

Title: Technological Developments Driving Industrial Natural Product Discovery

Authors: Richard Lewis ^a, Richard Hammond ^b, Mark Wilkinson^c & Nick Allenby^d

Affiliation:

^aAxitan Ltd, 1 Arlington Court, Whittle Way, Stevenage, SG1 2FS, UK.

^bSphere Bio Ltd., Building One, Granta Centre, Granta Park, Great Abington, Cambridgeshire CB21 6AL

^cBactobio, Vox Studios, 1-45 Durham St, London, SE11 5JH

^dFaculty of Health Sciences and Wellbeing, University of Sunderland, SR1 3SD

Abstract

Bacterial natural products have long been the foundation for many therapeutic agents. However, traditional culture-based approaches to discovering these products have been deprioritised by pharmaceutical companies, primarily due to the high rates of rediscovery. To revitalise the pipeline of new drugs, especially antibiotics - an area where natural products have historically played a crucial role - new technologies are essential. Culture-independent, or metagenomic, techniques combined with long-read sequencing technologies are now enabling the identification of novel biosynthetic gene clusters (BGCs). When paired with the heterologous expression of DNA extracted directly from environmental samples (eDNA), these approaches may provide access to untapped microbial biosynthetic diversity.

This review explores industrial screening for new compounds and examines how modern technologies such as metagenomics, *in situ* cultivation, and pico-droplet-based screening are advancing the search for novel natural products. These approaches have the potential to greatly expand the discovery of new bioactive compounds, helping to address the growing need for new therapeutic agents.

1. Introduction

1.1 Industrial Natural Product Discovery Endeavours

2. Metagenomics

2.1 Sequencing Technologies

2.2 Functional Metagenome assisted natural product discovery

2.3 Single Read BGC identification

2.4 Compound Discovery

3. New approaches to compound identification

4. The growth of unculturable bacteria

4.1 In situ cultivation

4.2 Diffusion chambers & iChips

5. Droplet-based microfluidic cultivation and analysis

5.1 High-throughput screening

5.2 Microbial engineering

5.3 Direct editing approaches

5.4 Directed evolution approaches

5.5 Technology access

5.6 Digital microfluidic systems

5.7 Surfactants

5.8 Future Directions

6. Conclusion

7. Conflict of interest

8. Acknowledgements

9. References

1. Introduction

Microbes have long been exploited for their natural products and through extensive industrial efforts provide many of the compounds currently used in medicine and agriculture. These compounds exhibit diverse chemical structures and possess a wide range of biological activities that have significantly contributed to the development of numerous drug classes, such as antibacterial, anticancer, and anti-cholesterol agents. Approximately one-quarter of all FDA-approved drugs are either nature-derived, or nature-inspired compounds (1). Of these natural products, bacteria and fungi have provided numerous clinically used antibiotic scaffolds. Most of these were discovered between 1940 and the early 1960s, a period now referred to as the “Golden age” of antibiotic discovery. Actinobacteria, and especially the genus *Streptomyces*, have been an especially prolific source of compounds producing well-known antibiotics such as tetracycline, erythromycin, chloramphenicol, and streptomycin (Figure 1) (2, 3). Most of these compounds were isolated from microbes residing in soil or sediments using “traditional” fermentation and plate-based methods. These methods were for decades an excellent source of diverse and potent new molecules (4). This productive era was followed by a ~40-year innovation gap, during which no new antibiotic scaffolds were introduced. The frequent rediscovery of known compounds led to the belief that this discovery pipeline was exhausted. This contributed to a decline in interest from pharmaceutical companies, particularly as synthetic and combinatorial chemistry techniques emerged, promising faster and more targeted drug discovery approaches (5). Unfortunately, these approaches have largely failed and given the growing global issue of antimicrobial resistance and its serious implications for human health, the discovery of new natural products is essential to address this global health challenge (6).

In this review we aim to discuss several related sectors where technological developments could have a profound impact on the discovery of natural products and hopefully lead to the re-entry of industry back into the natural product sector.

1.1 Industrial Natural Product Discovery Endeavours

The discovery of antibiotics dates back to 1928 when Alexander Fleming first reported the serendipitous discovery of the antimicrobial potential of penicillin from *Penicillium notatum* against *Staphylococcus aureus* (7). The development of it as a “miracle drug” was however, pioneered by the pathologist Howard Florey and the biochemist Ernst Chain (8), who shared the 1945 Nobel Prize with Fleming. The search by Florey and Heatley in collaboration with the US Department of Agriculture for new sources of penicillin led to the discovery of *Penicillium chrysogenum*, on a mouldy cantaloupe melon, which produced 200 times more penicillin than any source tested previously. Shortly after, the methodology underpinning the “Golden Age” of antibiotic discovery was developed by Selman Waksman who reported the discovery of streptomycin from a soil bacterium (9).

Antibiotics were defined by Selman Waksman in 1947 as “*a chemical substance, produced by micro-organisms, which has the capacity to inhibit the growth of and even to destroy bacteria and other micro-organisms*” (10). Following their widespread use in medicine, a more modern definition evolved “*a medical chemical compound that at low doses is able to selectively kill or inhibit growth of targeted bacteria with little to no impact on the patient*”. Antibiotics are a cornerstone of modern medicine, demonstrated by the partitioning of the history of medicine into pre- and post-antibiotic eras. Streptomycin’s introduction in 1945, is a clear example of this: the mortality rates associated with tuberculosis decreased from 39.9 deaths per 100,000 population in 1945 to 9.1 deaths per 100,000 between 1945 to 1955 (11).

The “Waksman Platform” involves the plating out of soil isolates onto agar plates and the selection of bacterial colonies capable of producing antibiotics. The presence of antibiotics is inferred through visual identification of zones of inhibition/clearance zones affecting bacterial colonies surrounding producer strains. More advanced assay formats incorporate Ehrlich’s idea of identifying compounds capable of targeting specific pathogens, so that the producer strains are screened against a pathogen of interest, generally by spotting extract from the producer strain onto a growing lawn of a target microorganism.

The simple binary live/dead assay format readout means that the platform is empirical *i.e.*, target neutral, and is thus unlimited by antibiotic mode of action. Consequently, it has been used to identify many classes of compounds with different modes of action which may be determined after the compound has been identified and isolated. Although it is relatively low innovation, low-throughput, and “low tech” the “Waksman Platform” was widely adopted and subsequently formed the basis of the pharmaceutical industry’s antibiotic discovery programmes for the next 30 years. The factor limiting the approach was largely due to the lack of available shaker space used to culture liquid bacterial cultures e.g. Eli Lilly was limited to screening 35,000 cultures per year due to this bottleneck (4). Most of this research was undertaken by pharmaceutical companies meaning much of the data has never been published (4, 12). However, an informed estimate suggests 10-20 million compounds were screened (4). Eli Lilly is reported as having screened 1,000,000 samples over a 30 year period (5, 6) and by the mid 1990’s Abbott had screened ~400,000 (4). Over a period of 30 years, it is estimated 1,000-10,000 natural products that had antibacterial or antifungal activity were identified, including 50-100 compounds that were ultimately approved for therapeutic use (12). These include some of the most important classes of present-day antibiotics.

83
84 However, by the 1970's it was however clear that the pipeline was running dry and
85 already known compounds formed an increasingly large proportion of identified hits (12)
86 necessitating dereplication screens to filter out already identified antibiotics (13, 14).
87 Fewer than two dozen novel natural products, or derivatives thereof, were developed as
88 drugs from 1970-2000 and of these only one, daptomycin, which was discovered in
89 1987, (15) and introduced into clinical use 2007, represented a novel antibiotic
90 scaffold. Consequently, the industry switched to screening methods which were
91 essentially the reverse of the Waksman Platform where libraries of natural product
92 compounds, and synthetic compounds synthesized by combinatorial chemistry, were
93 tested for their ability to inhibit particular cellular functions, as proposed by Cohen
94 (16). This approach was facilitated by the availability of bacterial genome sequences
95 starting with *Haemophilis influenzae* in 1995, which allowed the identification of
96 essential genes conserved across bacterial species which were thought to be likely
97 targets for antibiotics.

98
99 The effort invested in this approach was enormous: it is estimated that between 1996
100 and 2004 34 companies carried out 125 screens on 60 different bacterial targets (17).
101 Between 1995-2001, GSK alone ran 67 high-throughput programmes against
102 antibacterial targets screening a library of between 260,000-530,000 compounds
103 generating only 16 hits that produced 5 leads none of which were clinically
104 developed (18). This lack of success was replicated elsewhere and despite these
105 strenuous efforts by Big Pharma no new antibiotics were approved for clinical use
106 (12). There are thought to be several reasons behind the failure of the combinatorial
107 chemistry/high-throughput screening approach which have been well reviewed
108 elsewhere (4, 12, 18), but briefly: firstly, the low chemical diversity of compound
109 libraries generated by combinatorial chemistry, many of which exhibit similar
110 structural/physical characteristics and conform to Lipinski's "Rule of Five" which
111 predicts the likelihood of a molecule being able to likely to work as orally
112 deliverable drug (18, 19). Natural product compounds are generally more diverse and
113 often do not follow to Lipinski's "Rule of Five" (18, 20). Secondly, re-evaluations of
114 the vital bacterial genetic targets indicated that these targets were frequently
115 redundant or were poorly conserved across species (18). Thirdly, when hits were
116 identified, they often turned out to be non-specific membrane-active agents e.g.
117 detergents or uncoupling agents, and lastly compounds that worked well *in vitro*
118 often turned out to be ineffective *in vivo*, due to membrane effects *i.e.*
119 either them being unable to penetrate the bacterial cell membrane or being pumped
120 out of the cytoplasm by efflux pumps (18, 21).

121
122 The failure of the combinatorial chemistry/high-throughput screening approach and
123 the diminishing returns along with the relatively limited profit margin of antibiotics
124 meant that many of the major pharmaceutical companies largely abandoned
125 antibiotics discovery pipelines altogether (22). While smaller biotechnology firms and
126 academic researchers have continued, the overall pace of antibiotic discovery has
127 slowed. The rise of antibiotic resistance and the reduced number of new antibiotics
128 entering the market have created a significant public health challenge, as new classes
129 of antibiotics are urgently needed to combat multidrug-resistant bacterial infections.
130 This has refocused antibiotic attention on application of the Waksman Platform to
131 hitherto untouched subject material and its combination with new technologies.
132 We now know that traditional culturing methods can only grow a fraction of microbial

species found in nature. Many microbes are unculturable due to their specific environmental requirements or symbiotic relationships with other organisms. Also, with the advent of whole genome sequencing it became evident that the genomes of many organisms and especially Actinobacteria are rich in biosynthetic gene clusters (BGCs) (23). A recent bioinformatic study of ~170,000 known bacteria genomes and ~45,000 metagenome assembled genomes (MAGs) identified a vast number of novel BGCs and estimated that only 3% of the natural products encoded in these genomes have been experimentally characterised, let alone investigated for medical potential (24). Therefore, the majority of BGC diversity remains unexplored, representing a significant untapped resource for drug discovery.

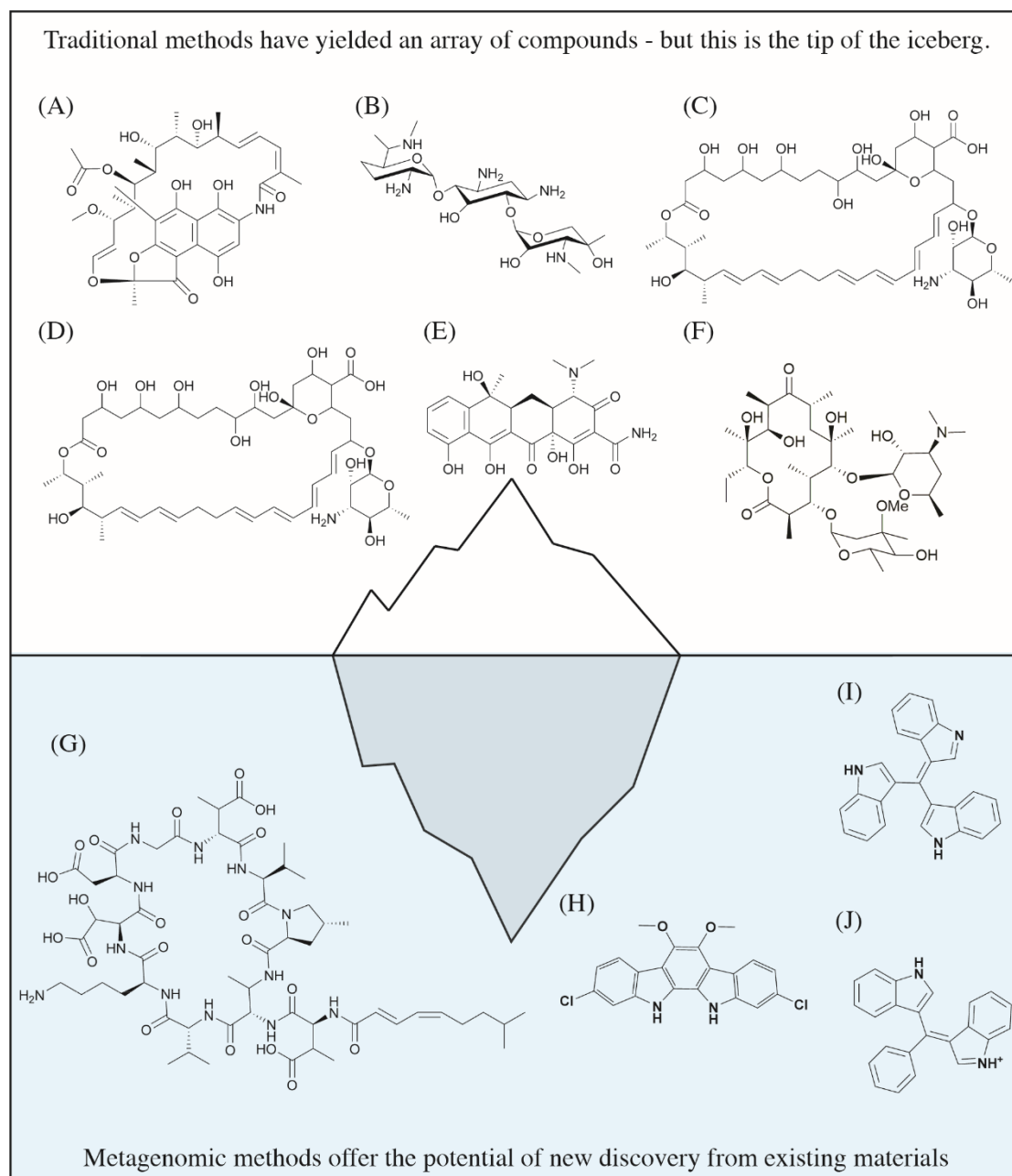
2. Metagenomics

Metagenomics is the study of genetic material recovered directly from environmental samples, such as soil, water, or human-associated microbiomes, and bypasses the necessity of isolating and culturing individual organisms. Microbial communities, particularly those in diverse and underexplored environments, are likely to harbour a wealth of bioactive molecules with therapeutic potential. This realisation has spurred significant interest from biotech companies, which recognise the potential of these microbial-derived compounds. This approach is particularly valuable because the vast majority of microbes in natural environments, estimated to be over 99%, are either difficult or impossible to culture in laboratory conditions (25-27).

Metagenome analysis has typically been performed by amplifying and sequencing the 16S ribosomal RNA (rRNA) gene, first utilised in the 1970s by Carl Woese to classify organisms based on evolutionary relationships (28). The 16S rRNA gene is present in all prokaryotes and is highly conserved across species, making it an ideal marker for identifying and classifying bacteria. In metagenomic studies, hypervariable regions of the 16S rRNA gene (~200 bp) are often amplified and then sequenced using accurate short-read sequencing technologies. This method has become a foundation of metagenomic phylogenetic analysis and has significantly advanced our understanding of microbial community structures and their functional roles across multiple microbiomes from the human microbiome to extreme habitats (29-31). A rough estimate of bacterial diversity using 16S rRNA gene sequencing suggests that there may be between 10,000 to 50,000 bacterial species in a gram of soil, with some estimates going even higher, depending on the source and complexity of the soil sample (32-34). A recent study of the ocean microbiome estimates that about one-quarter of the population of microbial cells on Earth belong to phyla with no cultured relatives (35). With only around 23,000 unique bacterial species currently listed in the NCBI database it is clear that there is vast and as yet untapped diversity of bacteria which can be utilised for biotechnology purposes and that most bacterial species remain uncharacterised and/or unculturable. While 16S rRNA metagenomic studies are crucial for large scale diversity estimation, they do not provide information regarding BGC diversity or prevalence and the short sequencing reads cannot provide identification to the species level.

