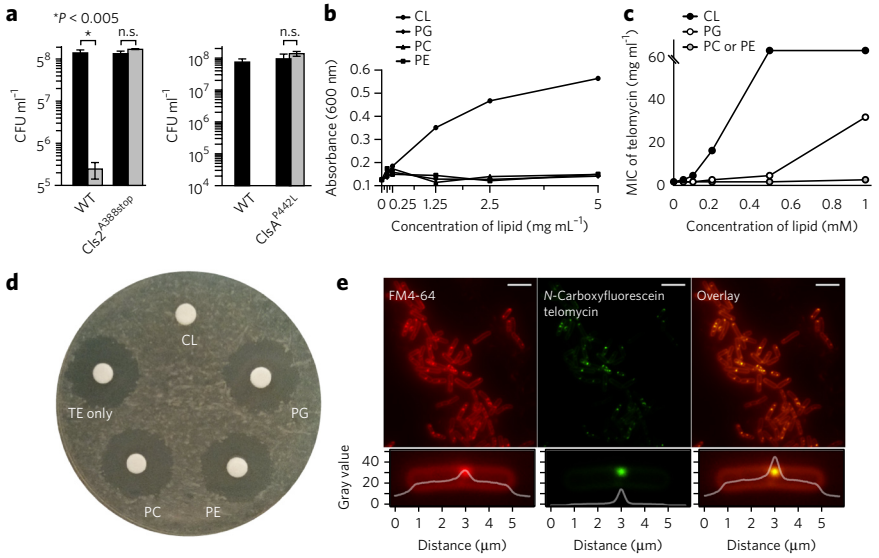

Figure 4 | Telomycin exerts bactericidal activity by interacting with cardiolipin. (**a**) Mutating cardiolipin synthase provides spontaneous telomycin resistance. Wild-type and spontaneously resistant *S. aureus* (left) and *B. subtilis* (right) were exposed to 32× MIC concentrations of telomycin for 90 min. CFU counts were taken before treatment (black) and after 90 min (gray). No viable wild-type *B. subtilis* cells were observed after treatment. Results are shown as mean ± s.d.; $n = 4$; two-tailed Student's *t*-test; $P = 0.0018$; n.s., nonsignificant. Data are representative of three independent experiments. (**b**) Telomycin-cardiolipin mixtures uniquely and rapidly acquire turbidity. Graph shows turbidity (absorbance at 600 nm) of mixtures of telomycin (final concentration 5 mM; 6.3 mg mL^{-1}) with cardiolipin (CL), phosphatidylglycerol (PG), phosphatidylethanolamine (PE) or phosphatidylcholine (PC) at concentrations ranging from 0.125 to 5 mg mL^{-1} after 5-min incubation. (**c**) Cardiolipin reduces telomycin's antibacterial activity in a dose-dependent manner. Graph shows MICs for wild-type *B. subtilis* with mixtures of telomycin (0.1 mM) and lipids at concentrations ranging from 1 to 0.05 mM. Identical results were observed for PE and PC (gray). (**d**) Cardiolipin abolishes telomycin's antibacterial activity. Shown is *B. subtilis* grown overnight with diffusion discs containing mixtures of telomycin (10 mM) and 2× molar equivalent of CL (II), PG (III), PE (IV) or PC (V). (**e**) Telomycin localization in 1-d cultures of *B. subtilis* labeled with the membrane dye FM4-64 (left) and a fluorescent conjugate of telomycin (middle). Summed pixel intensities of a selected bacterium are depicted below. Scale bars, 5 μm.

Although telomycin possesses a unique mechanism of action of and has demonstrated value *in vivo*[40], the emergence of spontaneous resistance could be problematic in a clinical setting. Interestingly, LL-AO341β shows improved activity relative to telomycin[42], but it lacks the N-terminal aspartate residue as well as a hydroxylation on the C-terminal proline, suggesting that alterations to the telomycin scaffold or tailoring may improve activity. To assess the impacts of structural alterations on antibacterial activity, we isolated six telomycin congeners and elucidated their structures using high-resolution MS, MS/MS sequencing and two-dimensional (2D) NMR experiments (**Supplementary Note**). The planar structures of telomycin B–G (**2**–**7**) were similar to those of the telomycin A scaffold but demonstrated incomplete tailoring, including loss of the *cis*-3-hydroxyproline hydroxylation (B–G), of the *trans*-3-hydroxyproline hydroxylation (C, E–G) or the erythro-hydroxyleucine hydroxylation (F and G) or saturation of the double bond in Z-$\Delta_{2,3}$-tryptophan. Although Z-$\Delta_{2,3}$-tryptophan saturation negatively affected bioactivity, the successive loss of each hydroxylation improved it, indicating that general hydrophobicity and tryptophan orientation may be important in driving the interaction with membranes in general and cardiolipin in particular. To assess the role of tryptophan in initiating this interaction we generated additional telomycin A analogs through directed biosynthesis, incorporating methyl-, methoxyl- or hydroxyl groups at indole position 5 in both tryptophan residues (**Supplementary Note**). Consistent with a role for the indole rings in embedding in the membrane and initiating an interaction, hydroxylation completely abolished telomycin activity, methoxy groups caused a modest decrease in activity, and methylation increased activity (**Supplementary Table 4**). Importantly, telomycin F and di-5-methyltryptophan telomycin (**8**) possessed activity against resistant strains that was comparable with the activity of telomycin A against sensitive strains, suggesting that engineered variants with superior bioactivity can overcome potential resistance and take advantage of cardiolipin as a promising antibacterial target.

