Mining potential drugs and natural products from microbial genomes via improved Red/ET recombineering

Dissertation
Zur Erlangung des Grades
Des Doktors der Naturwissenschaften
Der Naturwissenschaftlich-Technischen Fakultät
Der Universität des Saarlandes

von
Qiang Tu
Saarbrücken
2016
Tag des Kolloquiums: Freitag, 10. März 2017

Dekan: Prof. Dr. rer. nat. Guido Kickelbick

Berichterstatter: Prof. Dr. R. Müller

Prof. Dr. A. Luzhetskyy

Vorsitz: Prof. Dr. R. W. Hartmann

Akad. Mitarbeiter: Dr. Y. Rebets
Acknowledgements

First and foremost I would like to give my deepest gratitude to my scientific supervisor Prof. Dr. Rolf Müller for all of the precious guidance and the opportunity to work in this remarkable research environment in the laboratory of Department of Pharmaceutical Biotechnology at Saarland University. His inspirational support and encourage during these years are invaluable, such that I have learned so much from his scientific ability and attitude from which I will benefit for my whole research career in my future career.

I also thank Prof. Dr. Jörn Walter for his kind suggestions during my Ph.D. studies.

I owe my sincere gratitude to my co-advisor Prof. Dr. Youming Zhang from State Key Laboratory of Microbial Technology at Shandong University in China, who practically co-advised me to go through the tough research process, especially at the beginning of the Ph.D. study. He taught me many principles and techniques of molecular biology. He provided me very valuable advices and guidance for my research and career. Furthermore, I will also thank Prof. Dr. Xiaoying Bian in Shandong University for his constructive comments and friendly help. In addition, I also would like to show my gratitude to Dr. Youhua Deng for her genuine help during my stay in Germany.

I would like to thank Prof. Dr. A Francis Stewart, Dr. Jun Fu and Dr. Jia Yin from BioInnovationsZentrum, Technische Universität Dresden. Over the past years I had very enjoyable experience to work with them in several collaborated projects. Their ambitions for technology development always inspire me. They are very generous to provide recombineering reagents and advices.

At the same time I am thankful to Dr. Jennifer Herrmann from Helmholtz Institute for
Acknowledgements

Pharmaceutical Research Saarland for bioactivity test of disorazol and writing the manuscript together.

I am grateful to Dr. Stephan Hüttel and Dr. Ritesh Raju in our group who directed me in the isolation, purification and structural elucidation of natural products. Their rich experience in NMR measurement is also very helpful.

I wish to thank all the lab members for their critical and open-minded scientific discussion during the last five years and their contributions to the good working atmosphere. I thank Dr. Silke C. Wenzel for her extremely useful suggestions of my research, Thomas Hoffmann, Eva Luxenburger for LC-MS measurement, and Liujie Huo for protein expression and translation, Dr. Daniel Krug for LC/MS analysis and computer assistance. The administrative assistance of Birgitta Lelarge, Natja Mellendorf, Claudia Thiele and Ellen Merckel is very greatful. I also thank my colleagues Dr. Wei Ding, Dr. Shengbiao Hu, Dr. Yixin Ou, Dr. Shiping Shan, Dr. Chengzhang Fu, Dr. Ronald Garcia, Dr. Hilda Sucipto, Fu Yan, Ying Tang, Srikanth Duddela and Viktoria Schmitt for all their support in various ways as both scientific collaborators and friends.

Last but not least, heartfelt thanks to my parents for their moral support all the years.

Qiang Tu

February, 20th, 2016
Vorveröffentlichen der Dissertation

Teile dieser Arbeit wurden vorab mit Genehmigung der Naturwissenschaftlich-Technischen Fakultät III, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen veröffentlicht oder sind derzeit in Vorbereitung zur Veröffentlichung:

**Publikationen:**


Jia Yin, Michael Hoffmann, Xiaoying Bian, **Qiang Tu**, Fu Yan, Liqiu Xia, Xuezhi Ding, A. Francis Stewart, Rolf Müller, Jun Fu and Youming Zhang. Direct cloning and heterologous expression of the salinomycin biosynthetic gene cluster from *Streptomyces albus* DSM41398 in *S. coelicolor* A3 (2). *Scientific Reports*. 2015, 5: 15081.

(*: Authors contributed equally to the work)

**Patents**

**Qiang Tu**, Jun Fu, Francis A. Stewart, Rolf Mueller and Youming Zhang.

A new method for DNA transformation.


International patent application No: PCT/IB2013/071546 (2013)
Conference Contributions


Abstract

A great deal of natural product biosynthetic pathways have been detected following the development of bacterial genome-sequencing projects. Reconstruction, genetic modification and heterologous expression of the biosynthetic gene clusters provide an effective approach to discover novel natural product derivatives, while also improving productivities and yields, which sets the stage to expand understanding of their structural diversity and biological activity.

The present thesis copes with several natural product biosynthetic gene clusters of large size (>50kb) and severe complexity with high GC content and repetitive sequences. Typically, it is more challenging to engineer such colossal and complex gene clusters using traditional approaches. Conversely, Red/ET recombineering, without the size and site limitation of DNA engineering established in *E. coli*, appears to enrich the toolbox of molecular biology and break the size and site limitations of DNA manipulation. Here, we have improved the Red/ET recombination method by changing the long-standing protocol of electrocompetent cell preparation from cold conditions to room temperature, such that the DNA transformation efficiency has been greatly increased. This interesting finding should facilitate the cloning of large fragments from genomic DNA preparations and metagenomic samples.

To exemplify this improvement, the core disorazol biosynthetic pathway (~58kb) from *Sorangium cellulosum* So ce12 was cloned by improved Red/ET recombineering and heterologously expressed in *Myxococcus xanthus* DK1622, resulting in an appreciable increase of disorazol production by promoter exchange.

The full-length salinomycin gene cluster (106kb) from *Streptomyces albus* DSM41398 was isolated from the genome by direct cloning and stitching, and this gene cluster was transferred into the heterologous host *S. coelicolor* A3 (2).

As a third example of the efficacy of this technique, the magnetosomes gene clusters within a large genomic magnetosome island (~115kb) from magnetotactic bacteria *Magnetospirillum gryphiswaldense* was cloned and expressed in *Rhodospirillum*
Abstract

rubrum. A visible red spot near the pole of a permanent magnet could be observed at
the edge of a culture flask of heterologous mutants, signifying that the magnetosome
products had been successfully expressed in the pink-colored heterologous host.
Zusammenfassung


Der Disorazol Biosyntheseweg (disA-D, ~58kb) von Sorangium cellulosum So ce12 wurde durch diese verbesserte Red/ET Recombineering Methode kloniert und in Myxococcus xanthus DK1622 exprimiert. Durch Promoter Austausch war die Erhöhung der Disorazol Produktionausbeute möglich.

Das gesamte Salinomycin-Gencluster (salO-orf18, 106kb) aus Streptomyces albus DSM41398 wurde durch direct-cloning ebenfalls kloniert. Das Gencluster wurde in den heterologen Wirt S. coelicolor A3 (2) übertragen.

Table of contents

Acknowledgements..............................................................................................................I
Abstract................................................................................................................................II
Zusammenfassung..................................................................................................................III
Table of contents...................................................................................................................IV

A. Introduction........................................................................................................................1
   1. Natural products in drug discovery.................................................................1
   2. Biosynthetic logics of microbial natural products.................................................2
      Polyketide natural products.............................................................................7
      Nonribosomal peptide natural products..........................................................10
      PKS/NRPS hybrid systems..........................................................................10
      Magnetic particle production in microorganisms.......................................11
   3. Red/ET Recombineering.......................................................................................12
      Development of recombineering.................................................................14
      Application of recombineering.................................................................16
      Improvement of recombineering...............................................................19
   4. Heterologous expression of natural products in surrogate hosts......................20
   5. Outline of the dissertation....................................................................................27

B. Publications.....................................................................................................................30
   I. Genetic engineering and heterologous expression of the disorazol biosynthetic
      gene cluster via Red/ET recombineering.
   II. Room temperature electrocompetent bacterial cells improve DNA
      transformation and recombineering efficiency.
   III. Biosynthesis of magnetic nanostructures in a foreign organism by transfer of
      bacterial magnetosome gene clusters.
   IV. Direct cloning and heterologous expression of the salinomycin biosynthetic
      gene cluster from Streptomyces albus DSM41398 in S. coelicolor A3 (2).
Table of contents

C. Final discussion........................................................................................................31
   1. Development of a novel improved Red/ET recombination method........31
   2. Several applications in mining microbial genomes via Red/ET technology,
      including disorazol, salinomycin and magnetosome product..............33
   3. New molecular technologies are required for construction and modification of
      large gene cluster in microbial genomes..............................................38

D. References..............................................................................................................40

E. Appendix.................................................................................................................55
A. Introduction

1. Natural products in drug discovery

Nature has been an immense source of natural products for millennia, with various useful compounds developed from plants and animals.\textsuperscript{1-3} These molecules have played a considerable and meaningful feature for treating and inhibiting human diseases and agricultural pests across the world over a long time.\textsuperscript{4,5} Following the discovery of penicillin by A. Fleming, H. W. Florey in 1928,\textsuperscript{6} pharmaceutical research began to focus on microbes, triggering a subsequent “golden era” in massive screening efforts to identify a substantial amount of new microbial antibiotics (1945-1960). Over 23,000 natural products have been characterized, many of which are still used in clinic to date.\textsuperscript{7,8} Currently, approximately 40% of pharmaceuticals in clinical use are either natural products or their derivatives.\textsuperscript{9} Moreover, they are broadly applied not only in clinics, as antibiotics, lipid control agents, immunosuppressive agents, or as compounds exhibiting antitumor or antimalarial properties, but also in agriculture and veterinary applications, including pesticide, insecticides, herbicides, miticides, feed additives, antiparasitic agents and so on (Figure 1-1).\textsuperscript{10-14}

Natural products from microorganisms are believed to be a dominant source in searching for drug candidates, in comparison with other natural compound producers, such as plants or insects.\textsuperscript{15} This perspective is substantiated by the impression that microbial natural products describe the blueprint of an evolutionary optimization process comprising selection for a specific biological activity.\textsuperscript{16} However, the diversity of microorganisms is enormous and only a very minor portion (as low as 1%) of bacteria and fungi has been described so far.\textsuperscript{11} Thus, the identification of novel bacterial secondary metabolite producers is expected to result in the discovery of novel natural products with interesting biological activities in the near future. Moreover, the increasing insights of microbial secondary metabolite biosynthesis and regulation coupled with progresses in molecular genetics drive us to focus on the
discovery and the optimization of natural products in the post-genomic era.

Figure 1-1. Selected examples of natural products derived from microbes used in the clinic/clinical trials, agriculture, and veterinary applications. The structure, name, original producer and type of application are given.

2. Biosynthetic logics of microbial natural products

A wide array of natural products with useful bioactivities and diverse structures are produced by microorganisms via various complex biosynthetic pathways.\(^\text{17}\) So far, several important major groups of natural products have already been established by multimodular megasynthase systems such as polyketide synthases (PKS), non-ribosomal peptide synthetases (NRPS), or hybrids thereof.\(^\text{18-20}\) Additionally, there are also multitudinous bacterial natural products which are derived from other different metabolic routes, such as shikimate or isoprenoid pathway.\(^\text{21-23}\)

A great deal of microbial natural products are biosynthesized from simple monomeric building blocks (e.g. amino acids or short chain carboxylic acids) into enormous multifunctional enzymes via PKSs and NRPSs, respectively (Figure 1-2).\(^\text{24,25}\) In PKS
systems, the vast majority of monomer units incorporated during chain elongation are malonyl-CoA and methylmalonyl-CoA. In some cases, the acyltransferase (AT) domain directly selects the thioesters of monoacyl groups such as acetyl-, propionyl-, and benzoyl-CoAs, or structural variants, such as malonamyl-CoA or methoxymalonyl-CoA to be the start units (Figure 1-3A).\(^2\) Meanwhile, in NRPS systems, as well as the typical 20 proteinogenic amino acids, a broad variety of nonproteinogenic amino acids such as aryl acids are utilized as monomer building blocks for oligomerization and diversification during chain elongation and termination (Figure 1-3B).\(^2\) Recent studies exhibit that NRP chains can be modified with both C-capped and N-capped functionalities, such as C-capped amines as in bleomycin,\(^2\) or N-capped acyl groups such as the long-chain β-OH fatty acid in daptomycin.\(^2\) Although PKS and NRPS systems employ different building blocks, they manifest remarkable similarities in the modular architecture of various catalytic domains and assembly-like mechanisms.

![Figure 1-2. Core PKS and NRPS domains.](image)

Schematic abbreviations used throughout the whole text are shown\(^3\).

In enzymatic machinery, PKS possesses catalytic core domains along with a carrier protein that not only organizes the minimal machinery for chain assembly but also can be post-assembled with additional catalytic domains (Figure 1-3C).\(^1\) In principle, a typical PKS module includes three essential domains: (i) the C–C bond-forming ketosynthase (KS) domain with 45kDa, which performs a decarboxylative claisen thioester condensation of the extender unit (generally malonyl-CoA or methylmalonyl-CoA) with the acyl thioester intermediate from the upstream module,
(ii) the 50kDa AT domain for extender unit selection and loading, and (iii) the acyl carrier protein (ACP) domain, also known as a thiolation (T) domain with around 8-10kDa, where the acyl chain is assembled and elongated (Figure 1-3C). The typical order of domains is KS-AT-ACP. Additionally, several optional domains could be also involved to modify the polyketide intermediate, such as ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER) domains.\textsuperscript{19,31,32} This process is usually completed with a thioesterase (TE) domain that releases and controls the form of the biosynthetic intermediate, such as in linear, cyclic or branched cyclic forms (Figure 1-2).\textsuperscript{33}

Similarly, in an NRPS elongation minimal module, the basic machinery also consists of three necessary domains: (i) a 50kDa condensation (C) domain that catalyzes the peptide bond formation, (ii) a 50kDa adenylation (A) domain responsible for amino acid recognition and activation, and also (iii) a 8-10kDa peptidyl carrier protein (PCP) or thiolation (T) domain to which the activated amino acid is covalently attached as thioester (Figure 1-3D).\textsuperscript{19} The typical order of domains is C-A-PCP. The specificity of A domains towards corresponding amino acids act crucially in the peptide sequence of the final natural product and the biological activity as well. Nevertheless, the optional domains for modification in NRPS cover epimerase (E), methyltransferase (MT), heterocyclization (HC) and oxidase (Ox) domains.\textsuperscript{19} In the same fashion as the PKS system, the TE domain turns out to be the last biosynthetic step in most cases of NRPS by catalyzing the hydrolysis or cyclization of nascent products to constitute linear or cyclic peptides (Figure 1-2).
A. Introduction

Figure 1-3. Use of thioester monomers and core domains of PKS and NRPS system. (A) Common acyl-CoA thioesters used for chain initiation and elongation by PKSs. (B) Shared amino and aryl acid monomers used for chain initiation and elongation by NRPSs. (C) Core domains of PKS system. A minimal module consists of a KS domain (45 kDa), an AT domain (50 kDa) and an ACP domain (8-10 kDa). (D) Core domains of NRPS system. A minimal module consists of a C domain (50 kDa), an A domain (50 kDa) and a PCP domain (8-10 kDa).

In summary, the core domains of PKS (KS, AT, CP) possess similar counterparts in NRPS (C, A, PCP) which catalyze three main steps as explained previously, covering two catalytic domains and one thiolation domain. The possible combinations of these various domains present thousands of different PKS and NRPS biosynthetic gene clusters in the bacterial chromosome, offering a rich and valuable source of natural products.\(^3^4,3^5\) According to the current sequenced genomes of abundant bacteria, it should be noted that PKS and NRPS clusters are infrequent in bacteria with genomes less than 3 Mb. Conversely, in the bacteria with genomes more than 5 Mb, a linear correlation between the genome size and the number of PKS and NRPS gene clusters is finally discovered.\(^3^0\

The biosynthetic machinery of PKS and NRPS systems can be explained in four strategies: (1) A serine side chain in the carrier protein facilitates the pyrophosphate linkage of coenzyme A. In this case, an important post-translational modification by a phosphopantetheinyl transferase (PPTase) is required for both PKS and NRPS, converting the enzyme from the inactive apo- to the active holo-form (Figure 1-4).\(^3^6,3^7\) (2) Monomers are supplied into the assembly lines to generate elongated intermediates, and in PKS assembly lines, benzoyl-, butyryl-, cyclohexyl or propionyl-CoAs are employed as starter units\(^2^6\) and malonyl-CoA, methylmalonyl-CoA, alkyl-, hydroxyl or aminomalonyl-CoAs are utilized as the chain extender blocks.\(^3^8\) In contrast, several hundred nonproteinogenic amino acids participate in the biosynthesis of NRP microbial metabolites.\(^3^9\) (3) Chain initiation (including monomer selection, loading and acylation), chain elongation (including claisen, amide and ester condensations) and chain termination are carried out subsequently during the assembly (see detailed description below). (4) Post
assembly-line tailoring of nascent released products (Figure 1-5).

In addition, chain elongation in the step 3 includes three basic reactions:

(i) Substrate recognition, involving recognition of acyl-CoA thioesters by PKS and aminoacyl-adenylate by NRPS;

(ii) Covalent binding, e.g. thioester bound to a carrier protein, whether ACP in PKS or PCP in NRPS;

(iii) Condensation with the acyl- or peptidyl- residues from the upstream module.

Each of these enzymatic modification steps are performed on individual catalytic domains within the module (Figure 1-5).

Figure 1-4. Post-translational activation of carrier proteins (CPs) by a phosphopantetheinyl transferase (PPTase).
A. Introduction

Figure 1-5. Schematic exhibitions of PKS (up) and NRPS (down) biochemistry. Domains involved in the particular reaction steps are highlighted in grey.

2.1 Polyketide natural products

Polyketides are a wide group of natural products occurring in bacteria, fungi, animals and plants, which represent numerous clinical compounds such as macrolide antibiotic erythromycin, polyeone immunosuppressant rapamycin, antibiotic tetracycline, anti-cholesterol lovastatin, polyeone antifungal amphotericin B and mycotoxin carcinogen aflatoxin B1. All of them are biosynthesized from acyl-CoA precursors by PKS enzymes. Much of the research in polyketide biosynthesis in last two decades has focused on biochemistry and genetic studies characterizing that three types of bacterial PKSs. Type I PKSs are large multifunctional enzymes in which both catalytic domains and carrier ACP domains are linked in cis to constitute modules that are joined together to form a multimodular assembly line (Figure 1-5 up). Type II PKSs are multienzyme complexes that carry a single set of iteratively acting activities. Contrary to type I PKSs, the catalytic domains and the carrier ACP domains interact in trans to achieve acyl- chain elongations in type II PKSs. Lastly, type III PKSs can be distinguished from type I and type II PKSs by this groups preference to utilize malonyl-CoA rather than malonyl-S-pantetheinyl-T species as substrates, and
A. Introduction

thus type III PKS do not possess ACP domains. Type I and type II PKSs utilize ACP domain to activate the acyl-CoA substrates and to channel the growing polyketide intermediates, whereas type III PKSs, independent of ACP, act directly on the acyl-CoA substrates.

Type I PKSs are further subdivided into two groups: modular type I PKSs and iterative type I PKSs (iPKSs). iPKSs have nearly been discovered in fungi with reusing domains in a cyclic style. In iPKSs, the KS domain repetitively modifies the growing polyketide chain after each round of condensation and terminates polyketide biosynthesis at a mathematical perform. Interestingly, the same catalytic domains for multiple rounds of chain extension are repetitively utilized in both iPKSs, type II PKSs and type III PKSs systems.

Type II PKSs are composed of partite enzymes harboring the minimal PKS module as well as a set of tailoring enzymes. Each minimal PKS module consists of two KS domains (KS_α and KS_β) and an ACP domain to catalyze the individual reactions of chain assembly and modification. The subunit KS_β has also been known as “chain length factor”. Additionally, post assembly of the aromatic polyketides in type II PKSs systems require several specialized enzymes which exist in the genetic sequencing scaffold, such as aromatase (ARO), cyclase (CYC) and ketoreductase (KR).

Type III PKSs catalyze the condensation of one to several molecules of extender substrate onto a starter substrate, and these are mostly found in plants, including stilbene synthases and naringenin-chalcone synthases. However, a lot of studies of have already revealed the fact that type III PKSs can also exist in bacteria, such as quinolone alkaloid aurachins from the myxobacterium Stigmatella aurantiaca.

It is well known that polyketide elongation depends on at least three enzymatic domains: a KS, an AT and an ACP. Of late the conventional model of complex polyketide biosynthesis was broadened by discovery a special type of modular system named trans-AT PKS (Figure 1-6). The first report of this trans-AT PKS system was from pederin that was produced by an uncultured bacterial symbiont of Paederus
Several multimodular systems containing PKS modules lacking proper AT domains were discovered not only in actinomycetes, but also in a wide range of bacteria group. Previous studies of trans-AT phylogeny had emphasized these PKSs undergo a remarkable capacity to develop by horizontal gene transfer, allowing substantial recombination of PKS gene fragments. Furthermore, with the continuous detections of an increasing number of trans-AT PKS-derived polyketides demonstrating significant bioactivities, and a great deal of research had focused on a detailed understanding of trans-AT PKS function, including mechanistic and structural insights. Recent studies indicated that trans-AT PKSs possessed several distinct domains only existed in these systems. Some domains were proved to be useless, skipped or even iterative practice. Moreover, another visible characteristic of trans-AT PKS was named β-branching which present a variety of complex modifications in the β-keto position. Nowadays, with increasing discovery of trans-AT PKS-derived compounds from varying microorganisms, more details and specifics in this exceptional system would be understood and disclosed in the near future.

Figure 1-6. Domain organization of cis-AT (A) and trans-AT (B) PKS. (A) Protein...
domains are connected in cis. (B) Protein domains interact in trans. The figure was reproduced from reference 19.

2.2 Nonribosomal peptide natural products

NRPSs are linear, cyclic or branched peptides with no more than 20 amino residues after various modifications, such as acylation, epimerization, glycosylation, heterocyclization or N-methylation of the amide nitrogen. Hence, the monomers for NRPS assembly are not only proteinogenic but also nonproteinogenic amino acids as well as other carboxylic acids. There are also many well-known NRP antibiotics including penicillin G, cephamecin C and vancomycin that have already been developed to clinically used drugs.

According to their sphere of activity, the nonribosomal peptides have been separated into several classifications: (1) Antibiotics, such as vancomycin, which is antibiotic of critical importance used when typical antibiotics are found to be ineffective. (2) Antibiotic precursors, such as ACV-tripeptide, which is a precursor of the penicillin and cephalosporine families. (3) Cytostatics. (4) Immunosuppressive agents. (5) Nitrogen storage polymers. (6) Phytotoxins, such as victorin, a chlorinated cyclic pentapeptide made by the pathogenic fungus Cochliobolus victoriae. (7) Pigments, such as indigoidine. (8) Siderophores, such as Enterobactin and Vibriobactin are iron-chelating compounds. (9) Toxins, such as cyanotoxins from Cyanobacteria.