Shotgun metagenomics overcomes many of these problems and unlike 16S rRNA sequencing, which focuses on a specific marker gene and uses PCR amplification, shotgun metagenomics sequences all the genetic material in a sample without amplification (35, 36). Although unbiased, the sequencing depth needs to be high and is therefore significantly more expensive than traditional approaches (37), and either long or short read sequencing technologies can be utilised. The main bioinformatic steps are read assembly and classification. Often the sequencing reads are assembled using tools such as metaSPADES (38) for short reads, or metaFlye for long reads (39). The resulting contigs can be classified by tools such as Kraken (40), Diamond (41) and visualised with MEGAN (42). When sequencing depth is sufficient this approach allows the reconstruction of entire microbial genomes called “metagenome assembled genomes” (MAGs) and the function of the coding genes can be predicted. Tools like antiSMASH and Prism (43, 44) enable the identification of BGCs within metagenomic datasets, and these can be compared to the MIBiG database (45) to enable dereplication, and gene networking tools such as BiG-SCAPE (46) and BiG-SLiCE (47) can be used to assess similarity and relationships to related BGCs. Bioinformatic analysis of metagenomic sequence data using these tools has indicated that there is a huge amount of BGC diversity which has not been discovered through traditional routes (24). It is clear that we do not have a good understanding of the role of these BGCs in the environment, for example a study looking at BGC diversity across prokaryotic genomes unexpectedly found that gene clusters responsible for producing aryl polyene carboxylic acids constitute the largest BGC family in the sequence databases (48), and exactly what biological function these compounds have is still yet to be determined.

A recent study exemplifies the power of a long read BGC-centred approach, particularly in challenging environments like Antarctica (49). Waschulin et al. (2020) conducted a long-read metagenomic study of Antarctic soil, an environment known for its microbial diversity, but also for being largely underexplored in terms of BGCs. By combining long-read Oxford Nanopore sequencing technology, with genome mining techniques, the researchers were able to identify over 1,400 BGCs from largely uncultured and underexplored phyla (49). This discovery was particularly impressive given the relatively low amount of sequencing depth involved, highlighting the efficiency of long-read metagenomic approaches.



216

217 **Figure 1.** Tip of an iceberg, example structures of compounds derived by traditional
 218 and metagenome methods; (A) rifamycin, (B) gentamicin, (C) nystatin, (D) FK-506,
 219 (E) tetracycline, (F) erythromycin, (G) malacidin, (H) borregomycins, (I) turbomycin
 220 A, (J) turbomycin B.

221

222 2.1 Sequencing Technologies

223

224 There have been vast improvements in the sequencing technologies involved for BGC
 225 discovery with improvements in accuracy and cost reduction. The choice between

short-read (mainly Illumina) and long-read (Oxford Nanopore/PacBio) sequencing technologies profoundly impacts the accuracy and completeness of metagenomic assemblies, and consequently, BGC detection. Illumina short-read sequencing is the dominant technology for metagenomics due to its high accuracy, low error rate, and cost-effectiveness. It generates short reads of 150–300 base pairs (bp) and is particularly well-suited for taxonomic profiling and functional annotation of microbial diversity in complex environments such as soil, marine, and human microbiomes (36). The massive output of data from Illumina platforms makes it ideal for large-scale surveys of environmental samples, providing a broad overview of the microbial community structure and gene content. Short-read sequencing, however, faces significant limitations when it comes to the detection and assembly of BGCs. BGCs are often complex operons which contain repetitive elements, that are difficult to resolve with the short reads. This often results in fragmented assemblies, which can lead to incomplete or incorrect annotation predictions (50), especially for large biosynthetic gene clusters like those encoding for polyketides or non-ribosomal peptides. This fragmentation significantly complicates efforts to reconstruct full BGCs from metagenomic datasets, which in turn limits the discovery of novel secondary metabolites.

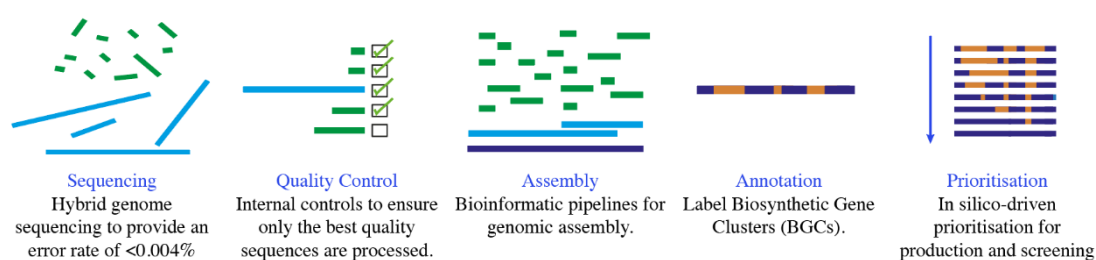


Figure 2: Hybrid next generation sequencing. Long-read technologies are used to for assembly and polished with short-read sequencing technologies for low error, high completeness genomes.

Long-read technologies are much more suited to BGC discovery, especially for smaller scale studies. PacBio’s single-molecule real-time (SMRT) sequencing generates reads that are typically 10–20 kilobases (kb) long, which can span entire BGCs in many cases. However, it could be argued that the largest improvement in long-read technology has been made by Oxford Nanopore Technology (ONT). Recent developments by ONT have dramatically decreased the error rates this has led to the improvement of read accuracy from around 96% in 2018 to >99% currently using the “Super-accurate” base-calling method.

To leverage the strengths of both short-read and long-read sequencing technologies, hybrid sequencing approaches have been adopted. By combining the high accuracy of short reads with the long-range continuity of long reads more accurate and complete metagenomic assemblies can be produced (Figure 2) (51, 52).

2.2 Single Read BGC identification

Oxford Nanopore technology allows the sequencing of ultra-long reads, and with the improvements in error rates, now supports the analysis of single reads for BGC identification. Typical metagenome assembly often eliminates reads which do not have a partner and this is especially important with smaller scale metagenome studies as the much of the data may be unique and this can lead to a substantial loss of information. However, this method requires careful eDNA isolation to produce high-molecular weight DNA to be sequenced. This approach was used by our lab with soil metagenome data which identified a single read of around 100Kb in length and within this read antiSMASH identified a BGC encoding a siderophore (Figure 3). This read was lost when performing standard assembly-based approaches. Although many of the genes are truncated indicating errors many of the smaller PFAM domains are correctly identified by antiSMASH. While this method has significant merit the computational requirements are considerable, and this type of analysis should be limited to only the largest reads.

1. Genomic or long-read sequences



2. Biosynthetic Gene Cluster (BGC) annotation



3. BGC prediction



4. BGC alignment and comparison



94% similarity

Practical example

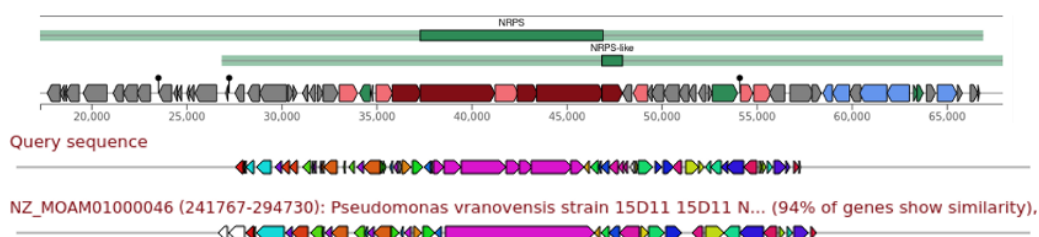


Figure 3. Biosynthetic Gene Cluster (BGC) analysis. Whole genomes or long-read sequences (1) are annotated for BGCs using bioinformatic tools (2). These predict BGC boundaries, which can then be compared to known BGCs (4). The practical example shows a single ONT long-read analysed by antiSMASH identifying a siderophore BGC.

2.3 Functional Metagenomic assisted natural product discovery

Functional metagenomics enables access to the vast genetic and biosynthetic diversity of uncultured microbes through the cloning and screening of eDNA (Figure 4). This approach bypasses the need for microbial cultivation, allowing the metabolic potential of microbial communities to be accessed directly from environmental samples.

However, the DNA encoding secondary metabolite biosynthesis is estimated to comprise less than 2% of many bacterial genomes (53), meaning that for natural product discovery randomly cloning BGCs from eDNA poses a significant challenge; cloning efficiencies need to be high to recover intact clusters. The low abundance of BGCs in most genomes necessitates the screening of large eDNA libraries to identify clones carrying complete pathways.

The first critical step in functional metagenomics is the extraction and purification of high-molecular-weight eDNA from environmental samples. Maintaining the integrity of eDNA is crucial for capturing entire BGCs (54). DNA extraction and separation, often by pulse-field electrophoresis is often needed to remove smaller contaminating DNA fragments. The high molecular weight DNA is then cloned into suitable vectors, often cosmid/fosmids, or bacterial artificial chromosomes (BACs) are used that can accommodate large DNA fragments and have low copy number which minimises genetic recombination and helps preserve BGCs integrity. Standard high efficiency *E. coli* cloning strains are used for the initial cloning and once verified the recombinant vectors are transferred into amenable heterologous hosts, such as *Streptomyces coelicolor*, where they often integrate into the genome for stability and facilitate the expression of cloned BGCs. Transferring eDNA-derived genes into an appropriate host organism allows for their expression and the production of natural products. The host can then be screened for desired phenotypes or enzymatic activities indicative of natural product biosynthesis. Using an expression host from the same genus is preferable but cross genera expression is often possible. Following this approach, many new compounds have been discovered and will be discussed below.

Heterologous expression can also be used for activating BGCs that remain silent or poorly expressed in their native hosts. Heterologous strains offer the advantages of being genetically amenable, faster growing and can often produce increased yields, particularly when expressing BGC which are silent in the natural host. Several species have been used as heterologous of BGC including the Streptomyces *Streptomyces coelicolor*, *lividans*, *albus* and *venzuelae*, the Gram-positive bacterium *Bacillus subtilis* and the Gram-negative bacterium *Pseudomonas putida* and the yeast *Saccharomyces cerevisiae*. Streptomyces hosts are most commonly used for the expression of Actinomycete BGC generally because they possess similar GC ratios, gene regulatory networks and transcriptional and translational machinery to the natural host and can also supply the necessary precursor molecules.

To mitigate interference from endogenous antibiotics it is crucial to use genome-minimised strains which often have several endogenous antibiotic gene clusters deleted. As background antibiotic activity from the heterologous host can interfere with the screening and identification of clones from ex-conjugants and a number are available for *S. coelicolor* (55, 56) *S. lividans* (57, 58), *S. albus* (59) and *S. venezuelae* (60, 61). Recently, *Pseudomonas putida* has become a popular choice of BGC expression host (62, 63) as it is fast growing, can utilize lignin derived compounds, it expresses a promiscuous phosphopantetheinyl transferase (PPTase) capable of activating heterologous acyl- and peptidyl-carrier proteins it is amenable to being cultured on an industrial scale, its genome has a relatively high GC content (61.5%) and a number of synthetic biology tools have been developed for it. *Bacillus subtilis* has also been used as a heterologous host, mostly to express BGC from *Bacilli* and *Firmicutes* (64). The preferred host for heterologous expression of eukaryotic fungal and plant BGC is the *Saccharomyces cerevisiae* (65, 66).

Cosmid libraries can be prepared relatively efficiently and quickly especially for small BGCs (67), however the size limitations of the bacteriophage head mean they have a DNA size limit of around 40 kb this means that for many larger BGCs these cannot be cloned in a single vector. As a result, multiple vectors with overlapping sequences need to be identified and combined, which is labour-intensive and technically challenging. Alternative cloning systems, such as transformation-associated recombination (TAR) cloning in yeast (68), or the use of larger capacity vectors like BACs (69), have been developed to overcome these limitations (68). A recent advance in this area is the CRAGE-CRISPR system, which combines Chassis-Independent Recombinase-Assisted Genome Engineering (CRAGE) with CRISPR-based activation (70). CRAGE facilitates the insertion of large BGCs into diverse bacterial hosts without the need for species-specific tools, while CRISPR activation enables precise transcriptional control of silent gene clusters. Using this approach, Ke et al. (2022) successfully activated multiple polyketide and non-ribosomal peptide BGCs in *Photorehabdus luminescens*. This system represents a significant advancement in BGC cloning and offers a method to rapidly activate and characterise BGCs (70).

2.4 Compound Discovery

An early example of untargeted metagenome natural product discovery was identification of the novel compounds turbomycin A and B, which were discovered by screening soil-derived metagenomic libraries using *Escherichia coli*. Turbomycins exhibit activity against both Gram-positive and Gram-negative bacteria, including *Staphylococcus aureus* (71). The identification of these triaryl cation compounds, demonstrated the feasibility of using functional metagenomics to identify novel antibiotics from uncultured soil bacteria (Figure 1. I & J).

Targeted homology-based screening of metagenomes can be used to identify and recover natural product biosynthetic gene sequence motifs by amplifying conserved genes through PCR from environmental DNA libraries. The borregomycins were discovered through the cloning, and heterologous expression of the indolotryptoline-

based *bor* biosynthetic gene cluster (72). (Figure 1. H).
Hover et al. (2018) identified malacidins by mining the metagenomes of around 2000 unique soils using degenerate PCR primers targeting the conserved regions of adenylation domains (AD) found in nonribosomal peptide synthetase genes. After sequencing the PCR products, a desert soil rich in novel regions was chosen to make a cosmid library. The malacidin BGC spanned 72 kb across three cosmids which were recombined using yeast TAR system and heterologously expressed using *Streptomyces albus* J1074 (73). Clone specific metabolites, malacidin A and B, were isolated and structural elucidation identified cyclic lipopeptide antibiotics which were effective against multidrug-resistant Gram-positive pathogens, including methicillin-resistant *Staphylococcus aureus* (MRSA) (Figure 1. G).

While culture-independent metagenomic approaches hold great promise for natural product discovery, particularly antibiotics, the number of novel compounds identified through cloning and expressing eDNA-derived biosynthetic gene clusters remains relatively small due to several challenges. Extracting high-molecular-weight DNA from soil is technically difficult due to fragmentation of DNA during extraction, and cloning large BGCs often requires the assembly of multiple overlapping clones (54). The presence of various contaminants such as humic acids, can interfere with downstream applications. Even when BGCs are fully assembled, their expression in heterologous hosts is frequently problematic. Successful expression requires compatible regulatory elements and metabolic compatibility between the BGC and the host organism, which is not always achievable. The native promoters and regulatory sequences of the BGCs may not be recognised by the host's transcriptional machinery, leading to non-existent or low expression levels and hampers the production of sufficient quantities of the natural product for structural characterisation.

Numerous efforts have been made to boost the production of cloned BGCs by genetic modification. One common approach is the insertion of strong, constitutive promoters to drive gene expression within the BGC, replacing native promoters that may not function efficiently in the host organism (74). For example, well-characterised promoters like *ermE** or *kasOp** in *Streptomyces* species have been used to enhance gene expression levels (75). Another strategy involves the removal or inactivation of repressor regulators that silence BGC expression, thereby de-repressing the pathway and enhancing metabolite production (76).

Recent developments have used CRISPR-Cas9 technology to activate silent BGCs by inserting strong promoters or removing repressors. For example, a CRISPR-Cas9 knock-in strategy successfully activated multiple BGCs in *Streptomyces* species, leading to the production of unique metabolites, including a novel aromatic pentangular type II polyketide in *Streptomyces viridochromogenes* (77).