DISCUSSION

The chemical and mechanistic diversity of antibacterial natural products has provided privileged scaffolds for drug development along with chemical probes of cellular processes. Here, we systematically profiled antibiotic scaffolds, using a new algorithm to sort these natural products according to their biosynthetic building blocks. Using this approach we could accurately group these diverse chemical structures into families and subfamilies for downstream analysis. Identified subfamilies could be profiled with extant metadata and passed through our comprehensive resistance pipeline to enrich for rare chemical families with unknown mechanisms of

action. As an example of this approach, telomycin was identified as an uncommon scaffold with a potentially novel mechanism of action, which could be validated by *in silico* resistance profiling and subsequent experimental work. Substrate and tailoring promiscuity is a hallmark of these modular antibiotic assembly lines[5], and we used these natural tendencies to produce improved telomycin analogs with superior activities against resistant isolates that were raised in the laboratory.

While this paper was undergoing review, a biosynthetic study of telomycin was published[43], detailing the sequencing and analysis of the telomycin biosynthetic gene cluster as well as the presentation of new telomycin congeners, including two which appear identical to telomycin B and C. This work delineated the boundaries of the telomycin biosynthetic gene cluster, revealing a number of elaborate tailoring modifications that take place to generate the mature telomycin scaffold. Most interestingly, this work demonstrated that telomycin undergoes transient acylation during biosynthesis, and noted that acylated variants had improved activity against antibiotic-resistant bacteria including vancomycin-resistant *Enterococci*. This finding is consistent with our results demonstrating that hydrophobic telomycin variants have improved activity, which is likely to be due to an increased association with bacterial lipid membranes and cardiolipin in particular.

The selectivity of telomycin for cardiolipin is intriguing, as cardiolipin is a unique phospholipid dimer found in the membranes of nearly all bacteria and archaea[46–49]. Telomycin is not hemolytic and is not known to possess significant toxicity to mammals[40], and we observed a lack of cytotoxicity at concentrations up to 128 μg mL^{-1}, in keeping with previous reports using cultured mammalian cell lines[43], in which cardiolipin is exclusively sequestered to the inner mitochondrial membrane[46–49]. Unlike many other peptide antibiotics, some members of the telomycin family of natural products have been shown to possibly be orally bioavailable[50], and it remains possible that future structural optimization may yield useful agents for combating bacterial infections.

In this work, we collected a comprehensive library of antibacterial natural products for the first time and quantified the diversity that modular natural products possess not just in their chemical scaffolds but also in their evolved targets. By collecting and analyzing these molecules with a retrobiosynthetic algorithm, we quantified what natural products chemists have long suspected: that novel natural product scaffolds with specific activity often have novel targets. Assembly line–like enzymes can produce large numbers of unique scaffolds, and new scaffolds have nearly a 40% chance of inhibiting a new target. This unparalleled frequency of success in hitting distinct targets yields obvious conclusions for the future of antibiotic research. By charting and navigating this valuable collection of small molecules, we can readily identify unique chemical scaffolds that will have a high likelihood of possessing distinct antibacterial targets. By taking up the approach, candidates (**Supplementary Data Set 1**) and resources (**Supplementary Fig. 1**) provided here, we hope the scientific community can meet the demand for antibiotics with new mechanisms of action to overcome antibiotic resistance.

Received 15 July 2015; accepted 9 December 2015;
published online 1 February 2016

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. GenBank: The telomycin gene cluster was deposited in GenBank under accession number KT881498.