2.3 PKS/NRPS hybrid systems

PKS and NRPS possess such similar catalytic activity and structural organization, that there could therefore be significant potential in the cooperation of these two types of multimodular enzymes, representing an enormous variety of potential structures. For instance, functional interactions between PKS and NRPS modules mean that a PKS-bound acyl intermediate could be directly transferred and subsequently elongated by a NRPS module or a NRPS-bound peptidyl intermediate could be directly transferred and elongated further by a PKS module. Up until to now, a great number of natural products have been expressed following such a hybrid
A. Introduction

In this case, the intermodular communication is essential for the functionality of the hybrid systems. Although PKS and NRPS have distinct domains to yield metabolic compounds, they share a similar modular organization, and both systems use carrier proteins, such as ACP for PKS and PCP for NRPS to manage the growing chain. In PKS/NRPS hybrid systems, the critical domains are the ACP domain of PKS and the C domain of NRPS, wherein the C domains in PKS/NRPS hybrids must accept an acyl-S-ACP for chain elongation, instead of an aminoacyl-S-PCP. On the other hand for a NRPS/PKS interface, the key domains are the PCP domain of NRPS and the KS domain of PKS. The KS domain must approve the peptidyl intermediate bound to upstream PCP and condense it with the downstream acyl-S-ACP.

Moreover, the availability of a PPTase with broad substrate specificity is of paramount importance if both species of CPs (ACPs and PCPs) involve in a functional hybrid system are to be activated (Figure 1-4). To date, previous studies have already stated that several biochemically characterized PPTases (e.g. MtaA, Sfp, Svp) can also activate both ACPs and PCPs.

2.4 Magnetic particle production in microorganisms

Besides PKSs and NRPSs compounds, microbial natural products arise from a multitude of different types of complex biosynthetic pathways. Among these products, the magnetosome is famous for its magnetotaxis and possible commercialization in bright future. Magnetosome chains are membranous prokaryotic structures, consisting of 15 to 20 magnetite crystals that together act like a compass needle to orient magnetotactic bacteria in geomagnetic fields.

Magnetotactic bacteria are a polyphyletic group of bacteria first reported in 1975 which can utilize the iron from the surrounding environment to assemble an internal chain of nanomagnetic particles within lipid vesicles, allowing the bacteria to orient along the magnetic field lines of Earth's magnetic field. Magnetotactic bacteria usually mineralize either iron oxide magnetosomes, which contain crystals of
magnetite (Fe₃O₄), or iron sulfide magnetosomes, which contain crystals of greigite 
(Fe₃S₄). Several other iron sulfide minerals have also been identified in iron sulfide 
magnetosomes — including mackinawite (tetragonal FeS) and a cubic FeS - which 
are thought to be precursors of Fe₃S₄.⁷⁹ Magnetosome crystals are typically 35-120 
nm long, which make them single-domain crystals, which have the maximum possible 
magnetic moment per unit volume for a given composition.⁷⁶ Magnetosome 
biosynthesis in magnetotactic bacteria is controlled by a set of about 30 genes within a 
large (~115kb) genomic magnetosome island.⁸⁰ However, these microbes are 
challenging with respect to strain cultivation or genetic modification on a molecular 
level.

3. **Red/ET Recombineering**

In modern natural products research, metabolic engineering intrigues scientists to 
manipulate biosynthetic pathways to yield new analogues as drug candidates. 
However, it is usually very challenging to manipulate the DNA in the native 
producing strains, and Red/ET recombineering, which is a highly efficient and 
straightforward technology established in *E. coli*, has recently emerged as a promising 
new technique for working with large genomes.⁸¹-⁸³ 

Red/ET recombineering, as a research hotspot studied by several different groups for 
many years, which is also known as “λ Red recombination”⁸⁴,⁸⁵, “recombineering”⁸⁶,⁸⁷, 
“ET recombination”⁸¹,⁸² or “ET cloning”⁸³, relies on *in vivo* homologous 
recombination mediated by bacteriophage recombinases. The Red operon of the λ 
phage and the RecET operon of the Rac prophage are basically equivalent in terms of 
enzymatic function. As can be seen in Figure 1-7,⁸⁸ Redα and RecE are dsDNA 
exonucleases, while Redβ and RecT are single strand DNA annealing proteins 
(SSAP). In the Red operon the Redγ is an inhibitor of RecBCD which is the major 
host exonuclease, aggressively degrading dsDNA fragments.⁸⁹ 

Redα or RecE bind to double-stranded DNA (dsDNA) ends, and progressively 
degrades linear dsDNA in a 5’ to 3’ direction, leaving long 3’ single strand DNA
(ssDNA) overhangs. Redβ or RecT bind to the ssDNA to form ssDNA-SSAP nucleoprotein filaments, which are able to pair their complementary ssDNA. Once aligned, the 3’- end becomes a primer for DNA replication.\textsuperscript{88} The host factors then act an important part in finalizing the recombinant DNA by DNA replication and repair machinery (Figure 1-7).\textsuperscript{90-92}

Under the effect of recombinases, the chosen DNA regions are amplified in \textit{E. coli}, not in vitro, and thus are subject to review by the \textit{E. coli} replication machinery by a new simple, efficient and more flexible approach. This methodology breaks through the size and site limitation of DNA engineering and allows an extensive range of modifications by insertion, deletion, substitution, fusion, point mutation, subcloning and direct cloning for almost any given DNA molecule at any chosen position. Notably, the homology sequence at the ends of the linear DNA can be as short as 35 bp for efficient homologous recombination, and single-strand oligonucleotides as short as 70 bases can be transformed for direct mutagenesis. These short homology sequences can be easily integrated into synthetic oligonucleotides, greatly expanding the utility of the technology.

Redβ was associated in a complex with Redα when it was isolated from the cell.\textsuperscript{93,94} Some of the Redβ mutations were found to be defective of not only Redβ function, but also the exonuclease activity of Redα.\textsuperscript{95,96} These imply that redα recruits Redβ loading onto the single-stranded 3’ overhang. A specific protein-protein interaction between RecE and RecT has also been shown in coimmunoprecipitation experiments, suggesting that recombination requires cooperative functions between the exonucleases and the respective DNA annealing proteins.\textsuperscript{88,97-99} It was also shown that substitution of the protein from one system to another can impede function, further confirming protein-protein interaction specificity of each pair.\textsuperscript{88} In this thesis, both Redα/Redβ and RecE/RecT systems were investigated for application, and hereby defined as Red/ET recombineering.
A. Introduction

3.1 Development of Red/ET recombinering

Conventional methods to generate recombinant DNA use restriction enzymes and T4 DNA ligases,¹⁰⁰ and invention of PCR technology¹⁰¹ greatly expanded the capacity and flexibility of this technology. However, these methods are disable to precisely manipulate DNA fragments larger than 15kb, such as cosmids, fosmids, P1 vectors, bacterial artificial chromosomes (BACs) and bacterial chromosomes.¹⁰²-¹⁰⁶

Recombination-mediated genetic engineering was first demonstrated in *Saccharomyces cerevisiae*, using a gapped plasmid for cloning genes off the chromosome and a linear plasmid for targeting genes,¹⁰⁷,¹⁰⁸ with further experiments using single-strand oligonucleotides as short as 20 bp for efficient site-directed mutagenesis.¹⁰⁹ The homology sequence at the ends of the linear DNA were as short

Figure 1-7. Mechanism of Red/ET recombinering. The diagram was reproduced from www.genebridges.com.
A. Introduction

as 35 bp, which could be easily incorporated in oligonucleotide synthesis, and this simplicity encouraged the biologists to establish an equivalent system in *E. coli* strain. However the host RecA-mediated homologous recombination requires much larger homologous sequence and usually resulted in very low efficiency. The first attempt at exploring λ phage Red recombinases used long homology arms (about 1kb) and the recombination was more efficient than that from RecA.\(^{110}\) Independently, Zhang and colleagues in the Stewart laboratory published that either RecE/RecT or Redα/Redβ could mediate efficient homologous recombination between very short homology arms (40-50 bps), and the Redαβ system was much more efficient than the RecET system.\(^{83}\) Thus, Red recombineering has become a popular technology for rapid and precise modification of DNA, including insertion, deletion or replacement by a drug selectable marker,\(^{81,83}\) subcloning by gap repair,\(^{82}\) oligonucleotide-directed mutagenesis\(^{111,112}\) and counter-selection mediated by variable negative selectable markers to recombine DNA sequence without selection.\(^{113-115}\)

Subsequently, the differences between the Red system and RecET system were distinguished by biologists. Specifically, when the plasmid replication was required in the homologous recombination, the Redαβ system was preferred. Conversely, when an unreplicable linear vector was used to recombine with another DNA fragment to form a circular plasmid, RecET was found to be dramatically more efficient. Consequently, a novel RecET-based direct cloning method was invented, using PCR-amplified linear vectors to retrieve sequences of interest from digested genomic DNA, by passing the tedious DNA library construction.\(^{116}\)

To date, recombineering has been established not only in *E. coli*, but also in several other species, such as *Agrobacterium tumefaciens*,\(^{117}\) *Mycobacterium tuberculosis*,\(^{118,119}\) *Photorhabdus luminescens*,\(^{120}\) *Xenorhabdus*,\(^{120}\) *Salmonella enterica*,\(^{121}\) *Yersinia pseudotuberculosis*,\(^{122}\) *Burkholderia thailandensis* and *Burkholderia pseudomallei*.\(^{123}\)
3.2 Application of Red/ET recombineering

Recombineering is now an alternative for conventional recombinant DNA technology, and it is especially favourable for large DNA molecular engineering.\textsuperscript{116,124} Based on the DNA substrate to be recombined, recombineering can be divided into two categories, with each utilizing different phage recombinases. Linear plus circular homologous recombination (LCHR) is typically catalysed by Red\textsuperscript{α}/Red\textsuperscript{β},\textsuperscript{125} and is principally used for engineering various vectors and bacterial chromosome. Linear plus linear homologous recombination (LLHR) utilizes RecE/RecT,\textsuperscript{116} which is primarily exploited for linear DNA cloning (PCR cloning)\textsuperscript{82} and direct cloning.\textsuperscript{116}

The initial step of Red\textsuperscript{αβ} recombineering is oligonucleotide synthesis, in which the homologous sequence is attached to the PCR primer. The PCR-generated DNA fragment usually contains an antibiotic selectable marker and a terminal sequence identical to the target. After transformation of this DNA fragment in the Red\textsuperscript{αβ} expressed cell containing the target plasmid, the linear fragment recombines into the plasmid as positioned by the homologous sequence, endowing the cell with new antibiotic resistance (Figure 1-8). Each oligonucleotide for use in recombineering consists of at least two parts, yet could include a third or even fourth part to incorporate restriction enzyme recognition sites, site specific recombinase recognition sites (loxP, FRT, Rox, etc.) or sequence for a short protein tag (His, Strep, Flag, etc.). The target DNA can be a gene locus on the \textit{E. coli} chromosome or any other stretch of DNA in a BAC or plasmid vector.\textsuperscript{81-83}
A. Introduction

Figure 1-8: Redαβ recombineering for a circular DNA modification.
This figure illustrates the principle of a circular DNA modification. Sm, selectable marker; the small blue arrow indicates a prokaryotic promoter. Parts A, B and C are the oligonucleotide regions. Part A is the required homology arm in the range of 30-50 bp shared by the target circular molecule and the linear molecule. The sequence of the homology regions can be chosen freely according to the locus on the target circular DNA. Part B is optional (protein tags, restriction sites, site-specific recombination target sites, etc.). Part C is the required primer for PCR, which is usually 18 to 24 nucleotides according to the provided template.\textsuperscript{232}

LLHR mediated by RecET is capable of multiple linear DNA fragment assembly, if each fragment has an overlapping region of sequence identity with its neighbour (Figure 1-9A). The final recombination product should contain a plasmid origin for replication and an antibiotic resistance gene for selection. Another merit of RecET recombineering is direct cloning, with application of this technique to the 15.6 kb plu3263 gene providing a good example (Figure 1-9B). Specifically, the genomic
DNA of *Photorhabdus luminescens* is digested with NdeI and PacI to release a fragment containing plu3262. The PCR-generated linear vector is composed of a pBR322 origin, with an ampicillin selectable marker, tetO promoter and its repressor tetR. One homology sequence of the PCR is from the ATG start codon extending 70 bp downstream and another is the PacI site extending 70 bp upstream. After linear plus linear homologous recombination in *E. coli* expressing RecET, the plu3262 is directly cloned under the tetO promoter. Notably, the gene is being directly cloned under a regulated promoter to avoid potential toxicity to the host, and the chosen homologous sequence is designed 552 base pairs inside of the dsDNA end. An additional interesting fact is that RecET is able to align the identical sequences even when they are not exposed.\textsuperscript{116}

![Figure 1-9. RecET recombineering for a circular DNA modification.](image)

(A) Schematic illustration of multiple linear DNA recombination to generate a circular plasmid. (B) An example of direct cloning of plu3263 from the *Photorhabdus luminescens* genome. The figure was reproduced from reference\textsuperscript{116}.

Natural microbial metabolites with promising medicinal and agricultural values, are always expressed from secondary metabolite pathway gene clusters. These gene clusters are often too large to engineer via conventional arrangements. Hence, components of the gene clusters were reconstituted into a single vector with several essential elements for gene transfer by using recombineering-direct cloning techniques.\textsuperscript{116} The gene clusters could be shaped further to ensure the successful expression in suitable heterologous hosts.\textsuperscript{124,126,127} Many “unusual” novel compounds have already been synthesized and identified via Red/ET recombineering technology.
A. Introduction

in the past few years.128-130

3.3 Improvement of Red/ET recombineering

Along with the application of recombineering, technical improvements have been always made in different ways. For instance, the temperature-sensitive pSC101 plasmid was used to express the recombinases. After a temperature shift the expression plasmid will be lost and the recombinant DNA product will not be contaminated.129 Another example is that transient co-expression of RecA results in higher survival rate of the cells after electroporation, consequently promoting recombinant yield.85

Genetic engineering relies on the transfer of foreign DNA into cells, and optimization of DNA transformation efficiency methods are important. Compared to various methods to prepare chemical competent cells for DNA transformation, electroporation is much more efficient, especially when the DNA molecule is large.131-133 In E. coli electroporation procedure, the cells will turn exceptionally competent through high-voltage pulse treatment after rinsing with ice-cold water or 10% glycerol.133-135 The foreign DNA material will easily and rapidly enter into the permeabilized cell through the holes in the membrane transiently caused by the high voltage (1200-1350V).136

According to previous research, electrocompetent cell preparation has to be made at ice-cold conditions and therefore both equipment and washing solutions need to be kept at low temperatures all times.137-139 Holding the cells at chilling conditions is the pivotal aim in the most of the protocols for electroporating Gram-negative bacterial strains involving E. coli without an adequate account.137 Furthermore, the Red/ET recombineering has also been coordinated with other genetic tools to widen its application scope, for instance site specific recombination120 and transposition.126 Recently an efficient RNA-guided site-specific DNA cleavage technology has been developed based on the Streptococcus pyogenes type II CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) adaptive immune
A. Introduction

The system contains the CRISPR associated protein Cas9, a trans-activating CRISPR RNA (tracrRNA), and a programmable CRISPR targeting RNA (crRNA), which processes a Cas9-mediated double-stranded break (DSB) at almost any target DNA locus. CRISPR/Cas9 has caught massive attention for its feasibility, and pioneer data show that combination of CRISPR-Cas9 and Red/ET recombineering could lead to efficient multigene editing of not only in the *E. coli* genome but also in *Tatumella citrea* with very high efficiency.

4. Heterologous expression of natural products in surrogate hosts

Natural products, exhibiting a broad range of distinct structures and biological activity, are usually synthesized by large multienzymes mega synthetases, such as PKSs, NRPs, hybrid or Ripps. Hence, it is obviously much more economical and time-saving to produce these specified natural products during fermentation. Nevertheless, the ideal approach is not easy to achieve. For instance, the yield of target compounds in native producer is too low to achieve the minimum demands of industrialization or environmentally unfriendly byproducts are generated in the reactions. Furthermore, many native producers harboring these biosynthetic pathways show slow growth rates even in optimized laboratory conditions and are resistant to genetically manipulation. Thus, heterologous expression of the corresponding secondary metabolite pathways into a more amenable host organism can play an important role in developing novel derivatives and potential drugs. It also affords a high-level investigative platform for the detailed research of complex biosynthetic mechanisms, allowing careful identification of the products from silent biosynthetic gene clusters, and the generation of fresh analogs through biosynthetic engineering in suitable heterologous hosts which are genetically more tractable and easily cultivated.

With respect to heterologous expression of increasing number of complex natural products, it is obvious that a general overview of the workflow for cloning and production of large natural product assembly lines has been concluded (Figure 20).
A. Introduction

According to this approach, several general considerations and procedures should be assumed into account:

1) **Identification and isolation of the corresponding biosynthetic pathway in the native producer.** In the past dozen years, more and more natural product biosynthetic gene clusters have been investigated, with an accompaniment of a variety of increasing published genome sequencing data. For heterologous expression of a biosynthetic gene cluster, the genes should first be isolated so that they can be mobilized into an appropriate heterologous host, either using traditional ways (a BAC or cosmid library) or by direct cloning, and the clusters from genomic DNA should be transferred to a vector.

2) **Identification of a suitable vector for transfer and integration into the chromosome of heterologous host, functional promoter architecture and the corresponding regulatory factors in the related host.** For our direct cloning of large gene clusters, the replication origin and the copy number of the plasmid are two main considerations. We often use several types of vectors: (i) pBR322 plasmids presenting 15-20 copies per cell with an insert capacity of ~50 kb. (ii) p15A plasmids presenting 10-12 copies per cell with an insert capacity of ~70 kb. (iii) pBAC2015 plasmids presenting a single copy per cell with an insert capacity of ~150kb. Afterwards, the tetracycline inducible promoter (tetR-Ptet) is included in these plasmids to regulate the expression of the gene clusters in order to minimize the potential toxic effect of natural product gene clusters to the host.

3) **Heterologous host selection.** An amenable heterologous host is the most crucial aspect to be considered. Firstly, the phylogenetic distance between the native producer and heterologous host is a key consideration. In general, the more closely related the heterologous host is to the native strain, the high success rate of functional transcription of the biosynthetic pathway. Additionally, the codon usage may be similar in comparison between with the closely-related species and the original producer, which may increase the efficiency. In the case of PKSs and
NRPSs heterologous expression, the host must principally contain PPTase, which is required for post-translational activation of PKS-NRPS proteins. Meanwhile, various necessary precursors or substrates, such as CoA-activated short chain carboxylic acids, both proteinogenic and non-proteinogenic amino acids, and short-chain fatty acids must be provided satisfactorily in the parasitifer. Additionally, the genetic engineering availability and cultivation conditions are also very important for heterologous host selection. Higher titer production and lower product cost are the pursuant goals to be achieved through heterologous expression.

4) Genetic engineering of expression constructs. Following the successful heterologous expression of biosynthetic gene cluster, yield of heterologous expression may be still not satisfactory. Due to this view, promoter exchange, reconstitution of essential clusters and insertion of regulatory regions are entailed for the expression construct. Depending on the host system selected, there are a number of alternative designs for the conservation of the clusters.

5) Transformation into the heterologous hosts. In the successful heterologous expression, the maintenance of a foreign biosynthetic gene cluster stability in a surrogate host is foundational for sustainable product expression. In principle, there are only two available options for exogenous gene cluster, wherein one is to maintain a steady host-compatible plasmid episomally, and another is direct integration into the chromosome. Biosynthetic gene clusters can be expressed from self-replicating plasmids, either from a single plasmid with the entire biosynthetic pathway, or from multiple plasmids individually expressing modules in cis. For example, the 36kb echinomycin gene cluster from St. lasaliensis was expressed in E. coli BL21 (DE3) using three pET-derived vectors with the yield of 0.3mg/L. During the plasmid-based expression, increasing the copy number of the biosynthetic gene cluster resulted in improvement of compound production due to an ample supply of precursors. An additional convenience is afforded by the use of shuttle
vectors for the concurrent expression of biosynthetic gene cluster in multiple hosts. However, plasmid-based expression is not suitable to all the heterologous hosts as it depends on the availability of plasmid replicons capable of functioning in the host, and this is a risk for recombination-competent hosts and also for prolonged expression.

In order to secure a stable heterologous expression, the biosynthetic gene clusters should be integrated into the host chromosome. Furthermore, such integration is essential approaches using the plasmid replicons fail in the host. Homologous recombination, transposition and phage-mediated integration are the most popular and frequent applications in chromosomal integrations.

6) Product analysis and optimization. The successful expression of target compounds in suitable surrogate hosts is not the destination for our purpose. Often times, an ocean of novel compounds have already been discovered and identified and are in the progress of heterologous expression as are drug candidates. Nowadays, wide array of advanced techniques, involving liquid chromatography-coupled mass spectrometry (LC-MS), fast chromatographic separation (UPLC), electrospray ionization (ESI), high-resolution (HR) time-of-flight mass spectrometry (ToF-MS) and nuclear magnetic resonance (NMR), are employed in the hunt for new natural products to cope with the increasing appearance of novel infectious diseases.

Furthermore, to maximally harvest target metabolites in actual production process, several efficient approaches can be carried out to solve difficulties in the fermentation, such as precursor feeding and growth conditions optimization (in control of concentrations of CO₂ and O₂, pH, stirring rate, reaction time, etc.).
Heterologous expression is a time-saving method, especially for biosynthetic pathways from slow-growing bacteria, and the biosynthetic pathways can be expressed in a bacterium which cultivates relatively fast. Over the past decade, several complete biosynthetic pathways from fastidious bacteria have already been reassembled by re-construction via Red/ET recombineering and successfully expressed in several different heterologous hosts, such as epothilones in \textit{M. xanthus},\textsuperscript{126} human alpha-defensin 5 mature peptide in \textit{Pichia pastoris},\textsuperscript{166} nikkomycin in \textit{Streptomyces ansochromogenes},\textsuperscript{167} pretubulysin in \textit{P. putida} and \textit{M. xanthus},\textsuperscript{168} myxochromide S and myxothiazol in \textit{Pseudomonas putida} and \textit{M. xanthus},\textsuperscript{126,146,165,169} salinomycin in \textit{Streptomyces coelicolor},\textsuperscript{170} luminmycin and glidobactin in \textit{E. coli} Nissle1917,\textsuperscript{129,130} and magnetosome products in \textit{Rhodospirillum rubrum}.\textsuperscript{171}

Additionally, transformation-associated recombination (TAR) cloning, another classic recombineering method, widely uses \textit{in vivo} recombination in \textit{Saccharomyces cerevisiae} to directly and specifically isolate desired large gene clusters from complex genomes.\textsuperscript{172,173} This great technology was discovered in late 1990s but had been
A. Introduction

modified and received a high degree of development during the last ten years.\textsuperscript{174-176} Fragments up to 250kb can be directly fished out in yeast from multiple samples within 14 days via an optimized TAR cloning approach.\textsuperscript{175} Many applications based on this powerful method have been already established in the post-genomic era. According to the protocol, the 56kb Colibactin gene cluster from \textit{Citrobacter koseri} was directly cloned into the vector for heterologous expression.\textsuperscript{177} Another good example of the application of TAR cloning is the reconstruction of the evolutionary history of the SPANX gene family in primates.\textsuperscript{178,179} Unfortunately, an obvious disadvantage of TAR cloning is that cloning efficiency drops with a high chromosomal GC content, in both yeast and bacteria.\textsuperscript{175}

Furthermore, several practical synthetic biological technologies for gene clusters assembly have also been developed recently years, such the Gibson assembly and DNA assembler techniques.\textsuperscript{180-182} In comparison with Red/ET recombineering and TAR cloning, they only are capable of assembling small genetic segments prepared by standard PCR or direct chemical synthesis, but for insertions, deletions and point mutations, these methods are flexible and variable during DNA engineering.