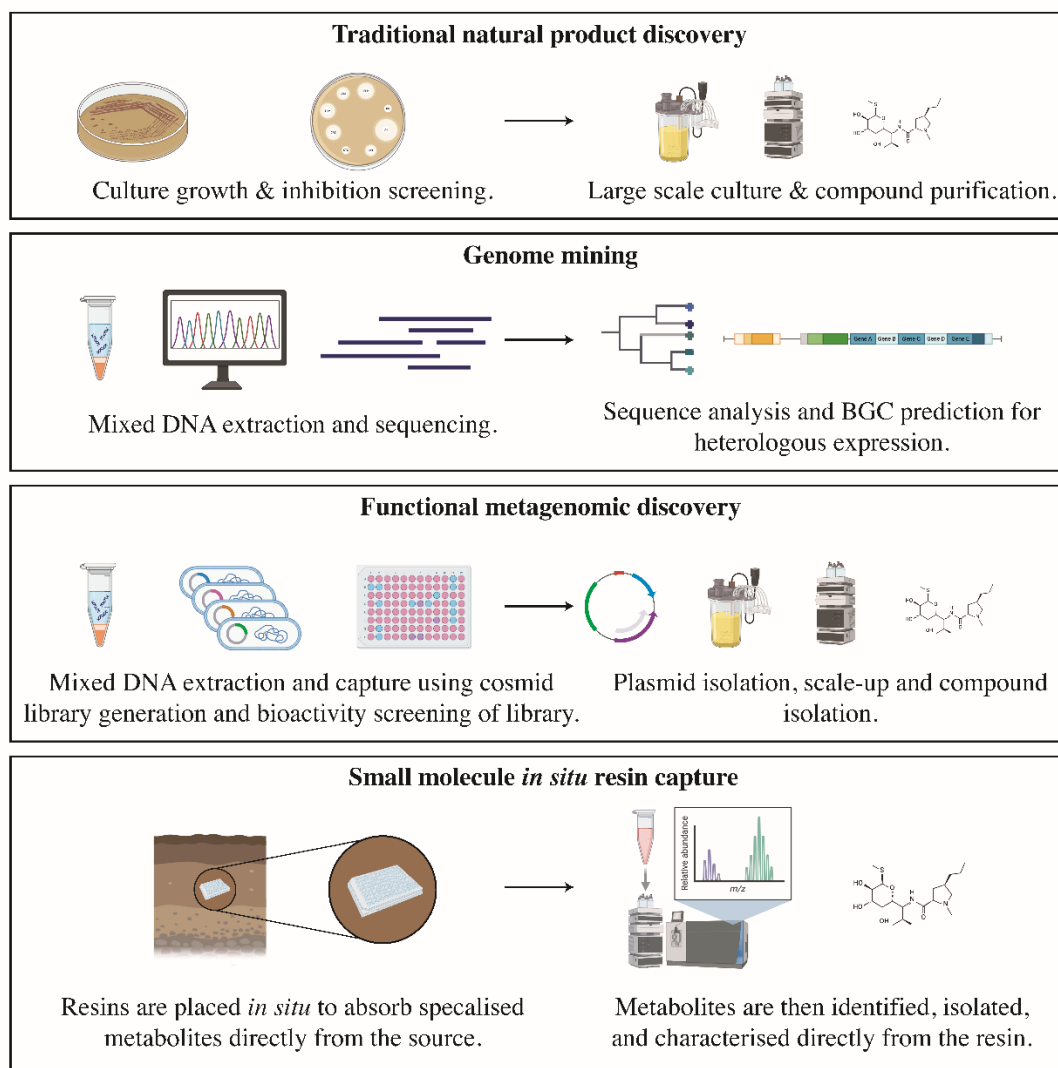


Figure 4. Traditional and new approaches to Natural Product discovery.

3. New approaches to compound identification

A new approach for the discovery of natural products is called Small Molecule *in situ* Resin Capture (SMIRC)(78). This method addresses the limitations of existing strategies by bypassing the need for microbial cultivation and directly capturing compounds from natural environments using an adsorbent resin such as HP20 (Figure 4). The authors describe deploying the resin in diverse ecological niches, where it adsorbs a wide array of naturally occurring chemical entities. These compounds are subsequently extracted from the resin and examined using traditional analytical techniques, such as mass spectrometry (MS) to identify novel compounds. After purification nuclear magnetic resonance (NMR) is used to elucidate their structures. This method has already proven effective in capturing novel compounds with unique structural features (78). The main advantage of SMIRC is its ability to uncover new compounds without the prerequisite knowledge of the microbial producers or their growth conditions. This is particularly significant in the context of soil and marine environments, where most of the microbial diversity remains unexplored. The

implications of SMIRC extend beyond natural product discovery. The technique offers a new method to view and understand the chemical interactions and biodiversity of various ecosystems. By analysing the metabolomic signatures captured from different environments insights may be gained into the ecological roles of natural products (78).

4. The growth of unculturable bacteria

In London's Science Museum displayed alongside the original *Penicillium* mould donated by Fleming to a colleague (Object:1997-731) is a peculiar object (Object: E2017.0416). This is instantly recognisable to anyone who has worked in a modern molecular biology laboratory as a disposable, grey, plastic Thermo Fisher pipette tip rack (pictured in Figure 4 of Berdy et al., 2017)(79). Why is a cut-down, throw-away laboratory consumable displayed alongside the museums scientific crown jewels? The caption to the exhibit explains why. It reads "*Finding new antibiotics- the iChip 2010: bacteria can produce chemicals that are useful as antibiotics, but they are hard to grow in the laboratory. Innovations such as the iChip allow bacteria to grow quickly in soil. Bacteria from a soil sample are placed into the tiny chambers of the iChip, which is then returned to the soil, so the cells can grow from nutrients in the ground. Without technologies such as the iChip, up to 90% of bacteria would be inaccessible for research*". This section aims to provide a background to this object, how it is used and why it is important.

It has long been known that the number of cells observed under a microscope in an environmental sample far exceeded the number of colonies that would grow on an agar plate (80). Later investigations established the presence of metabolically active cells in the environment that do not replicate on laboratory media, later supported by additional molecular evidence (81, 82). These observations gave rise to the concept of the "Great Plate Count Anomaly". The discrepancy between the numbers of observed and culturable bacteria depends on the environment but can reach several orders of magnitude (83). This suggested that the antibiotics discovered during the Golden Age came from mining a thin seam of culturable bacteria and that unculturable bacteria (84-86);— so called microbial "dark matter", represented a vast, untapped source of new antibiotics which could help to fill the "discovery void " (12). Although standard approaches continue to yield novel strains (87-90) the idea that the reason why most bacteria cannot be cultured in the laboratory is simply that essential aspects of their natural environments are not replicated closely enough was pursued, and two main approaches which attempt to address this have been the subject of much research (91). The first focuses on synthetically replicating the essential conditions required by the unculturable bacteria in the laboratory. For example, the growth of unculturable bacteria on standard media may be limited by lack of necessary siderophores and co-cultivation with species able to provide these, or supplementation with commercially available molecules, resulted in their growth (92). The second approach essentially seeks to move the laboratory to the environment, or *vice versa*, and so grow unculturable bacteria in their native conditions. One methodology which this approach encompasses is the "dilution to extinction" method (93) which has been particularly well applied to seawater (94). This approach was adopted by Rappé and co-workers who diluted seawater samples so that each well of a 96 well microtitre plates supplemented seawater received 22 cells and so were able to culture for the first time

the SAR11 clade, a major constituent of the bacterioplankton (95). The technique was extended to isolate four unique cell lineages that belong to previously uncultured or undescribed marine *Proteobacteria* clades (96, 97).

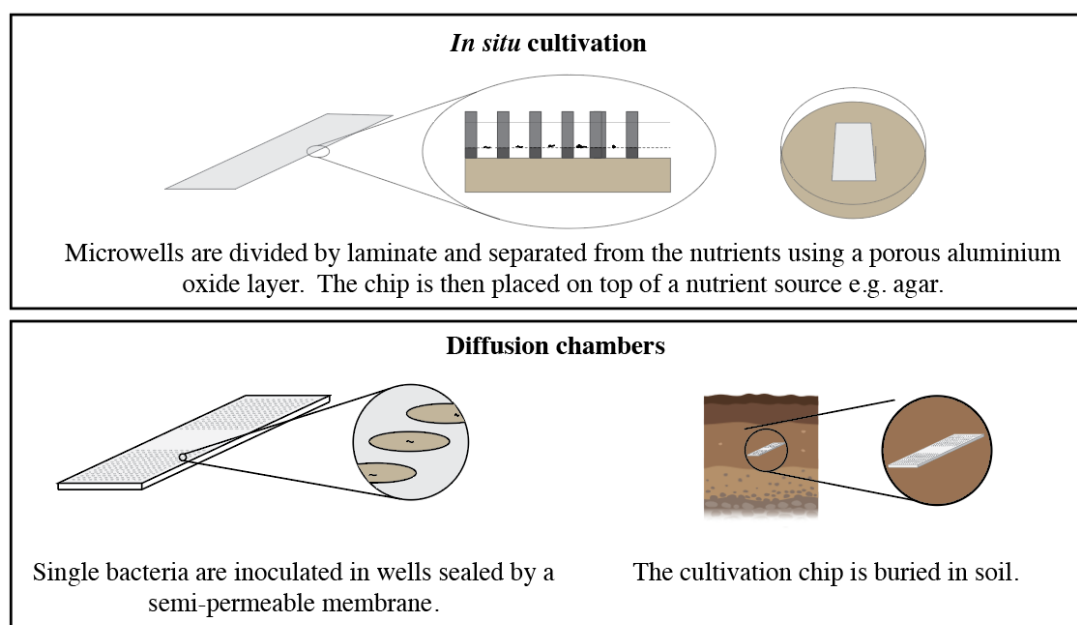


Figure 5: Two types of cultivation discussed below.

4.1 *In situ* cultivation

An alternative approach by Ferrari and co-workers used a solid support to immobilise and culture microcolonies of bacteria from soil slurry contained in a tissue culture insert, placed in contact with a polycarbonate membrane (98). A follow-up study indicated they could be picked and manipulated for further cultivation (99). Use of a “micro petri dish” was used by Ingham (2007) where a miniaturized, disposable microbial culture chip fabricated by microengineering a highly porous ceramic sheet was used. 180,000 growth compartments per 8 x 36mm chip, were constructed, filled with gellified river water as a culture medium, and used to culture and screen oligotrophic bacteria isolated from river water on the basis of organic phosphate metabolism (100). The results indicated 22 colonies, identified from a total of 207,000 microcolonies cultured, were capable of taking up fluorescein 3,6-diphosphate (FDP) and rapidly converting it to fluorescein. This sophisticated experimental methodology demonstrated that this was a high throughput system, (HTS) capable of identifying rare bacteria which occur in natural isolates at low frequency (100).

4.2 Diffusion chambers & iChips

A simpler, less technologically advanced approach where the bacteria were in contact with, but separated from, the “environment” was pioneered by the laboratories of Kim Lewis and Slava Epstein who designed a diffusion chamber where bacteria isolated from marine sediment were encapsulated in agar enclosed within a semi-permeable membrane and the sealed chamber was incubated in a tank of seawater, on a sand bed, to replicate the natural habitat (101). The device resembled an agar disc sandwiched between two membrane sheets and sealed with large steel washers (Figure 5).

Bacterial microcolonies were found to grow within the agar and comparison of their numbers with the number of bacteria present in the inoculum indicated 40% were able to grow, in the diffusion chamber, as opposed to 0.05% which grew on standard plates. Several microorganisms from the diffusion chamber microcolonies were grown in pure culture and found to be previously uncultured bacteria and interestingly, growth of several strains, MSC1 and MSC2, was found to be dependent on their co-culture suggesting interspecies signalling molecules were essential for growth (101). This was confirmed by further work which investigated the nature of the dependence of strain MSC33 on being co-cultured with MSC105 and found that a 5 amino acid peptide (LQPEV) was necessary for the *Psychrobacter* sp. strain MSC33 to grow on standard media, the low concentration (3.5 nM) necessary confirmed the peptide was acting as a signalling molecule, as opposed to a nutrient (102). Similar experiments involving passaging of bacteria from freshwater sediment and soil through four successive diffusion chambers whilst also being plated on agar plates, indicated that multiple diffusion chamber cultivations were necessary to obtain isolates of certain genera, including representatives of rarely cultivated phyla, such as *Deltaproteobacteria*, *Verrucomicrobia*, *Spirochaetes* and *Acidobacteria*, on plates (103, 104). It was uncertain at the time whether this was due to a “bulking up” effect necessary to produce a detectable number of cells on plates or whether the passaging resulted in “domestication” of the strains and a newfound ability to grow on plates (103). More recently, the “microbial scout” hypothesis has been proposed which suggests the stochastic *i.e.* random, low-frequency, awakening of dormant cells may at least partly, explain the awakening of “uncultivable” strains after multiple passages/cultivations (105-108). However, the “domestication” phenomenon has also been reported in other situations, for example Nichols and co-workers identified a *Psychrobacter* sp. strain variant, MSC33c, which unlike its parent, did not require co-culture to grow on standard media (102).

A modification of the environmental bacterial growth approach involved the use of uninoculated diffusion chambers constructed using membranes with 0.2 µm diameter pores, to “trap” filamentous bacteria which are able to penetrate the diffusion chamber when it is placed on a soil surface (109). The experiment proved successful at selectively isolating filamentous actinomycetes including, for example, some unusual species from the genera *Dactylosporangium*, *Catellatospora*, *Catenulispora*, *Lentzea* and *Streptacidiphilus* (109), and also indicated that the use of gellan gum was more effective than agar, as had been noted by other investigators (88, 89). The diffusion chamber concept was later scaled up in the “diffusion bioreactor” where the large size means extended incubations can be maintained without requiring additional substrate supplementation in the chamber, although the results were similar to those obtained with small-scale devices (110). The early work using diffusion chambers typically used a very basic apparatus consisting of a single chamber and so produced a mixed culture that required considerable time and effort to process, *i.e.*, to isolate, purify and re-inoculate individual colonies. Unsurprisingly, there were many attempts to improve the efficiency and scalability of the methodology, which focussed on culturing individual bacteria from isolates each in their own separate micro-culture chambers

Aoi and co-workers used a device, named the “hollow fibre membrane chamber”, (HFMC), where the chambers consisted of multiple porous PVDF tubes which were seeded with diluted bacterial isolates from aqueous environments and the whole assembly was cultured in liquids mimicking the original environments (111). The

results were broadly consistent with those obtained from the iChip (see below) although it is difficult to see how the HFMC technology may be applied to non-aqueous situations. Ben-Dov and co-workers describes the use of GMD (gel microdroplet) technology to encapsulate bacteria from coral mucus, wastewater and soil in 1-3 mm diameter polysulfone coated agar spheres and their growth when the spheres are cultivated in the lab in environments mimicking the original habitats (112). Zengler and co-workers also applied a GMD based approach to encapsulate bacteria from soil and seawater in $\sim 10^7$ gel droplets which were cultured in columns through which appropriate media were pumped (113). GMD containing microcolonies were identified using flow cytometry to sort them into wells of 96 well microtitre plates (113). A similar proof of concept approach by Akselband and co-workers used an artificial mixture of slow and fast growing marine bacteria as inoculum and a combination of GMD and sorting the cultured droplets according to bacterial growth by FACS (114).

Technology	Advantages	Disadvantages
iChip	Enables growth of previously unculturable microbes; maintains natural environmental conditions	Limited scalability; may not replicate all environmental conditions precisely
Microfluidics	Precise control of microenvironments; high-throughput screening of growth conditions; minimal sample usage	Complex setup and operation; requires specialized equipment and expertise
Co-cultivation	Mimics natural microbial interactions; can induce production of novel secondary metabolites	Not all microbial interactions can be replicated; potential for dominance by faster-growing organisms
High-throughput culturing	Rapid testing of multiple conditions; large-scale cultivation with automation; good for discovering new strains	May miss microbes with specific or unknown growth requirements; potential bias towards fast-growing strains
Dilution to extinction	Selective enrichment of slow-growing or rare microbes; avoids competition from fast-growing species	Can be time-consuming; requires careful dilution and monitoring; may not capture all species in a sample
Trap Method	Captures microbes directly from the environment; simple and low-cost; effective for isolating rare or slow-growing organisms	May not support growth of fastidious organisms; limited control over microenvironment; potential contamination issues
Hollow Fibre Membrane Chamber	Mimics natural environmental conditions; allows for nutrient diffusion; effective for	Complex setup and maintenance; potential issues with membrane fouling;

studying slow-growing or unculturable microbes	requires careful monitoring of conditions
---	--

Table 2. Comparison of Microbial Cultivation Technologies

Development work on the diffusion chamber itself succeeded in producing a more scalable, high-throughput technology platform for microbial isolation, in parallel, on a large scale. The Isolation Chip, or iChip for short (115), consists of a single device comprising hundreds of miniature diffusion chambers that can be loaded with an average of one cell per chamber. A proof of concept study looking at bacterial 16S rRNA generated from libraries produced from iChips and Petri dishes inoculated with bacteria from soil and seawater indicated that more species were cultured using the iChips than conventional plate cultures *i.e.*, 525 from soil using iChips versus 314 from soil using Petri dishes, and 635 from seawater using iChips versus 265 from seawater using Petri dishes. There was virtually no overlap between the species isolated using the iChip and plate culture for each environmental source (115).

As the iChip's developers acknowledge, it is an idea rather than a device *per se* and although a detailed protocol for iChip use has been published (79) alternative modified forms are possible, including an "iPlate" compatible with its use in combination with standard 96 well microtitre plates (116). The simplicity of the technology has facilitated its uptake, for example, in undergraduate practical classes (117) and its ease of use in both solid and liquid environments has also contributed to its adoption by many research laboratories. Despite the widespread use of the iChip it is unclear whether it is necessarily the most effective method, as a direct comparison of the dilution to extinction and diffusion chamber techniques (and an agar modification technique) using intertidal sediment samples indicated that the diffusion chamber resulted in the lowest number of isolates from the three methods (118). Nevertheless, the iChip has been used to isolate hitherto unculturable bacteria from many environments/experimental situations, for example, mangroves, (119), estuarine sediment (120, 121), marine sediment (122, 123), marine sponges, (124, 125), decaying seaweed (126), various marine samples (127), paddy soil (128), crude oil contaminated soil (129), rocksalt (130), hot springs (131), the porcine intestine (132) and the human oral cavity (133).