References

1. Bush, K. *et al.* Tackling antibiotic resistance. *Nat. Rev. Microbiol.* **9**, 894–896 (2011).
2. Fischbach, M.A. & Walsh, C.T. Antibiotics for emerging pathogens. *Science* **325**, 1089–1093 (2009).
3. Newman, D.J. & Cragg, G.M. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J. Nat. Prod.* **75**, 311–335 (2012).
4. Vining, C.L. Roles of secondary metabolites from microbes. in *CIBA Foundation Symposium 171—Secondary Metabolites: Their Function and Evolution* 184–198 (Wiley, Chichester, 1992).
5. Fischbach, M.A. & Clardy, J. One pathway, many products. *Nat. Chem. Biol.* **3**, 353–355 (2007).
6. Payne, D.J., Gwynn, M.N., Holmes, D.J. & Pompliano, D.L. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat. Rev. Drug Discov.* **6**, 29–40 (2007).
7. Baumann, S. *et al.* Cystobactamids: myxobacterial topoisomerase inhibitors exhibiting potent antibacterial activity. *Angew. Chem. Int. Edn. Engl.* **53**, 14605–14609 (2014).
8. Lin, A.H., Murray, R.W., Vidmar, T.J. & Marotti, K.R. The oxazolidinone eperezolid binds to the 50S ribosomal subunit and competes with binding of chloramphenicol and lincomycin. *Antimicrob. Agents Chemother.* **41**, 2127–2131 (1997).
9. Keller, S., Schadt, H.S., Ortel, I. & Süssmuth, R.D. Action of atrop-abyssomicin C as an inhibitor of 4-amino-4-deoxychorismate synthase PabB. *Angew. Chem. Int. Edn. Engl.* **46**, 8284–8286 (2007).
10. Li, J.W. & Vederas, J.C. Drug discovery and natural products: end of an era or an endless frontier? *Science* **325**, 161–165 (2009).
11. Cundliffe, E. & Demain, A.L. Avoidance of suicide in antibiotic-producing microbes. *J. Ind. Microbiol. Biotechnol.* **37**, 643–672 (2010).
12. D'Costa, V.M., McGrann, K.M., Hughes, D.W. & Wright, G.D. Sampling the antibiotic resistome. *Science* **311**, 374–377 (2006).
13. Thaker, M.N. *et al.* Identifying producers of antibacterial compounds by screening for antibiotic resistance. *Nat. Biotechnol.* **31**, 922–927 (2013).
14. Bibikova, M.V., Ivanitskaia, L.P. & Singal, E.M. [Directed screening of aminoglycoside antibiotic producers on selective media with gentamycin]. *Antibiotiki* **26**, 488–492 (1981).
15. Ivanitskaia, L.P., Bibikova, M.V., Gromova, M.N., Zhdanovich, IuV. & Istratov, E.N. [Use of selective media with lincomycin for the directed screening of antibiotic producers]. *Antibiotiki* **26**, 83–86 (1981).
16. Forsberg, K.J. *et al.* The shared antibiotic resistome of soil bacteria and human pathogens. *Science* **337**, 1107–1111 (2012).
17. Boucher, H.W. *et al.* Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clin. Infect. Dis.* **48**, 1–12 (2009).
18. Doroghazi, J.R. *et al.* A roadmap for natural product discovery based on large-scale genomics and metabolomics. *Nat. Chem. Biol.* **10**, 963–968 (2014).
19. Donadio, S., Maffioli, S., Monciardini, P., Sosio, M. & Jabes, D. Antibiotic discovery in the twenty-first century: current trends and future perspectives. *J. Antibiot.* **63**, 423–430 (2010).
20. Walsh, C.T. & Wencewicz, T.A. Prospects for new antibiotics: a molecule-centered perspective. *J. Antibiot.* **67**, 7–22 (2014).
21. Goss, R.J.M., Shankar, S. & Fayad, A.A. The generation of "unnatural" products: synthetic biology meets synthetic chemistry. *Nat. Prod. Rep.* **29**, 870–889 (2012).
22. Ling, L.L. *et al.* A new antibiotic kills pathogens without detectable resistance. *Nature* **517**, 455–459 (2015).
23. Cociancich, S. *et al.* The gyrase inhibitor albicidin consists of *p*-aminobenzoic acids and cyanoalanine. *Nat. Chem. Biol.* **11**, 195–197 (2015).
24. Hamamoto, H. *et al.* Lysocin E is a new antibiotic that targets menaquinone in the bacterial membrane. *Nat. Chem. Biol.* **11**, 127–133 (2015).
25. Weber, T. *et al.* antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* **43**, W1, W237-43 (2015).
26. Medema, M.H. *et al.* Minimum Information about a Biosynthetic Gene cluster. *Nat. Chem. Biol.* **11**, 625–631 (2015).
27. Hadjithomas, M. *et al.* IMG-ABC: a knowledge base to fuel discovery of biosynthetic gene clusters and novel secondary metabolites. *MBio.* **6**, e00932–15 (2015).
28. Bérdy, J. *et al. Handbook of Antibiotic Compounds* Vols. I–X (CRC Press, Boca Raton, Florida, USA, 1980–1982).
29. Bérdy, J. Thoughts and facts about antibiotics: where we are now and where we are heading. *J. Antibiot.* **65**, 385–395 (2012).
30. Koch, M.A. *et al.* Charting biologically relevant chemical space: a structural classification of natural products (SCONP). *Proc. Natl. Acad. Sci. USA* **102**, 17272–17277 (2005).
31. Over, B. *et al.* Natural-product-derived fragments for fragment-based ligand discovery. *Nat. Chem.* **5**, 21–28 (2013).
32. Wilson, D.N. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nat. Rev. Microbiol.* **12**, 35–48 (2014).