Recent research indicates that the challenges of genetic engineering on biosynthetic pathways have been generally solved by Red/ET recombineering plus TAR cloning.\textsuperscript{176,183} The following strategy is to select a suitable surrogate host for expression. An ideal heterologous host should grow fast, be genetically tractable, provide all necessary precursors in sufficient quantities, ensuring functional expression of the required proteins and a low background for native secondary metabolites.\textsuperscript{184} It is usually arduous to meet all the requirements above for all natural products, but with better growth and genetic characteristics are the basic goals for our research. According to previous data, several specific strains are good models for expression of heterologous genes, including \textit{E. coli},\textsuperscript{17} \textit{Myxococcus xanthus},\textsuperscript{126} several \textit{Streptomyces} species (\textit{S. albus},\textsuperscript{185} \textit{S. lividans}\textsuperscript{186} and \textit{S. coelicolor}\textsuperscript{187}), \textit{Pseudomonas putida}\textsuperscript{165} and \textit{Rhodospirillum rubrum}.

While \textit{E. coli} has been proven to be a powerful and cogent model microorganism for
genetic and metabolic engineering with a quick growth rate and its use for biological products overproduction are also well established. However, original \textit{E. coli} often lacks the essential intracellular machinery to produce natural products. Thus, before heterologous expression, important modifications are required in \textit{E. coli} strain. For example, in order to express lumininlde via direct cloning, a PPTase from \textit{Stigmatella aurantiaca} \cite{73} was introduced into our standard recombineering host, \textit{E. coli} GB05 to form \textit{E. coli} GB05-MtaA. \cite{116} Coincidentally, in the heterologous epothilone production, the methylmalonyl-CoA decarboxylase gene (\textit{ygfG}) was replaced by the \textit{pcc} genes to cumulate MM-CoA. \cite{189} All the evidence illuminate that \textit{E. coli} could successfully express large PKS/NRPS gene clusters after selective alterations. \textit{M. xanthus} is the best characterized myxobacterium and can be genetically engineered much more easily than any other myxobacteria. \cite{126} Moreover, in native producers, a variety of PKS/NRPS secondary metabolites are manufactured, such as myxochromid A$_2$, myxalamid, myxovirescin, myxochelin and DKxanthene, \cite{190} demonstrating the appearances of PPTases, MM-CoA and self-resistance related to bioactive compounds. Furthermore, demonstrating an amazing performance, \textit{M. xanthus} presents a much shorter generation time (5h) than \textit{S. cellulosum} (16h), another major species of myxobacteria harboring miscellaneous antibiotics, such as epothilone. Up until now, the Müller group have established heterologous production of many myxobacterial secondary metabolites involving pretubulysin, myxochromide S, myxothiazol in \textit{M. xanthus}. \cite{126,168,169} \textit{M. xanthus} can heterologously express not only myxobacterial biosynthetic gene clusters, but also streptomycete-derived biosynthetic gene clusters, e.g. Oxytetracycline yield in \textit{M. xanthus} reached up to a high titer 10 mg/L. \cite{191} Those results suggest that \textit{M. xanthus} is a promising candidate for heterologous host to express PKS/NRPS biosynthetic pathways from both myxobacteria and actinomycetes. Ultimately, the majority of antibiotics and other pharmaceutically relevant bioactive compounds were isolated from actinomycetes (mainly belonging to the genus \textit{Streptomyces}) that had been best analyzed for dozens of years. Several \textit{Streptomyces}
A. Introduction

species, such as *S. avermitilis*, *Saccharopolyspora erythraea* and *S. coelicolor* have been sequenced and harnessed as heterologous hosts to produce their own bioactive compounds. Intense efforts of heterologous expression the streptomycete-derived secondary metabolites in these *Streptomyces* hosts have already been reviewed by Baltz.

Rhodospirillum rubrum (*R. rubrum*) is a photosynthetic model organism because of its pink color in the medium. The photosynthesis of *R. rubrum* differs from that of plants as it possesses not chlorophyll, but bacteriochlorophylls. For past decade, *R. rubrum* has already been developed into a modified heterologous host for several natural products, such as *Pseudomonas aeruginosa* membrane protein MscL, various PHA (polyhydroxyalkanoates) synthase genes and magnetosome biosynthetic gene cluster from *Magnetospirillum gryphiswaldense*. Based on these achievements, it is believed that more and more natural products are planning to be developed in *R. rubrum*.

5. Outline of the dissertation

The goal of the work described in this thesis is to improve heterologous transformation of genomic DNA, which is related to bacterial natural products, based on the increasing numbers of sequenced microbial genomes. The large gene clusters encoding secondary metabolite biosynthesis pathways can now be rapidly cloned and modified by Red/ET recombineering. In order to enhance the efficiency of recombineering, technical improvement for electrocompetent bacterial cell preparation was described in this thesis. Compared to long-standing ice-cold preparation method, the room temperature electrocompetent bacterial cells have resulted in increased transformation efficiency, which consequently promote the efficiency of Red/ET recombineering. Afterwards, three typical bioactive compound groups had been analyzed: the cytotoxic disorazols from myxobacterium *Sorangium cellulosum* So ce12, the magnetotactic magnetosomes from alphaproteobacterium *Magnetospirillum gryphiswaldense* and the antibacterial and antitumor salinomycin.
from *Streptomyces albus* DSM41398.

An important direction of the thesis is to improve the efficiency of Red/ET recombineering by optimization of the conditions for electrocompetent cell preparation. DNA transformation is the routine work in most molecular biology laboratories, and the electroporation has been proven to be more efficient than the chemical transformation. The long-standing protocol for electrocompetent cell preparation is rinsing the cells by ice-cold water or 10% glycerol. However, we found that the electrocompetent *E. coli* cells prepared at room temperature was 5 folds more efficient for transformation of plasmids with different sizes, replication origins and selectable markers. The room temperature electrocompetent *E. coli* cells also resulted in 6-10 folds increased efficiency of RecET recombineering. The beneficial effect of room temperature preparation of electrocompetent cells was not been only shown in *E. coli* but also in several other gram-negative bacteria. Ultimately, these results and data were presented in publication II.

Disorazols represent a family of 29 structurally complex macrocyclic polyketides, which were first isolated from the myxobacterium *Sorangium cellulosum* So ce12. These compounds inhibit cancer cell proliferation at picomolar concentrations by preventing tubulin polymerization and inducing destabilization of microtubules, which ultimately leads to the induction of apoptosis. According to the biosynthetic model four core genes *disA-D* encode eleven PKS modules and one NRPS module, with an acyltransferase domain is missing in the polyketide synthase modules. The last separate gene *disD* encodes an AT protein revealing the disorazol gene cluster falls into trans-AT PKS family. Thus, we intend to clone and engineer disorazol biosynthetic gene cluster by Red/ET recombineering. Followed by successful expression in heterologous host *Myxococcus xanthus* DK1622, the minor product disorazol A2 was changed to be prominent in several derivate without significant modifications. Interestingly, the yields of disorazols increased seven times after overexpression the solitary AT domain (*disD* gene) by insertion of a stronger promoter. These results were reported in publication I.
A. Introduction

The magnetosomes, which are membrane-enclosed crystals of a magnetic iron mineral, are biosynthesized by magnetotactic bacterium *Magnetospirillum gryphiswaldense*. Magnetotactic bacteria are a peculiar natural phenomenon that can absorb iron (mostly Fe$_3$O$_4$) from the surrounding solution and take it to produce an interior chain of nanomagnetic particles within lipid vesicles which are regarded as magnetosome chains. However, it is difficult to cultivate the bacterium in the optimized laboratory conditions. Through the use of Red/ET recombineering, we direct cloned a set of large genomic magnetosome island (~115kb) into several vectors and transferred them into the photosynthetic bacterium *Rhodospirillum rubrum*, respectively. Small magnetosome particles could be detected by transmission electron microscopy (TEM) and purified from disrupted cells by magnetic separation. Four main gene clusters, either essential *mamAB* or regulatory *mamGFDC, mms6* and *mamXY* in magnetosome formation, have been identified and defined through heterologous expression in *Rhodospirillum rubrum*. This biosynthesis of magnetosomes within other organisms enlarges the probability of commercialization of tailored magnetosome production within microorganisms, and these results were recently described in publication III.

The natural compound salinomycin, produced by *Streptomyces albus*, has a potent and selective activity against cancer stem cells and is therefore a potential anti-cancer drug. The salinomycin biosynthetic gene cluster (*salO-orf18*) from *Streptomyces albus* DSM41398 was separately cloned into three plasmids and stitched into an intact gene cluster (106kb) under the native promoter using Red/ET recombineering in *E. coli*. The large gene cluster was transferred in to *Streptomyces coelicolor A3* (2) for heterologous expression, and this was the first report of such a large genomic region directly cloned from a Gram-positive strain. These studies suggested a new approach to characterization of the relevant functional genes to identify novel analogues by module exchange, and the data and analysis were presented in publication IV.
B. Publications

I. Genetic engineering and heterologous expression of the disorazol biosynthetic gene cluster via Red/ET recombineering.

II. Room temperature electrocompetent bacterial cells improve DNA transformation and recombineering efficiency.

III. Biosynthesis of magnetic nanostructures in a foreign organism by transfer of bacterial magnetosome gene clusters.

IV. Direct cloning and heterologous expression of the salinomycin biosynthetic gene cluster from *Streptomyces albus* DSM41398 in *S. coelicolor* A3 (2).
Genetic engineering and heterologous expression of the disorazol biosynthetic gene cluster via Red/ET recombineering.

Qiang Tu\textsuperscript{1,2}, Jennifer Herrmann\textsuperscript{1}, Shengbiao Hu\textsuperscript{3}, Ritesh Raju\textsuperscript{1,†}, Xiaoying Bian\textsuperscript{2}, Youming Zhang\textsuperscript{2*} & Rolf Müller\textsuperscript{1*}

\textsuperscript{1}Department of Microbial Natural Products, Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection Research and Department of Pharmaceutical Biotechnology, Saarland University, Campus E8.1, 66123 Saarbrücken, Germany.

\textsuperscript{2}Shandong University – Helmholtz Joint Institute of Biotechnology, State Key Laboratory of Microbial Technology, School of Life Science, Shandong University, Jinan 250100, People’s Republic of China.

\textsuperscript{3}College of Life Science, Key Laboratory of Microbial Molecular Biology of Hunan Province, Hunan Normal University, Changsha 410081, Hunan Province, People’s Republic of China.

\textsuperscript{†}Present address: University of Western Sydney, Department of Pharmacology, School of Medicine, Campbelltown, NSW 2560, Australia.

\textsuperscript{*}Correspondence and requests for materials should be addressed to Y.Z. (email: zhangyouming@sdu.edu.cn) or R.M. (email: rom@helmholtz-hzi.de)
Genetic engineering and heterologous expression of the disorazol biosynthetic gene cluster via Red/ET recombineering

Qiang Tu1,2, Jennifer Herrmann1, Shengbiao Hu3, Ritesh Raju1,†, Xiaoying Bian2, Youming Zhang2 & Rolf Müller1

Disorazol, a macrocyclic polyketide produced by the myxobacterium Sorangium cellulosum So ce12 and it is reported to have potential cytotoxic activity towards several cancer cell lines, including multi-drug resistant cells. The disorazol biosynthetic gene cluster (dis) from Sorangium cellulosum (So ce12) was identified by transposon mutagenesis and cloned in a bacterial artificial chromosome (BAC) library. The 58-kb dis core gene cluster was reconstituted from BACs via Red/ET recombineering and expressed in Myxococcus xanthus DK1622. For the first time ever, a myxobacterial trans-AT polyketide synthase has been expressed heterologously in this study. Expression in M. xanthus allowed us to optimize the yield of several biosynthetic products using promoter engineering. The insertion of an artificial synthetic promoter upstream of the disD gene encoding a discrete acyl transferase (AT), together with an oxidoreductase (Or), resulted in 7-fold increase in disorazol production. The successful reconstitution and expression of the genetic sequences encoding for these promising cytotoxic compounds will allow combinatorial biosynthesis to generate novel disorazol derivatives for further bioactivity evaluation.

Natural products from microorganisms, fungi, plants and insects display a broad spectrum of biological activities. Currently, approximately 49% of anti-infectives compounds and 61% of anticancer pharmaceutical agents in clinical use are natural products or their derivatives1. Over the last decades, myxobacteria have become well known producer organisms, offering a rich and valuable source of natural products2,3. Most of these compounds are biosynthesized by multifunctional megasynthetases, such as polyketide synthases (PKSs)4, nonribosomal peptide synthetases (NRPSs)5 and hybrids thereof6. Genes encoding these PKSs and NRPSs in bacteria are often clustered together on the chromosome, so a gene cluster can be cloned into a vector and then transferred to a heterologous host for functional expression7. Recent studies demonstrate the usefulness of heterologously expressed secondary metabolite pathways for the production of natural products8. Heterologous expression can improve fermentation yields and generate new natural or synthetic products that can be evaluated as potential pharmacological agents in the course of targeted derivatization or structure-activity relationship studies9.

Reconstructing biosynthetic gene clusters in various vectors for heterologous expression in more productive hosts can help show how newly discovered biosynthetic gene clusters function. Derivatives of the new available pharmacologically active compounds can then be produced by biomolecular re-engineering and combinatorial biosynthesis10.

1 Department of Microbial Natural Products, Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection Research and Department of Pharmaceutical Biotechnology, Saarland University, Campus E8.1, 66123 Saarbrücken, Germany. 2 Shandong University – Helmholtz Joint Institute of Biotechnology, State Key Laboratory of Microbial Technology, School of Life Science, Shandong University, Jinan 250100, People’s Republic of China. 3 College of Life Science, Key Laboratory of Microbial Molecular Biology of Hunan Province, Hunan Normal University, Changsha 410081, Hunan Province, People’s Republic of China. † Present address: University of Western Sydney, Department of Pharmacology, School of Medicine, Campbelltown, NSW 2560, Australia. Correspondence and requests for materials should be addressed to Y.Z. (email: zhangyouming@sdu.edu.cn) or R.M. (email: rom@helmholtz-hzi.de)
Large clusters of genes that encode the enzymes for natural product biosynthesis have been difficult to engineer using conventional technology. Red/ET recombineering in combination with TAR (Transformation-associated recombination) cloning is necessary for large biosynthetic gene clusters to be engineered conveniently\(^1\).\(^1\)\(^2\)\(^2\).

Red/ET recombineering is based on \textit{in vivo} homologous recombination in \textit{E. coli}\(^1\)\(^3\)\(^4\). The greatest advantage of the technology is that it works regardless of restriction enzyme binding sites or the size of DNA fragments to be shuffled. This technology has made it much easier to genetically manipulate complex biosynthetic pathways in bacteria. Over the past decade, several complete biosynthetic pathways from fastidious bacteria have been heterologously expressed via Red/ET recombineering, e.g. myxochromide S and myxothiazol in \textit{Pseudomonas putida} and \textit{M. xanthus}\(^1\)\(^6\)\(^7\), epothilones in \textit{M. xanthus}\(^1\)\(^9\), human alpha-defensin 5 mature peptide in \textit{Pichia pastoris}\(^2\)\(^0\), nikkomycin in \textit{Streptomyces ansochromogenes}\(^2\)\(^1\), pretubulysin in \textit{P. putida} and \textit{M. xanthus}\(^2\)\(^2\), luminmycin and glidobactin in \textit{E. coli}\(^2\)\(^3\)\(^4\), salinomycin in \textit{Streptomyces coelicolor}\(^2\)\(^5\) and even a minimal set of genes for magnetosome biosynthesis from the magnetotactic bacterium in \textit{Rhodospirillum rubrum}\(^2\)\(^6\).

Secondary metabolite gene clusters in microbes express natural products with potential medicinal and agricultural qualities\(^2\)\(^7\). However, many of the microorganisms hosting these biosynthetic pathways grow slowly even in optimized laboratory conditions and cannot be genetically manipulated\(^2\). Heterologous expression of complete secondary metabolite pathways plays a significant role in hunting for new natural products and developing them into useful drugs\(^4\). Many heterologous expression instruments for secondary metabolite pathways have been reported so far, ranging from targeted expression by shuttle vectors to the random expression of large DNA fragments from chromosomes by transposition\(^9\).

Disorazols, a family of structurally complex macrocyclic polyketides, are produced by the myxobacterium \textit{Sorangium cellulosum} \textit{So ce}12 (Fig. S1) and firstly isolated in 1994\(^2\)\(^8\). Disorazols inhibit cancer cell proliferation at low picomolar concentrations by preventing tubulin polymerization and inducing destabilization of microtubules, which ultimately leads to the induction of apoptosis\(^2\)\(^9\)\(^-\)\(^3\)\(^1\). The extraordinary potency of disorazols fostered their development as peptide-conjugates for cancer therapy\(^3\)\(^2\)\(^5\)\(^,\)\(^3\)\(^3\)\(^7\) and encouraged the generation of new and simplified disorazol derivatives by means of chemical synthesis\(^3\)\(^4\)\(^-\)\(^3\)\(^6\). However, there are no reports to date on genetic engineering approaches for the production of new analogs of the disorazol compound class.

The \textit{dis} biosynthetic gene cluster was identified by transposon mutagenesis. In 2005, the cluster was cloned into a BAC or cosmid library of \textit{S. cellulosum} \textit{So ce}12 by two independent groups\(^3\)\(^7\)\(^-\)\(^3\)\(^8\). The clusters showed the anticipated \textit{disA-C} genes encoding hybrid \textit{trans}-AT type I PKS/NRPS megaenzymes, and also another gene, \textit{disD}, that encoded an additional acyl transferase protein (Fig. 1).

According to the biosynthetic model, seven malonyl-CoA units and one serine are incorporated as extender units, forming half of the disorazol bis-lactone core unit. Two polyketide monomers may dimerize to form disorazol via the thioesterase (TE) domain\(^8\), possibly requiring an esterase encoded by \textit{orf3}\(^3\)\(^7\) (Table S4).

The native strain produces only small amounts of disorazols (~1 mg per liter fermentation medium) and is difficult to cultivate\(^3\)\(^8\). Consequently, it is challenging to produce large quantities of disorazol for further
development. Using an amenable heterologous host should be a rational way to assure higher and stable disorazol yields and possibly optimize its structure by molecular engineering.

Here we report the Red/ET recombineering of the dis biosynthetic gene cluster into a stable vector containing a p15A replication origin and a MycoMar transposase element. When the dis gene cluster was transposed into the chromosome of the heterologous host M. xanthus DK1622 several disorazol derivatives were produced. Subsequent gene deletions proved that only the disA-D genes and not orf9 or the putative esterase gene orf3′ were needed for disorazol production⁵⁷. Further, we also improved disorazol production in the heterologous host M. xanthus DK1622 by replacing the native promoter of the disD gene encoding a discrete AT protein with an artificial synthetic promoter.

Results and Discussion

Reconstitution of the disorazol A biosynthetic gene cluster. The disorazol A biosynthetic gene cluster has been cloned, sequenced and identified previously from a BAC library of So ce1²⁷. The BAC contained most of the dis gene cluster from disA to disD. However, the BAC pBeloBAC11-dis was a large and low copy vector and very difficult to transfer between hosts for heterologous expression. To construct a more efficient expression vector and very difficult to transfer for expression in different heterologous hosts, we sequentially modified the original BAC (pBeloBAC11-dis) by Red/ET recombineering⁵⁷. The backbone of pBeloBAC11-dis was replaced by a cassette containing the p15A replication origin (p15A ori), the origin of transfer (oriT) for conjugation purposes, two inverted repeats (IRs), a MycoMar transposase gene (Tps) for transposition, an inducible promoter tetr-P₄₃ for driving the dis gene cluster upstream of disA and a kanamycin resistance gene for selection in M. xanthus DK1622.

In the resulting construct p15A-dis, the dis gene cluster (containing disA-D and orf9) is in a relatively high copy number vector (20–30 copies per cell in E. coli). Instead of the native promoter, expression in this vector is controlled by a tetracycline inducible promoter in this vector works in several heterologous hosts, e.g. E. coli, M. xanthus and P. putida⁴⁹ (Figs 2 and S2).

We previously found that disorazol production was no longer detectable when an esterase gene (orf3′) was mutated by transposon insertion in mutant strain So12_EXI_IE-3⁷. This mutated esterase gene was implicated in bis-lactone formation during disorazol biosynthesis. We recovered plasmid pTrn-Rec_IE2 (Fig. S6), which contained several genes near the transposition in the mutant So12_EXI_IE-3. The transposon was found in the middle of the carboxyl esterase gene orf3′ (only 6.7 kb upstream of the disA start codon).

The plasmid pTrn-Rec_IE2 also included a S-adenosyl methionine (SAM) dependent methyl transferase gene orf2⁵⁷. As the product of orf2 may O-methylate the OH group at C-6′ adjacent to the orf3′ gene, it might also be essential for disorazol biosynthesis (Fig. S6, Table S4). Hence, we inserted both, the repaired carboxyl esterase gene orf3′ and the SAM-dependent methyl transferase gene orf2′ together into p15A-dis to form p15A-dis-est by Red/ET recombineering. To gain the fusion plasmid p15A-dis-est, firstly, two separate PCR cassette “cmR-orf2′ gene” with suitable homologous arms to the region (containing two Hind III restriction sites in both sides) between the orf9 and the disD genes were introduced into the vector, respectively. After digestion by Hind III restriction enzyme in correct clones, the linear fragment “cmR-orf2′-spectR” was integrated to obtain the final construct p15A-dis-est. By this, the cmR gene was introduced to drive orf2′ and orf3′ genes. Likewise, the spectR gene was introduced to drive the disD gene (Figs 2 and S2).

Certain gene products may be toxic to the host cell, potentially limiting the nature of downstream applications when introduced into E. coli directly at high copy number⁴⁰. All E. coli strains containing the dis gene cluster with the native promoter were found to carry mutations after recombineering. Therefore, it was very challenging to obtain the expression construct containing the dis gene cluster directly in E. coli because the growth of the host was impeded. We reasoned that one of the dis proteins interfered with a primary metabolic pathway in E. coli to disrupt growth. To address this issue, an inducible promoter P₄₃ was used to regulate gene expression. P₄₃ is a versatile tetracycline-based regulatory system that is usually used to selectively control expression of downstream genes⁵⁹. No other promoter system is suitable for so many diverse hosts, including E. coli, M. xanthus and P. putida⁵⁰,²²,²³,⁴¹. Besides, P₄₃ had already enabled several mixed PKS/NRPS natural products to be produced in heterologous hosts unrelated to the native producing organisms, such as myxochromide S from myxobacterium Stigmatella aurantiaca, which has been engineered into P. putida⁴⁹.