Although use of the iChip has greatly enhanced and facilitated the culturing of previously uncultivable bacteria the approach has not yet realised its full potential in terms of reinvigorating the Waksman Platform. An early indication of its promise in discovering new natural products was its role in the discovery of neocitreamicin I and II which are novel members of the previously discovered citreamicin class of antibiotics with activity against methicillin-resistant *S. aureus* and vancomycin-resistant enterococci (134). Lassomycin, an inhibitor of the ClpP1P2C1 protease, and amycobactin, an inhibitor of the SecY protein exporter in *Mycobacterium tuberculosis* were also identified during screens of iChip cultured bacteria (135, 136). A more high-profile success which may be attributed to the iChip is its significant role in the discovery of teixobactin. Ling and co-workers screened extracts from 10,000 isolates obtained by growth in iChips for antimicrobial activity on plates overlaid with *S. aureus*. An extract from a new species of β -proteobacteria, provisionally named *Eleftheria terrae*, showed good activity and

contained a compound with a molecular mass of 1,242 Da, determined by mass spectrometry, which had not been previously reported. The compound was subsequently isolated and its complete chemical structure determined by NMR and advanced Marfey's analysis (137). The molecule, teixobactin, proved to be an unusual depsipeptide which contains enduracididine, methylphenylalanine, and four D-amino acids. Teixobactin had excellent activity against Gram-positive pathogens, including drug-resistant strains and potency against most species, including difficult-to-treat enterococci and *M. tuberculosis* was below 1 µg/ml.

Resistance was not developed to this compound, suggesting that the target is not a protein. And as treatment of whole cells of *S. aureus* with teixobactin ($1-5 \times \text{MIC}$) resulted in significant accumulation of the soluble cell wall precursor undecaprenyl-N-acetylmuramic acid-pentapeptide (UDP-MurNAc-pentapeptide), similar to the vancomycin-treated control cells, showing that one of the membrane-associated steps of peptidoglycan biosynthesis is blocked (137). Teixobactin inhibited peptidoglycan biosynthesis reactions *in vitro* in a dose-dependent manner with either Lipid I, Lipid II or undecaprenyl-pyrophosphate as a substrate and conversely the addition of purified Lipid II prevented teixobactin from inhibiting growth of *S. aureus*. These experiments clearly showed that teixobactin specifically interacts with the peptidoglycan precursor itself, rather than interfering with the activity of one of the enzymes. Purified precursors were incubated with teixobactin at different molar ratios, followed by extraction and subsequent thin-layer chromatography analysis and Lipid I and Lipid II were fully trapped in a stable complex that prevented extraction of the lipid from the reaction mixture (137). Moreover, teixobactin efficiently bound to the wall teichoic acid (WTA) precursor undecaprenyl-PP-GlcNAc (Lipid III). Although WTA is not essential *per se*, inhibition of late membrane-bound WTA biosynthesis steps is lethal due to accumulation of toxic intermediates. Moreover, teichoic acids anchor autolysins, preventing uncontrolled hydrolysis of peptidoglycan. Inhibition of teichoic acid synthesis by teixobactin would therefore help liberate autolysins, contributing to the excellent lytic and killing activity of this antibiotic (137). However, later work indicated the binding of teixobactin to Lipid II is weak and suggested formation of clusters of the antibiotic on membrane surfaces and entanglement of cell wall precursors also played a role in mediating its killing effect (138). This was confirmed by further work where it was shown teixobactin operates by settling onto the bacterial membrane where the Ile and D-*allo*-Ile residues anchor it, whilst the enduracidine binds to the PPi and MurNAc of Lipid II (139). The presence of L and D-amino acids in teixobactin allow it to form an antiparallel β -sheet structures which expand to form large-scale fibrillar structures (139). These compromise the integrity/functioning of the cell membrane in several ways: firstly, by sequestering Lipid II, secondly, by displacing phospholipids leading to thinning of the membrane and thirdly the concentration of the C₅₅ hydrophobic tails of Lipid II into hydrophobic patches causes ion leaks and loss of membrane potential, ultimately leading to cell death (139).

A second isolate from the same screen, *Lysobacter* sp. K5869, produced a compound, hypeptin, which exhibited a striking structural similarity and similar mode of action with teixobactin (140). Additionally, clovibactin, another depsipeptide antibiotic, produced by *E. terrae* ssp *carolina* binds to the PPi of multiple cell wall precursors and associates to form large-scale fibrillar structures in a similar way to teixobactin (141).

Teixobactin is the first member of a new class of Lipid II binding antibiotics, structurally distinct from glycopeptides, lantibiotics, and defensins, and lack of resistance to it, coupled with promising *in vivo* efficacy data, suggested it is a suitable candidate for clinical development (142). It is to be hoped that the iChip mediated discovery of teixobactin, (and other Lipid II binding depsipeptides), may not be an isolated event but the beginning of a new era of natural product drug discovery based on applying the Waksman Platform to newly mined microbial “dark matter” (143, 144).

Although the iChip has gained widespread acceptance by the scientific community in most embodiments it cannot be said to be a truly high throughput system. Indeed, it is debateable whether a cultivation system in which difficult to grow bacteria are cultured for several weeks will ever be described as “high-throughput” irrespective of the technology used simply because of the extended cultivation period. Although in basic embodiments iChips commonly possess only 96 compartments, attempts have been made to scale up the format so as to increase the number of compartments. For example, Liu describes the use of 384 well iChips (128), Nakai describes an “iChip” comprising 900 compartments (145) and the most common format of iChip described by Ingram possessed 180,000 compartments (100). However, it seems likely that significant improvements to iChip methodology will come not merely from increasing the compartment number but from the application of small-scale sophisticated, highly controllable liquid/particle handling technologies to the iChip format, in particular, the application of microfluidics/gel microdroplet (GMD) technology and FACS (146-148). For example, a simple improvement to iChip methodology which may be solved by use of a use of FACS relates to the problem of controlling the cell load per compartment. Conventionally, the one-cell-per-compartment achieved by dilution of the starting bacterial sample, leads to some compartments receiving more than one cell and others none. However, this uncertainty may be eliminated by the use of FACS to load a single cell per compartment (128, 149). The use of FACS may be also applied to more complex situations for example, Ohan conducted co-culture experiments where loaded multiple cell types and numbers of particular cell types were sorted by FACS into GMD which after culture were analysed by FACS (150). The use of microfluidics coupled with GMD is well known and has enabled many sophisticated high-throughput bacterial culture experiments for example, including studies to investigate antimicrobial resistance (151), co-culture (152) induction of natural product biosynthesis using co-culture (153) and growth of rare bacterial species isolated from complex samples (154). Although in most published studies the GMD are either held in an oil-water emulsion for a period before analysis or are cultured as a mixture in an array/ chamber in a conventional culture situation the additional advantage conferred by iChips *i.e.* the ability to separate and culture the GMD in a natural or nature-mimicking experimental situation, has been appreciated (155, 156). The application of microfluidics to natural product discovery will be discussed herein more detail in the next section.

5 Droplet-based microfluidic cultivation and analysis

Although *in situ* cultivation approaches are making significant progress the throughput is low. A complementary approach is microfluidic cultivation which

allows precise control over microenvironmental conditions and facilitates high-throughput experimentation (157), by encapsulating cells in microdroplets within an immiscible carrier fluid, typically oil, these systems create monodisperse emulsions that serve as isolated microreactors (158).

Droplet-based microfluidic systems have matured to become an important tool in biotechnology R+D for high-throughput experimentation (159). The underlying principle is as follows: by using two immiscible liquids, individual, separate, volumes of a first fluid are trapped in a second carrier fluid in a monodisperse emulsion. In a biological context, the first fluid phase is usually aqueous and thus biocompatible to allow encapsulation of live cells; the second fluid phase is an oil to separate the aqueous droplets.

Isolating individual fluid volumes in this way brings new approaches to research and development activities. Droplets are typically at sub-millimetre scale, with volumes in the nanolitre to picolitre range. This allows creation of millions of separate reaction volumes using small quantities of reagent extending the scale of experimentation. Each volume can contain a single cell, which allows separate assessment of each member of the cell population giving access to understanding of heterogeneity. Furthermore, droplets are amenable to process automation, allowing complex workflows to be performed quickly and repeatably on a large experimental space.

Droplets can be generated in a variety of ways, but typically microfluidic techniques are used. By tightly controlling geometry, flow rate, and liquid properties uniform monodisperse picolitre droplets can be generated at a high rate, up to 5 kHz (160) Microfluidics also then allows straightforward integration of other processes using droplets to create experimental workflows in the droplet domain. Examples of these processes include:

1. Splitting, whereby a single droplet is split into two droplets in a controlled volume ratio
2. Merging, whereby two droplets are combined into a single droplet
3. Selection, whereby a droplet can have a force applied and thus separated from other droplets. The force application can be controlled by a measurement of the droplet, thus allowing real-time selection of droplets and population enrichment.

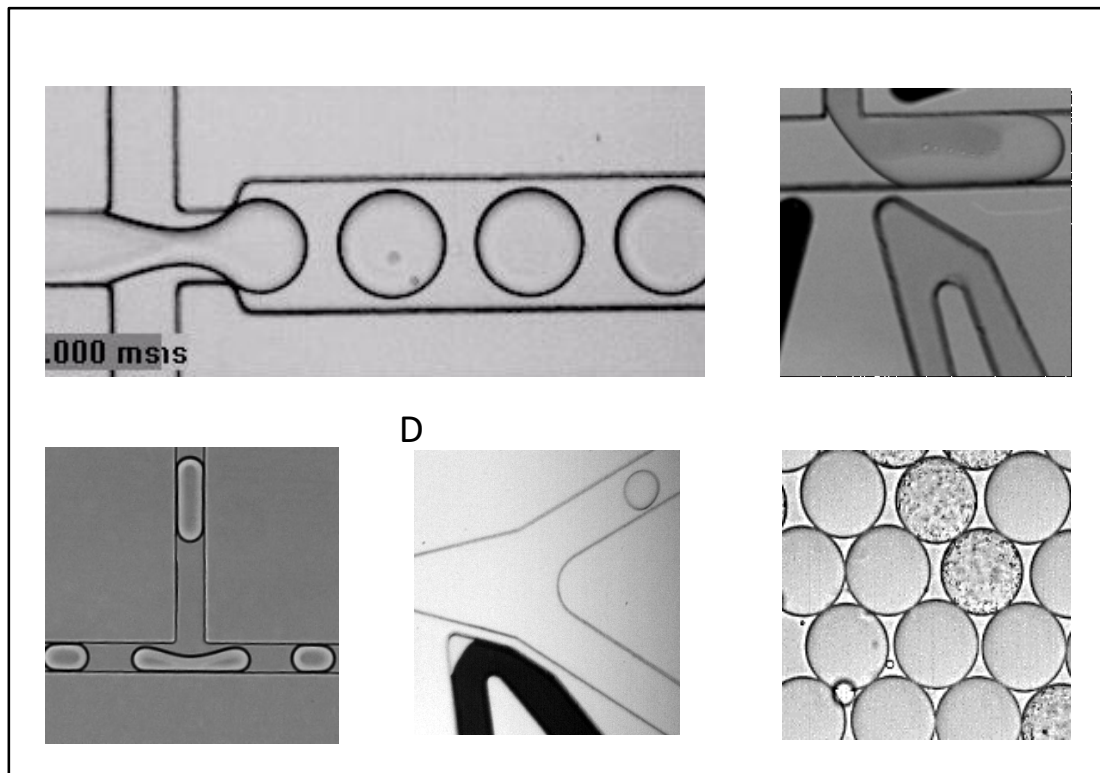


Figure 6: example unit operations in droplet microfluidic systems. [A] droplet generation: the aqueous phase enters from the left, oil phase with surfactant from top and bottom, under continuous constant flow the aqueous phase is broken down into droplets [B] injection into droplets: additional liquid can be added to droplets as they flow past the injection point. Droplet flow left to right, additional material is provided in the side channel at the top, electrode provides a timed pulse to disrupt the droplet surface to allow injection [C] droplet splitting: droplets can be split. Droplets flow from top and split into two equal parts at the junction. The volume split can be controlled through geometry (not shown) [D] droplet sorting: droplets flow left to right, when the electrode is switched on the droplet is diverted from the bottom channel to the top channel. If the droplet is measured just upstream (e.g. fluorescence signal) then droplets can be sorted into positive and negative. [E] proliferation in droplets. Image shows *E. coli* proliferation in droplets encapsulated at 1.1×10^6 CFU/ml concentration.

A critical part of the technique is the composition of the two phases and stability of the droplets. The aqueous phase needs to contain all materials necessary to maintain cell viability and subsequent outgrowth, and to support any characterization or measurement method such as binding of fluorescent markers. The oil phase needs to be non-toxic to cells and allow gas exchange to support respiration. Surfactants are also needed to maintain the interface between the two phases so the droplets do not collapse; furthermore, these interfacial molecules also need to minimize leakage of biomolecules from the aqueous phase into the oil phase and support biochemical assays (161). Leakage is often a limitation of the approach such that small, hydrophobic, molecules are not able to be retained in the aqueous droplet. Techniques such as double-emulsions (water-oil-water or oil-water-oil) can be used to address this situation (162). Hydrogels can also be used within the droplet structure to

replace the inner aqueous phase; this brings additional functionality to the droplet structure such as mimicking the extracellular environment (163).

Another critical part of the technique is the method used to characterize the content of each droplet. Typically, this involves a bioassay with an optical signal such as fluorescence which detects the presence of a target molecule in the droplet (164). In a cell context, such assays can detect either surface-bound markers or secreted markers, as the droplet retains the secreted molecules near the source. Furthermore, the small volume of the droplet ensures a rapid rise in concentration of the secreted target molecule allowing strong signal generation in relatively short time periods. The ability to utilize cell-secreted molecules as a marker and not just cell surface markers is an important advantage of droplet-based systems, particularly for natural product research.

A third critical part of the technique is controlling the composition of the droplet, particularly the number of cells in each droplet. This is done through controlling concentration of the aqueous phase such that the number of cells in each drop follows the Poisson probability distribution defined by λ (the average number of cells per unit volume) and k (the desired number of cells per droplet). By adjusting the starting concentration, the distribution of cells per droplet can be controlled to suit the particular application or workflow (165).

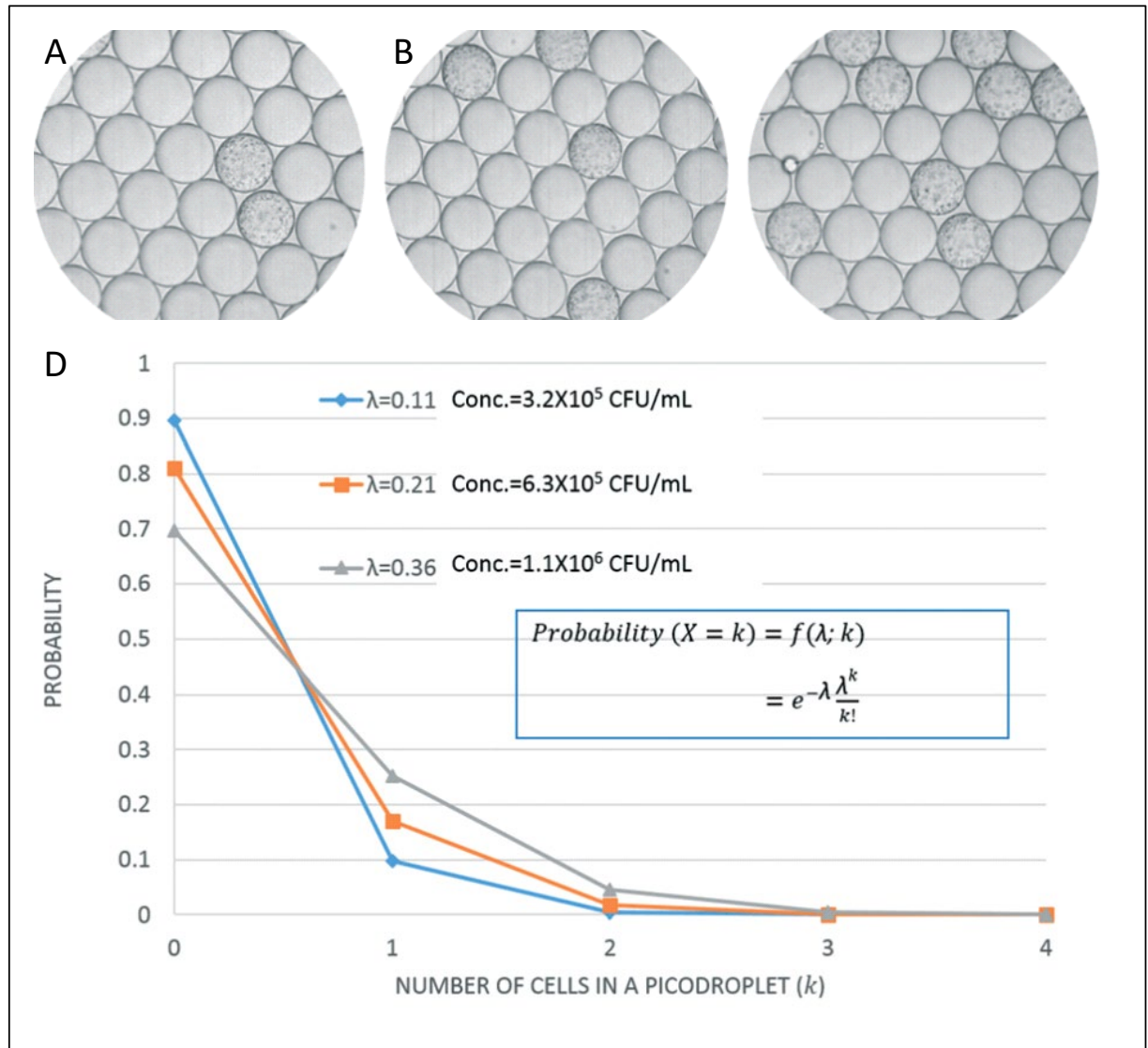


Figure 7: Poisson distribution in droplets. [A] *E. coli* encapsulated at 3.2×10^5 CFU/ml [B] *E. coli* encapsulated at 6.3×10^5 CFU/ml [C] *E. coli* encapsulated at 1.1×10^6 CFU/ml. Note the increasing number of droplets containing proliferating colonies in each image [D] calculated probability of number of cells in each droplet at encapsulation. Note as concentration rises more droplets have single cells but also two cells. The input concentration is a key control parameter for system operation.