33. Srivastava, A. *et al.* New target for inhibition of bacterial RNA polymerase: 'switch region'. *Curr. Opin. Microbiol.* **14**, 532–543 (2011).
34. Skinnider, M.A. *et al.* Genomes to natural products PRediction Informatics for Secondary Metabolomes (PRISM). *Nucleic Acids Res.* **43**, 9645–9662 (2015).
35. Gibson, M.K., Forsberg, K.J. & Dantas, G. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *ISME J.* **9**, 207–216 (2015).
36. Nakao, M. *et al.* Pyloricidins, novel anti-*Helicobacter pylori* antibiotics produced by *Bacillus* sp. I. Taxonomy, fermentation and biological activity. *J. Antibiot.* **54**, 926–933 (2001).
37. Nakajima, M. *et al.* Mycoplanecins, novel antimycobacterial antibiotics from *Actinoplanes awajinensis* subsp. *mycoplanecinus* subsp. nov. II. Isolation, physico-chemical characterization and biological activities of mycoplanecin A. *J. Antibiot.* **36**, 961–966 (1983).
38. Misiek, M. *et al.* Telomycin, a new antibiotic. *Antibiot. Annu.* **5**, 852–855 (1957–1958).
39. Gourevitch, A. *et al.* Microbiological studies on telomycin. *Antibiot. Annu.* **5**, 856–862 (1957–1958).
40. Tisch, D.E., Huftalen, J.B. & Dickison, H.L. Pharmacological studies with telomycin. *Antibiot. Annu.* **5**, 863–868 (1957–1958).
41. Sheehan, J.C., Mania, D., Nakamura, S., Stock, J.A. & Maeda, K. The structure of telomycin. *J. Am. Chem. Soc.* **90**, 462–470 (1968).
42. Oliva, B., Maiese, W.M., Greenstein, M., Borders, D.B. & Chopra, I. Mode of action of the cyclic depsipeptide antibiotic LL-AO341β1 and partial characterization of a *Staphylococcus aureus* mutant resistant to the antibiotic. *J. Antimicrob. Chemother.* **32**, 817–830 (1993).
43. Fu, C. *et al.* Biosynthetic studies of telomycin reveal new lipopeptides with enhanced activity. *J. Am. Chem. Soc.* **137**, 7692–7705 (2015).
44. Tsai, M. *et al. Staphylococcus aureus* requires cardiolipin for survival under conditions of high salinity. *BMC Microbiol.* **11**, 13 (2011).
45. Machaidze, G., Ziegler, A. & Seelig, J. Specific binding of Ro 09-0198 (cinnamycin) to phosphatidylethanolamine: a thermodynamic analysis. *Biochemistry* **41**, 1965–1971 (2002).
46. Kawai, F. *et al.* Cardiolipin domains in *Bacillus subtilis* Marburg membranes. *J. Bacteriol.* **186**, 1475–1483 (2004).
47. Mileykovskaya, E. & Dowhan, W. Cardiolipin membrane domains in prokaryotes and eukaryotes. *Biochim. Biophys. Acta* **1788**, 2084–2091 (2009).
48. Oliver, P.M. *et al.* Localization of anionic phospholipids in *Escherichia coli* cells. *J. Bacteriol.* **196**, 3386–3398 (2014).
49. Arias-Cartin, R., Grimaldi, S., Arnoux, P., Guigliarelli, B. & Magalon, A. Cardiolipin binding in bacterial respiratory complexes: structural and functional implications. *Biochim. Biophys. Acta* **1817**, 1937–1949 (2012).
50. Stolpnik, V.G., Solovena, Y.V. & Antonenko, L.I. [Concentration of neothelomycin in the blood of rabbits following intramuscular or oral administration]. *Antibiotiki* **11**, 567–568 (1966).

Acknowledgments

This work was funded through a Natural Sciences and Engineering Research Council (NSERC) of Canada Discovery grant (RGPIN 371576-2014) (N.A.M.) and a Joint Programme Initiative on Antimicrobial Resistance funded through the Canadian Institutes of Health Research (CIHR; N.A.M.). C.W.J. is funded through a CIHR Doctoral Research Award. E.D.B and N.A.M. are supported by the Canada Research Chairs Program. The authors acknowledge and thank R. Epand for valuable communications.

Author contributions

C.W.J. performed telomycin studies, assisted in antibacterial library curation and development of scoring strategies, performed Antibioticome analysis, contributed to study design and wrote the manuscript. M.A.S. developed PRISM, developed the Antibioticome web application, assisted in resistance gene collection and contributed to study design. C.A.D. developed the retrobiosynthetic algorithm, devised scoring strategies for the Antibioticome, developed the Antibioticome web application and contributed to study design. P.N.R. curated the antibacterial library. G.M.C. developed the retrobiosynthetic algorithm and devised scoring strategies for the Antibioticome. C.W. collected resistance genes. S.F. performed microscopy studies. E.D.B. edited the manuscript. J.B. curated the antibacterial library. D.Y.L. assisted in resistance gene collection and curation of the antibacterial library. N.A.M. contributed to study design and wrote the manuscript.

Competing financial interests

The authors declare no competing financial interests.

Additional information

Any supplementary information, chemical compound information and source data are available in the online version of the paper. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Correspondence and requests for materials should be addressed to N.A.M.

ONLINE METHODS

Cataloging the natural antibiotic collective. We compiled a comprehensive database of all known antibiotics isolated from microbial sources. The *Handbook of Antibiotic Compounds*[28] was supplemented by an exhaustive review of published antibiotic literature and patents. For each antibiotic, chemical structures were generated in SMILES format, and both the known or suspected mechanism of action of the compound and relevant toxicity data were also recorded. In total, 10,343 microbially produced antibiotics were identified. Of these, 3,159 were selective to bacteria, meaning they had no broad cytotoxicity or activity against eukaryotic cells or organisms. A number of computational and manual curation processes were performed in order to ensure the high quality data. When chemical structures of the same molecule within two different data sources differed, the molecule was manually inspected and its SMILES redrawn. Publicly available SMILES were observed to frequently represent tautomers of their natural products, so an algorithm was implemented to redraw enol and iminol tautomers of ester and amide bonds, respectively. Finally, SMILES representing salts of their corresponding natural products were redrawn in their uncharged states.

Development of a retrobiosynthetic similarity scoring algorithm for natural products. In order to organize antibacterial natural products based on the biosynthetic history of these evolved small molecules, we developed a software package consisting of two algorithms. The first algorithm, termed GRAPE, performs *in silico* retrobiosynthesis of peptide, polyketide, and carbohydrate-containing compounds. GRAPE first cleaves defined biosynthetic bonds (chemical bridges [disulfide bonds, aromatic ethers, etc.], heterocycles, core bonds [esters, amides, lactones, etc.], and tailoring modifications [glycosylation, methylation, sulfation, etc.]), and then attempts to match remaining monomers with a comprehensive library of known biosynthetic units (amino acids, carbohydrates, etc.). Remaining units are parsed through a polyketide retrobiosynthetic module which scans the carbon backbone and assesses likelihood of polyketide origin, and predicts the architecture of the parent polyketide synthase (PKS). This collection of ordered monomers for each natural product deconstructed by GRAPE is then passed to a second algorithm for pairwise analysis based on order and composition. The second algorithm, termed GARLIC, aligns units of biosynthetic information identified by the retrobiosynthetic algorithm (i.e., proteinogenic and non-proteinogenic amino acids, ketide units, sugars, halogens, nonribosomal peptide and polyketide starter moieties, and selected tailoring modifications) to all molecules within the database of targeted antibiotics using a modified Needleman-Wunsch algorithm[51]. GRAPE and GARLIC implement chemical abstractions developed by the Chemistry Development Kit[52], version 1.4.19. A more comprehensive description of the GRAPE and GARLIC algorithms can be found in C.A.D., G.M.C., H. Li, C.W.J., P.N.R., M.A.S. *et al.* (unpublished observations).