The transposon method, which was also applied in this study, is clearly more stable and efficient than using shuttle vectors⁴⁹. Several indispensable elements were inserted into the target vectors, for instance Tps and oriT. The mariner transposon MycoMar is frequently used in Gram-negative hosts for genetic modification⁴⁵,⁴³ and to transfer and integrate a gene cluster into the chromosome of heterologous host strains⁵⁹,²². The transformation efficiency of large gene sets is higher when using the MycoMar transposon than using homologous recombination, as has been described for the heterologous expression of epothilone and myxochromide S⁵⁸. This powerful tool for transforming large genes was used in the disorazol heterologous expression system to make it easier to integrate the dis gene cluster into the genome of host strains. The oriT was also incorporated for conjugation in other heterologous hosts strains, such as P. putida⁴⁴.

Heterologous expression of dis gene cluster in M. xanthus DK1622. Both expression constructs p15A-dis and p15A-dis-est (Fig. S2) were introduced into the heterologous host M. xanthus DK1622 by electroporation as previously described¹⁹. The dis gene cluster was randomly transposed into the chromosome of M. xanthus. Transformants were screened on CTT agar containing kanamycin to select for M. xanthus::p15A-dis and M. xanthus::p15A-dis-est mutants. Six randomly chosen colonies of each mutant were verified by PCR¹⁹, which confirmed that the dis gene cluster had been integrated into the M. xanthus chromosome in each case. All the checked mutants contained the whole disorazol gene clusters. Several resulting mutants M. xanthus::p15A-dis and M. xanthus::p15A-dis-est were cultivated (both induction by anhydrotetracycline (AHT)) for compound
extraction and detection. All the mutants produced detectable amounts of disorazols by the analysis of high performance liquid chromatography-tandem mass spectrometry (HPLC-MS) 45. We have found small amounts of various disorazol compounds (including disrazols A1, A2, A3, A4, B2, B4 and F2) in both extracts of M. xanthus::
p15A-dis and M. xanthus:: p15A-dis-est (Figs 3 and S3, Table S2), upon comparing the secondary metabolite profiles from M. xanthus wild type strain and mutants. As expected, these results indicate that the chosen set of genes is sufficient to produce the polyketide-nonribosomal peptide skeleton of the disorazols.

Unexpectedly, without the orf2′ and the orf3′ genes, M. xanthus:: p15A-dis can also produce disorazols. The overall yields of disorazols in M. xanthus:: p15A-dis (averagely were 0.4 mg/L) match with that in M. xanthus:: p15A-dis-est (averagely were 0.42 mg/L). Result exhibited that the orf3′ gene is dispensable in the disorazol biosynthesis in the chosen heterologous host. There might be an enzyme that can substitute for the similar function of the orf3′ gene product in M. xanthus host. The dis gene cluster could be inactivated in the transposon mutant So12_EXI_IE-3 due to a strong polar effect37 because it is adjacent to the disA gene, possibly preventing downstream genes in an operon from being transcribed46,47.

The HPLC-MS and NMR data showed that the major compound in both mutants M. xanthus:: p15A-dis and M. xanthus:: p15A-dis-est was disorazol A2 which constituted 55% of final product after purification from crude extracts (Figs S3 and S4, Table S2), whereas disorazol A1 was 20%. But in the native host So ce12, disorazol A1 was the chief component (nearly 70% after purification, 10 times higher than disorazol A2) produced among the 29 derivatives28. The most probable explanation was that an O-methyl transferase that methylates the OH group at C-6′ was absent in the heterologous expression of dis gene cluster. This methyl transferase gene could be possibly located elsewhere in the chromosome of the native producer So ce12, which still needs further investigation. Only small amounts of the C-6′ methylated disorazols A1, A3 and A4 were produced in M. xanthus (Fig. S3), which might be due to partial methylation by a nonspecific M. xanthus O-methyl transferase. After 5 L fermentation of mutant strain M. xanthus::p15A-dis, the yield of disorazol A2 was approximately 0.24 mg/L, which is 5-fold higher than described in the native producer strain So ce1228,48. The result unambiguously demonstrated again that secondary metabolites can be produced in heterologous hosts under the control of the versatile Ptet Promoter which encouraged further investigation of disorazol formation.

Figure 2. Diagram of disorazol A gene cluster engineering. Firstly, the backbone of plasmid pBeloBAC11-dis (i) was replaced by p15A ori-tps cassette to form p15A-dis (ii) which containing an original MycoMar transposon by Red/ET recombineering. In this way dis gene cluster was driven by Ptet promoter. Then, the interrupted esterase gene orf3′ in pTn-Rec_IE2 plasmid (iii) from transposon mutant So12_EXI_IE-3 plasmid was recovered, repaired and engineered to form the vector p15A-amp-cm-orf2′-orf3′-spec (iv) that contained the whole length of the esterase gene orf3′. Next, linear DNA fragment released by Hind III was integrated into disorazol vector p15A-dis (ii) to get the final construct p15A-dis-est (v) via Red/ET recombineering. Finally, two types of modified vectors p15A-dis (ii) and p15A-dis-est (v) were electroporated into M. xanthus respectively and kanamycin-resistant colonies were selected for further analysis. Hind III restriction sites used for releasing linear fraction “orf2′-orf3′-specR were indicated in ↑. The insertion site of the linear fragment DNA containing orf2′ and orf3′ gene was marked with ↓.
Biological activity of disorazol compounds. After having isolated disorazols from our heterologous host \textit{M. xanthus}, biological studies revealed exceptional high cytotoxicity of disorazol A2 on eukaryotic cells. We determined IC\textsubscript{50} values against several established human cancer cell lines from different origin and disorazol A2 strongly inhibited the growth of these cell lines with IC\textsubscript{50} values between 0.05 and 4.9 nM (Table 1). However, compared to disorazol A1, the antiproliferative activity of disorazol A2 was less pronounced on most cell lines, except for human U-937 histiocytic lymphoma. Most likely, the higher IC\textsubscript{50} values for disorazol A2 are due to the lack of a methyl group at C-6' compared to disorazol A1, which in turn might lead to a less favourable binding to the target structure tubulin. Nevertheless, when compared to other anticancer drugs, such as epothilone B or vinblastine, disorazol A2 is still much more effective \textit{in vitro}.

Optimized production with biomolecular technology. An unusual feature of the disorazol biosynthetic gene cluster is that it has only one discrete AT domain on the DisD module, and hence it is called a \textit{trans}-AT type of PKS. In recent years, \textit{trans}-AT PKSs have been found in an important group of biosynthetic enzymes that produce bioactive natural products, including pederin, rhizoxin, leinamycin, myxovirescin, chivosazol and psymberin. Accessing functionally-optimized polyketides by modifying PKSs through targeted synthase re-engineering is an encouraging approach to optimize natural products for application. However, in contrast to ATs from \textit{cis}-AT PKSs, the mechanisms and structures of \textit{trans}-acting ATs are still unexplored.

The \textit{disD} gene has been modified here to show how \textit{trans}-acting ATs affect the disorazol biosynthesis pathway. In order to enhance the expression of the solitary AT domain, we introduced another strong promoter \textit{Pcp25} upstream of the \textit{disD} gene. \textit{Pcp25} is a highly active, constitutive lactococcal consensus promoter, whose sequence has already been reported. Previous studies have illustrated that overexpression of single genes or multigene transcriptional units by promoter exchange in myxobacteria can improve the production of secondary metabolites.

On the other hand, the role of \textit{orf9} gene (showing similarity to hypothetical proteins), which separates the \textit{disC} and \textit{disD} genes, in the \textit{dis} gene cluster has not been defined. To discover the actual function of the \textit{orf9} gene in disorazol biosynthesis, we inactivated it on the expression construct \textit{p15A-dis} and then performed heterologous production in \textit{M. xanthus}.

The PCR cassette ”\textit{P}_{\textit{gpd}}\textit{spect}\textit{RE}” (P18–P20 in Table S1), containing promoter \textit{P}_{\textit{gpd}} and a spectinomycin resistance gene (\textit{spect}), with two different pairs of homologous arms, was inserted into p15A-dis by Red/ET recombineering to form two plasmids p15A-dis-\textit{P}_{\textit{gpd}} and p15A-dis-\textit{P}_{\textit{gpd}-\textit{orf9}} (Figs 4A and S5). In the first plasmid p15A-dis-\textit{P}_{\textit{gpd}}, the promoter \textit{P}_{\textit{gpd}} was inserted directly upstream of the \textit{disD} gene. \textit{P}_{\textit{gpd}} is a highly active, constitutive lactococcal consensus promoter, whose sequence has already been reported. Previous studies have illustrated that overexpression of single genes or multigene transcriptional units by promoter exchange in myxobacteria can improve the production of secondary metabolites.

The \textit{disD} gene was thereby controlled by the \textit{P}_{\textit{gpd}} Promoter in both expression constructs. The recombinants were analyzed after growth on low-salt Luria-Bertani (LB) broth plates plus spectinomycin. The verified constructs were transformed into \textit{M. xanthus} DK1622 and three randomly picked positive transformants of each type of strain were cultivated to analyze the production by HPLC-MS. To clearly identify disorazol, retention times (RT) and the MS\textsuperscript{2} fragmentation pattern were compared to authentic reference substances. The concentration of disorazol A\textsubscript{1} in the culture was determined by UPLC–HRMS. A standard curve between peak area and concentration was established from serial dilutions for disorazol A\textsubscript{1} down to 0.01 μg/mL.
The peak area of disorazol A2 (base peak chromatograms, BPC +759.3 ± 0.1, RT = 18.2 min) was calculated by BrukerDaltonics compass data analysis 4.0. The yields of all disorazols were estimated from their relative peak areas in the HPLC-MS chromatogram by comparison with the standard curve for each derivative. All the resulting host strains still produced disorazols with growing production titres based on HPLC-MS analysis. The generated M. xanthus:: p15A-dis-Pcp25 expression host produced on average seven times more disorazol A2 compared to M. xanthus:: p15A-dis and mutant strain M. xanthus:: p15A-dis-Pcp25Δorf9 produced approximately 2.5-fold when compared to M. xanthus:: p15A-dis (Figs 4B and 5). Hence, the orf9 gene ablation did reduce disorazol production although it was described as having “no functional prediction” in BLAST analysis37. The orf9 gene, following the TE domain, might affect the biosynthetic formation of the final product by incorporation and cyclization of two sides of the disorazol bis-lactone. The successful enhancement of disorazol heterologous production suggested that re-engineering trans-AT PKSs domains on the molecular level was a feasible and practicable approach in investigating the characteristic enzymes.

Trans-AT PKSs are an important but still less known family of biosynthetic systems in comparison to cis-AT PKSs58,59. There are significant differences in the existing biosynthetic protocols between trans-AT and cis-AT PKSs. A single discrete AT DisD recognize and load all molonyl-CoAs for all the disPKS modules. Here we change the native promoter of disD gene with a stronger and artificial synthetic Pcp25 promoter which it would increase the transcription of disD gene and most likely raises the amount of DisD protein. Sufficient ATs could

<table>
<thead>
<tr>
<th>Human Cell line</th>
<th>Origin</th>
<th>IC50 [nM]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Disorazol A1</td>
</tr>
<tr>
<td>A-431</td>
<td>epidermoid carcinoma</td>
<td>1.866</td>
</tr>
<tr>
<td>A-549</td>
<td>lung carcinoma</td>
<td>0.072</td>
</tr>
<tr>
<td>HCT-116</td>
<td>colon carcinoma</td>
<td>0.032</td>
</tr>
<tr>
<td>HepG2</td>
<td>hepatocellular carcinoma</td>
<td>0.002</td>
</tr>
<tr>
<td>HL-60</td>
<td>acute myeloid leukemia</td>
<td>0.058</td>
</tr>
<tr>
<td>K-562</td>
<td>chronic myeloid leukemia</td>
<td>0.074</td>
</tr>
<tr>
<td>KB-3.1</td>
<td>cervix carcinoma</td>
<td>0.025</td>
</tr>
<tr>
<td>SW480</td>
<td>colomadenocarcinoma</td>
<td>0.030</td>
</tr>
<tr>
<td>U-2 OS</td>
<td>osteosarcoma</td>
<td>0.038</td>
</tr>
<tr>
<td>U-87 MG</td>
<td>glioblastoma-astrocytoma</td>
<td>0.072</td>
</tr>
<tr>
<td>U-937</td>
<td>histiocytic lymphoma</td>
<td>0.293</td>
</tr>
</tbody>
</table>

Table 1. Activity of disorazol A1 and disorazol A2 against human cancer cell lines. IC50 values refer to antiproliferative activities.
SalI and primer P5 (see Table S1), were co-transformed into digested by more stable and higher copies of DNA (P3, P4 in Table S1). Then two linear fragments, p15A-amp-orf2′-Tn-hyg′ were used: GB2005, derived from DH10B by deletion of fhuA, pTn-Rec_IE2, we first changed the backbone of pTn-Rec_IE2 into p15A-amp-orf2′-orf3′ and then digested the new cassette at the gbaA locus41,62; GB05-dir, derived from GB2005 by the PBAD-ETgA operon, which was integrated into the ybcC locus in GB200541. The integration ablates expression of ybcC, which encoded a putative exonuclease similar to that encoded by Redα. The heterologous host for PKS/NRPS gene cluster expression was M. xanthus DK1622 grown at 30 °C in CTT medium (1% casitone, 8mM MgSO4, 10mM Tris-HCl, pH 7.6, and 1mM potassium phosphate, pH 7.6)63 with or without km (50 μg/ml) before or after introduction of the disorazol gene cluster.

**Red/ET recombineering.** All methods were essentially as described previously62. By using Red/ET recombineering, 0.3 μg of one linear DNA fragment (either a PCR product or a fragment obtained from restriction enzyme digestion) was electroporated into 50 μl Red/ET-competent E. coli cells (such as GB-red cells). After electroporation, colonies grew on the LB-agar plates under selection for the antibiotic resistance gene and then were examined for the intended Red/ET recombination product by restriction analysis with a set of different suitable enzymes.

All PCR reactions carried out using Taq polymerase (Invitrogen GmbH, Karlsruhe, Germany) according to the manufacturer’s protocol. For the amplification of the ~1000 bp cassette with high GC content, DMSO was added to a final concentration of 3%. The conditions using an Eppendorf master cycler were as follows: 10 min at 95 °C to activate the polymerase, denaturation at 95 °C (30 s), annealing at 58 °C (30 s), and extension at 72 °C (35 s); 35 cycles. The PCR product was directly used without any purification.

**Reconstitution of dis gene cluster.** To harvest the full length of the esterase gene orf3′ in plasmid pTn-Rec_IE2, we first changed the backbone of pTn-Rec_IE2 into p15A-amp-orf2′-Tn-hyg in order to get more stable and higher copies of DNA (P3, P4 in Table S1). Then two linear fragments, p15A-amp-orf2′-Tn-hyg′ digested by SalI and primer P5 (see Table S1), were co-transformed into E. coli GB05-dir cells41 to remove R6K-Tn-hyg′ genes and recover the whole size of orf3′ gene. Thus, we obtained plasmid p15A-amp-orf3′-orf3′ harboring the full-length esterase gene. In order to insert the whole length orf2′ and orf3′ genes into the disorazol plasmid, we inserted two single PCR cassettes “spectR” and “cmR” with suitable homologous arms (P6-P9 Table S1) into the vector to engineer plasmid p15A-amp-cm-orf2′-orf3′-spect and then digested the new construct with HindIII to release the linear cassette “cmR-orf2′-orf3′-spectR” whose HindIII restriction site were homologous to p15A-dis vector. In the last step, the “cmR-orf2′-orf3′-spectR” cassette was transformed into strain GB-red::p15A-dir to generate the final plasmid (Figure S7). Two expression constructs p15A-dis and p15A-dis-est were obtained, containing four core-large genes from the disorazol A pathway (ten PKSs and one NRPS, ~58kb), with the Pnet promoter located upstream of the first PKS domain (Fig. 2).

**Figure 5.** Production of disorazol A after insertion of Pp25 promoter. Quantification of disorazol heterologous production by HPLC-MS analysis of the culture extracts from different M. xanthus DK1622 mutant strains. All LC-MS- derived area values are normalized to the crude extracts of each sample by method of standard curves. The depicted values are mean values from three independent mutants. Error bars show calculated SDs, yield, control strain.
Electroporation of *M. xanthus* DK 1622. The engineered gene clusters were introduced into the chromosome of *M. xanthus* DK1622 by electroporation. Briefly, *M. xanthus* cells from 1.7 ml of overnight culture with OD600 ~ 0.6 were collected and electrocompetent cells were prepared after washing twice with ice-cold water. A mixture of 50 μl cell suspension in cold water and 3 μg DNA was electroporated (Electroporator 2500, Eppendorf AG, Hamburg, Germany) at 1300 V using a 0.1 cm cuvette. After electroporation, the cells were resuspended in 1.7 ml fresh CTT medium, and incubated at 30 °C in a 2 ml Eppendorf tube with a hole punched in the lid on a Thermomixer (Eppendorf) at 11000 r.p.m. for 6 h. Then 1 ml 1.5% CTT agar solution at 42 °C was added to the tube and the cells were plated in soft agar for selection on CTT agar plates supplemented with km (50 μg/ml). Km-resistant colonies appeared after 4 days and were checked by colony PCR as follows. Part of a single colony was washed once in 1 ml H₂O and resuspended in 100 μl H₂O. Then, 2 μl of the resulting suspension was used as a PCR template using Taq polymerase according to the manufacturer’s protocol. The disorazol-specific primers used to check the integration of the disC gene into the *M. xanthus* chromosome were the same as used in a previous study. For PCR amplification, primers 10 and 11 were used (see Table S1).

Expression and analysis of disorazol production. Plasmids harboring a core-region or reconstituted dis gene cluster were introduced into *M. xanthus* DK1622 by electroporation. The resulting mutants (M. xanthus DK1622:: p15A-dis) were cultivated in 30 ml CTT medium. The medium was inoculated with 0.5 ml of the overnight culture and incubated at 30 °C on a rotary shaker at 180 rpm. After induction (anhydrotetracycline, final concentration 0.5 μg/mL) and addition of XAD adsorber resin (2%, 24 h), incubation was continued for 2 more days. The cells and the resin were harvested by centrifugation and extracted with methanol. The extracts were evaporated and then redissolved in 1 mL MeOH. A 5 μlL solution was analyzed by HPLC-MS and analysis was performed on an Agilent 1100 series solvent delivery system that was equipped with a photodiode array detector and coupled to a Burker HCTultra ion trap mass spectrometer. Chromatographic conditions were as follows: Luna RP-C₁₈ column, 100 × 2 mm, 2.5 μm particle size, and precolumn C₁₈, 8 × 3 mm, 5 μm. Solvent gradient (with solvents A [water and 0.1% formic acid] and B [CH₃CN and 0.1% formic acid]): 20% B from 0 to 20 min, 20% B-95% B within 10 min, followed by 5 min with 95% B at a flow rate of 0.4 mL/min. Detection was carried out in positive ion mode, auto MS². Disorazols were identified by comparison to the retention times and the MS² data of disorazols identified from the original producer in our myxo-database (target screening, Table S2). The relative production of disorazols was calculated from the peak areas of the extracted ion chromatograms (EICs) of each derivative.

High-resolution mass spectrometry was performed on an Accela UPLC-system (Thermo-Fisher) coupled to a linear trap-FT-Orbitrap combination (LTQ-Orbitrap), operating in positive ionization mode. Separation was achieved on a Waters BEH RP-C₁₈ column (50 × 2.1 mm; 1.7 μm particle diameter; flow rate 0.6 mL/min, Waters), with a mobile phase of H₂O/CH₃CN (each containing 0.1% formic acid) and a gradient of 5–95% CH₃CN over 9 mins. UV and MS detection were performed simultaneously. Coupling of HPLC to MS was supported by an Advion Triversa Nanomate nano-ESI system attached to a Thermo Fisher Orbitrap. Mass spectra were acquired in centroid mode at 200–2000 m/z with a resolution of R = 30000.

Target screening method. The HPLC-HR-MS data of crude extracts were further analyzed to identify the known compounds present in the extracts using the software Target Analysis (Bruker Daltonik GmbH). The known compounds were identified on the basis of their high resolution mass, isotope pattern and retention time according to the known method. With this approach, re-isolation of known but less interesting compounds could be avoided whereas unknown compounds with potential bioactivity could be identified easily.

Isolation of disorazol A₂. *M. xanthus* containing p15A-dis was cultivated in 5 L CTT medium supplemented with 30 μg/ml kanamycin and 2% XAD 16 resin (after 2 days of incubation) at 30 °C for 5 days. The resin was collected by sieving, washed with H₂O twice, and then extracted stepwise with acetoacetate (5 L). The extract was concentrated in vacuo and then extracted stepwise with acetoacetate (5 L) and then with hexane to defat. The resulting MeOH extract (0.87 g) was fractionated initially on a Sephadex LH-20 column (100 × 2 cm) using MeOH as a mobile phase, and 55 fractions were obtained. Fractions containing disorazol A₂ were subjected to semi-preparative reversed-phase HPLC system (Jupiter Proteo C₁₂, 250 × 21.2 cm, 4 μm, DAD at 254 nm) with an isocratic system of 75% MeOH/H₂O with 0.05% TFA to yield (1.2 mg, tR ≈ 22 min).

NMR. NMR spectra were recorded in CD₂OD on a DRX 500 MHz spectrometer (1H at 500 MHz, 13C at 125 MHz) equipped with a 5-mm probe and a Bruker Ascend 700 MHz spectrometer (1H at 700 MHz, 13C at 175 MHz) equipped with a 5-mm TXI cryoprobe system (Bruker Biospin GmbH, Germany). Chemical shift values of 1H- and 13C-NMR spectra are reported in ppm relative to the residual solvent signal given as an internal standard. Multiplicities are described using the following abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, b = broad; corrected coupling constants are reported in Hz.

References


Acknowledgements
The authors would like to thank Eva Luxenburger, Dr. Stephan Hüttel and Dr. Thomas Hoffmann (HIPS) for expert assistance with various analytical techniques, Viktoria Schmitt for assistance in biological function experiments and Dr. Jun Fu (Dresden University of Technology) for expert technical assistance in Red/ET recombineering. Research in the laboratory of R.M. was funded by the Deutsche Forschungsgemeinschaft (DFG) and the Bundesministerium für Bildung und Forschung (BMBF). The work in the laboratory of Y. Z. was supported by funding from the Recruitment Program of Global Experts. The authors acknowledge Dr. MA Meredyth Stewart and Dr. Vinothkannan Ravichandran’s help in proofreading this manuscript.

Author Contributions
Q.T., S.H. and Y.Z. planned and performed cloning experiments. Q.T. and X.B. performed genetic transfers, cultivation experiments and data analysis. Q.T. and R.R. performed HPLC and compound isolation. R.R. performed NMR experiments and data analysis. J.H. performed biological functional studies. Q.T., Y.Z. and R.M. designed the study and wrote the paper. All authors discussed the results and commented on the manuscript.

Additional Information
Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Tu, Q. et al. Genetic engineering and heterologous expression of the disorazol biosynthetic gene cluster via Red/ET recombineering. Sci. Rep. 6, 21066; doi: 10.1038/srep21066 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/
Supplemental Information
Genetic Engineering and Heterologous Expression of the Disorazol Biosynthetic Gene Cluster via Red/ET Recombineering

Qiang Tu, Jennifer Herrmann, Shengbiao Hu, Ritesh Raju, Xiaoying Bian, Youming Zhang and Rolf Müller

Inventory of Supplemental Information

Supplemental data
Figure S1. Chemical structures of disorazols.
Structures of all disorazol derivatives mentioned in this paper. On Page 3.