Droplet-based systems fall into three classes based on the automation approach taken – passive array based, active array based, and flow based. In passive arrays, the droplets are generated in a two-dimensional array where they form a contiguous single layer of droplets. This array can then be interrogated via two-dimensional imaging techniques to assess which droplets have generated a signal (166).

In active arrays, also known as digital microfluidics, the droplets are generated and fed into an array that can be actively moved in two-dimensional space using techniques such as electrowetting (167). This array approach allows precise manipulation of many droplets in parallel to perform functions such as merging and splitting of droplets, thus enabling more complex workflows such as library generation or cell-cell pairing. The droplets are then assessed within the array to find

specific droplets with the required markers and then these droplets are separated for further analysis off-system.

In flow-based systems, the droplets are generated and then continue to move through microfluidic channels in single file. Similar to active arrays, further functions can be carried out on droplets as they move through the microfluidic channels including assessment and selection of droplets with the desired markers.

Droplet-based systems have two main application areas for natural products: high throughput screening (searching for naturally occurring molecules of interest and the associated organism producing the molecule, particularly 'non culturable' organisms) and microbial engineering (synthesizing molecules of interest via genetic manipulation). These will now be discussed in turn, focussing on active array and flow-based droplet systems.

5.1 High-throughput screening

Droplet-based systems provide a high-speed, robust platform for screening that can address large starting populations and libraries. By integrating the basic functions of droplet generation, characterisation and separation every member of the population can be assessed individually to isolate the sub-population of interest. This approach can be applied both to the microorganisms themselves and extracts from the microorganisms (168).

An early exemplar for the field is work by Najah and co-workers in 2014 (169). Here they used a droplet microfluidic system to screen uncultured bacteria from a wheat stubble field to identify microorganisms that can efficiently hydrolyse cellulosic biomass. Bacteria were encapsulated in 20 pL droplets containing a fluorogenic cellobiohydrolase substrate and sorted based on activity. The resulting selected population had 17-fold higher cellobiohydrolase activity and 7-fold higher endoglucanase activity when compared to selection based on growth on a starch-containing medium. Their system used custom-made polydimethylsiloxane (PDMS) chips manufactured with soft lithography methods and custom optics and electronics for detection and sorting. By assessing 100,000 bacteria in twenty minutes, this work demonstrated the underlying capability of droplet methods for rapid cell-by-cell screening for natural products.

Terekhov and co-workers developed their MDE-FACS system for high throughput screening (170). In this system, double emulsion droplets of ~500pL are used to encapsulate cells with a functional marker. These are screened using FACS and the selected cells either directly sequenced (if not amenable to culture) or plated and cultivated for LC-MS analysis and sequencing to give detailed genotype/phenotype descriptions. Using this approach, they demonstrate both high-efficiency screening of biocatalysts displayed on the surface of yeast and selection of bacteria from the human oral microbiota that inhibit growth of *S. aureus*. A synergistic mixture of pyocyanin, phenazine-1-carboxylic acid, and heptyl-4-hydroxyquinoline produced by *P. aeruginosa* was found to inhibit *S. aureus* growth indicating the potential of this approach in natural product discovery.. The latter in particular shows the power of droplets to allow pairwise interactions between cells to eliminate interference during

884 screening. In this work, double-emulsion droplets were used to allow the usage of a
885 traditional flow cytometry techniques for fluorescence-based screening. In this flow-
886 based system, the droplet generation chips were custom made using soft-lithography
887 techniques with PDMS.

888
889 Oberpaul and co-workers used droplet-based fluorescent sorting to select growth
890 medium to support isolation of microbial diversity from soil samples and subsequent
891 bioactivity-guided discovery of natural products (171). In the first part of this work,
892 droplets of ~300pL were generated using solidified agarose with eight different media
893 types and a mixed population of soil microbes. Encapsulation was performed at $\lambda=0.1$
894 to give a low probability of co-culture (<0.5%). The droplet populations were sorted
895 based on membrane potential using a carbocyanine dye DiOC₂(3), cultured,
896 sequenced for diversity (1071 colonies) and compared to the original population. This
897 approach clearly identified one medium capable of supporting a much more diverse
898 population at genus level, and the cultured microbes were then used to screen for
899 bioactivity in a well-plate format. Droplet generation was performed using a
900 commercially-available continuous flow device and fluorescence-based screening
901 using traditional flow cytometer techniques, facilitated by the solidified agarose
902 droplets.

903
904 Theodorou and co-workers show a novel approach to overcome the problem of
905 natural compound extraction in a high-throughput screening system with mammalian
906 cells to identify pro-apoptotic compounds (172). A metagenomic plasmid library is
907 prepared by extracting DNA from environmental samples and transformed into a non-
908 pathogenic laboratory bacterial strain that is rendered invasive. This library is then
909 placed into droplets at single bacterium level and allowed to proliferate avoiding
910 fitness competition issues while maintaining clonality. Single droplets containing
911 amplified metagenomic clones are then merged using electro-fusion with previously
912 made single droplets containing mammalian cells to create a library of unique
913 combinations of bacteria and cell. Some bacteria invade the mammalian cells, are
914 lysed and release their newly synthesised bioactive compounds. The emulsion is
915 broken, the cells washed and screened using flow cytometry. Apoptosis is identified
916 through cleavage of fluorescent cell-permeable reporter peptides specific to early-
917 stage apoptosis. Cells are subsequently lysed to recover the bacterium and determine
918 the structure of the pro-apoptotic natural product. Droplets are an essential part of the
919 workflow to allow non-biased library amplification and controlled merging of single
920 bacterial clones and target cells. As the screening is done using in-cell signals,
921 droplets are no longer needed at this point. For this work, custom glass/PDMS
922 microfluidic chips were used to generate droplets at ~220pL volumes for bacteria and
923 700pL for mammalian cells.

924
925 Similarly, Mahler and co-workers used droplet-based techniques to screen and
926 cultivate antibiotic-producing microbes from natural communities (153). Again,
927 droplets were used to provide isolation between individual microbes to avoid bias or
928 competition and allow analysis of difficult-to-culture organisms. ~200pL droplets
929 were used at $\lambda = 0.4$ to give just over 6% probability of non-single cells in droplets
930 with an average of 9×10^6 droplets. The droplets were incubated using dynamic
931 droplet incubation to ensure aerobic and homogenous cultivation (173). Post
932 incubation, droplets were either directly deposited on agar plates or first screened for
933 antimicrobial activity using a fluorescent reporter cell line. The plating out uses an

interesting technique whereby droplets are dispensed at 10-50 droplets per second from a nozzle following a spiral pattern across the agar surface thus providing efficient spatial separation for further colony growth and selection. Using this approach, Mahler demonstrates droplet approaches provide higher CFU concentration compared to convention plating and a more diverse set of bacterial species. They also then identified and selected a novel soil bacterium producing antimicrobial compounds, *i.e.* bacillaenes and surfactins so proving the concept that the methodology/equipment could be used in natural product discovery. This work was done using custom-made PDMS microfluidic devices and custom-made optics, electronics and software for droplet detection and sorting.

Another application of droplet microfluidics for natural product selection is isolation of microbial biosynthetic gene clusters (BGCs). BGCs represent a valuable source of bioactive molecules but are sparsely distributed in the genome and thus difficult to characterize. Xu and coworkers developed a method called MAPLE – Microfluidic Automated Plasmid Library Enrichment (174). In this approach, a metagenomic library is created and single cells encapsulated in droplets. The droplets are incubated to allow cell proliferation, then merged with a second droplet containing primers and probes for PCR. After thermocycling, droplets are sorted using fluorescence to identify presence of the target gene. The selected droplets are then de-emulsified to allow next-generation sequencing with greatly enriched target DNA content. In this work, complex custom microfluidic devices for encapsulation, merging and sorting were manufactured using PDMS casting techniques. The technique was then used to characterize microbial polyketide synthesis in an Antarctic soil metagenome. The type 1 polyketide synthase enzyme cluster was targeted using a degenerate primer set targeting the ketosynthase gene. The technique was shown to outperform both shotgun and amplicon sequencing, detecting more KS genes (366 hits compared to 32 and 43 respectively) and giving deeper coverage of these genes (170 hits over 1000-fold compared to 3 hits). The authors note the application of this technique to narrow down a plasmid library to gene clusters associated with specific drug resistance.

Droplet microfluidic methods typically use fluorescence reporter assays to identify cells of interest. However, not all target molecules are amenable to this approach. To overcome this limitation, and allow assessment of more complex mixtures, other measurement techniques are being developed. Of particular interest is coupling of droplets microfluidics to mass spectrometry to allow real-time spectral analysis. Murray et al provide a recent review of this nascent field (175) showing how both electrospray ionisation (ESI) and matrix-assisted laser desorption ionization (MALDI) can be coupled into droplet-based systems.

5.2 Microbial engineering

While screening of naturally occurring molecules is a powerful approach, it can be enhanced by using techniques to *de novo* engineer microbes. In a droplet context, these fall into two classes: direct genome editing (whereby pre-defined changes are made to the gene to achieve a desired outcome) and directed evolution (whereby the microbe is subjected to specific evolutionary pressure to promote the desired outcome).

5.3 Direct editing approaches

There are two approaches to gene editing methods. In the first approach, the edits are done to cells in bulk outside the droplet system using standard techniques such as chemical transfection or bulk electroporation, and then droplets used to screen for successful edits. In the second approach, the editing is also done in the droplet. This approach facilitates easier integration of the overall workflow, much reduced reagent volumes, and can increase the editing efficiency (176).

As an example of this approach, Yun and co-workers used droplet microfluidics to engineer promoters and fine-tune expression to improve erythromycin production in *Saccharopolyspora erythraea* (177). Two plasmid libraries of synthetic promoters were generated using random mutagenesis and used to transform *E. coli*, then mobilised to *S. erythraea* by conjugation to give ~50,000 variants. Spores from the library were then encapsulated into droplets of ~400pL volume and incubated for three days along with wild-type spores. Screening was performed using green fluorescence expressed by the cells at approximately 15Hz. Of note is the use of relative fluorescent signal to rank promoter performance within the library rather than a binary yes/no assessment typically associated with screening. This places much more stringent precision and linearity performance requirements on the fluorescence measurement method to give robust quantification. Using this approach erythromycin production was improved up to 137-fold. Droplet generation and sorting was performed using custom microfluidic devices and electronics in a flow-through manner developed during earlier work on *Streptomyces*.

Iwai and co-workers used an array-based digital droplet microfluidic approach to automate CRISPR-based strain engineering (178). They developed a custom 10 x 10 array whereby the electrodes are used both for droplet movement and merging via electrowetting and for electroporation. Loading of the array is done using acoustic droplet ejection techniques dispensing ~5nL volume droplets which provides an interesting alternative droplet generation method, albeit at nL rather than pL volume range. Initial work demonstrated >80% successful transformation of GFP into *E. coli* in a highly parallel way. This was then extended to implement Multiplexed Automated Genome Engineering (MAGE) (179) with CRISPR to make small genome edits in *E. coli*. They demonstrate implementing a point mutation in the galactokinase gene with the same yield as the cuvette-based bench process and creating a stable strain for production of indigoidine. The chips used in this work were custom-made using lithography techniques.

Wippold and co-workers utilise droplets in a flow-based system to perform high-efficiency DNA transfer between species (180). They took a previously developed conjugation-optimised strain of *B. subtilis* (the XPORT system (181)) and co-encapsulated this into ~120pL volume droplets with a target cell, in this case another strain of *B. subtilis*, at around 200 droplets per second. Such DNA transfer is reliant on close contact between the donor and receiver cells, and thus usually requires the donors in great excess to the receiver, around $1 \times 10^4:1$. Bringing the donor and receiver cell in close proximity within a picolitre-volume droplet was found to improve conjugation process efficiency by up to two orders of magnitude with ratios of only 3:1. The droplet microfluidics in this work were custom made using soft

lithography techniques in PDMS. The two cells are mixed prior to droplet formation and the encapsulation ratios controlled using the flow rates of the two inputs.

5.4 Directed evolution approaches

Directed evolution brings both microbial engineering and high-throughput screening together into a single workflow to provide evolutionary pressure on an organism. Droplet techniques are well suited to directed evolution as they provide a simple way to execute the necessary recursive multi-step workflow with large numbers of cells. Several recent reviews have been published including Chiu and Stavrakis in 2019 (182), Manteca et al in 2021(183), and Gantz et al in 2023 (184).

As an example of the workflow integration possible using droplets, Rosenthal and co-workers recently demonstrate a process where cells are cycled between growth, mutagenesis and screening to evolve enzymatic activity, in this case alditol oxidase (185). Using *E. coli*, the process integrates mutagenesis using a fusion of nicking CAS9 and a mutagenic polymerase (Poll), a fluorescence-based activity assay, encapsulation of bacteria and reagents, incubation, selection and merging of positive droplets back into growth medium. The process is then repeated after ~20 hours to allow selected cells to grow and continue mutagenesis. They achieved a 10.5-fold improvement compared to wild type in a few days, compared with a 2.4-fold improvement using traditional multi-step error-prone PCR. The process is run in small droplets (~5pL volume) at low occupation ($\lambda = 0.08$) to give very low probability of non-single cells in droplets ($< 0.3\%$) and uses high speeds of 5×10^5 cells per hour to compensate for the large number of empty droplets.

5.5 Technology access

Much of the work reviewed in this paper uses custom-built microfluidics designed and made by academic microfluidics research teams. However, in an industrial context, the use of droplet technology for natural product identification and optimization requires the availability of robust equipment and tools to support commercial exploitation. This includes instrumentation, reagents (particularly surfactants), assays and software. Several companies have emerged in the last decade providing such tools as the value and capabilities of droplet-based methods have become established.

Vendor	Platform	Specifications
Atrandi Biosciences https://atrandi.com/	Onyx	Hydrogel-based droplets Droplet generation Droplet merging Droplet splitting Flow control with up to four syringe pumps High speed camera for process monitoring

Vendor	Platform	Specifications
	Styx	Hydrogel based droplets Droplet characterization and sorting Fluorescent detection, up to four detection channels High speed camera for process monitoring
LiveDrop Bio https://livedrop-bio.com/	Modaflow	Droplet generation Droplet merging Droplet pairing Droplet characterization and sorting Fluorescent detection, up to six detection channels
Samplix https://samplix.com/	Xdrop	Double emulsion droplets Droplet generation Up to eight parallel samples
	Xdrop Sort	Double emulsion droplets Droplet generation Droplet characterization and sorting Fluorescent detection, single channel Up to eight parallel samples
Dolomite Microfluidics https://www.dolomite-microfluidics.com/microfluidic-systems/	μ Encapsulator System	Droplet generation Flow controls with up to three pressure pumps High speed camera for process monitoring

Table 5: commercially available flow-based droplet microfluidic platforms.