Development of PRISM. We recently described the development of PRISM[34] (PRediction Informatic for Secondary Metabolomes), a web application designed to identify biosynthetic gene clusters and predict the structures of genetically encoded nonribosomal peptides and type I and II polyketides. PRISM implements a library of nearly 500 hidden Markov models to identify conserved biosynthetic genes and the substrates of adenylation, acyltransferase, and acyl-adenylating enzymes. A simple greedy algorithm is used to identify plausible biosynthetic gene clusters, and rules specific to modular, *trans*-acyltransferase, and iterative type I polyketides, type II polyketides, and nonribosomal peptides are used to define true clusters. A library of 57 virtual tailoring reactions is leveraged in order to generate a combinatorial library of chemical structures when multiple potential substrates are biosynthetically plausible for one or more tailoring enzymes, including macrocyclization, heterocyclization, aromatization, halogenation, C- and O-glycosylation, O-, N-, and C-methylation, carbamoylation, amination, formylation, phosphorylation, sulfonation, oxidation and reduction, mono- and dioxygenation, Baeyer-Villager rearrangement, and acyl group transfer. A set of hidden Markov models for conserved deoxysugar biosynthesis genes and a BLAST database of natural product glycosyltransferases is used to predict potential combinations of hexose and deoxysugars which tailor the natural product scaffold, and a library of models for type II polyketide cyclases is used to predict the scaffolds of type II polyketides. In addition to generating combinatorial libraries of chemical structures, biosynthetic information detected by PRISM can be aligned to biosynthetic information from other clusters, or to retrobiosynthetic information generated by GRAPE, within the GARLIC alignment package. The PRISM web application can be accessed at http://magarveylab.ca/prism.

Generation of the antibacterial tree. A subset of our antibacterial small molecule collection, consisting of 1,908 targeted antibiotics, was retrobiosynthetically decomposed using GRAPE, and the decomposed natural products were aligned to one another using GARLIC. The similarity matrices generated by GARLIC alignments of biosynthetic information were converted to dissimilarity matrices. The generated dissimilarity matrices were used to hierarchically cluster the antibiotics using the hclust function within the stats R package. A tree was generated with the ape R package[53]. The tree was visualized and clades were collapsed in Dendroscope based on observed families of natural products[54].

Development of a database of hidden Markov models for antibiotic resistance genes. In order to identify resistance determinants within biosynthetic gene clusters, we compiled a comprehensive database of 257 profile hidden Markov models for genes associated with antimicrobial resistance. 166 hidden Markov models were obtained from the Resfams antibiotic resistance profile hidden Markov model database[35]. A single model (PFAM12847, a generic methyltransferase model) was removed as it was observed not to be specific to antibiotic resistance. The Resfams database was supplemented by the development of an additional 91 profile hidden Markov models associated with antibiotic resistance (**Supplementary Data Set 2**). Sequences were manually collected based on homology to experimentally annotated sequences, aligned using MUSCLE[55], and trimmed using trimAl[56] to remove gaps present in fewer than 50% of sequences. Hidden Markov models were generated from the resulting alignments using the hmmbuild program, version 3.1b1, from the HMMER3 software package[57]. Bitscore cutoffs for each hidden Markov model were determined by manual analysis of the results of a search of the UniProtKB database[58], using the HMMER web server[59]. The resulting library of hidden Markov models created in this study, and the aligned sequences used to construct them, are available at http://magarveylab.ca/resistance/.

Development of a web application to search the antibacterial chemical space. In order to make our results accessible to the broader scientific community, we developed a web application capable of querying our comprehensive antibacterial collection in order to predict the molecular target of real or predicted molecular structures. The web application accepts as input a single molecular structure, in SMILES format, or a line- or tab-delimited file of SMILES. User-submitted molecular structures undergo *in silico* retrobiosynthesis, and units of biosynthetic information are aligned to all molecules within a database of targeted antibiotics. The database implemented within the Antibioticome search web application is an annotated subset of the antibacterial library consisting of 1,868 compounds. Each compound is grouped into a larger family of natural products associated with a molecular target. For each user-submitted molecule, the web application reports the single highest-scoring targeted antibiotic as determined by retrobiosynthetic alignment, in addition to the natural product family to which that antibiotic belongs, its putative molecular target, and a score indicating the quality of the alignment.

A complete guide to the Antibioticome search web application is available at http://magarveylab.ca/antibioticome/#!/help. The Antibioticome search web application is written in Java 7 for the Tomcat 7 web server. To ensure that all results will remain confidential, discrete user logins will be made available in the near future.