Figure S2. related to Figure 2. Heterologous expression constructs of two type expression plasmids p15A-dis and p15A-dis-est. On Page 9.

Figure S3. related to Figure 3. HPLC-MS analysis of target screening of extracts from M. xanthus:: p15A-dis. On Page 10.

Figure S4. related to Figure 3. 1H NMR spectrum data of disorazol A2. On Page 11.

Figure S5. related to Figure 5. Modify disD gene through Red/ET recombineering. On Page 13.

Figure S6. Construct of the recovered plasmid pTn-Rec_IE2. On Page 7.

Table S1 Oligonucleotides used in this study.

Table S2 related to Figure 3. Target screening analysis data of extracts from M. xanthus::p15A-dis. On Page 10.

Table S3 related to Figure 3. NMR data for disA2 comparison with the natural product. On Page 11.

Table S4 Proteins encoded on the recovered plasmid pTn-Rec_IE-2 and their putative function in disorazol biosynthesis. On Page 5.

Supplemental experimental procedures

Supplemental references
**Figure S1.** Chemical structures of disorazols.

1. Disorazol A₁
2. Disorazol A₂
3. Disorazol A₃
4. Disorazol A₄
5. Disorazol A₅
6. Disorazol A₆
7. Disorazol A₇
8. Disorazol B₁
9. Disorazol B₂
10. Disorazol B₃
11. Disorazol B₄
12. Disorazol C₁
13. Disorazol C₂
14. Disorazol D₁
15. Disorazol D₂
16. Disorazol D₃
17. Disorazol D₄
18. Disorazol D₅
19. Disorazol E₁
20. Disorazol E₂
21. Disorazol E₃
22. Disorazol F₁
23. Disorazol F₂
24. Disorazol F₃
25. Disorazol G₁
26. Disorazol G₂
27. Disorazol G₃
28. Disorazol H
29. Disorazol I
Figure S2. Heterologous expression constructs of two type of expression plasmids p15A-dis and p15A-dis-est.

Construct 1 is p15A-dis. Then we insert the repaired carboxyl esterase gene orf3' and the SAM-dependent methyl transferase gene orf2' together into p15A-dis by Red/ET recombineering to form construct 2 p15A-dis-est. After insertion, two genes are in the middle of orf9 gene and disD gene.
Figure S3. related to Figure 3. HPLC-MS analysis (BPC m/z 720-780) of extracts from *M. xanthus:: p15A-dis.*

The peaks are disB4 (1), m/z 761 [M+H]+; dis762 (2), m/z 763 [M+H]+; disB2 (3), m/z 779 [M+H]+; disA2 (4), m/z 745 [M+H]+; disA1 (5), m/z 759 [M+H]+; disF2 (6), m/z 729 [M+H]+.
Figure S4. $^1$H NMR spectrum of disorazol A$_2$. 
Figure S5. Modify disD gene through Red/ET recombineering.

1: only insert promoter P_{cp25} in front of disD gene.

2: insert promoter P_{cp25} in front of disD gene by deletion orf9 gene.
Figure S6. Construct of the recovered plasmid pTn-Rec_IE2. (data according to Kopp et al.)

Diagram of the genes encoded adjacent to the transposition site of mutant So12_EXI_IE-3 that was cloned into the recovered plasmid pTn-Rec_IE-2. ↓ is the location of the transposon insertion site.
<table>
<thead>
<tr>
<th>name</th>
<th>sequences (5'-3')</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>AATGTCCGCACAATGTGCAGCCATTATTTTC&lt;br&gt;ACATTCAAGTCAACGCTACGCGTGA&lt;br&gt;CGCTGTCCGAGACCGACTTGAGGC&lt;br&gt;TTCCTACCTGCGGATCTCCTGATTGATG&lt;br&gt;GCGGGATCGTTG</td>
<td>Used to replace the backbone of disBAC plasmid</td>
</tr>
<tr>
<td>P2</td>
<td>AATGAATAGTTTGACAAAAATCTAGA&lt;br&gt;TAACGAGGATCAAGATGGACAGGAGCAAG&lt;br&gt;CGCCATTTGCGATCATCGGCATG&lt;br&gt;GCGGAGTTCCCCGAGATCTGTTGAGGT</td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td>TCCGCTGCATCTGGATCTGGATATGG&lt;br&gt;ACAGGAAGGCCTTGAAGAAGCTGAG&lt;br&gt;TTCGAACAAAAAATCTAG</td>
<td>Used to replace the backbone of pTn-Rec_IE2 plasmid</td>
</tr>
<tr>
<td>P4</td>
<td>CTCAGCCCAGTGTCAGGGCGTAGACGC&lt;br&gt;GCCGTTGGGCCGTTGATGACGAG&lt;br&gt;TCTTGACATGCTTAAATACGTTG</td>
<td></td>
</tr>
<tr>
<td>P5</td>
<td>TACCGTCGCGGCCACGCGCAGCAG&lt;br&gt;GCCCGGCTCGATGACGAG&lt;br&gt;TACAACCCGCCGCCAGCTCCACGA&lt;br&gt;GCGGCTGGCGGCCAGGGCTA AGCA&lt;br&gt;CC</td>
<td>Used to replace R6k-Tn-hyg genes in order to form full length of esterase gene</td>
</tr>
<tr>
<td>P6</td>
<td>CAGATCGCTGAGATAGTGCTCCTCAG&lt;br&gt;GATTACGATAAGCTGTTAGAGCAC&lt;br&gt;GGTGCCGGAAAAACATGATGCTTCA&lt;br&gt;TCGGTTGCTTCTCTTATAATTTTTAA&lt;br&gt;TCTGTTA</td>
<td>Used to form the homology arm to p15A-dis plasmid</td>
</tr>
<tr>
<td>P7</td>
<td>CCTCAGCCCAGTGTCAGGGCGTAGAC&lt;br&gt;GCCGTTGGGCCGTTGATGACGAG&lt;br&gt;TCTTGACATGCTTAAATACGTTG</td>
<td></td>
</tr>
<tr>
<td>P8</td>
<td>TGACCTCCGCTGACGACGAGCGC&lt;br&gt;AGCGACATGCTGCACTCTGTTCTGGCTTT&lt;br&gt;ACGGCCCGCCCTGCCACTC</td>
<td>Used to form the homology arm to p15A-dis plasmid</td>
</tr>
<tr>
<td>P9</td>
<td>TCGGCGCTGGCCAGCGCGGAG&lt;br&gt;TGGAGCGACGAGCGGAGCTCG&lt;br&gt;GGCGGCGGCGGCCTGCTGTT&lt;br&gt;TGATACC</td>
<td></td>
</tr>
<tr>
<td>P10</td>
<td>TGGAAGAGCGGATGACCATGAGAG&lt;br&gt;GTGGCAAACTCTGGCTGGGAG</td>
<td>Used to verify M. xanthus::p15A-dis mutant.</td>
</tr>
<tr>
<td>P11</td>
<td>TGTTGGAGGGCCAGGCTGCTGAG&lt;br&gt;AGGGTGAGGAGGCGGCGTTCG</td>
<td></td>
</tr>
</tbody>
</table>
P12  TGGAGGTCCGCCCGATCGCCGAGGGCCGAGCTGCACGAGCGCCTCGCGCGGCA
     GAGCCCTTACGCCCGCCCTGCCACTCATCGCAG
     Used to inactivate disA gene

P13  TGATAGAGAAAAATGGAATGATATGTCGACAAAAATCTAGGAGGATGATGAGC
     CGCCAACTTTTGGCGAAAATGAGAC

P14  CGTAGGACACCCGGTTGGCGATCGACGAGTAATCGGCGCTCGCGATCACGGG
     GTTTCCCTTTGAGCTCGTCCTGTTA
     Used to inactivate disB gene

P15  TGCCCGCGCTCAGGCTATTACCAATCGACCTGGACCAGAAGCGCGCTTTGAACGT
     CGGGGTAACCTCCAAACCTTTTGGGCGAAAAATGAGAC
     Used to inactivate disC gene

P16  CGGCTCGGGTAGCGAGACCGAGCCCGAGCTCAAGAAGGGCCACTGATCCAGGGGG
     AATACCCTTTACGCCCGCCCTGCCACTCATCGCAG

P17  TGGCCGGCGTACCAGGCGAGGAGCTGACTCGGCTCTACGCCATCCTGCAAGAGGAATGATGACCACTTTTGGCGAAAATGAGAC

P18  TGAGACCCCCGGAACATGATGTATCTTCATCGGCTTTCTTTCCTCTCTATCA
     CTGATAGGGAGTGGTAAAATAACTCTATGTGATTATACC
     Used to insert P_{cp25} promoter in front of disD gene

P19  CGCGCCGAGCGAGCCCGCCCTGCCCGGCACGTACGTGGCGGCGCCTCCTCCCCGC
     GAGCCACGTGTGGCGAACAAATTAATCC

P20  TGCGTTTGTAGATAGCGAGCGATCCGATG
     ATAGACGACCCCGCGCTGAAACCGAGCAGTACGTTGACAAATTAATC
     Used to delete orf9 gene by insertion of P_{cp25}
     promoter in front of disD gene
Table S2 related to Figure 3. Target screening analysis data of extracts from *M. xanthus::p15A-dis*.

<table>
<thead>
<tr>
<th>RT [min]</th>
<th>m/z</th>
<th>Compound</th>
<th>sum formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.36</td>
<td>335.16205</td>
<td>Ar001-23-2</td>
<td>C_{14}H_{27}N_{2}O_{5}S_{1}</td>
</tr>
<tr>
<td>12.48</td>
<td>384.28846</td>
<td>AS_DK1622_383-1</td>
<td>C_{25}H_{38}N_{3}O_{2}</td>
</tr>
<tr>
<td>12.48</td>
<td>384.28846</td>
<td>AS_DK1622_383-2</td>
<td>C_{25}H_{38}N_{3}O_{2}</td>
</tr>
<tr>
<td>5.1</td>
<td>631.27543</td>
<td>Cittilin A</td>
<td>C_{34}H_{39}N_{4}O_{8}</td>
</tr>
<tr>
<td><strong>12.81</strong></td>
<td>729.37322</td>
<td><strong>Disorazol F2</strong></td>
<td>C_{42}H_{52}N_{2}O_{9}</td>
</tr>
<tr>
<td>11.58</td>
<td>745.36879</td>
<td><strong>Disorazol A2</strong></td>
<td>C_{42}H_{53}N_{2}O_{10}</td>
</tr>
<tr>
<td><strong>12.24</strong></td>
<td>759.38345</td>
<td><strong>Disorazol A1</strong></td>
<td>C_{43}H_{55}N_{2}O_{10}</td>
</tr>
<tr>
<td><strong>12.24</strong></td>
<td>759.38345</td>
<td><strong>Disorazol A3</strong></td>
<td>C_{43}H_{55}N_{2}O_{10}</td>
</tr>
<tr>
<td><strong>12.24</strong></td>
<td>759.38345</td>
<td><strong>Disorazol A4</strong></td>
<td>C_{43}H_{55}N_{2}O_{10}</td>
</tr>
<tr>
<td><strong>7.77</strong></td>
<td>779.37324</td>
<td><strong>Disorazol B2</strong></td>
<td>C_{42}H_{55}N_{2}O_{12}</td>
</tr>
<tr>
<td><strong>9.49</strong></td>
<td>761.36162</td>
<td><strong>Disorazol B4</strong></td>
<td>C_{42}H_{53}N_{2}O_{11}</td>
</tr>
<tr>
<td><strong>10.04</strong></td>
<td>763.37894</td>
<td><strong>Disorazol 762</strong></td>
<td>C_{42}H_{54}N_{2}O_{11}</td>
</tr>
<tr>
<td>7.44</td>
<td>519.25799</td>
<td>DKxanthen-518</td>
<td>C_{29}H_{35}N_{4}O_{5}</td>
</tr>
<tr>
<td>7.8</td>
<td>519.25783</td>
<td>DKxanthen-518</td>
<td>C_{29}H_{35}N_{4}O_{5}</td>
</tr>
<tr>
<td>7.18</td>
<td>535.25311</td>
<td>DKxanthen-534</td>
<td>C_{29}H_{35}N_{4}O_{6}</td>
</tr>
<tr>
<td>7.54</td>
<td>535.25281</td>
<td>DKxanthen-534</td>
<td>C_{29}H_{35}N_{4}O_{6}</td>
</tr>
<tr>
<td>8.07</td>
<td>549.26825</td>
<td>Dkxanthen-548</td>
<td>C_{30}H_{37}N_{4}O_{6}</td>
</tr>
<tr>
<td>3.56</td>
<td>183.09128</td>
<td>Marinoquinoline A</td>
<td>C_{12}H_{11}N_{2}</td>
</tr>
<tr>
<td>5.36</td>
<td>225.13794</td>
<td>Marinoquinoline B</td>
<td>C_{15}H_{17}N_{2}</td>
</tr>
<tr>
<td>5.84</td>
<td>259.12181</td>
<td>Marinoquinoline C</td>
<td>C_{18}H_{15}N_{2}</td>
</tr>
<tr>
<td>2.55</td>
<td>197.12852</td>
<td>Media compound-Amb-001</td>
<td>C_{10}H_{17}N_{2}O_{2}</td>
</tr>
<tr>
<td>3.56</td>
<td>245.12887</td>
<td>Media compound-Amb-002</td>
<td>C_{14}H_{17}N_{2}O_{2}</td>
</tr>
<tr>
<td>4.01</td>
<td>245.12862</td>
<td>Media compound-Amb-003</td>
<td>C_{14}H_{17}N_{2}O_{2}</td>
</tr>
<tr>
<td>3.51</td>
<td>211.14406</td>
<td>Media compound-Amb-004</td>
<td>C_{11}H_{19}N_{2}O_{2}</td>
</tr>
<tr>
<td>13.01</td>
<td>416.31467</td>
<td>Myxalamid A</td>
<td>C_{26}H_{42}N_{1}O_{3}</td>
</tr>
<tr>
<td>12.59</td>
<td>402.2989</td>
<td>Myxalamid B</td>
<td>C_{25}H_{40}N_{1}O_{3}</td>
</tr>
<tr>
<td>11.84</td>
<td>388.28299</td>
<td>Myxalamid C</td>
<td>C_{24}H_{38}N_{1}O_{3}</td>
</tr>
<tr>
<td>13.83</td>
<td>430.33012</td>
<td>Myxalamid-430</td>
<td>C_{27}H_{44}N_{1}O_{3}</td>
</tr>
<tr>
<td>5.4</td>
<td>405.16444</td>
<td>Myxochelin A</td>
<td>C_{20}H_{25}N_{2}O_{7}</td>
</tr>
<tr>
<td>3.93</td>
<td>404.18101</td>
<td>Myxochelin B</td>
<td>C_{20}H_{26}N_{3}O_{6}</td>
</tr>
<tr>
<td>13</td>
<td>624.4456</td>
<td>Myxovirescin A</td>
<td>C_{35}H_{62}N_{1}O_{8}</td>
</tr>
<tr>
<td>12.73</td>
<td>622.42858</td>
<td>Myxovirescin B</td>
<td>C_{35}H_{60}N_{1}O_{8}</td>
</tr>
<tr>
<td>16.86</td>
<td>632.44647</td>
<td>Myxovirescin C</td>
<td>C_{35}H_{63}N_{1}Na_{1}O_{7}</td>
</tr>
<tr>
<td>11.03</td>
<td>642.45472</td>
<td>Myxovirescin Variante KP641</td>
<td>C_{35}H_{64}N_{1}O_{9}</td>
</tr>
<tr>
<td>2.3</td>
<td>261.12324</td>
<td>Tyr-Pro Dioxopiperazin</td>
<td>C_{14}H_{17}N_{2}O_{3}</td>
</tr>
<tr>
<td>2.41</td>
<td>261.1234</td>
<td>Tyr-Pro Dioxopiperazin</td>
<td>C_{14}H_{17}N_{2}O_{3}</td>
</tr>
</tbody>
</table>

The shadow part are disorazol and its derivatives (in bold in the table).
Table S3 related to Figure 3. NMR data for disA$_2$ comparison with the natural product.

<table>
<thead>
<tr>
<th>pos</th>
<th>$\delta$H, mult ($J$ in Hz) Disorazol$_1$ A$_2$</th>
<th>$\delta$H, mult ($J$ in Hz) Disorazol$_2$ A$_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-H</td>
<td>8.34, s br</td>
<td>8.42, s</td>
</tr>
<tr>
<td>5-H</td>
<td>6.18, d br $(11.7)$</td>
<td>6.17, d $(11.9)$</td>
</tr>
<tr>
<td>6-H</td>
<td>6.47, dd $(11.7, 11.9)$</td>
<td>6.41, m</td>
</tr>
<tr>
<td>7-H</td>
<td>7.36, dd $(11.9, 15)$</td>
<td>7.16, dd $(11.4, 14.7)$</td>
</tr>
<tr>
<td>8-H</td>
<td>5.71, m</td>
<td>5.75, m</td>
</tr>
<tr>
<td>9-H</td>
<td>3.66, dd $(4.2, 9.9)$</td>
<td>3.64, dd $(4.4, 9.9)$</td>
</tr>
<tr>
<td>10-H</td>
<td>4.06, dd $(4.2, 9.7)$</td>
<td>3.96, dd $(4.2, 10.0)$</td>
</tr>
<tr>
<td>11-H</td>
<td>5.34, dd $(9.7, 11.5)$</td>
<td>5.28, dd $(9.0, 11.2)$</td>
</tr>
<tr>
<td>12-H</td>
<td>5.86, ddd $(5.5, 11, 11)$</td>
<td>5.77, m</td>
</tr>
<tr>
<td>13-Ha</td>
<td>2.88, m</td>
<td>2.71, m</td>
</tr>
<tr>
<td>13-Hb</td>
<td>2.46, ddbr</td>
<td>2.43, dd $(3.2, 14.6)$</td>
</tr>
<tr>
<td>14-H</td>
<td>5.35, dd $(2.5, 11)$</td>
<td>5.15, m</td>
</tr>
<tr>
<td>16-H</td>
<td>3.88, d $(7.5)$</td>
<td>3.73</td>
</tr>
<tr>
<td>17-H</td>
<td>5.64, m</td>
<td>5.74, m</td>
</tr>
<tr>
<td>18-H</td>
<td>5.71, m</td>
<td>5.15, m</td>
</tr>
<tr>
<td>19-H$_3$</td>
<td>1.74, d $(1, 6)$</td>
<td>1.65</td>
</tr>
<tr>
<td>20-H$_3$</td>
<td>1.07$^a$, s</td>
<td>0.94$^a$, s</td>
</tr>
<tr>
<td>21-H$_3$</td>
<td>1.03$^a$, s</td>
<td>0.93$^a$, s</td>
</tr>
<tr>
<td>3'-H</td>
<td>8.49, s</td>
<td>8.62, s</td>
</tr>
<tr>
<td>5'-Ha</td>
<td>3.11, dd $(5.4, 14.9)$</td>
<td>2.95, dd $(6.0, 14.7)$</td>
</tr>
<tr>
<td>5'-Hb</td>
<td>2.65, dd $(3.8, 14.9)$</td>
<td>2.43, dd $(3.3, 14.4)$</td>
</tr>
<tr>
<td>6'-H</td>
<td>4.42, m</td>
<td>4.66, m</td>
</tr>
<tr>
<td>7'-H</td>
<td>5.88, dd $(9, 15.1)$</td>
<td>5.93, dd $(9.4, 12.0)$</td>
</tr>
<tr>
<td>8'-H</td>
<td>6.43, dd $(11, 15.1)$</td>
<td>6.39, m</td>
</tr>
<tr>
<td>9'-H</td>
<td>6.01, dd $(11, 11)$</td>
<td>6.17, d $(11)$</td>
</tr>
<tr>
<td>10'-H</td>
<td>6.37, dd $(11, 11)$</td>
<td>6.36, m</td>
</tr>
<tr>
<td>11'-H</td>
<td>6.48, dd $(11, 11)$</td>
<td>6.40, m</td>
</tr>
<tr>
<td>12'-H</td>
<td>5.55, ddd $(5.5, 11, 11)$</td>
<td>5.54, m</td>
</tr>
<tr>
<td>13'-Ha</td>
<td>2.88, m</td>
<td>2.72, m</td>
</tr>
<tr>
<td>13'-Hb</td>
<td>2.36, m</td>
<td>2.26, m</td>
</tr>
<tr>
<td>14'-H</td>
<td>5.35, dd $(2.5, 11.6)$</td>
<td>5.15, m</td>
</tr>
<tr>
<td>16'-H</td>
<td>3.88, d $(7.5)$</td>
<td>3.72, m</td>
</tr>
<tr>
<td>17'-H</td>
<td>5.64, m</td>
<td>5.54, m</td>
</tr>
<tr>
<td>18'-H</td>
<td>5.71, m</td>
<td>5.50, m</td>
</tr>
<tr>
<td>19'-H$_3$</td>
<td>1.74, d</td>
<td>1.65</td>
</tr>
<tr>
<td>20'-H$_3$</td>
<td>1.02$^b$, s</td>
<td>0.85$^b$, s</td>
</tr>
<tr>
<td>21'-H$_3$</td>
<td>1.01$^b$, s</td>
<td>0.87$^b$, s</td>
</tr>
</tbody>
</table>

$^1$NMR data taken in MeOH-$d_4$, $^2$DMSO-$d_6$. $^a,b$ overlapping signals.
Table S4 Proteins encoded on the recovered plasmid pTn-Rec_IE-2 and their putative function in disorazol biosynthesis (data according to Kopp et al.)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Size (bp)</th>
<th>Proposed function of the similar protein</th>
<th>Similarity/ Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>orf1'</td>
<td>522</td>
<td>arylesterase-related protein</td>
<td>29%/ 43%</td>
</tr>
<tr>
<td>orf2'</td>
<td>591</td>
<td>SAM-dependent methyl transferase</td>
<td>48%/ 58%</td>
</tr>
<tr>
<td>orf3'</td>
<td>1284</td>
<td>putative esterase β-lactamase</td>
<td>35%/ 51%</td>
</tr>
<tr>
<td>orf4'</td>
<td>1782</td>
<td>adenylate cyclase</td>
<td>31%/ 51%</td>
</tr>
<tr>
<td>orf5'</td>
<td>854</td>
<td>outer membrane protein (incomplete)</td>
<td>36%/ 46%</td>
</tr>
</tbody>
</table>
Genetic inactivation of the disorazol biosynthetic genes

As mentioned earlier, ten PKS modules and one NRPS module are encoded in the genes *disA-C* in the conserved disorazol biosynthetic gene cluster. To confirm that *disA-C* is involved in the biosynthesis of disorazol, the module was inactivated by disrupting the gene, including the PKS modules 1 and 5 and the NRPS module, on the p15A-dis expression construct (Figure 1). A 1100 bp fragment conferring chloramphenicol resistance (cm) to the linker region between these modules was separately inserted into the p15A-dis construct by homologous recombination in *E. coli*. The modified deletion constructs pDisA, pDisB and pDisC were screened using low-salt LB plates plus chloramphenicol and then verified by restriction analysis (Table S1). The successfully modified constructs were then transformed into *M. xanthus* DK1622, and the production profile of positive recombinants was analyzed by HPLC-MS (Figure S6-I).