5.6 Digital microfluidic systems

Digital microfluidic systems are less mature than flow systems. However, some commercial systems are now appearing on the market, as shown in Table 6. Applying such approaches to natural product discovery could provide novel highly parallelised workflows and remains an area ripe for further development.

Vendor	Platform	Specifications
Sci-bots https://sci-bots.com/	DropBot	68 electrodes, each 2.2 x 2.2mm square, 8 dispensing electrodes, 8 reservoir electrodes

Vendor	Platform	Specifications
Nicoya https://nicoyalife.com/	Alto	Integrated digital microfluidics and surface plasmon resonance
Lightcast Bio https://www.lightcast.bio/	Single cell platform	Beta access programme only to assess technology

Table 6: commercially available digital microfluidic platforms.

5.7 Surfactants

Various surfactants are available commercially to support droplet microfluidics. All systems use hydrofluoroether oils such as 3M™ Novec™ 7500 or Fluorinert™ FC-40 as these are inert, biologically compatible and compatible with PDMS which is widely used to manufacture the microfluidic devices.

Vendor	Name	Oil phase	Specifications
Bio-Rad https://www.bio-rad.com	Droplet generation oil	HFE 7500	Optimised for ddPCR using probes
Dolomite Microfluidics	Fluoro-phase	HFE 7500	Fluorinated silica nanoparticles (Pickering emulsion), reduced leakage from droplets
Emulseo https://www.emulseo.com/	FluoSurf-C™, Fluosurf-O™ and -Fluosurf-S™	HFE 7500	O series optimized for low autofluorescence, S series optimized for extreme thermocycling
Fluigent https://www.fluigent.com/	dSurf	HFE 7500	Suitable for dPCR and cell culture / analysis
RAN Biotechnologies https://www.ranbiotechnologies.com	FluoroSurfactant 008	HFE 7500, FC-40	Suitable for encapsulation of DNA, RNA, proteins and single cells
Sphere Fluidics Ltd.	Pico-Surf™	HFE 7500	Suitable for encapsulation of

Vendor	Name	Oil phase	Specifications
			DNA, RNA, proteins and single cells

Table 7: commercially available surfactants and oil phases for droplet microfluidics

Surfactant selection is driven by the specifics of the workflow, particularly the target molecules of interest and the measurement modality. Small, hydrophobic, molecules are more difficult to retain in droplet-based systems and nano-particle based surfactants may offer advantages in this scenario. If PCR techniques form part of the workflow, such as a preparatory step for DNA sequencing, then the surfactants optimized for thermal stability are preferred to avoid droplet breakdown. Often cell viability is a critical aspect of the workflow to retain cells for further analysis or outgrowth, the surfactant should be assessed for long-term biological compatibility before committing to extensive experimental programmes.

It should also be noted that these oils are classified as per and polyfluoroalkyl substances (PFAS) which are coming under increased scrutiny due to their long-term environmental impact. Alternative non-PFAS materials are starting to be developed as replacements which will enter the market in the coming years and become the default materials for droplet-based methods.

5.8 Future directions

Droplet microfluidics is now a well-established technique for rapid single-cell functional assessment in the context of natural product discovery. Future work will continue to integrate this core technology into screening pipelines. For example, droplets are uniquely suited to metagenomic studies as the compartmentalization fundamentally supports maintaining phenotype-genotype linkage. Droplets can also support synthetic biology approaches whereby large perturbation libraries are generated (either rationally or through random approaches) and then screened to better understand genetic pathways.

These benefits are further amplified in the context of machine learning approaches. AI-type deep neural network models rely on large quantities of high-quality, unbiased data for training: droplets provide this repeatable, high-throughput, low-cost platform to generate such data sets. For example, using AI models to predict functional BGCs could be supported using droplet-based methods to establish large robust training data sets and carry out post-prediction functional validation. Zhou et al (186) provide an overview of potential synergies between AI models and droplet microfluidics, both using droplets to generate data and using AI models to improve droplet assessment and screening.

While droplets are typically used for isolating and assessing single cells, an interesting technique is cell-cell pairing, whereby droplets are assembled with one of each type of cell. This approach allows high-throughput techniques to assess true in-context

biological function, such as directly identifying novel bioactive compounds from environmental samples at scale.

6 Conclusion

The original “Golden Age” of natural product discovery was powered by a single screening paradigm: the “Waksman Platform.” By the late 1970s, however, the limitations of this approach became clear, contributing to its decline and the pharmaceutical industry’s subsequent retreat from natural product exploration. Today, a suite of emerging technologies has the potential to rekindle interest in natural products, particularly if integrated into a unified, high-throughput discovery platform.

Metagenomic sequencing has revealed a far greater diversity of BGCs than previously appreciated, suggesting that many organisms—both cultured and uncultured—have the capacity to produce a broader range of natural products than currently realised. Yet significant challenges remain. Successfully cloning and expressing large DNA fragments (>50 kb) from metagenomic libraries is still technologically demanding, and activating silent or poorly expressed BGCs represents another major bottleneck.

Beyond metagenomics, the iChip and its variants now enable access to previously uncultivable microbial “dark matter,” which constitutes more than 99% of bacterial species. Although early implementations are relatively small-scale, combining the iChip with high-throughput techniques—such as droplet-based microfluidics—could provide platforms that meet modern pharmaceutical demands. Droplet microfluidics enables ultra-high-throughput, low-volume screening workflows, facilitating the rapid identification of cells producing compounds of interest and supporting downstream directed evolution and gene editing strategies. Moreover, complementary approaches like SMIRC promise to further illuminate the chemical diversity available, providing additional pathways for discovering novel compounds.

Taken together, these advances have the potential to create a modern successor to the Waksman Platform. By integrating metagenomic approaches, iChip-enabled cultivation, droplet-based screening, and other emerging technologies, we may usher in a new era of natural product discovery.

7 Conflicts of interest

MDW is Chief Scientific Officer for Bactobio, a UK-based biotechnology company that is focussed on antimicrobial drug discovery.
RJH is Chief Technical Officer for Sphere Bio Ltd., a UK-based biotechnology company that sells instruments, consumable and reagents for performing droplet-based workflows.
NA and RL declare no conflict of interests.

8 Acknowledgements

We thank Dr Brad Hover for reading the manuscript and his comments.

9 References

1. Newman DJ, Cragg GM. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J Nat Prod*. 2020;83(3):770-803.
2. Berdy J. Bioactive microbial metabolites. *J Antibiot (Tokyo)*. 2005;58(1):1-26.
3. Watve MG, Tickoo R, Jog MM, Bhole BD. How many antibiotics are produced by the genus *Streptomyces*? *Arch Microbiol*. 2001;176(5):386-90.
4. Katz L, Baltz RH. Natural product discovery: past, present, and future. *J Ind Microbiol Biotechnol*. 2016;43(2-3):155-76.
5. Baltz RH. Marcel Faber Roundtable: Is our antibiotic pipeline unproductive because of starvation, constipation, or lack of inspiration? *Journal of Industrial Microbiology and Biotechnology*. 2006;33(7):507-13.
6. Baltz RH. Natural product discovery and development at Eli Lilly and Company: one scientist's view. *SIM News*. 2005;55(1):5-16.
7. Fleming A. On the antibacterial action of cultures of a *Penicillium*, with special reference to their use in the isolation of *B. influenzae*. *British Journal of Experimental Pathology*. 1929;10(3):226-36.
8. Abraham EP, Chain E, Fletcher CM, Gardner AD, Heatley NG, Jennings MA, et al. Further observations on penicillin. *The Lancet*. 1941;238(6164):177-89.
9. Schatz A, Bugle E, Waksman SA. Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. *Proceedings of the Society for Experimental Biology and Medicine*. 1944;55:66-9.
10. Waksman SA. What is an antibiotic or an antibiotic substance? *Mycologia*. 1947;39(5):565-9.
11. Grove RD, Hetzel AM. *Vital Statistics Rates in the United States, 1940-1960*. Washington, D.C.: US Government Printing Office; 1968.
12. Silver LL. Challenges of antibacterial discovery. *Clin Microbiol Rev*. 2011;24(1):71-109.
13. Stapley EO. A comparative study of fermentation conditions for the production of antibiotics by submerged and surface methods. *Applied Microbiology*. 1958;6(6):392-8.
14. Ackermann BL, Regg BT, Colombo L, Stella S, Coutant JE. Rapid analysis of antibiotic-containing mixtures from fermentation broths by using liquid chromatography-electrospray ionization-mass spectrometry and matrix-assisted laser desorption ionization-time-of-flight-mass spectrometry. *J Am Soc Mass Spectrom*. 1996;7(12):1227-37.
15. Debono M, Barnhart M, Carrell CB, Hoffmann JA, Occolowitz JL, Abbott BJ, et al. A21978C, a complex of new acidic peptide antibiotics: isolation, chemistry, and mass spectral structure elucidation. *J Antibiot (Tokyo)*. 1987;40(6):761-77.
16. Cohen SS. A strategy for the chemotherapy of infectious disease. *Science*. 1977;197(4302):431-2.

- 1216 17. Chan PF, Holmes DJ, Payne DJ. Finding the gems using genomic discovery:
1217 antibacterial drug discovery strategies – the successes and the challenges. *Drug*
1218 *Discovery Today: Therapeutic Strategies*. 2004;1(4):519-27.
- 1219 18. Payne DJ, Gwynn MN, Holmes DJ, Pompliano DL. Drugs for bad bugs:
1220 confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov*.
1221 2007;6(1):29-40.
- 1222 19. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and
1223 computational approaches to estimate solubility and permeability in drug discovery
1224 and development settings. *Advanced Drug Delivery Reviews*. 1997;23:3-25.
- 1225 20. Lipinski CA. Lead- and drug-like compounds: the rule-of-five revolution.
1226 *Drug Discov Today Technol*. 2004;1(4):337-41.
- 1227 21. Benedetto Tiz D, Kikelj D, Zidar N. Overcoming problems of poor drug
1228 penetration into bacteria: challenges and strategies for medicinal chemists. *Expert*
1229 *Opin Drug Discov*. 2018;13(6):497-507.
- 1230 22. Projan SJ. Why is big Pharma getting out of antibacterial drug discovery? *Curr*
1231 *Opin Microbiol*. 2003;6(5):427-30.
- 1232 23. Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR,
1233 James KD, et al. Complete genome sequence of the model actinomycete *Streptomyces*
1234 *coelicolor* A3(2). *Nature*. 2002;417(6885):141-7.
- 1235 24. Gavriilidou A, Kautsar SA, Zaburannyi N, Krug D, Muller R, Medema MH, et
1236 al. Compendium of specialized metabolite biosynthetic diversity encoded in bacterial
1237 genomes. *Nat Microbiol*. 2022;7(5):726-35.
- 1238 25. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng JF, et al.
1239 Insights into the phylogeny and coding potential of microbial dark matter. *Nature*.
1240 2013;499(7459):431-7.
- 1241 26. Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, et al.
1242 A new view of the tree of life. *Nat Microbiol*. 2016;1:16048.
- 1243 27. Handelsman J. Metagenomics: application of genomics to uncultured
1244 microorganisms. *Microbiol Mol Biol Rev*. 2004;68(4):669-85.
- 1245 28. Woese CR, Fox GE. Phylogenetic structure of the prokaryotic domain: the
1246 primary kingdoms. *Proc Natl Acad Sci U S A*. 1977;74(11):5088-90.
- 1247 29. Almeida A, Mitchell AL, Boland M, Forster SC, Gloor GB, Tarkowska A, et
1248 al. A new genomic blueprint of the human gut microbiota. *Nature*.
1249 2019;568(7753):499-504.
- 1250 30. Lloyd-Price J, Arze C, Ananthakrishnan AN, Schirmer M, Avila-Pacheco J,
1251 Poon TW, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel
1252 diseases. *Nature*. 2019;569(7758):655-62.
- 1253 31. Gilbert JA, Quinn RA, Debelius J, Xu ZZ, Morton J, Garg N, et al.
1254 Microbiome-wide association studies link dynamic microbial consortia to disease.
1255 *Nature*. 2016;535(7610):94-103.
- 1256 32. Fierer N, Leff JW, Adams BJ, Nielsen UN, Bates ST, Lauber CL, et al. Cross-
1257 biome metagenomic analyses of soil microbial communities and their functional
1258 attributes. *Proc Natl Acad Sci U S A*. 2012;109(52):21390-5.
- 1259 33. Schloss PD, Girard RA, Martin T, Edwards J, Thrash JC. Status of the
1260 Archaeal and Bacterial Census: an Update. *mBio*. 2016;7(3).
- 1261 34. Roesch LF, Fulthorpe RR, Riva A, Casella G, Hadwin AK, Kent AD, et al.
1262 Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J*.
1263 2007;1(4):283-90.

1264 35. Paoli L, Ruscheweyh HJ, Forneris CC, Hubrich F, Kautsar S, Bhushan A, et
1265 al. Biosynthetic potential of the global ocean microbiome. *Nature*.
1266 2022;607(7917):111-8.

1267 36. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun
1268 metagenomics, from sampling to analysis. *Nat Biotechnol*. 2017;35(9):833-44.

1269 37. Sharpton TJ. An introduction to the analysis of shotgun metagenomic data.
1270 *Front Plant Sci*. 2014;5:209.

1271 38. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new
1272 versatile metagenomic assembler. *Genome Res*. 2017;27(5):824-34.

1273 39. Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, Shin SB, et
1274 al. metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nat*
1275 *Methods*. 2020;17(11):1103-10.

1276 40. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence
1277 classification using exact alignments. *Genome Biol*. 2014;15(3):R46.

1278 41. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using
1279 DIAMOND. *Nat Methods*. 2015;12(1):59-60.

1280 42. Huson DH, Auch AF, Qi J, Schuster SC. MEGAN analysis of metagenomic
1281 data. *Genome Res*. 2007;17(3):377-86.

1282 43. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, et al.
1283 antiSMASH 7.0: new and improved predictions for detection, regulation, chemical
1284 structures and visualisation. *Nucleic Acids Res*. 2023;51(W1):W46-W50.

1285 44. Skinnider MA, Merwin NJ, Johnston CW, Magarvey NA. PRISM 3: expanded
1286 prediction of natural product chemical structures from microbial genomes. *Nucleic*
1287 *Acids Res*. 2017;45(W1):W49-W54.

1288 45. Kautsar SA, Blin K, Shaw S, Navarro-Munoz JC, Terlouw BR, van der Hooft
1289 JJJ, et al. MIBiG 2.0: a repository for biosynthetic gene clusters of known function.
1290 *Nucleic Acids Res*. 2020;48(D1):D454-D8.

1291 46. Navarro-Munoz JC, Selem-Mojica N, Mullowney MW, Kautsar SA, Tryon
1292 JH, Parkinson EI, et al. A computational framework to explore large-scale
1293 biosynthetic diversity. *Nat Chem Biol*. 2020;16(1):60-8.

1294 47. Kautsar SA, van der Hooft JJJ, de Ridder D, Medema MH. BiG-SLiCE: A
1295 highly scalable tool maps the diversity of 1.2 million biosynthetic gene clusters.
1296 *Gigascience*. 2021;10(1).

1297 48. Cimermancic P, Medema MH, Claesen J, Kurita K, Wieland Brown LC,
1298 Mavrommatis K, et al. Insights into secondary metabolism from a global analysis of
1299 prokaryotic biosynthetic gene clusters. *Cell*. 2014;158(2):412-21.

1300 49. Waschulin V, Borsetto C, James R, Newsham KK, Donadio S, Corre C, et al.
1301 Biosynthetic potential of uncultured Antarctic soil bacteria revealed through long-read
1302 metagenomic sequencing. *ISME J*. 2022;16(1):101-11.

1303 50. Albertsen M, Hugenholtz P, Skarshewski A, Nielsen KL, Tyson GW, Nielsen
1304 PH. Genome sequences of rare, uncultured bacteria obtained by differential coverage
1305 binning of multiple metagenomes. *Nat Biotechnol*. 2013;31(6):533-8.

1306 51. Bertrand D, Shaw J, Kalathiyappan M, Ng AHQ, Kumar MS, Li C, et al.
1307 Hybrid metagenomic assembly enables high-resolution analysis of resistance
1308 determinants and mobile elements in human microbiomes. *Nat Biotechnol*.
1309 2019;37(8):937-44.

1310 52. Koren S, Schatz MC, Walenz BP, Martin J, Howard JT, Ganapathy G, et al.
1311 Hybrid error correction and de novo assembly of single-molecule sequencing reads.
1312 *Nat Biotechnol*. 2012;30(7):693-700.

1313 53. Garcia JA, Fernandez-Guerra A, Casamayor EO. A close relationship between
1314 primary nucleotides sequence structure and the composition of functional genes in the
1315 genome of prokaryotes. *Mol Phylogenet Evol.* 2011;61(3):650-8.