General chemical procedures. 1D (^{1}H and ^{13}C) and 2D (^{1}H-^{13}C HMBC, HSQC, ^{1}H-^{1}H NOESY, TOCSY, and COSY) NMR spectra for telomycins were recorded on a Bruker AVIII 700 MHz NMR spectrometer in d_6-DMSO (Sigma-Aldrich). High-resolution MS spectra were collected on a Thermo LTQ OrbiTrap XL mass spectrometer (ThermoFisher Scientific) with an electrospray ionization source (ESI) and using CID with helium for fragmentation. LC-MS data was collected using a Bruker AmazonX ion trap mass spectrometer coupled with a Dionex UltiMate 3000 HPLC system, using a Luna C18 column (150 mm × 4.6 mm, Phenomenex) for analytical separations, running acetonitrile with 0.1% formic acid and ddH_2O with 0.1% formic acid as the mobile phase.

Microbial strains and telomycin production. *Streptomyces canus* was purchased from the American Type Culture Collection (ATCC, ATCC no. 12647).

doi:10.1038/nchembio.2018

S. canus was maintained on Bennett's agar at 28 °C. *Staphylococcus aureus* Newman was maintained on cation adjusted Mueller Hinton broth (CAMHB) agar at 37 °C. *Bacillus subtilis* 168 was maintained on CAMHB agar at 28 °C.

Fresh colonies of *S. canus* were used to inoculate 50 mL cultures of GYM media containing 0.5% glycine (GGYM), and then grown for 72 h at 28 °C and 250 r.p.m. 10 mL of starter culture was used to inoculate 500 mL of the same media, followed by growth for 72 h at 28 °C and 250 r.p.m. Cultures were harvested by centrifugation, followed by a methanol extraction of the pellet and extraction of the supernatant with 2% absorbent HP-20 resin (Diaion). Resins were eluted with excess methanol, and the eluent was pooled with the pellet methanol extract and evaporated to dryness. The extract was resuspended in a small volume of methanol and separated on an open gravity column of LH-20 size exclusion resin (Sephadex) with methanol as a mobile phase. Fractions containing telomycin were pooled, evaporated to dryness, and resuspended in methanol. Telomycin was isolated by preparative scale LC-MS using a Luna 5 µm C_{18} column (250 × 15 mm, Phenomenex) with water (0.1% formic acid) and acetonitrile (0.1% formic acid) as the mobile phase, at a flow rate of 10 mL/min. After 4 min, acetonitrile was increased in a linear manner (curve 5) from 5% to 30% at 14 min, then increased 31% by 20 min, then to 40% by 40 min, followed by a wash of 100% methanol. Telomycins eluted at the following retention times: A - 26 min, B - 28 min, C - 31.5 min, D - 32.5 min, E - 34 min, F - 35 min, G - 38 min.

Directed biosynthesis of new telomycins. Fresh colonies of *S. canus* were used to inoculate 50 mL cultures of GYM media containing 0.5% glycine (GGYM), and then grown at 28 °C and 250 r.p.m. After 24 h, non-natural amino acids were added by sterile syringe filtration, to a final concentration of 4 mM. Cultures were grown for an additional 48 h and harvested, following the same isolation and purification procedure of natural telomycins. Unnatural amino acids were purchased from Sigma-Aldrich. Individual analog retention times were recorded and candidate peaks selected for purification and testing. Chemical structures were assigned using NMR and high-resolution mass spectrometry (see **Supplementary Note** on structural characterization).

Determination of antibacterial activity. Minimum inhibitory concentrations (MICs) for telomycins were determined using broth microdilution in cation-adjusted Mueller Hinton broth (CAMHB). *B. subtilis* strains were cultured at 28 °C, and *S. aureus* strains were cultured at 37 °C. The MIC was determined as the lowest concentration of drug at which no growth was observed after 16 h. For assessing the impact of treating strains with sodium chloride concentrations before telomycin, *S. aureus* was grown overnight in CAMHB containing 0, 0.5, 1, or 1.5 M NaCl and then MICs were recorded as described above. Broth microdilution assays were performed in fresh media with identical NaCl concentrations and again MIC was determined as the lowest concentration of drug at which no growth was observed after 16 h.

Measuring turbidity of telomycin-lipid mixtures. To measure the turbidity of telomycin-lipid mixtures, telomycin was dissolved in methanol at a concentration of 12.8 mg/mL, and cardiolipin (Sigma-Aldrich; C0563; ≥98% pure), phosphatidylglycerol (Sigma-Aldrich; 63371; ≥98% pure), phosphatidylcholine (Sigma-Aldrich; P3556; ≥99% pure), and phosphatidylethanolamine (Sigma-Aldrich; P7943; ≥97% pure) were dissolved in methanol at a concentration of 10 mg/mL. 20 µL of telomycin was mixed with 20, 10, 5, 1, or 0.5 µL of each lipid (final reaction volume 40 µL) in a flat-bottom polystyrene 96-well plate. Optical density at 600 nm (OD_{600}) was measured after 10 min.

Colony-forming unit (CFU) assays. Cultures of *S. aureus* and *B. subtilis* were grown overnight in CAMHB. Starter cultures were used to inoculate fresh media and were grown to OD_{600} = 0.2, after which they were transferred to 96-well plates. Cultures were serially diluted and plated on CAMHB agar to determine initial colony-forming units. Telomycin dissolved in DMSO was added to each well (1:100) and incubated for 90 min, followed by serial dilution and plating to determine colony-forming units. Results are shown as ± s.d.; $n = 4$; two-tailed Student's *t*-test.