The mutant strains no longer produced disorazols. The missing peaks indicate that all these modules and genes are vital for the disorazol biosynthetic pathway. Without any one of these genes, no disorazols are produced. It is, however, still possible that the significant changes in the cluster architecture that were caused by the gene inactivation may have affected the expression of the biosynthetic enzymes (Figure S6-II).
Biological evaluation

Cell lines were obtained from the German Collection of Microorganisms and Cell Cultures (*Deutsche Sammlung für Mikroorganismen und Zellkulturen*, DSMZ) or were part of our internal collection and were cultured under conditions recommended by the depositor. Half-inhibitory concentrations (IC\textsubscript{50}) in terms of growth inhibition were determined as described previously\textsuperscript{1}. In brief, cells were treated in 96-well plates with serial dilutions of disorazol A\textsubscript{1} and A\textsubscript{2} for 5 d. Cell viability was assessed via tetrazolium salt reduction and average IC\textsubscript{50} values were obtained in two independent experiments by sigmoidal curve fitting.
Supplemental references

Publication I

Genetic engineering and heterologous expression of the disorazol biosynthetic gene cluster via Red/ET recombineering.

Qiang Tu, Jennifer Herrmann, Shengbiao Hu, Ritesh Raju, Xiaoying Bian, Youming Zhang & Rolf Müller

Author Contributions

Q.T., S.H. and Y.Z. planned and performed cloning experiments. Q.T. and X.B. performed genetic transfers, cultivation experiments and data analysis. Q.T. and R.R. performed HPLC and compound isolation. R.R. performed NMR experiments and data analysis. J.H. performed biological functional studies. Q.T., Y.Z. and R.M. designed the study and wrote the paper. All authors discussed the results and commented on the manuscript.

Signatures:

Qiang Tu: (Signature)

Jennifer Herrmann: (Signature)

Shengbiao Hu: (Signature)

Ritesh Raju: (Signature)

Xiaoying Bian: (Signature)

Youming Zhang: (Signature)

Rolf Müller: (Signature)
Room temperature electrocompetent bacterial cells improve DNA transformation and recombineering efficiency.

Qiang Tu1,2,4,*, Jia Yin1,3,4,*, Jun Fu1,3, Jennifer Herrmann2, Yuezhong Li1, Yulong Yin4, A. Francis Stewart3#, Rolf Müller2# & Youming Zhang1#

1Shandong University–Helmholtz Institute of Biotechnology, State Key Laboratory of Microbial Technology, School of Life Sciences, Shandong University, Shanda Nanlu 27, 250100 Jinan, People’s Republic of China.

2Department of Microbial Natural Products, Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection Research and Department of Pharmaceutical Biotechnology, Saarland University, Campus E8.1, 66123 Saarbrücken, Germany.

3Department of Genomics, Dresden University of Technology, BioInnovations-Zentrum, Tatzberg 47-51, 01307 Dresden, Germany.

4Animal Nutrition and Human Health Laboratory, College of Life Science of Hunan Normal University, 410081 Changsha, People’s Republic of China.

*These authors contributed equally to this work.

#Correspondence and requests for materials should be addressed to A.F.S. (email: francis.stewart@biotec.tu-dresden.de) or R.M. (email: Rolf.Mueller@helmholtz-hzi.de) or Y.M.Z. (email: zhangyouming@sdu.edu.cn)
Room temperature electrocompetent bacterial cells improve DNA transformation and recombineering efficiency

Qiang Tu1,2,4,*, Jia Yin1,3,4,*, Jun Fu1,3, Jennifer Herrmann2, Yuezhong Li3, Yulong Yin4, A. Francis Stewart3, Rolf Müller2 & Youming Zhang1

Bacterial competent cells are essential for cloning, construction of DNA libraries, and mutagenesis in every molecular biology laboratory. Among various transformation methods, electroporation is found to own the best transformation efficiency. Previous electroporation methods are based on washing and electroporating the bacterial cells in ice-cold condition that make them fragile and prone to death. Here we present simple temperature shift based methods that improve DNA transformation and recombineering efficiency in E. coli and several other gram-negative bacteria thereby economizing time and cost. Increased transformation efficiency of large DNA molecules is a significant advantage that might facilitate the cloning of large fragments from genomic DNA preparations and metagenomics samples.

Usage of various competent cells in different molecular biology techniques such as cloning, amplification of plasmid DNA, construction of genomic libraries, gene expression, and mutagenesis are the routine procedures in each laboratory. Most commonly and extensively used bacterial strain is the Gram-negative bacterium Escherichia coli1,2. E. coli cells can be made competent by washing with divalent cations such as Ca2+ at 0 °C or under ice-cold conditions3,4. However, such metal ion liquids washed competent cells would have lower transformation efficiency than using the electroporation method. In the electroshock methods (electroporation transformation), high-voltage pulse treated E. coli cells become exceptionally competent after washing with ice-cold 10% glycerol or water4–8. The high voltage causes the cellular membrane to be transiently permeabilized, allowing the foreign material to enter into the cells9. High efficient electrocompetent cells are mainly used in library construction, mutagenesis and recombineering10. Protocols for electroporating Gram-negative bacteria including E. coli have already been described by many researchers4–6,9,11,12. Generally, cells are grown up to a suitable density, harvested, and followed by a series of washes to remove culture medium. Several factors have been identified to cause potential impact on the efficiency of electroporation transformation process. These factors include the electrical field strength, pulse decay time, pulse shape, temperature, type of cell, type of suspension buffer, concentration and size of the nucleic acid to be transferred9,13. According to the methods reported earlier, electrocompetent cell preparation have to be performed at ice-cold temperature and the equipment and washing solutions have to be maintained at the same temperature as well14–16.

Recombineering is now an alternative technology for conventional recombinant DNA engineering, a unique tool for large size DNA engineering, as well as the most appealing method of choice for bacterial genome engineering17,18. There are two main recombineering activities: one is based on linear plus circular homologous

1. Shandong University–Helmholtz Institute of Biotechnology, State Key Laboratory of Microbial Technology, School of Life Sciences, Shandong University, Shanda Nanlu 27, 250100 Jinan, People’s Republic of China. 2. Department of Microbial Natural Products, Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection Research and Department of Pharmaceutical Biotechnology, Saarland University, Campus E8.1, 66123 Saarbrücken, Germany. 3. Department of Genomics, Dresden University of Technology, BioInnovations-Zentrum, Tatzberg 47-51, 01307 Dresden, Germany. 4. Animal Nutrition and Human Health Laboratory, College of Life Science of Hunan Normal University, 410081 Changsha, People’s Republic of China. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to A.F.S. (email: francis.stewart@biotec.tu-dresden.de) or R.M. (email: Rolf.Mueller@helmholtz-hzi.de) or Y.M.Z. (email: zhangyouming@sdu.edu.cn)
recombination (LCHR) initiated by the Red operon from λ phage, and the other is the linear plus linear homologous recombination (LLHR) which is initiated by RecE/RecT from Rac phage. LCHR is mainly applied to engineer plasmids which includes Bacterial Artificial Chromosome (BAC) while LLHR is primarily applied to linear DNA cloning (PCR cloning) and direct cloning. Direct cloning is a shortcut for cloning of a large DNA fragments from genomic DNA without library construction and screening. For accomplishing the direct cloning, the DNA segment of interest should meet the linear cloning vector in one cell and then recombine each other. Therefore, the transformation efficiency and homologous recombination efficiency in the RecET proficient cells become the major limitation.

Keeping the cells cold was the pivotal point in the most of the protocols for electroporating Gram-negative bacterial strains including *E. coli* but there was no detailed explanations why this is important. However, an improved transformation efficacy in the pathogen *Pseudomonas aeruginosa* when cells were washed at room temperature (RT) had previously been reported. We surprisingly discovered that electrocompetent cells could be prepared at room temperature so that the cooling steps would be omitted. This was really astonishing because the conventional preparation method of electrocompetent cells for Gram-negative bacteria must be performed at 4°C or preferably at 0–2°C. Additionally, we found that the efficiency of direct cloning which was mediated by RecET recombinering would be dramatically improved by using the electrocompetent cells prepared at room temperature (named as room temperature competent cells). This astonishing discovery permitted the preparation and distribution of electrocompetent cells at a higher temperature. Here we present a novel DNA transformation method that is simplified, fast, efficient, convenient, and cost effective. This simple procedure does not only improve electroporation transformation efficiency in *E. coli* but also has implications for other bacterial hosts, e.g. *Agrobacterium*, *Burkholderia*, *Photorhabdus* and *Xenorhabdus*.

**Results**

**Effect of temperature shift on electrocompetent cells.** It was inconvenient to maintain low temperature conditions for preparation, storage and transport of the electrocompetent cells. We intended to test the transformation efficiency of the electrocompetent cells prepared at room temperature. A large plasmid pGB-amp-Ptet-plu1880 (27.8 kb) was transformed into *E. coli* GB2005 strain at various temperature. The warm electrocompetent cells showed 10 times higher transformation efficiency than the cold electrocompetent cells (Fig. 1a). After placing the cold electrocompetent cells at room temperature for 15 minutes, the transformation efficiency increased by 5 folds (Fig. S1a). In contrast, after the room temperature electrocompetent cells were placed on ice for 15 minutes before electroporation, there was a significant decrease in transformation efficiency (Fig. S1b).

The room temperature in our laboratory was set at 24°C. To determine the range of optimum temperature for the preparation of competent cells, we prepared the cells at different temperature ranges and revealed that the best temperature for electrocompetent cell preparation was in the range of 24°C–28°C (Fig. S2).

**Effect of different plasmids on electrocompetent cells.** Plasmids were varied in the size, selection marker and origins of replication. Initially we tested three plasmids with different sizes. Two of them were p15A origin plasmids with ampicillin (amp) or chloramphenicol (cm) resistance. Another one was a pBR322 origin plasmid with ampicillin resistance. All the plasmids gained higher transformation efficiency with room temperature electrocompetent cells (Fig. 1b, column 1–3). We also tested BAC vectors with different size and selection markers. All BACs gained higher transformation efficiency when room temperature electrocompetent *E. coli* GB2005 cells were used (Fig. 1b, column 4–6). These results indicated that for electrocompetent cell transformation, room temperature electrocompetent cells were more efficient than cold electrocompetent cells irrespective of their size, selection marker and origins of replication. Therefore the room temperature electrocompetent cells could be a better candidate for gene cloning, construction of DNA libraries and mutagenesis than cold electrocompetent ones.

**Effect of different strains on electrocompetent cells.** The *E. coli* GB2005 was an optimized strain for plasmid transformation and propagation along with this strain, several other commonly used *E. coli* strains were also tested for room temperature transformation as well. Results revealed that although different *E. coli* strains varied in their relative transformation efficiencies, all of them exhibited higher transformation efficiency when their electrocompetent cells were prepared at room temperature (Fig. 1c). We also tested the improving approach in a few of other Gram-negative bacterial strains. *Burkholderia glumae* PG1 was an industrial strain for detergent lipase production, which could also be the heterologous host used for PKS/NRPS gene clusters expression (unpublished data). An oriV origin plasmid pRK- apra-kan based on plasmid pBC301, was utilized for transformation. When PG1 competent cells were prepared at room temperature, the electroporation efficiency of RK; plasmid was around three times higher than the cells prepared on ice (Fig. S3a). Other Gram-negative bacterial strains, such as *Agrobacterium*, *Burkholderia*, *Photorhabdus* and *Xenorhabdus*, were tested for RK; plasmid transformation by using room temperature and cold temperature protocols. All the results indicated that room temperature competent cells had higher transformation efficiency than cold competent cells (Fig. S3b).

**Improvement of recombinering by using room temperature electrocompetent cells.** The transformation plasmid transformation efficiency significantly increased by room temperature electrocompetent cells was not the destination. It was necessary to evaluate the improvement of the room temperature protocol on lambda Red or Rac RecET mediated recombinering. A simple assay using a PCR product of linear vector (p15A ori plus cm or pBR322 ori plus cm) and a PCR product with kanamycin (kan) was built to test LLHR efficiency. *E. coli* strain GB05-dir with recET on its chromosome was used for LLHR test. The results showed that LLHR in room temperature competent cells was 6–10 times more efficient than the cells prepared on ice. Both p15A origin and
pBR322 origin plasmids gained the same fold increase (Fig. 2a, b). A direct cloning experiment to fish out the thailandepsin gene cluster (~39 kb) from *Burkholderia thailandensis* had been performed, around 150 colonies were obtained by using cold electrocompetent cells, but by using room temperature electrocompetent cells more than 600 colonies were obtained. This improvement leads to a higher chance to clone large DNA fragments from genomic DNA pools directly.

PCR cloning is a routine exercise in every molecular biology laboratory. It is thus our interest to find out an easy and inexpensive way to clone PCR products. Since the electrocompetent cells prepared at room temperature improves the LLHR efficiency around 10 folds, it is essential to find out the minimum homology sequences needed for recombineering. Previously, we identified 20 bp as the minimum length of sequence homology required for recombineering. To test whether the minimal length could be further shortened, pBAD24 vector was digested with EcoR I/Hind III as linear recipient, and PCR product cassette (Tn5-neo) flanked with short homology arms to the ends of digested pBAD24 vector was used as linear donor fragment (Fig. S4). Seven PCR products with different sizes of homology arms (HA) were designed to test the LLHR efficiency. Results revealed that only 8 bp of terminal homology was sufficient via room temperature protocol (Fig. S5). When ice-cold cells were used, the minimum homology arms required for recombineering were found 12 bp. These data indicated that LLHR might be used to generate a kit for PCR product or small DNA fragment cloning by using homology arms as short as 8 bp.

In contrast to the LLHR experiment, LCHR efficiency was not increasing in the room temperature protocol when compared to the cold protocol (Fig. 2c). However, we discovered that LCHR efficiency would be significantly raised when freshly prepared cold electrocompetent cells were placed at room temperature for 3 minutes (Fig. 2d), suggesting that transient swelling of the cells had a beneficial effect.
Stability of room temperature electrocompetent cells. Normally, after 2.5–3.0 hours cultivation at 37 °C, *E. coli* GB2005 reached OD<sub>600</sub> 0.4–0.6 which was in the log phase, the period with the best transformation efficiency of the cells. When bacterial cells were overgrown, the transformation efficiency dropped down (Fig. S6a), and the transformation efficiency of the cold electrocompetent cells was completely lost after 4 hours or 6 hours (only 18 and 5 colonies respectively) (Fig. S6a). But room temperature electrocompetent cells still kept relatively high efficiency even after 4 or 6 hours cultivation. *E. coli* GB2005 cultured for 4 hours at 37 °C reached OD<sub>600</sub> 1.0 ~ 1.2 and cultured for 6 hours reached OD<sub>600</sub> > 1.8 which was at the plateau phase. It was noteworthy that over-grown or even overnight cultured bacterial cells could still be used for transformation when room temperature protocol was used for preparing competent cells.

To predigest the transformation process, we had tested whether the recovery step could be omitted. For simple plasmid transformation, the recovery step could be omitted when the electrocompetent cells were prepared by using room temperature protocol (Fig. S6b). Although the transformation efficiency in the un-recovery room temperature group was around 30% less than in the recovery room temperature group, it was still at least 5 times higher comparing to the cold temperature group, either recovery or not. Results suggested that plasmid or ligation transformation could be performed in a few minutes after electroporation by using room temperature competent cells. Previous results concluded that the room temperature electrocompetent cells had much better transformation efficiency than cold electrocompetent cells. Furthermore we wanted to know how long the competent cells could stay at room temperature without any significant loss of transformation efficiency. Results showed that room temperature competent cells lost around 30% efficiency after 1 hour storage at room temperature, around 60% lost after 4 hours and around 80% lost after one day (Fig. S7). These results indicated that the room temperature competent cells lost their transformation efficiency to the maximum when stored in room temperature more than one day. To avoid this efficiency loss, room temperature competent cells were prepared by using 10% glycerol<sup>11</sup> and dried by vacuum and stored at 4 °C till three days. Result showed that dried room temperature competent cells prepared in 10% glycerol lost their 55% efficiency as compared to the room temperature competent cells without dry (Table S2). But interestingly, the dried competent cells prepared in 10% glycerol could keep the LLHR efficiency up to 3 days without any further efficiency loss (Table S3). This ability gives us an opportunity in the future to deliver the competent cells in routine cooling pack, which is easier and cost effective.

Electron microscopy analysis of competent cells. To find the reasons of higher efficiency in room temperature protocol, electron microscopy was used for comparative analysis of the morphological shapes of cold competent cells and room temperature competent cells of *E. coli*. Their comparative analysis showed that cold competent cells appeared to shrink more than room temperature cells, and the surface of room temperature competent cells was found smoother (Fig. S6a–d). Shrunken cells might be more difficult to transform, and the bacterial cell membrane and wall could be more permeable for foreign DNA entry at a higher temperature. Additionally, it may be difficult for the shrunken cells to form pores that allow DNA transfer through the cell membrane under
electroporation conditions, and after electroporation most of the cold competent cells were found to be lysed. (Fig. 3c,f). From this we assume that the bacterial cell membrane/cell wall might have better permeability for foreign material to enter into the cell.

**Discussion**

The ability to introduce exogenous DNA molecules into the cells plays key role in the development of molecular biology techniques, such as mutagenesis and genetic engineering of microorganisms. Several methods have been reported to introduce exogenous DNA molecules into the cells which includes chemical treatment, electroporation, utilization of a biolistic gun, polyethylene glycol, ultrasound, microwave, and hydrogel. In those methods, electroporation has been often demonstrated to be more efficient and convenient way to transform a large number of microorganisms used for genetic studies, and many efforts have been performed to increase its efficiency.

The phage-derived homologous recombination systems have been developed into very useful DNA engineering technologies, well known as recombineering which has also been performed in electrocompetent cells. This suggests that a crucial step in recombineering is the transformation of *E. coli* by electroporation.

In conventional electroporation transformation, the electrocompetent cells were prepared on ice and the other supplies were also in cold environment, including pre-chilled cuvette, buffer, and centrifuge. The cells must be repeatedly washed before electroporation to remove conductive solutes. If the conductivity of a cell mixture is too high, then arcing will occur during electroporation, which will ruin the experiment. The washing process can elicit a stress response that can lead to decrease in transformation efficiency. If the cells are kept at 4 °C then they are inactive and this stress response is prevented. However, current studies reveals that the electroporation transformation efficiency is decreased at ice-cold temperature (Fig. 1). This decreased efficiency might be due to ice cold temperature which alters the cell membrane topology. The cell membrane mainly consists of phospholipids and proteins and the phospholipid bilayer forms a stable barrier between two aqueous compartments. Embedded proteins of phospholipid bilayer carry out the specific functions of the plasma membrane, including the selective transportation of molecules across the membrane and cell-cell recognition. At ice-cold temperature, the fatty acid tails of the phospholipids become more rigid. This affects the fluidity, permeability, and the cell's ability to live. Therefore, cold temperatures may be not favourable for the survival and thus decreases the transformation efficiency. Additionally, during electroporation process due to externally applied electric field there is a significant increase in the permeability of cell's plasma membrane, which is used to introduce exogenous DNA into bacterial cell. Previous reports revealed that if the environmental conditions were changed, including the temperature, the cell membranes undergoes a gross morphological changes. These structural perturbations were associated with characteristic disturbances of functions such as loss of selective permeability. Similar results were observed in this study that cold competent cells were appeared to shrink more than room temperature cells, and additionally more cold competent cells were found lysed after electroporation (Fig. 3).

The temperature effects on electroporation transformation could be explained by thermal effects during electro pore formation. According to the electroporation theory, hydrophobic pores in the cell membrane were formed spontaneously by lateral thermal fluctuations of the lipid molecules, which suggested that hydrophobic pores formation would be enhanced by increased temperature conditions. To improve recombination efficiency many parameters had been described previously except the transformation efficiency. This study showed that LLHR efficiency in room temperature competent cells was higher than in the same cells prepared on ice (Fig. 2a,b), but the room temperature protocol did not increase LCHR efficiency when compared to the cold protocol (Fig. 2c). The Red recombinases (Red alpha and beta) might be not stable while preparing the competent cells at room temperature.

In conclusion, this study reports an unexpected finding, that is contrary to common assumption, that it is better to prepare bacterial cells at room temperature than on ice for electroporation. In addition, this study also
shows that this is not only efficient for *E. coli* but also for several other gram-negative and gram-positive hosts. However, further research will be essential to confirm the transfer and principle of membrane in competent cells.

**Methods**

**Strains, plasmids and reagents.** The bacterial strains and plasmids used in this study were listed in Table S1. The antibiotics were purchased from Invitrogen. *E. coli*, *Agrobacterium*, *Photorhabdus* and *Xenorhabdus* were cultured in Luria–Bertani (LB) broth or on LB agar plates (1.2% agar) with ampicillin [amp] (100 μg/mL), kanamycin [kan] (15 μg/mL) or chloramphenicol [cm] (15 μg/mL) as required. *Burkholderia glumae* PG1 was cultured in MME medium (5 g/L K2HPO4, 1.75 g/L Na(NH4)HPO4 × 4H2O, 1 g/L Citrate, 0.1 g/L MgSO4 × 7H2O, 8 g/L Glucose; pH 7.0). *Burkholderia* DSM7029 was cultured in CYCG medium (6 g/L Casitone, 1.4 g/L CaCl2 × 2H2O, 2 g/L Yeast extract and 20 mL/L Glycerol).

**Preparation of electrocompetent cell at cold and room temperature conditions.** The electrocompetent cells at cold temperature were prepared according to the protocol established previously in our lab. For electrocompetent cells at room temperature, overnight culture were diluted into 1.4 mL LB medium and again cultured at 37 °C at 900 rpm in an Eppendorf ThermoMixer. After 2 hours of incubation when OD600 was approximately reached up to 0.6, the bacterial cells were centrifuged at 9000 rpm at room temperature (24 °C). The supernatant was then discarded and the cells were resuspended in 1 mL of dH2O at room temperature, and washing step was repeated. The bacterial cells were again resuspended in about 30 μL of dH2O (24 °C) and the tubes were placed at room temperature. 300 ng of the each plasmid DNA or PCR products were added into the prepared cells. The DNA-cell mixture were then transferred into 1 mm-gap cuvette (24 °C) for electroporation at 1250 volts. The cuvette was then flushed with 1 mL fresh medium and the cells were recovered by the incubation at 37 °C for 1 hour. In the end, the culture was streaked on the LB plates containing appropriate antibiotics.

**Preparation of electrocompetent cell to test the effect of different temperature range.** *E. coli* GB2005 strain was cultured at 37 °C till OD600 was reached at 0.6. The cells were pelleted and washed by dH2O at different temperature range (2, 15, 20, 22, 24, 26, 28, 30, 32, 34 and 37 °C). The cuvettes were also kept at these temperatures. After electroporation with pGB-amp-Ptet-plu1880 plasmid, 1 mL LB was added into the cuvette to recover the transformed cells and then incubated at 37 °C. After 1 hour incubation 0.004 μL of cells (diluted by fresh LB) were streaked on LB plates containing ampicillin (100 μg/mL). The colonies were counted after 24 hours of cultivation.

**Recombineering assays.** In LCHR assay we used a 2 kb p15A-cm plasmid carrying the chloramphenicol resistance gene and a 2 kb kan-PCR product carrying the kanamycin resistance gene. Each end of the kan-PCR product had a 50-bp homology arm to the p15A-cm plasmid between the chloramphenicol gene (cm) and the p15A origin. The circular plasmid (200 ng) and the PCR product (200 ng) were co-electroporated into *E. coli* GB2005 expressing the lambda Red recombinase to generate the chloramphenicol plus kanamycin-resistant plasmid p15A-cm-kan (4 kb). The expression plasmid was pSC101-BAD-gbaA-tet. The recombinants were selected on LB plates with double antibiotic selection.