1316 54. Brady SF. Construction of soil environmental DNA cosmid libraries and
1317 screening for clones that produce biologically active small molecules. *Nat Protoc.*
1318 2007;2(5):1297-305.

1319 55. Gomez-Escribano JP, Bibb MJ. Engineering *Streptomyces coelicolor* for
1320 heterologous expression of secondary metabolite gene clusters. *Microb Biotechnol.*
1321 2011;4(2):207-15.

1322 56. Thanapipatsiri A, Claesen J, Gomez-Escribano JP, Bibb M, Thamchaipenet A.
1323 A *Streptomyces coelicolor* host for the heterologous expression of Type III polyketide
1324 synthase genes. *Microb Cell Fact.* 2015;14:145.

1325 57. Peng Q, Gao G, Lu J, Long Q, Chen X, Zhang F, et al. Engineered
1326 *Streptomyces lividans* Strains for Optimal Identification and Expression of Cryptic
1327 Biosynthetic Gene Clusters. *Front Microbiol.* 2018;9:3042.

1328 58. Ahmed Y, Rebets Y, Estevez MR, Zapp J, Myronovskyi M, Luzhetskyy A.
1329 Engineering of *Streptomyces lividans* for heterologous expression of secondary
1330 metabolite gene clusters. *Microb Cell Fact.* 2020;19(1):5.

1331 59. Fazal A, Thankachan D, Harris E, Seipke RF. A chromatogram-simplified
1332 *Streptomyces albus* host for heterologous production of natural products. *Antonie Van*
1333 *Leeuwenhoek.* 2020;113(4):511-20.

1334 60. Jung WS, Lee SK, Hong JS, Park SR, Jeong SJ, Han AR, et al. Heterologous
1335 expression of tylosin polyketide synthase and production of a hybrid bioactive
1336 macrolide in *Streptomyces venezuelae*. *Appl Microbiol Biotechnol.* 2006;72(4):763-9.

1337 61. Hong JS, Park SH, Choi CY, Sohng JK, Yoon YJ. New olivosyl derivatives of
1338 methymycin/pikromycin from an engineered strain of *Streptomyces venezuelae*.
1339 *FEMS Microbiol Lett.* 2004;238(2):391-9.

1340 62. Cook TB, Jacobson TB, Venkataraman MV, Hofstetter H, Amador-Noguez D,
1341 Thomas MG, et al. Stepwise genetic engineering of *Pseudomonas putida* enables
1342 robust heterologous production of prodigiosin and glidobactin A. *Metab Eng.*
1343 2021;67:112-24.

1344 63. Weihmann R, Domrose A, Drepper T, Jaeger KE, Loeschke A. Protocols for
1345 yTREX/Tn5-based gene cluster expression in *Pseudomonas putida*. *Microb*
1346 *Biotechnol.* 2020;13(1):250-62.

1347 64. Liu Q, Shen Q, Bian X, Chen H, Fu J, Wang H, et al. Simple and rapid direct
1348 cloning and heterologous expression of natural product biosynthetic gene cluster in
1349 *Bacillus subtilis* via Red/ET recombineering. *Sci Rep.* 2016;6:34623.

1350 65. Liu Z, Lin Z, Nielsen J. Expression of fungal biosynthetic gene clusters in *S.*
1351 *cerevisiae* for natural product discovery. *Synth Syst Biotechnol.* 2021;6(1):20-2.

1352 66. Kealey JT, Craig JP, Barr PJ. Identification of a lichen depside polyketide
1353 synthase gene by heterologous expression in *Saccharomyces cerevisiae*. *Metab Eng*
1354 *Commun.* 2021;13:e00172.

1355 67. Brady SF, Chao CJ, Clardy J. New natural product families from an
1356 environmental DNA (eDNA) gene cluster. *J Am Chem Soc.* 2002;124(34):9968-9.

1357 68. Kouprina N, Larionov V. Transformation-associated recombination (TAR)
1358 cloning and its applications for gene function; genome architecture and evolution;
1359 biotechnology and biomedicine. *Oncotarget.* 2023;14:1009-33.

1360 69. Park HS, Park JH, Kim HJ, Kang SH, Choi SS, Kim ES. BAC cloning and
1361 heterologous expression of a giant biosynthetic gene cluster encoding antifungal
1362 neotetrafibricin in *streptomyces rubrisoli*. *Front Bioeng Biotechnol.* 2022;10:964765.

1363 70. Ke J, Robinson D, Wu ZY, Kuftin A, Louie K, Kosina S, et al. CRAGE-
1364 CRISPR facilitates rapid activation of secondary metabolite biosynthetic gene clusters
1365 in bacteria. *Cell Chem Biol.* 2022;29(4):696-710 e4.

1366 71. Gillespie DE, Brady SF, Bettermann AD, Cianciotto NP, Liles MR, Rondon
1367 MR, et al. Isolation of antibiotics turbomycin a and B from a metagenomic library of
1368 soil microbial DNA. *Appl Environ Microbiol.* 2002;68(9):4301-6.

1369 72. Chang FY, Brady SF. Discovery of indolotryptoline antiproliferative agents by
1370 homology-guided metagenomic screening. *Proc Natl Acad Sci U S A.*
1371 2013;110(7):2478-83.

1372 73. Hover BM, Kim SH, Katz M, Charlop-Powers Z, Owen JG, Ternei MA, et al.
1373 Culture-independent discovery of the malacidins as calcium-dependent antibiotics
1374 with activity against multidrug-resistant Gram-positive pathogens. *Nat Microbiol.*
1375 2018;3(4):415-22.

1376 74. Yamanaka K, Reynolds KA, Kersten RD, Ryan KS, Gonzalez DJ, Nizet V, et
1377 al. Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster
1378 yields the antibiotic taromycin A. *Proc Natl Acad Sci U S A.* 2014;111(5):1957-62.

1379 75. Li S, Wang J, Li X, Yin S, Wang W, Yang K. Genome-wide identification and
1380 evaluation of constitutive promoters in streptomycetes. *Microb Cell Fact.*
1381 2015;14:172.

1382 76. Laureti L, Song L, Huang S, Corre C, Leblond P, Challis GL, et al.
1383 Identification of a bioactive 51-membered macrolide complex by activation of a silent
1384 polyketide synthase in *Streptomyces ambofaciens*. *Proc Natl Acad Sci U S A.*
1385 2011;108(15):6258-63.

1386 77. Zhang MM, Wong FT, Wang Y, Luo S, Lim YH, Heng E, et al. CRISPR-Cas9
1387 strategy for activation of silent *Streptomyces* biosynthetic gene clusters. *Nat Chem*
1388 *Biol.* 2017.

1389 78. Bogdanov A, Salib MN, Chase AB, Hammerlindl H, Muskat MN, Luedtke S,
1390 et al. Small molecule in situ resin capture provides a compound first approach to
1391 natural product discovery. *Nat Commun.* 2024;15(1):5230.

1392 79. Berdy B, Spoering AL, Ling LL, Epstein SS. In situ cultivation of previously
1393 uncultivable microorganisms using the ichip. *Nat Protoc.* 2017;12(10):2232-42.

1394 80. Amann J. Die direkte Zählung der Wasserbakterien mittels des
1395 Ultramikroskops. *Centralblatt für Bakteriologie.* 1911;29:381-4.

1396 81. Roszak DB, Colwell RR. Survival strategies of bacteria in the natural
1397 environment. *Microbiol Rev.* 1987;51(3):365-79.

1398 82. Amann RI, Ludwig W, Schleifer KH. Phylogenetic identification and in situ
1399 detection of individual microbial cells without cultivation. *Microbiological Reviews.*
1400 1995;59(1):143-69.

1401 83. Staley JT, Konopka A. Measurement of in situ activities of nonphotosynthetic
1402 microorganisms in aquatic and terrestrial habitats. *Annu Rev Microbiol.* 1985;39:321-
1403 46.

1404 84. Goh KM, Shahar S, Chan KG, Chong CS, Amran SI, Sani MH, et al. Current
1405 Status and Potential Applications of Underexplored Prokaryotes. *Microorganisms.*
1406 2019;7(10).

1407 85. Pham VH, Kim J. Cultivation of unculturable soil bacteria. *Trends Biotechnol.*
1408 2012;30(9):475-84.

1409 86. Stewart EJ. Growing unculturable bacteria. *J Bacteriol.* 2012;194(16):4151-60.

1410 87. Joseph SJ, Hugenholtz P, Sangwan P, Osborne CA, Janssen PH. Laboratory
1411 cultivation of widespread and previously uncultured soil bacteria. *Appl Environ*
1412 *Microbiol.* 2003;69(12):7210-5.

1413 88. Sait M, Hugenholtz P, Janssen PH. Cultivation of globally distributed soil
1414 bacteria from phylogenetic lineages previously only detected in cultivation-
1415 independent surveys. *Environ Microbiol.* 2002;4(11):654-66.

1416 89. Janssen PH, Yates PS, Grinton BE, Taylor PM, Sait M. Improved culturability
1417 of soil bacteria and isolation in pure culture of novel members of the divisions
1418 Acidobacteria, Actinobacteria, Proteobacteria, and Verrucomicrobia. *Appl Environ*
1419 *Microbiol.* 2002;68(5):2391-6.

1420 90. Mu DS, Liang QY, Wang XM, Lu DC, Shi MJ, Chen GJ, et al.
1421 Metatranscriptomic and comparative genomic insights into resuscitation mechanisms
1422 during enrichment culturing. *Microbiome.* 2018;6(1):230.

1423 91. Keller M, Zengler K. Tapping into microbial diversity. *Nat Rev Microbiol.*
1424 2004;2(2):141-50.

1425 92. D'Onofrio A, Crawford JM, Stewart EJ, Witt K, Gavrish E, Epstein S, et al.
1426 Siderophores from neighboring organisms promote the growth of uncultured bacteria.
1427 *Chem Biol.* 2010;17(3):254-64.

1428 93. Button DK, Schut F, Quang P, Martin R, Robertson BR. Viability and
1429 isolation of marine bacteria by dilution culture: theory, procedures, and initial results.
1430 *Appl Environ Microbiol.* 1993;59(3):881-91.

1431 94. Yang SJ, Kang I, Cho JC. Expansion of Cultured Bacterial Diversity by Large-
1432 Scale Dilution-to-Extinction Culturing from a Single Seawater Sample. *Microb Ecol.*
1433 2016;71(1):29-43.

1434 95. Rappe MS, Connon SA, Vergin KL, Giovannoni SJ. Cultivation of the
1435 ubiquitous SAR11 marine bacterioplankton clade. *Nature.* 2002;418(6898):630-3.

1436 96. Connon SA, Giovannoni SJ. High-throughput methods for culturing
1437 microorganisms in very-low-nutrient media yield diverse new marine isolates. *Appl*
1438 *Environ Microbiol.* 2002;68(8):3878-85.

1439 97. Giovannoni SJ, Foster RA, Rappé MS, Epstein S. New Cultivation Strategies
1440 Bring More Microbial Plankton Species into the Laboratory. *Oceanography.*
1441 2007;20(2):62-9.

1442 98. Ferrari BC, Binnerup SJ, Gillings M. Microcolony cultivation on a soil
1443 substrate membrane system selects for previously uncultured soil bacteria. *Appl*
1444 *Environ Microbiol.* 2005;71(12):8714-20.

1445 99. Ferrari BC, Gillings MR. Cultivation of fastidious bacteria by viability
1446 staining and micromanipulation in a soil substrate membrane system. *Appl Environ*
1447 *Microbiol.* 2009;75(10):3352-4.

1448 100. Ingham CJ, Sprenkels A, Bomer J, Molenaar D, van den Berg A, van
1449 Hylckama Vlieg JE, et al. The micro-Petri dish, a million-well growth chip for the
1450 culture and high-throughput screening of microorganisms. *Proc Natl Acad Sci U S A.*
1451 2007;104(46):18217-22.

1452 101. Kaeberlein T, Lewis K, Epstein SS. Isolating "uncultivable" microorganisms
1453 in pure culture in a simulated natural environment. *Science.* 2002;296(5570):1127-9.

1454 102. Nichols D, Lewis K, Orjala J, Mo S, Ortenberg R, O'Connor P, et al. Short
1455 peptide induces an "uncultivable" microorganism to grow in vitro. *Appl Environ*
1456 *Microbiol.* 2008;74(15):4889-97.

1457 103. Bollmann A, Lewis K, Epstein SS. Incubation of environmental samples in a
1458 diffusion chambers increases the diversity of recovered isolates. *Applied and*
1459 *Environmental Microbiology.* 2007;73(20):6386-90.

1460 104. Bollmann A, Palumbo AV, Lewis K, Epstein SS. Isolation and physiology of
1461 bacteria from contaminated subsurface sediments. *Appl Environ Microbiol.*
1462 2010;76(22):7413-9.

1463 105. Epstein SS. General model of microbial uncultivability. In: *Uncultivated*
1464 *Microorganisms, Microbiology Monographs*, Vol 10. 2009:131-50.

1465 106. Buerger S, Spoering A, Gavrish E, Leslin C, Ling L, Epstein SS. Microbial
1466 scout hypothesis, stochastic exit from dormancy, and the nature of slow growers. *Appl*
1467 *Environ Microbiol.* 2012;78(9):3221-8.

1468 107. Buerger S, Spoering A, Gavrish E, Leslin C, Ling L, Epstein SS. Microbial
1469 scout hypothesis and microbial discovery. *Appl Environ Microbiol.* 2012;78(9):3229-
1470 33.

1471 108. Lewis K. Persister cells, dormancy and infectious disease. *Nat Rev Microbiol.*
1472 2007;5(1):48-56.

1473 109. Gavrish E, Bollmann A, Epstein S, Lewis K. A trap for in situ cultivation of
1474 filamentous actinobacteria. *J Microbiol Methods.* 2008;72(3):257-62.

1475 110. Chaudhary DK, Khulan A, Kim J. Development of a novel cultivation
1476 technique for uncultured soil bacteria. *Sci Rep.* 2019;9(1):6666.

1477 111. Aoi Y, Kinoshita T, Hata T, Ohta H, Obokata H, Tsuneda S. Hollow-fiber
1478 membrane chamber as a device for in situ environmental cultivation. *Applied and*
1479 *Environmental Microbiology.* 2009;75(11):3826-33.

1480 112. Ben-Dov E, Kramarsky-Winter E, Kushmaro A. An in situ method for
1481 cultivating microorganisms using a double encapsulation technique. *FEMS*
1482 *Microbiology Ecology.* 2009;68(3):363-71.

1483 113. Zengler K, Toledo G, Rappe' M, Elkins J, Mathur EJ, Short JM, et al.
1484 Cultivating the uncultured: exploring microbial diversity using ichip technology.
1485 *Proceedings of the National Academy of Sciences of the United States of America.*
1486 2002;99(24):15681-6.

1487 114. Akselband Y, Cabral C, Castor TP, Chikarmane HM, McGrath P. Enrichment
1488 of slow-growing marine microorganisms from mixed cultures using gel microdrop
1489 (GMD) growth assay and fluorescence-activated cell sorting. *Journal of Experimental*
1490 *Marine Biology and Ecology.* 2006;329(2):196-205.

1491 115. Nichols D, Cahoon N, Trakhtenberg EM, Pham L, Mehta A, Belanger A, et al.
1492 Use of ichip for high-throughput in situ cultivation of "uncultivable" microbial
1493 species. *Appl Environ Microbiol.* 2010;76(8):2445-50.

1494 116. Braus MJ, Swader R, Petry G, Whitman T. Affordable Diffusion
1495 Microchamber Array Designs for Isolating Microbes in Classrooms and Laboratories.
1496 *J Microbiol Biol Educ.* 2021;22(2).

1497 117. Alessi AM, Redeker KR, Chong JPJ. A practical introduction to microbial
1498 molecular ecology through the use of isolation chips. *Ecol Evol.* 2018;8(24):12286-
1499 98.

1500 118. Jung D, Liu B, He X, Owen JS, Liu L, Yuan Y, et al. Accessing previously
1501 uncultured marine microbial resources by a combination of alternative cultivation
1502 methods. *Microb Biotechnol.* 2021;14(3):1148-58.

1503 119. Ding L, Xu P, Zhang W, Yuan Y, He X, Su D, et al. Three New
1504 Diketopiperazines from the Previously Uncultivable Marine Bacterium *Gallaecimonas*
1505 *mangrovi* HK-28 Cultivated by iChip. *Chem Biodivers.* 2020;17(7):e2000221.

1506 120. Vitorino IR, Klimek D, Calusinska M, Lobo-da-Cunha A, Vasconcelos V,
1507 Lage OM. *Stieleria sedimenti* sp. nov., a Novel Member of the Family *Pirellulaceae*
1508 with Antimicrobial Activity Isolated in Portugal from Brackish Sediments.
1509 *Microorganisms.* 2022;10(11).