Red blood cell (RBC) hemolysis assay. Hemolytic activity of telomycin was measured against a 0.25% sheep red blood cell (RBC, Fisher Scientific) suspension in phosphate-buffered saline. Telomycin was serially diluted from 256 to 2 µg/mL and incubated with the RBC suspension for 1 h at 37 °C in a polypropylene 96-well plate with conical wells. After 1 h, RBCs were pelleted (1,000*g* for 5 min) and the supernatant was transferred to a flat-bottom polystyrene 96-well plate, measuring absorbance at 540 nm. 1% Triton X-100 was used as a positive control, while DMSO alone and RBCs alone were used as negative controls.

Measuring bioactivity of telomycin-lipid mixtures. To assess the impact of various lipids on telomycin's antibacterial activity on solid agar, 10 µL of 10 mM telomycin was mixed with 10 µL of 20 mM cardiolipin (Sigma-Aldrich; C0563; ≥98% pure), phosphatidylglycerol (Sigma-Aldrich; 63371; ≥98% pure), phosphatidylcholine (Sigma-Aldrich; P3556; ≥99% pure), or phosphatidylethanolamine (Sigma-Aldrich; P7943; ≥97% pure). All compounds were dissolved in methanol. Mixtures were incubated for 10 min, then added to diffusion disks, allowed to dry, then placed on a CAMHB agar plate with *B. subtilis* as an indicator organism. The plate was incubated overnight at 28 °C.

To assess the impact of various lipids on telomycin's antibacterial activity in liquid media, telomycin was dissolved in methanol at 12.8 mg/mL (approx. 10 mM) and 1 µL was added to a 96-well plate. Lipids were dissolved in methanol and 10 µL of 10, 5, 2, 1, or 0.5 mM lipid was added to each telomycin-containing well. CAMHB inoculated with *B. subtilis* was added to each well and serially diluted to determine the minimum inhibitory concentration of each telomycin-lipid mixture. Plates were incubated shaking at 28 °C for 16 h.

Preparation of *N*-fluorescein labeled telomycin. To generated a telomycin-fluorescein conjugate, 2.5 mg of telomycin A was dissolved in 200 µL DMSO and mixed with 10 mg 5(6)-carboxyfluorescein *N*-hydroxysuccinimide ester (Sigma-Aldrich) dissolved in 180 µL DMSO, and 20 µL 0.5 M sodium bicarbonate. The reaction was allowed to proceed overnight at room temperature, after which the reaction was quenched by the addition of formic acid and N-labeled telomycin was purified by preparative scale LC-MS.

Measuring cardiolipin content of bacterial cells. Cardiolipin was extracted using an acidic Bligh Dyer method. A 50 mL culture of bacteria grown for 24 h was pelleted and resuspended in 1 mL 0.1 N HCl. 2.5 mL of methanol and 1.25 mL of chloroform was added to each sample, followed by 30 min incubation at room temperature. After this, 1.25 mL of 0.1 N HCl and 1.25 mL of chloroform was added to create a two-phase solution that was then centrifuged at 3,000*g* for 10 min. The bottom phase was recovered, evaporated to dryness, and then resuspended in methanol. An established LC-MS method was used to quantify cardiolipin content, using a reverse-phase Luna C18 column (150 mm × 4.6 mm, Phenomenex). The mobile phases were A (90% acetonitrile, 10% water, 0.5% glacial acetic acid, 0.5% triethylamine) and B (90% isopropanol, 10% water, 0.5% glacial acetic acid, 0.5% triethylamine), running at 0.8 mL/min. After 3 min, solvent B was increased from 50% to 100% by 22 min, then held for 15 min, before being returned to 50% by 38 min. Cardiolipin species eluted between 27 and 32 min. Relative cardiolipin content was assessed by summing the areas associated with cardiolipin species ions in simultaneously and identically extracted wild type and mutant bacteria, presenting percentage of cardiolipin signal detected in the mutant strain, relative to wild type.

Genome sequencing and analysis of antibiotic biosynthetic gene clusters. A single colony of *S. canus* (ATCC 12647) was used to inoculate a 50 mL culture of GYM media containing 0.5% glycine (GGYM), and then grown for 96 h at 30 °C and 250 r.p.m. 500 µL of culture was centrifuged at 12*g* for 5 min and resuspended in 500 µL SET buffer (75 mM NaCl, 25 mM EDTA pH 8.0, 20 mM Tris HCl pH 7.5, 2 mg/mL lysozyme) to lyse for 2 h at 37 °C. Proteinase K and SDS were added after lysis to final concentrations of 0.5 mg/mL and 1%, respectively. Lysis mixtures were incubated at 55 °C for 2 h before adjusting the concentration of NaCl to 1.25 M and extracting twice with phenol-chloroform. Isopropanol was added (equivalent to 60% the volume of the solution) to precipitate genomic DNA, which was subsequently washed twice with 70% ethanol and resuspended in sterile water for sequencing. For sequencing *S. aureus* and *B. subtilis* strains, single colonies were used to inoculate 3 mL overnight cultures in tryptic soy broth (TSB), incubated at 37 °C and 30 °C respectively. Genomic DNA was isolated using a GenElute Genomic DNA Extraction kit (Sigma-Aldrich). Genomic DNA for all strains was sent for library preparation and Illumina sequencing at the Farncombe Metagenomics Facility at McMaster University, using an Illumina HiSeq DNA sequencer. Contigs were assembled using the ABySS genome assembly program and with Geneious bioinformatic software.