One of LLHR assays was just like above mentioned setup except the p15A-cm plasmid was linearized between the two 50-bp homology arms. The linear plasmid backbone (200 ng) and the kan-PCR product (200 ng) were co-electroporated into *E. coli* GB2005 expressing the recET recombinase to generate the plasmid p15A-cm-kan. Another LLHR assay used EcoR I and Hind III digested 4.5 kb pBAD24 plasmid (450 ng) and 1.7 kb Tn5-neo PCR (170 ng) to generate the ampicillin plus kanamycin-resistant plasmid pBAD24-neo (6.2 kb). The expression plasmid was pSC101-BAD-ETgA-tet. The recombinants were selected on LB plates with double antibiotic selection.

Both kan-PCR and Tn5-neo-PCR were amplified from suicide R6K plasmid to avoid the selection background from the carryover of the PCR template (any residual circular plasmid). The negative control with DNA electroporation into uninduced cells was done to indicate the sufficient selection pressure. After colony counting, 8 clones from each of the triplicate experiments were picked up for plasmid DNA preparation and restriction analysis to prove the successful accomplishment of the recombineering experiment.

**Preparation of dried room temperature electrocompetent cell.** The electrocompetent cells were washed twice with dH2O or 10% glycerol at room temperature, cells were pelleted once again and remaining dH2O or 10% glycerol was removed by pipetting. Cell pellet was dried under the vacuum for 30 min and stored at 4 °C. For transformation, dried cells were resuspended in 25 μL dH2O (without glycerol) at room temperature and DNA was added into the cells. Cells and DNA were electroporated at 1250 v by using 1 mm-gap electroporation cuvette and Eppendorf electroporator as usual.

**Cells preparations for Electron Microscope.** *E. coli* cells were harvested by centrifugation and were fixed in 2% paraformaldehyde/1% glutaraldehyde for 20 min at room temperature. After repeated washing with ultrapure water the cell pellet was resuspended and a small aliquot of the samples in water was placed on a silicon wafer and was dried under ambient conditions. Next, the *E. coli* cells were investigated with secondary electrons under high-vacuum conditions in an ESEM type FEI Quanta400 FEG at 5 kV accelerating voltage.

**References**


Acknowledgements
This research was supported by the funding of Y. Z. from the Recruitment Program of Global Experts and Shandong Innovation and Transformation of Achievements Grant (2014ZZCX02601) and the Program of Introducing Talents of Discipline to Universities (B16030), funding of R.M. from the Deutsche Forschungsgemeinschaft (DFG) and the Bundesministerium für Bildung und Forschung (BMBF), funding of A. F. S. from the TUD Elite University Support the Best program, funding of J. F. from the International S&T Cooperation Program of China (ISTCP 2015DFE32850) and National Natural Science Foundation of China (31570094), funding of J. Y. from China/Shandong University International Postdoctoral Exchange Program and China Postdoctoral Science Foundation (2015M582081). The authors acknowledge Mr. Nazeer Abbasi’s help in proofreading this manuscript and Dr. Marcus Koch for ESEM scanning.

Author Contributions
Q.T. and J.Y. participated in the design of this study, performed data collection analysis, and drafted the manuscript; J.F, J.H, Y.L. and Y.Y. participated in interpretation data; A.F.S. and R.M. gave the advice for experimental design and discussed the data, also helped in the revision of the final manuscript. Y.Z. designed and oversaw the study, performed data interpretation and drafted the manuscript. All authors read and approved the final manuscript.

Additional Information
Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Tu, Q. et al. Room temperature electrocompetent bacterial cells improve DNA transformation and recombineering efficiency. Sci. Rep. 6, 24648; doi: 10.1038/srep24648 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/
Supplemental Information

Room temperature electrocompetent bacterial cells improve DNA transformation and recombineering efficiency

Qiang Tu, Jia Yin, Jun Fu, Jennifer Herrmann, Yuezhong Li, Yulong Yin, A. Francis Stewart, Rolf Müller and Youming Zhang

Inventory of Supplemental Information

Supplemental data

Figure S1. Temperature shift effect on the competent cells for transformation. On Page 5.

Figure S2. Effect of different temperature on electrocompetent cells. On Page 5.

Figure S3. Transformation efficiency comparison of warm and cold temperature in different gram-negative strains. On Page 6.

Figure S4. Diagram of LLHR by using short homology arms. On Page 7.

Figure S5. Effect of the length of homology arms on room temperature electrocompetent cells (warm cells). On Page 7.

Figure S6. Effect of over-grown cells on transformation efficiency and electroporation without recovery step. On Page 8.

Figure S7. Stability of room temperature electrocompetent cells stored at room temperature. On Page 8.

Table S1 Strains and plasmids.

Table S2 Transformation efficiency (colonies on plates with ampicillin (x10^4)) using cells prepared in dH2O or 10% glycerol. On Page 9.

Table S3 LLHR efficiency (colonies on plates with ampicillin and kanamycin) using cells prepared in dH2O or 10% glycerol. On Page 9.

Supplemental references
Figure S1. Temperature shift effect on the competent cells for transformation. (a) GB2005 cells transformed by ~0.1µg of pGB-Ptet-plu1880 (27.8kb) and plated on Amp plates. 1 -cells prepared at RT; 2 -cells prepared on ice; 3 -cells prepared on ice first then left at RT for 2.5min before electroporation; 4 -same as 3 but at RT for 4min; 5 -same as 4 but at RT for 10min; 6 -same as 3 but at RT for 15min. (b) 1 -cells prepared at RT; 2 -cells prepared at RT, then placed on ice for 15min before electroporation. Error bars, SD; n = 3.
Figure S2. Effect of different temperature on electrocompetent cells. GB2005 cells were transformed by ~0.1µg of pGB-Ptet-plu1880 (27.8kb) and plated on Amp plates. Ice to 37°C were used for preparing competent cells and electroporation. It shows the results in relative transformation efficiency using the transformants at 24°C (room temperature) as standard (100%). Transformants from different temperatures were divided by standard to give the relative transformation efficiency. It also shows that significant results were obtained by preparing competent cells between 24°C-28°C. This confirms that preparation of electrocompetent cells can be made as simple as possible. Error bars, SD; n = 3.
Figure S3. Transformation efficiency comparison of warm and cold temperature in different gram-negative strains. (a) pRK2-apra-km plasmid was used to transform into Burkholderia PG1. The transformants were Km resistant. (b) A few bacterial strains: Agrobacterium (G⁻), Burkholderia DSM7029 (G⁺), Photorhabdus (G⁻), and Xenorhabdus (G⁻) were used to perform the transformation experiment. Error bars, SD; n = 3.
Figure S4. Diagram of LLHR by using short homology arms. pBAD24 circle vector digested by EcoRI plus HindIII was used as linear recipient vector. The homology sequences are exactly exposed at the ends. Tn5-neo PCR product flanked with short homology arms to the ends of digested pBAD24 vector is used as linear donor fragment.
Figure S5. Effect of the length of homology arms on room temperature electrocompetent cells (warm cells). Seven PCR products with different homology arms (HA) were used for testing the LLHR efficiency. The homology arms can be as short as 8bp for LLHR to occur when cells were prepared at RT. When ice-cold cells used, the minimum homology arms are 12bp. Error bars, SD; n = 3.
Figure S6. Effect of over-grown cells on transformation efficiency and electroporation without recovery step. (a) 35µL overnight cultured GB2005 cells were diluted into 1.4mL LB medium and cultured at 37°C for different time courses. Electrocompetent cells were transformed by 0.1µg of pGB-Ptet-plu1880 and plated on LB plates plus amp. 1 -cells growing for 2.5 hours, OD600=0.4; 2 -cells growing for 4 hours, OD600=1.2; 3 -cells growing for 6 hours, OD600=1.8. (b) GB2005 cells transformed by 0.1µg of pGB-Ptet-plu1880 and plated on Amp plates. 1 -cells plated directly after electroporation; 2 - same as 1 but after 1 hour recovery at 37°C. Error bars, SD; n = 3.
Figure S7. Stability of room temperature electrocompetent cells stored at room temperature. The room temperature competent cells lost around 30% of efficiency after 1 hour of storage at room temperature, 60% after 4 hours and approximately 80% after 1 day. Error bars, SD; n = 3.
Table S1 Strains and plasmids.

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Characteristics</th>
<th>References or sources</th>
</tr>
</thead>
</table>
| E. coli GB05      | F-mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ rpsL
|                   |                 |                      |
|                   |                  |                      |
| E. coli GB05-dir  | GB2005, araC-BAD-ETγA | 2                     |
| E. coli GB05-red  | GB2005, araC-BAD-γβαA | 1                     |
| Burkholderia glumae PG1 | lipidase -producing wild-type strain, host for heterologous expression of PKS/NRPS gene clusters | 3                     |
| Agrobacterium tumefaciens | gram-positive strain | 4                     |
| Burkholderia DSM7029 | gram-negative strain | 3                     |
| Photorhabdus luminescens | gram-negative strain | 5                     |
| Xenorhabdus stockiae | gram-negative strain | 5                     |
| Plasmid           |                 |                      |
| pGB-amp-Ptet-plu1880 | pBR322 replicon, ampR | 6                     |
| pRK2-apra-km      | oriV origin, kmR | This study           |
| pBC301            | oriV origin     |                      |
| pBeloBAC11-dis    | BAC, kmR        | This study           |
| p15A-cm           | p15A replicon, cmR | 2                     |
| p15A-cm-km        | p15A replicon, cmR, kmR | 2                  |
| pBAD24            | pBR322 replicon, ampR | 9                    |
| pBAD24-neo        | pBR322 replicon, ampR, kmR | This study |
Table S2 Transformation efficiency (colonies on plates with ampicillin (x10^4)) using cells prepared in dH2O or 10% glycerol.

<table>
<thead>
<tr>
<th></th>
<th>Cells before dry</th>
<th>Dried cells day 0</th>
<th>Dried cells day 1</th>
<th>Dried cells day 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>dH2O</td>
<td>640</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10% glycerol</td>
<td>468</td>
<td>196</td>
<td>212</td>
<td>188</td>
</tr>
</tbody>
</table>

Table S3 LLHR efficiency (colonies on plates with ampicillin and kanamycin) using cells prepared in dH2O or 10% glycerol.

<table>
<thead>
<tr>
<th></th>
<th>Cells before dry</th>
<th>Dried cells day 0</th>
<th>Dried cells day 1</th>
<th>Dried cells day 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>dH2O</td>
<td>420</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10% glycerol</td>
<td>360</td>
<td>298</td>
<td>272</td>
<td>284</td>
</tr>
</tbody>
</table>

References

Publication II

Room temperature electrocompetent bacterial cells improve DNA transformation and recombineering efficiency.

Qiang Tu*, Jia Yin*, Jun Fu, Jennifer Herrmann, Yuezhong Li, Yulong Yin, A. Francis Stewart†, Rolf Müller† and Youming Zhang†

Author Contributions

Q.T. and J.Y. participated in the design of this study, performed data collection analysis, and drafted the manuscript; J.F, J.H, Y.L. and Y.Y. participated in interpretation data; A.F.S. and R.M. gave the advice for experimental design and discussed the data, also helped in the revision of the final manuscript. Y.Z. designed and oversaw the study, performed data interpretation and drafted the manuscript. All authors read and approved the final manuscript.

Signatures:

Qiang Tu: 

Jia Yin: 

Jun Fu: 

Jennifer Herrmann: 

Yuezhong Li: 

Yulong Yin: 

A. Francis Stewart: 

Rolf Müller: 

Youming Zhang: 
Biosynthesis of magnetic nanostructures in a foreign organism by transfer of bacterial magnetosome gene clusters.

Isabel Kolinko¹, Anna Lohße¹, Sarah Borg¹, Oliver Raschdorf³⁻², Christian Jogler¹, †, Qiang Tu³⁻⁴, Mihály Pósfai⁵, Éva Tompa⁵, Jürgen M. Plitzko²⁻⁶, Andreas Brachmann¹, Gerhard Wanner¹, Rolf Müller³, Youming Zhang⁴* and Dirk Schüler¹*

¹Ludwig-Maximilians-Universität München, Department of Biology I, Großhaderner Straße 2-4, 82152 Martinsried, Germany.
²Max Planck Institute of Biochemistry, Department of Molecular Structural Biology, Am Klopferspitz 18, 82152 Martinsried, Germany.
³Helmholtz Institute for Pharmaceutical Research Saarland, Helmholtz Centre for Infection Research and Department of Pharmaceutical Biotechnology, Saarland University, PO Box 151150, 66041 Saarbrücken, Germany.
⁴Shandong University – Helmholtz Joint Institute of Biotechnology, State Key Laboratory of Microbial Technology, Life Science College, Shandong University, Jinan 250100, China.
⁵University of Pannonia, Department of Earth and Environmental Sciences, Veszprém, H-8200 Hungary.
⁶Bijvoet Center for Biomolecular Research, Utrecht University, 3584 CH Utrecht, The Netherlands.

†Present address: Leibniz Institute DSMZ, Department of Microbial Cell Biology and Genetics, Inhoffenstraße 7B, 38124 Braunschweig, Germany.

*e-mail: dirk.schueler@lmu.de; zhangyouming@sdu.edu.cn
Biosynthesis of magnetic nanostructures in a foreign organism by transfer of bacterial magnetosome gene clusters

Isabel Kolinko, Anna Lohse, Sarah Borg, Oliver Raschdorf, Christian Jäger, Qiang Tu, Mihály Pósfai, Éva Tompa, Jürgen M. Plitzko, Andreas Brachmann, Gerhard Wanner, Rolf Müller, Youming Zhang and Dirk Schüler

The synthetic production of monodisperse single magnetic domain nanoparticles at ambient temperature is challenging. In nature, magnetosomes—membrane-bound magnetic nanocrystals with unprecedented magnetic properties—can be biomineralized by magnetotactic bacteria. However, these microbes are difficult to handle. Expression of the underlying biosynthetic pathway from these fastidious microorganisms within other organisms could therefore greatly expand their nanotechnological and biomedical applications. So far, this has been hindered by the structural and genetic complexity of the magnetosome organelle and insufficient knowledge of the biosynthetic functions involved. Here, we show that the ability to biomineralize highly ordered magnetic nanostructures can be transferred to a foreign recipient. Expression of a minimal set of genes from the magnetotactic bacterium *Magnetospirillum gryphiswaldense* resulted in the production of monodisperse single magnetic nanoparticles, as well as their chain-like organization and biocompatibility. The synthetic cassettes comprising the *M. gryphiswaldense* magnetosome biosynthesis within the photosynthetic model organism *Rhodospirillum rubrum*. Our findings will enable the sustainable production of tailored magnetic nanostructures in biotechnologically relevant hosts and represent a step towards the endogenous magnetization of various organisms by synthetic biology.

The alphaproteobacterium *M. gryphiswaldense* produces uniform nanosized crystals of magnetite (Fe₃O₄), which can be engineered by genetic and metabolic means and are inherently biocompatible. The stepwise biogenesis of magnetosomes involves the invagination of vesicles from the cytoplasmic membrane, magnetosomal uptake of iron, and redox-controlled biomineralization of magnetite crystals, as well as their self-assembly into nanochains along a dedicated cytoskeletal structure to achieve one of the highest structural levels in a prokaryotic cell. We recently discovered genes controlling magnetosome synthesis to be clustered within a larger (115 kb) genomic magnetosome island, in which they are interspersed by numerous genes of unrelated or unknown functions. Although the smaller *mamGFDC*, *mms6* and *mamXY* operons have accessory roles in the biomineralization of properly sized and shaped crystals, only the large *mamAB* operon encodes factors essential for iron transport, magnetosome membrane (MM) biogenesis, and crystallization of magnetite particles, as well as their chain-like organization and intracellular positioning. However, it has been unknown whether this gene set is sufficient for autonomous expression of magnetosome biosynthesis.

Using recombineering (recombinogenic engineering) based on phage-derived Red/ET homologous recombination, we stitched together several modular expression cassettes comprising all 29 genes (26 kb in total) of the four operons in various combinations (Supplementary Fig. 1), but lacking the tubulin-like *ftsZm* genes. This gene was omitted from its native *mamXY* operon because of its known interference with cell division during cloning. Regions 200–400 bp upstream of all operons were retained to ensure transcription from native promoters. Transposable expression cassettes comprising the MycoMar (*tps*) or Tn5 transposase gene, two corresponding inverted repeats, the origin of transfer *oriT*, and an antibiotic resistance gene were utilized to enable transfer and random chromosomal integration in single copy (Supplementary Tables 3 and 4). Chromosomal reintegration of all cassettes into different non-magnetic single-gene and operon deletion strains of *M. gryphiswaldense* resulted in stable wild type-like restoration of magnetosome biomineralization, indicating that transferred operons maintained functionality upon cloning and transfer (Supplementary Fig. 2).

We next attempted the transfer of expression cassettes to a foreign non-magnetic host organism (Fig. 1). We chose the photosynthetic alphaproteobacterium *R. rubrum* as a first model because of its biotechnological relevance and relatively close relationship to *M. gryphiswaldense* (16S rRNA similarity to *M. gryphiswaldense* = 90%). Although the *mamAB* operon alone has been shown to support some rudimentary biomineralization in *M. gryphiswaldense*, neither genomic insertion of the *mamAB* operon alone (pTps_AB) nor in combination with the accessory *mamGFDC* genes (pTps_ABG) had any detectable phenotypic effect (Supplementary Table 1). We also failed to detect a magnetic response (Cmag) in the classical light scattering assay after insertion of pTps_ABG6 (*mamAB + mamGFDC + mms6*). However, the cellular iron content of *R. rubrum* ABG6 increased 2.4-fold compared with the untransformed wild type (Supplementary Table 1). Transmission electron microscopy (TEM) revealed a loose chain
of small (~12 nm) irregularly shaped electron-dense particles (Fig. 2a,ii), identified as poorly crystalline hematite (Fe₂O₃) by analysis of the lattice spacings in high-resolution TEM images (Supplementary Fig. 3), much as in the hematite particles previously identified in *M. gryphiswaldense* mutants affected in crystal formation. To further enhance biomineralization, we next transferred pTPs_XYZ, an insertional plasmid harbouring *mamX*, *Y* and *Z* from the *mam*XY operon, into *R. rubrum_ABG6* (Supplementary Fig. 1). The resulting strain *ABG6* encompassed all 29 relevant genes of the magnetosome island except *ftsZm*. Intriguingly, cells of *ABG6* exhibited a significant magnetic response (Supplementary Table 1) and were ‘magnetotactic’, that is, within several hours accumulated a significant magnetic response (Supplementary Table 1) and were referred to as ‘magnetosome’ and ‘magnetosome’ was essential for magnetosome synthesis. *R. rubrum_ABG6X-dJ* failed to express magnetosome particles (Supplementary Fig. 10), which phenocopied a *mamJ* deletion in the related *M. magneticum* another tested example was *MamJ*, which is assumed to connect magnetosome particles to the cytoskeletal magnetosome filament formed by the actin-like *MamK*. Much as in *M. gryphiswaldense*, deletion of *mamJ* caused agglomeration of magnetosome crystals in ~65% of *R. rubrum_ABG6X-dJ* cells (Fig. 2a,v, Supplementary Fig. 10 and Table 1). Together, these observations indicate that magnetosome biogenesis and assembly within the foreign host are governed by very similar mechanisms and structures as in the donor, which are conferred by the transferred genes.

As magnetosomes in *R. rubrum_ABG6X* were still smaller than those of *M. gryphiswaldense*, we wondered whether full expression of biomineralization may depend on the presence of further auxiliary functions, possibly encoded outside the canonical magnetosome operons. For instance, deletion of *feoB1* encoding a constituent of a ferrous iron transport system specific for magnetotactic bacteria caused fewer and smaller magnetosomes in *M. gryphiswaldense*. Strikingly, insertion of *feoA1* into *R. rubrum* strain *ABG6X* resulted in even larger, single-crystalline and twinned magnetosomes and longer chains (440 nm) (Fig. 2a,vi, Table 1). The size (37 nm) of the crystals approached that of the canonical magnetosystems operons. For instance, deletion of *feoB1* encoding a constituent of a ferrous iron transport system specific for magnetotactic bacteria caused fewer and smaller magnetosomes in *M. gryphiswaldense*. Strikingly, insertion of *feoB1* into *R. rubrum* strain *ABG6X* resulted in even larger, single-crystalline and twinned magnetosomes and longer chains (440 nm) (Fig. 2a,vi, Supplementary Table 1). The size (37 nm) of the crystals approached that of the donor, and cellular iron content was substantially increased (0.28% of dry weight) compared with *R. rubrum_ABG6X* (0.18%), although still lower than in *M. gryphiswaldense* (3.5%), partly because of the considerably larger volume of *R. rubrum* cells (Fig. 2c).

Magnetosome particles could be purified from disrupted cells by magnetic separation and centrifugation and formed stable suspensions (Fig. 3). Isolated crystals were clearly enclosed by a layer of organic material resembling the MM attached to magnetosomes of *M. gryphiswaldense*. Smaller, immature crystals were surrounded by partially empty vesicles (Fig. 3c, inset), which were also seen in thin-sectioned cells (Supplementary Fig. 8) and on average were smaller (66 ± 6 nm) than the abundant photosynthetic intracytoplasmic membranes (ICMs) (93 ± 34 nm; Fig. 3a, Supplementary Fig. 8).

Organic material of the putative MM could be solubilized from isolated magnetite crystals of *R. rubrum_ABG6X* by various detergents (Fig. 3d), in a similar manner to that reported for MM of
Figure 2 | Phenotypes of R. rubrum strains expressing different magnetosome gene clusters and auxiliary genes. a, TEM images: R. rubrum wild type (i), containing a larger phosphate inclusion (P) and some small, non-crystalline, electron-dense particles; R. rubrum_ABG6 (ii); R. rubrum_ABG6X (iii); R. rubrum_ABG6X_ftsZm (iv); R. rubrum_ABG6X_dj (v); R. rubrum_ABG6X_feo (vi). Insets: Magnifications of non-crystalline electron-dense particles (i) or heterologously expressed nanocrystals (ii–vi). All insets are of the same particles/crystals as in their respective main images, except for (v). For further TEM micrographs see Supplementary Fig. 10. b, Unlike the untransformed R. rubrum wild type, cells of R. rubrum_ABG6X accumulated as a visible red spot near the pole of a permanent magnet at the edge of a culture flask. c, TEM micrograph of a mixed culture of the donor M. gryphiswaldense and the recipient R. rubrum_ABG6X_feo, illustrating characteristic cell properties and magnetosome organization. Insets: Magnifications of magnetosomes from M. gryphiswaldense and R. rubrum_ABG6X_feo. d, High-resolution TEM lattice image of a twinned crystal from R. rubrum_ABG6X, with Fourier transforms (i) and (ii) showing intensity maxima consistent with the structure of magnetite.