1510 121. Polrot A, Kirby JR, Olorunniji FJ, Birkett JW, Sharples GP. iChip increases
1511 the success of cultivation of TBT-resistant and TBT-degrading bacteria from estuarine
1512 sediment. *World J Microbiol Biotechnol.* 2022;38(10):180.

1513 122. He S, Li P, Wang J, Zhang Y, Lu H, Shi L, et al. Discovery of New Secondary
1514 Metabolites from Marine Bacteria *Hahella* Based on an Omics Strategy. *Mar Drugs*.
1515 2022;20(4).

1516 123. Ding H, Liu J, Yang C, Guo C, Ding L, Jung D, et al. *Hanstruepera marina* sp.
1517 nov. and *Hanstruepera flava* sp. nov., two novel species in the family
1518 Flavobacteriaceae isolated by a modified in situ cultivation technique from marine
1519 sediment. *Front Microbiol*. 2022;13:957397.

1520 124. MacIntyre LW, Charles MJ, Haltli BA, Marchbank DH, Kerr RG. An Ichip-
1521 Domesticated Sponge Bacterium Produces an N-Acyltyrosine Bearing an alpha-
1522 Methyl Substituent. *Org Lett*. 2019;21(19):7768-71.

1523 125. Steinert G, Whitfield S, Taylor MW, Thoms C, Schupp PJ. Application of
1524 diffusion growth chambers for the cultivation of marine sponge-associated bacteria.
1525 *Mar Biotechnol (NY)*. 2014;16(5):594-603.

1526 126. Ihua MW, Guiheneuf F, Mohammed H, Margassery LM, Jackson SA, Stengel
1527 DB, et al. Microbial Population Changes in Decaying *Ascophyllum nodosum* Result
1528 in Macroalgal-Polysaccharide-Degrading Bacteria with Potential Applicability in
1529 Enzyme-Assisted Extraction Technologies. *Mar Drugs*. 2019;17(4).

1530 127. Vitorino I, Santos JDN, Godinho O, Vicente F, Vasconcelos V, Lage OM.
1531 Novel and Conventional Isolation Techniques to Obtain Planctomycetes from Marine
1532 Environments. *Microorganisms*. 2021;9(10).

1533 128. Liu H, Xue R, Wang Y, Stirling E, Ye S, Xu J, et al. FACS-iChip: a high-
1534 efficiency iChip system for microbial 'dark matter' mining. *Mar Life Sci Technol*.
1535 2021;3(2):162-8.

1536 129. Liu X, Wang M, Nie Y, Wu XL. Isolation Chip Increases Culturable Bacterial
1537 Diversity and Reduces Cultivation Bias. *Curr Microbiol*. 2021;78(5):2025-32.

1538 130. Megaw J, Kelly SA, Thompson TP, Skvortsov T, Gilmore BF. Profiling the
1539 microbial community of a Triassic halite deposit in Northern Ireland: an environment
1540 with significant potential for biodiscovery. *FEMS Microbiol Lett*. 2019;366(22).

1541 131. Zhao J, Shakir Y, Deng Y, Zhang Y. Use of modified ichip for the cultivation
1542 of thermo-tolerant microorganisms from the hot spring. *BMC Microbiol*.
1543 2023;23(1):56.

1544 132. Moote PE, Bescucci DM, Ortega Polo R, Uwiera RRE, Inglis GD.
1545 Comparison of Strategies for Isolating Anaerobic Bacteria from the Porcine Intestine.
1546 *Appl Environ Microbiol*. 2021;87(9).

1547 133. Sizova MV, Hohmann T, Hazen A, Paster BJ, Halem SR, Murphy CM, et al.
1548 New approaches for isolation of previously uncultivated oral bacteria. *Appl Environ*
1549 *Microbiol*. 2012;78(1):194-203.

1550 134. Peoples AJ, Zhang Q, Millett WP, Rothfeder MT, Pescatore BC, Madden AA,
1551 et al. Neocitreomicins I and II, novel antibiotics with activity against methicillin-
1552 resistant *Staphylococcus aureus* and vancomycin-resistant *Enterococci*. *J Antibiot*
1553 (Tokyo). 2008;61(7):457-63.

1554 135. Gavrish E, Sit CS, Cao S, Kandror O, Spoering A, Peoples A, et al.
1555 Lassomycin, a ribosomally synthesized cyclic peptide, kills mycobacterium
1556 tuberculosis by targeting the ATP-dependent protease ClpC1P1P2. *Chem Biol*.
1557 2014;21(4):509-18.

1558 136. Quigley J, Peoples A, Sarybaeva A, Hughes D, Ghiglieri M, Achorn C, et al.
1559 Novel Antimicrobials from Uncultured Bacteria Acting against *Mycobacterium*
1560 tuberculosis. *mBio*. 2020;11(4).

1561 137. Ling LL, Schneider T, Peoples AJ, Spoering AL, Engels I, Conlon BP, et al. A
1562 new antibiotic kills pathogens without detectable resistance. *Nature*.
1563 2015;517(7535):455-9.

1564 138. Shukla R, Medeiros-Silva J, Parmar A, Vermeulen BJA, Das S, Paioni AL, et
1565 al. Mode of action of teixobactins in cellular membranes. *Nat Commun*.
1566 2020;11(1):2848.

1567 139. Shukla R, Lavore F, Maity S, Derks MGN, Jones CR, Vermeulen BJA, et al.
1568 Teixobactin kills bacteria by a two-pronged attack on the cell envelope. *Nature*.
1569 2022;608(7922):390-6.

1570 140. Wirtz DA, Ludwig KC, Arts M, Marx CE, Krannich S, Barac P, et al.
1571 Biosynthesis and Mechanism of Action of the Cell Wall Targeting Antibiotic
1572 Hypeptin. *Angew Chem Int Ed Engl*. 2021;60(24):13579-86.

1573 141. Shukla R, Peoples AJ, Ludwig KC, Maity S, Derks MGN, De Benedetti S, et
1574 al. An antibiotic from an uncultured bacterium binds to an immutable target. *Cell*.
1575 2023;186(19):4059-73 e27.

1576 142. Piddock LJ. Teixobactin, the first of a new class of antibiotics discovered by
1577 iChip technology? *J Antimicrob Chemother*. 2015;70(10):2679-80.

1578 143. Sherpa RT, Reese CJ, Montazeri Aliabadi H. Application of iChip to Grow
1579 "Uncultivable" Microorganisms and its Impact on Antibiotic Discovery. *J Pharm*
1580 *Pharm Sci*. 2015;18(3):303-15.

1581 144. Lewis K, Epstein S, D'Onofrio A, Ling LL. Uncultured microorganisms as a
1582 source of secondary metabolites. *J Antibiot (Tokyo)*. 2010;63(8):468-76.

1583 145. Nakai R, Kusada H, Sassa F, Makino A, Morigasaki S, Hayashi H, et al.
1584 *Roseiterribacter gracilis* gen. nov., sp. nov., a novel filterable alphaproteobacterium
1585 isolated from soil using a gel-filled microwell array device. *PLoS One*.
1586 2024;19(6):e0304366.

1587 146. Kaminski TS, Scheler O, Garstecki P. Droplet microfluidics for microbiology:
1588 techniques, applications and challenges. *Lab Chip*. 2016;16(12):2168-87.

1589 147. Long F, Guo Y, Zhang Z, Wang J, Ren Y, Cheng Y, et al. Recent Progress of
1590 Droplet Microfluidic Emulsification Based Synthesis of Functional Microparticles.
1591 *Glob Chall*. 2023;7(9):2300063.

1592 148. Ning R, Fan J, Kong L, Jiang X, Qian Y, Du T, et al. Recent developments of
1593 droplets-based microfluidics for bacterial analysis. *Chinese Chemical Letters*.
1594 2022;33(5):2243-52.

1595 149. Manome A, Zhang H, Tani Y, Katsuragi T, Kurane R, Tsuchida T.
1596 Application of gel microdroplet and flow cytometry techniques to selective
1597 enrichment of non-growing bacterial cells. *FEMS Microbiol Lett*. 2001;197(1):29-33.

1598 150. Ohan J, Pelle B, Nath P, Huang JH, Hovde B, Vuyisich M, et al. High-
1599 throughput phenotyping of cell-to-cell interactions in gel microdroplet pico-cultures.
1600 *Biotechniques*. 2019;66(5):218-24.

1601 151. Watterson WJ, Tanyeri M, Watson AR, Cham CM, Shan Y, Chang EB, et al.
1602 Droplet-based high-throughput cultivation for accurate screening of antibiotic
1603 resistant gut microbes. *Elife*. 2020;9.

1604 152. Park J, Kerner A, Burns MA, Lin XN. Microdroplet-enabled highly parallel
1605 co-cultivation of microbial communities. *PLoS One*. 2011;6(2):e17019.

1606 153. Mahler L, Niehs SP, Martin K, Weber T, Scherlach K, Hertweck C, et al.
1607 Highly parallelized droplet cultivation and prioritization of antibiotic producers from
1608 natural microbial communities. *Elife*. 2021;10:e64774.

1609 154. Tan JY, Wang S, Dick GJ, Young VB, Sherman DH, Burns MA, et al. Co-
1610 cultivation of microbial sub-communities in microfluidic droplets facilitates high-

1611 resolution genomic dissection of microbial 'dark matter'. *Integr Biol (Camb)*.
 1612 2020;12(11):263-74.
 1613 155. Liu W, Kim HJ, Lucchetta EM, Du W, Ismagilov RF. Isolation, incubation,
 1614 and parallel functional testing and identification by FISH of rare microbial single-
 1615 copy cells from multi-species mixtures using the combination of chemistode and
 1616 stochastic confinement. *Lab Chip*. 2009;9(15):2153-62.
 1617 156. Bai Y, Patil SN, Bowden SD, Poulter S, Pan J, Salmond GP, et al. Intra-
 1618 species bacterial quorum sensing studied at single cell level in a double droplet
 1619 trapping system. *Int J Mol Sci*. 2013;14(5):10570-81.
 1620 157. Weibel DB, Whitesides GM. Applications of microfluidics in chemical
 1621 biology. *Curr Opin Chem Biol*. 2006;10(6):584-91.
 1622 158. Mazutis L, Gilbert J, Ung WL, Weitz DA, Griffiths AD, Heyman JA. Single-
 1623 cell analysis and sorting using droplet-based microfluidics. *Nat Protoc*.
 1624 2013;8(5):870-91.
 1625 159. Droplet-based microfluidics. *Nature Reviews Methods Primers*. 2023;3(1):33.
 1626 160. The S-Y, Lin R, Hung L-H, Lee AP. Droplet microfluidics. *Lab on a Chip*.
 1627 2008;8:196-220.
 1628 161. Baret J-C. Surfactants in droplet-based microfluidics. *Lab on a Chip*.
 1629 2012;12(3):422-33.
 1630 162. Payne EM, Holland-Moritz DA, Sun S, Kennedy RT. High throughput
 1631 screening in droplet microfluidics: perspective into key challenges and future
 1632 prospect. *Lab on a Chip*. 2020;20:2247-62.
 1633 163. Zhu Z, Yang CJ. Hydrogel Droplet Microfluidics for High-Throughput Single
 1634 Molecule/Cell Analysis. *Acc Chem Res*. 2017;50(1):22-31.
 1635 164. Srisa-Art M, DeMello AJ, Edel JB. High throughput DNA droplet assays
 1636 using picolitre reaction volumes. *Analytical Chemistry*. 2007;79(17):6682-9.
 1637 165. Josephides D, Davoli S, Whitley W, Ruis R, Salter R, Gokkaya S, et al. Cyto-
 1638 Mine: An Integrated, Picodroplet System for High-Throughput Single-Cell Analysis,
 1639 Sorting, Dispensing, and Monoclonality Assurance. *SLAS Technol*. 2020;25(2):177-
 1640 89.
 1641 166. Du L, Liu H, Zhou J. Picoliter droplet array based on bioinspired microholes
 1642 for in situ single-cell analysis. *Microsyst Nanoeng*. 2020;6:33.
 1643 167. Teng P, Tian D, Fu H, Wang S. Recent progress of electrowetting for droplet
 1644 manipulation: from wetting to superwetting systems. *Materials Chemistry Frontiers*.
 1645 2020;4(1):140-54.
 1646 168. Wang BL, Ghaderi A, Zhou H, Agresti J, Weitz DA, Fink GR, et al.
 1647 Microfluidic high-throughput culturing of single cells for selection based on
 1648 extracellular metabolite production or consumption. *Nat Biotechnol*. 2014;32(5):473-
 1649 8.
 1650 169. Najah M, Calbrix R, Mahendra-Wijaya IP, Beneyton T, Griffiths AD, Drevelle
 1651 A. Droplet-based microfluidics platform for ultra-high-throughput bioprospecting of
 1652 cellulolytic microorganisms. *Chem Biol*. 2014;21(12):1722-32.
 1653 170. Terekhov SS, Smirnov IV, Stepanova AV, Bobik TV, Mokrushina YA,
 1654 Ponomarenko NA, et al. Microfluidic droplet platform for ultrahigh-throughput
 1655 single-cell screening of biodiversity. *Proc Natl Acad Sci U S A*. 2017;114(10):2550-5.
 1656 171. Oberpaul M, Brinkmann S, Marner M, Mihajlovic S, Leis B, Patras MA, et al.
 1657 Combination of high-throughput microfluidics and FACS technologies to leverage the
 1658 numbers game in natural product discovery. *Microb Biotechnol*. 2022;15(2):415-30.

172. Theodorou E, Scanga R, Twardowski M, Snyder MP, Brouzes E. A Droplet Microfluidics Based Platform for Mining Metagenomic Libraries for Natural Compounds. *Micromachines (Basel)*. 2017;8(8).
173. Mahler L, Tovar M, Weber T, Brandes S, Rudolph MM, Ehgartner J, et al. Enhanced and homogeneous oxygen availability during incubation of microfluidic droplets. *RSC Advances*. 2015;5:101871-8.
174. Xu P, Modavi C, Demaree B, Twigg F, Liang B, Sun C, et al. Microfluidic automated plasmid library enrichment for biosynthetic gene cluster discovery. *Nucleic Acids Res*. 2020;48(8):e48.
175. Murray BE, Penabad LI, Kennedy RT. Advances in coupling droplet microfluidics to mass spectrometry. *Curr Opin Biotechnol*. 2023;82:102962.
176. Gach PC, Iwai K, Kim PW, Hillson NJ, Singh AK. Droplet microfluidics for synthetic biology. *Lab Chip*. 2017;17(20):3388-400.
177. Yun K, Zhang Y, Li S, Wang Y, Tu R, Liu H, et al. Droplet-Microfluidic-Based Promoter Engineering and Expression Fine-Tuning for Improved Erythromycin Production in *Saccharopolyspora erythraea* NRRL 23338. *Front Bioeng Biotechnol*. 2022;10:864977.
178. Iwai K, Wehrs M, Garber M, Sustarich J, Washburn L, Costello Z, et al. Scalable and automated CRISPR-based strain engineering using droplet microfluidics. *Microsyst Nanoeng*. 2022;8:31.
179. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, et al. Programming cells by multiplex genome engineering and accelerated evolution. *Nature*. 2009;460(7257):894-8.
180. Wippold JA, Chu M, Renberg R, Li Y, Adams B, Han A. XPORT ENTRAP: A droplet microfluidic platform for enhanced DNA transfer between microbial species. *N Biotechnol*. 2024;81:10-9.
181. Brophy JAN, Triassi AJ, Adams BL, Renberg RL, Stratis-Cullum DN, Grossman AD, et al. Engineered integrative and conjugative elements for efficient and inducible DNA transfer to undomesticated bacteria. *Nat Microbiol*. 2018;3(9):1043-53.
182. Chiu FWY, Stavrakis S. High-throughput droplet-based microfluidics for directed evolution of enzymes. *Electrophoresis*. 2019;40(21):2860-72.
183. Manteca A, Gadea A, Van Assche D, Cossard P, Gillard-Bocquet M, Beneyton T, et al. Directed Evolution in Drops: Molecular Aspects and Applications. *ACS Synth Biol*. 2021;10(11):2772-83.
184. Gantz M, Neun S, Medcalf EJ, van Vliet LD, Hollfelder F. Ultrahigh-Throughput Enzyme Engineering and Discovery in In Vitro Compartments. *Chem Rev*. 2023;123(9):5571-611.
185. Rosenthal RG, Diana Zhang X, Durdic KI, Collins JJ, Weitz DA. Controlled Continuous Evolution of Enzymatic Activity Screened at Ultrahigh Throughput Using Drop-Based Microfluidics. *Angew Chem Int Ed Engl*. 2023;62(24):e202303112.
186. Zhou J, Dong J, Hou H, Huang L, Li J. High-throughput microfluidic systems accelerated by artificial intelligence for biomedical applications. *Lab Chip*. 2024;24(5):1307-26.