Identification of telomycin-resistance mutations in sequenced isolate genomes. To identify mutations that conferred resistance to telomycin, we

isolated colonies of *S. aureus* Newman that appeared inside the zone of inhibition observed during telomycin disk diffusion assays. Colonies were cultured overnight in TSB containing 50 μg/mL telomycin (~5× MIC), then used to inoculate a 96-well plates containing a serial dilution of telomycin, up to 128 μg/mL. Genomic DNA was extracted from highly resistant isolates and their sensitive parental strain, and sequenced with an Illumina MiSeq platform. Sequences were compared using BreSeq, which exclusively identified mutations in the predominant, house-keeping cardiolipin synthase gene (*cls2*). *Bacillus subtilis* telomycin-resistant mutants were identified in a similar manner, and mutations in cardiolipin synthase were confirmed by sequencing PCR products of the two cardiolipin synthase genes: *clsA* and *clsB*. A point mutation was observed next to the second HKD motif active site in *clsA*, which resulted in a 90% decrease in cardiolipin levels. PCR primers used for the amplification of *B. subtilis* cardiolipin synthase genes are as follows – clsAF: 5′-GTTTTAAAGAAATCTGCCCG-3′; clsAR: 5′-GCGAGACGGATTCTTTTATT-3′; clsBF: 5′-ATGAAGGTATTTAT CGTGAT-3′; clsBR: 5′-TTATAAGAAATAAGATAATG-3′.

Structure elucidation. The structure of telomycin A was confirmed by a series of 1D and 2D NMR spectroscopy experiments, high resolution mass measurement, and MS/MS fragmentation and annotation. Structures of new naturally occurring variants were elucidated by MS/MS fragmentation and annotation, high resolution mass measurements, and comparison of 1D and 2D NMR experiments to those of telomycin A. Structures of telomycins generated by directed biosynthesis were confirmed by MS/MS fragmentation and annotation, and high resolution mass measurements. The structure of di-5-methyltryptophan telomycin was also confirmed by 1D and 2D NMR experiments.

Figures including structures, MS/MS fragmentation, and NMR spectra, as well as tables including high resolution mass measurements and NMR chemical shifts, can be found in the **Supplementary Note**.

Cytotoxicity assay. HEK293 cells were obtained from the American Type Culture Collection (ATCC; ATCC CRL-1573) and maintained in Minimal Essential Media (MEM) Alpha modifications supplemented with 10% FBS. Cells were cultured in 96-well plates containing 200 μL of media and 5,000 cells per well. After 3 h of incubation, cells were treated with a serial dilution of telomycin A. Experiments were performed in duplicate with DMSO as a negative control and mitomycin C as a positive control. After 48 h incubation, 10% (22 μL) Alamar Blue (Life Technologies) was added and incubated with cells for 4 h at 37 °C. Fluorescence was measured at an excitation wavelength of 530 nm and an emission wavelength of 590 nm and compared to a no-drug control.

Fluorescence microscopy. A stationary phase culture of *B. subtilis* 168 was washed twice in LB medium to remove extracellular debris, centrifuging at 10,000 r.p.m. for 1 min each wash. Both *N*-carboxyfluorescein telomycin and 10-*N*-nonyl acridine orange (NAO) were added to cells to compare probe localization. *N*-carboxyfluorescein telomycin was added to a final concentration of about 1 μM, and NAO was added at a final concentration of 0.5 μM. Both probes were incorporated for 20 min in the dark. The suspensions were washed again to remove extracellular probe, then added to a poly-L-lysine–coated 0.17 μm glass bottom microplate (Brooks Automation). To these samples, the membrane dye FM 4-64 was added to a final concentration of 1 μg/mL. Cells were imaged using a Nikon Eclipse Ti inverted microscope at 1,000× magnification with a Nikon Plan Apo λ 100× oil-immersion objective. Overlays were prepared using the Nikon Elements software suite.

Profile plots were prepared in ImageJ[60] by cropping regions of interest, and converting to 8-bit greyscale images. No background subtractions were made, and the gray values across the length of a cell were plotted for FM 4-64, *N*-carboxyfluorescein telomycin, and overlay, to highlight probe localization.

Code availability. Source code for PRISM, our retrobiosynthetic algorithm, and the Antibioticome web application will be made available upon request.

Statistical analysis. Significance of CFU assays were calculated with two-tailed Student's *t*-tests, using Excel 2013 (Microsoft). All CFU experiments were performed at least twice in the laboratory, with four replicates ($n = 4$) for each condition.

URLs. The Antibioticome web application can be reached at http://magarveylab.ca/antibioticome/. The PRISM web application can be reached at http://magarveylab.ca/prism/.

51. Needleman, S.B. & Wunsch, C.D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48**, 443–453 (1970).
52. Steinbeck, C. *et al.* The Chemistry Development Kit (CDK): an open-source Java library for chemo- and bioinformatics. *J. Chem. Inf. Comput. Sci.* **43**, 493–500 (2003).
53. Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290 (2004).
54. Huson, D.H. & Scornavacca, C. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst. Biol.* **61**, 1061–1067 (2012).
55. Edgar, R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
56. Capella-Gutiérrez, S., Silla-Martínez, J.M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
57. Eddy, S.R. Accelerated profile HMM searches. *PLoS Comput. Biol.* **7**, e1002195 (2011).
58. Magrane, M. & Consortium, U. UniProt Knowledgebase: a hub of integrated protein data. *Database (Oxford)* **2011**, bar009 (2011).
59. Finn, R.D., Clements, J. & Eddy, S.R. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* **39**, W29–W37 (2011).
60. Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).

doi:10.1038/nchembio.2018