M. gryphiswaldense23. Proteomic analysis of the SDS-solubilized MM revealed a complex composition (Supplementary Fig. 6), and several genuine magnetosome proteins (MamKCJAFDMBYOE, Mms6, MmsF) were detected among the most abundant polypeptides (Supplementary Table 2). An antibody against MamC, the most abundant protein in the MM of M. gryphiswaldense23, also recognized a prominent band with the expected mass (12.4 kDa) in the MM of R. rubrum_ABG6X (Supplementary Fig. 6).

The subcellular localization of selected magnetosome proteins in R. rubrum depended on the presence of further determinants encoded by the transferred genes. For example, MamC tagged with a green fluorescent protein, which is commonly used as magnetosome chain marker in M. gryphiswaldense24, displayed a punctuate pattern in the R. rubrum wild type background. In contrast, a filamentous fluorescent signal became apparent in the majority of cells (79%) of the R. rubrum_ABG6X background, in which the full complement of magnetosome genes are present (Supplementary Fig. 7), reminiscent of the magnetosome-chain localization of these proteins in M. gryphiswaldense24.

Our findings demonstrate that one of the most complex prokaryotic structures can be functionally reconstituted within a foreign, hitherto non-magnetic host by balanced expression of a multitude of structural and catalytic membrane-associated factors. This also provides the first experimental evidence that the magnetotactic trait can be disseminated to different species by only a single event, or a few events, of transfer, which are likely to occur also under natural conditions by horizontal gene transfer as speculated before18,25,26.

The precise functions of many of the transferred genes have remained elusive in native magnetotactic bacteria, but our results will now enable the dissection and engineering of the entire pathway in genetically more amenable hosts. The approximately 30 transferred magnetosome genes constitute an autonomous expression unit that is sufficient to transplant controlled synthesis of magnetite nanocrystals and their self-assembly within a foreign organism. However, further auxiliary functions encoded outside the mam and mms operons are necessary for biomineralization of donor-like magnetosomes. Nevertheless, this minimal gene set is likely to shrink further as a result of systematic reduction approaches in different hosts.

Importantly, the results are promising for the sustainable production of magnetic nanoparticles in biotechnologically relevant photosynthetic hosts. Previous attempts to magnetize both prokaryotic and eukaryotic cells by genetic and metabolic means (for example, refs 27,28) resulted in only irregular and poorly crystalline iron deposits. This prompted ideas to borrow genetic parts of the bacterial magnetosome pathway for the synthesis of magnetic nanoparticles within cells of other organisms29. Our results now set the stage for synthetic biology approaches to genetically endow both uni- and multicellular organisms with magnetization by biomineralization of tailored magnetic nanostructures. This might be exploited for instance in nanomagnetic actuators or in situ heat
generators in the emerging field of magnetogenetics\textsuperscript{30}, or for endogenous expression of magnetic reporters for bioimaging\textsuperscript{31}.

**Methods**

**Bacterial strains, media and cultivation.** The bacterial strains are described in Supplementary Table 4. E. coli strains were cultivated as previously described\textsuperscript{13}. A volume of 1 mM \( \alpha\)-aminoadipic acid was added for the growth of auxotrophic strains BW29427 and WM3064. M. gryphiswaldense strains were cultivated in flask standard medium (FSM), in liquid or on plates solidified by 1.5% agar, and incubated at 30 °C under microoxic (1% \( \text{O}_2 \)) conditions\textsuperscript{33}. Cultures of \( R. \text{rubrum} \) strains were grown as specified (Supplementary Fig. 3).

**Construction of magnetosome gene cluster plasmids and conjugative transfer.** The oligonucleotides and plasmids used in this study are listed in Supplementary Tables 4 and 5. Red/ET (\text{Lambda red and RecET}) recombination was performed as described previously\textsuperscript{14}. Briefly, a cloning cassette was amplified by polymerase chain reaction (PCR) and transferred into electrocompetent \( E. \text{coli} \) cells (DH10b) expressing phage-derived recombinases from a circular plasmid (pSC101-BAD-gbaA). After transfer of the cassette, recombination occurred between homologous regions on the linear fragment and the plasmid.

To stitch the magnetosome gene clusters together into a transposon plasmid (Supplementary Fig. 1) we used triple recombination\textsuperscript{14} and co-transformed two linear fragments, which recombined with a circular plasmid. Recombinants harbouring the correct plasmids were selected by restriction analysis\textsuperscript{12}. Conjugations into \( M. \text{gryphiswaldense} \) were performed as described before\textsuperscript{12}. For conjugation of \( R. \text{rubrum} \), cultures were incubated in ATCC medium 112. Approximately 2 \( \times \) \( 10^8 \) cells were mixed with 1 \( \times \) \( 10^9 \) \( E. \text{coli} \) cells, spotted on American Type Culture Collection (ATCC) 112 agar medium and incubated for 15 h. Cells were flushed from the plates and incubated on ATCC 112 agar medium supplemented with appropriate antibiotics for 7–10 days (\( \text{Tc} = 10 \mu \text{g ml}^{-1} \), \( \text{Km} = 20 \mu \text{g ml}^{-1} \), \( \text{Gm} = 10 \mu \text{g ml}^{-1} \), where Tc, tetracycline; Km, kanamycin; Gm, gentamicin). Sequential transfer of the plasmids resulted in 1 \( \times \) \( 10^{-3} \) to 1 \( \times \) \( 10^{-2} \) antibiotic-resistant insertants per recipient, respectively. Two clones from each conjugation experiments were chosen for further analyses. Characterized insertants were indistinguishable from wild type with respect to motility, cell morphology or growth (Supplementary Fig. 5).

**Analytical methods.** The optical density of \( M. \text{gryphiswaldense} \) cultures was measured turbidimetrically at 565 nm as described previously\textsuperscript{33}. The optical density of \( R. \text{rubrum} \) cultures was measured at 660 nm and 880 nm. The ratio of 880/660 nm was used to determine yields of chromatophores within intact cells (Supplementary Fig. 4). Furthermore, \text{bacteriochlorophyll a} was assayed with a light scattering assay as described previously\textsuperscript{19}. Briefly, cells were washed and recorded with a Morgagni 268 microscope. Sizes of crystals and vesicles were measured with ImageJ software.

**Microscopy.** For TEM of whole cells and isolated magnetosomes, specimens were directly deposited onto carbon-coated copper grids. Magnetosomes were stained with 1% phosphotungstic acid or 2% uranyl acetate. Samples were viewed and recorded with a Morgagni 268 microscope. Sizes of crystals and vesicles were measured with ImageJ software.

High-resolution TEM was performed with a FEOL 3010 microscope.
operated at 297 kV and equipped with a Gatan Imaging Filter for the acquisition of energy-filtered compositional maps. For TEM data processing and interpretation, DigitalMicrograph and SingleCrystal software were used. Cryo-electron tomography was performed as described previously. Fluorescence microscopy was performed with an Olympus IX81 microscope equipped with a Hamamatsu Orca AG camera using exposure times of 0.12–0.25 s. Image rescaling and cropping were performed with Photoshop 9.0 software.

Received 9 September 2013; accepted 16 January 2014; published online 23 February 2014

References


Acknowledgements

This work was supported by the Human Frontier Science Foundation (grant RGP0652/2012), the Deutsche Forschungsgemeinschaft (grants SCHU 1080/12-1 and 15-1) and the European Union (Bio2MaN4MRI). The authors thank F. Kiemer for expert help with iron measurements and cultivation experiments.

Author contributions

I.K., D.S., Y.Z., Q.T., C.J. and R.M. planned and performed cloning experiments. I.K. and A.L. performed genetic transfers and cultivation experiments. G.W. prepared cryo- and chemically fixed cells. S.B., O.R. and G.W. performed TEM and I.K. analysed the data. J.P. and O.R. performed cryo-electron tomography experiments. E.T. and M.P. took high-resolution TEM micrographs and analysed the data. I.K. and A.L. took fluorescence micrographs and performed phenomenon experiments. I.K. performed western blot experiments and analysed proteomic data. A.B. performed Illumina genome sequencing and I.K. analysed the data. I.K. and D.S. designed the study and wrote the paper. All authors discussed the results and commented on the manuscript.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to Y.Z. and D.S.

Competing financial interests

I.K. and D.S. (LMU Munich) have filed a patent application on the process described in this work (Production of magnetic nanoparticles in recombinant host cells, EP13193478).
Biosynthesis of magnetic nanostructures in a foreign organism by transfer of bacterial magnetosome gene clusters

Construction of Tn5 transposon plasmids
For construction of translational (C-terminal) gene fusions, the mamDC promoter (XbaI, BamHI restriction sites added) was cloned in front of either the mamGFDC operon or the mamJ gene (NdeI, KpnI), which were followed by the egfp gene (KpnI, EcoRI). The resulting construct was cloned into pBAM1\(^1\) modified by a tetracycline resistance cassette (exchange of km\(^R\) against tc\(^R\) with SanDI and AatII). The replicative plasmid pFM211 (Frank Müller, unpublished) harboring ftsZm with a mCherry fusion under control of an inducible lac promoter was recombined with pBAM1 to construct pBAM-ftsZm_mcherry. The resident km\(^R\) was replaced by tc\(^R\) using ET-recombination. For construction of pBAM_feoAB1, a fragment with P\(_{mamH}\) and feoAB1 was amplified by PCR from pRU1feoAB (XbaI, EcoRI) and cloned into Tet-pBAM1.

Intracellular iron measurements
Cellular iron contents were determined after incubation under photoheterotrophic conditions in 10 ml Hungate tubes using a modified version of the ferrozine assay\(^2\). To this end, 4 ml cultures were centrifuged for 1 min at 11.000 rpm, resuspended in 90 µl HNO\(_3\) (65%) and incubated for 3 h at 99 °C.

Sequencing
For whole genome sequencing of strain \(R.\) rubrum_ABG6X a genomic DNA library was generated with the Nextera Kit (Illumina). Sequencing (1.25 Mio clusters, 2x 250 bp) was performed with a MiSeq sequencer (Illumina). Data analysis with CLC Genomics Workbench (CLCbio) confirmed single-site integration of both expression cassettes without mutations, except for a spontaneous deletion (aa 169-247) within the hypervariable non-essential CAR domain of mamJ which was shown to be irrelevant for protein function\(^3\).
Magnetosome isolation, electrophoresis and immuno-chemical detection

For magnetosome isolation and expression analysis, cultures of *R. rubrum* were grown photoheterotrophically in sealed 5 liter flasks illuminated by white light, 1000 lux intensity. Cells were harvested, washed and resuspended into HEPES buffer. Cell suspensions were lysed by sonication and cellular debris was removed by low-speed centrifugation. Magnetic separation of magnetosome particles, solubilization of the enclosing organic layer and fractionation of non-magnetic membrane fraction and soluble proteins were performed as previously described. Polyacrylamide gels were prepared according to the procedure of Laemmli. Protein samples from different cellular fractions (magnetosome membrane, soluble fraction, non-magnetic membrane fraction) were resuspended in electrophoresis sample buffer and denatured at 98 °C for 5 min. 10 µg of protein extracts were separated on a 15% SDS-polyacrylamide gel. Protein bands were visualized by Coomassie brilliant blue staining. Western blot analysis for detection of MamC was performed as previously described.

Mass spectrometry

For mass spectrometry 25 µg solubilised proteins were tryptically in-gel digested as described previously. The resulting fragments were separated on a C18 reversed-phase column and analyzed by nano-electrospray ionization-LC tandem MS (ESI-LC-MS/MS), recorded on an Orbitrap mass spectrometer. Spectra were analyzed via Mascot™ software using the NCBI nr Protein Database and a database from *M. gryphiswaldense*.


Biosynthesis of magnetic nanostructures in a foreign organism by transfer of bacterial magnetosome gene clusters
Fig. S1: Construction scheme of insertion cassettes for modular expression of the mam and mms operons. (a) Recombineering of a BAC containing the mamAB operon (blue) and a vector backbone (Km-p15A-Tps-oriT-Km, orange) harboring a MycoMar transposase gene (tps), inverted repeats (IR), origin of transfer (oriT), p15A origin of replication (p15A) and a kanamycin\(^R\) cassette (km\(^R\), orange). (b) Insertion of a spectinomycin\(^R\) cassette (spec\(^R\), pink) and the mamGFDC operon (green) into pTps_AB by triple recombination. (c & d) Stitching of pTps_ABG by insertion of the mms6 operon and a chloramphenicol\(^R\) cassette. (e) pTps_XYZ consisting of a Tps vector backbone (orange), mamXYZ (pale blue) and a gentamicin\(^R\) gene (gm\(^R\), purple) was constructed. (f) Plasmids were transferred by conjugation into R. rubrum. Transposition of the DNA-fragments within the IR sequences occurred at random positions at TA dinucleotide insertion sites by a “cut and paste” mechanism\(^1\). (g) Chromosomal insertion sites of the transposed constructs in R. rubrum_ABG6X are shown with adjacent genes (red) as revealed by whole genome sequencing performed with a MiSeq sequencer (Illumina) (accession number of R. rubrum ATCC 11170: NC_007643). pTps_ABG6 inserted within a gene encoding a putative aldehyd dehydrogenase (YP_426002), and pTps_XYZ inserted within rru_A2927, encoding a putative acriflavin resistance protein (protein accession number YP_428011). Sequences of inserted magnetosome operons matched those of the donor (M. gryphiswaldense) with no detectable mutations, except for a deletion (aa 169-247) within the hypervariable non-essential CAR domain of mamJ, which was shown to be irrelevant for protein function\(^2\).
Fig. S2: Transmission electron micrographs of MSR mutants expressing various insertional transposon constructs. The plasmids pTps_AB, pTps_ABG and pTps_ABG6 were transferred into the non-magnetic *M. gryphiswaldense* mutants ∆mamAB² and MSR-1B, the latter lacking most of the magnetosome genes except of the *mamXY* operon³⁴. After transfer of pTps_AB, a wt-like phenotype was restored in ∆mamAB_AB as revealed by C_{mag} (1.2 ± 0.2) and measured crystal sizes (37 ± 10 nm) in comparison with *M. gryphiswaldense* wt (36 ± 9 nm, C_{mag}=1.4 ± 0.2) (see also Table S1). Mutant MSR-1B was only partly complemented after insertion of pTps_AB and pTps_ABG, that is, C_{mag} and crystal sizes were still lower than in the wt (Table S1). Transfer of pTps_ABG6 restored nearly wt-like magnetosome formation in MSR-1B (35 ± 8 nm, C_{mag}=0.9± 0.1). ±= s.d. Scale bar: 1 μm, insets: 0.2 μm.
Fig. S3: HRTEM image of a poorly crystalline iron oxide particle from *R. rubrum_ABG6* with the corresponding Fourier transform (i) that shows diffuse, faint intensity maxima consistent with the structure of hematite.
Fig. S4: Growth, magnetic response and ICM/Bchl a production of *R. rubrum_AGB6X*. (a & b) Cells were grown in ATCC 112 (chemotrophic, 20% O$_2$), Sistrom A (phototrophic, anoxygenic) and M2SF (chemotrophic, 1%O$_2$) medium for 3 (30 °C), 4 (23 °C) or 10 (18 °C) days. Optical density at 660 nm (minimal Bchl a absorption, black diamonds), 880 nm (maximal Bchl a absorption) and magnetic response (grey diamonds) were measured. The ratio OD$_{880}$/OD$_{660}$ (white diamonds) correlates with the amount of chromatophores produced in the cells$^5$ (median values n=3, error bars indicate s.d.). No C$_{mag}$ was detectable under aerobic and microaerobic conditions at 30 °C. (c & d) Absorption spectra of extracted bacteriochlorophylls from *R. rubrum* wt (c) and *R. rubrum_ABG6X* (d) (phototrophic growth, 30 °C).
Fig. S5: Growth of *R. rubrum* wt and *R. rubrum_ABG6X* (OD$_{660}$). Cells of *R. rubrum* (Sistrom A medium, 1000 lux) were incubated in Sistrom A medium (1000 lux) for 3 days at 23 °C under anaerobic conditions. No major growth differences between wt (n=3) and mutant strain ABG6X (median values n=3, error bars indicate s.d.) were detectable.
Fig. S6: Proteomic analysis of magnetosomes from *R. rubrum_ABG6X*. (a) 1D SDS-PAGE of Coomassie blue stained proteins solubilized from isolated magnetosome particles of *M. gryphiswaldense* and *R. rubrum_ABG6X*. Bands of the same size are indicated (arrowheads). (b) Immunodetection of MamC (12.4 kDa) in blotted fractions of *M. gryphiswaldense* and *R. rubrum_ABG6X* using an anti-MamC antibody. A signal for MamC was detectable in the magnetic membrane fraction of *R. rubrum_ABG6X* (6), which was absent from the soluble fraction, but faintly present also in the non-magnetic membrane fraction (5), possibly originating from empty membrane vesicles or incomplete magnetic separation during isolation. Protein extracts from *M. gryphiswaldense*: 1. soluble fraction, 2. non-magnetic membrane fraction, 3. magnetosome membrane. Protein extracts from *R. rubrum_ABG6X*: 4. soluble fraction, 5. non-magnetic membrane fraction, 6. magnetic (“magnetosome”) membrane fraction. M: Marker.
Fig. S7: Fluorescence microscopy of *R. rubrum* wt and *R. rubrum_ABG6X* cells expressing different EGFP-tagged magnetosome proteins. For localization studies of fluorescently labeled magnetosome proteins, strains were cultivated in ATCC medium overnight at 30 °C with appropriate antibiotics (Table S3). (a & b) MamGFDC with a C-terminal MamC-EGFP fusion expressed in *R. rubrum* wt (n=151) (a), and *R. rubrum_ABG6X* (n=112) (b). In the transformed strain, a filamentous structure is visible for 79% of the cells (n=89). c & d, MamJ-EGFP expressed in *R. rubrum* wt (n=109) (c), and in *R. rubrum_ABG6X* (n=89) displaying a chain-like fluorescence signal in 63% of the cells (n=56) (d). Scale bar: 2 µm.
Fig. S8: TEM of cryo- or chemically fixed, thin sectioned \textit{R. rubrum} strains.

Cells were cultivated under photo-heterotrophic conditions. ICM sizes of cryo-fixed \textit{R. rubrum} \textit{wt} (93 \pm 34 \text{ nm}, n=95) and vesicles surrounding immature magnetosomes of cryo-fixed \textit{R. rubrum\textunderscore ABG6X} (66 \pm 6 \text{ nm}, n=6) were measured.
Fig. S9: Size distribution of magnetosome crystals in *M. gryphiswaldense* and different *R. rubrum* strains. Whereas crystals of *R. rubrum_ABG6* (n=303) and *R. rubrum_ABG6X* (n=306) were smaller than those of the donor *M. gryphiswaldense* (n=310), crystal sizes of *R. rubrum_ABG6X_feo* (n=301) were significantly larger, approaching those of the donor strain (see also Supplementary Table 1).
Figure S10: Transmission electron micrographs of whole cells of different *R. rubrum* strains expressing magnetosome gene clusters. Scale bar: 1 µm, inset: 0.2 µm.


Biosynthesis of magnetic nanostructures in a foreign organism by transfer of bacterial magnetosome gene clusters

Table S1: Summary of magnetic responses ("Cmag"), intracellular iron content and crystal size and number of various strains (median values, ± = standard deviation). If not indicated otherwise, cells were grown in the presence of 50 µM ferric citrate. Magnetic response and total iron content measurements were performed with (n) biological replicates under identical conditions (see also material & methods). For determination of crystal size and number per cell, cells of one clone were analyzed by TEM (n=sample size). The Mann-Whitney test (http://elegans.som.vcu.edu/~leon/stats/utest.html) was performed for crystal size comparison of R. rubrum_ABG6X and R. rubrum_ABG6X_feo: the difference was highly significant (p<0.001, two tailed test). Crystal size comparison of R. rubrum_ABG6X_feo and M. gryphiswaldense revealed no significant difference (p≥0.05, two tailed test).

<table>
<thead>
<tr>
<th>Strain</th>
<th>Magnetic response (&quot;Cmag&quot;)</th>
<th>Iron content (% dry weight)</th>
<th>Crystal size (nm)</th>
<th>Crystal number per cell</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. gryphiswaldense MSR-1</td>
<td>1.4 ± 0.2</td>
<td>3.5 (n=3)</td>
<td>36 ± 9 (n=310)</td>
<td>24 ± 8 (n=52)</td>
</tr>
<tr>
<td>M. gryphiswaldense ∆mamAB_AB</td>
<td>1.2 ± 0.2</td>
<td>n.d. (n=3)</td>
<td>37 ± 10 (n=112)</td>
<td>23 ± 7 (n=24)</td>
</tr>
<tr>
<td>M. gryphiswaldense MSR-1B_AB</td>
<td>0.2 (n=3)</td>
<td>n.d.</td>
<td>17 ± 6 (n=112)</td>
<td>16 ± 6 (n=20)</td>
</tr>
<tr>
<td>M. gryphiswaldense MSR-1B_ABG</td>
<td>0.6 ± 0.1</td>
<td>n.d. (n=3)</td>
<td>25 ± 6 (n=104)</td>
<td>13 ± 6 (n=20)</td>
</tr>
<tr>
<td>M. gryphiswaldense MSR-1B_ABG6</td>
<td>0.9± 0.2</td>
<td>n.d. (n=3)</td>
<td>35 ± 8 (n=103)</td>
<td>18 ± 8 (n=22)</td>
</tr>
<tr>
<td>R. rubrum ATCC 11170</td>
<td>0.07 ± 0.04 (n=3)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R. rubrum_AB</td>
<td>-</td>
<td>0.08 (n=3)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R. rubrum_ABG</td>
<td>-</td>
<td>0.10 ± 0.01 (n=3)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R. rubrum_ABG6</td>
<td>-</td>
<td>0.17 (n=4)</td>
<td>12 ± 6 (n=304)</td>
<td>26 ± 10 (n=50)</td>
</tr>
<tr>
<td>R. rubrum_ABG6X</td>
<td>0.3 ± 0.2 (n=3)</td>
<td>0.17 ± 0.02 (n=4)</td>
<td>24 ± 7 (n=307)</td>
<td>10 ± 4 (n=50)</td>
</tr>
<tr>
<td>R. rubrum_ABG6X 500 µM ferric citrate</td>
<td>0.3 (n=4)</td>
<td>n. d.</td>
<td>25 ± 7 (n=301)</td>
<td>11 ± 5 (n=51)</td>
</tr>
<tr>
<td>R. rubrum_ABG6X 100 µM ferrous sulfate</td>
<td>0.2 (n=4)</td>
<td>n.d.</td>
<td>24 ± 8 (n=312)</td>
<td>10 ± 5 (n=52)</td>
</tr>
<tr>
<td>Strain</td>
<td>cmag</td>
<td>cmag</td>
<td>cmag</td>
<td>cmag</td>
</tr>
<tr>
<td>------------------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td><strong>0.6 ± 0.1</strong></td>
<td>0.18 ± 0.03</td>
<td>26 ± 9</td>
<td>11 ± 4</td>
</tr>
<tr>
<td>R. rubrum_ABG6X_ftsZm</td>
<td>(n=3)</td>
<td>(n=3)</td>
<td>(n=300)</td>
<td>(n=51)</td>
</tr>
<tr>
<td>R. rubrum_ABG6X_dJ</td>
<td>0.2</td>
<td>0.18 ± 0.01</td>
<td>27 ± 9</td>
<td>9 ± 4</td>
</tr>
<tr>
<td></td>
<td>(n=3)</td>
<td>(n=3)</td>
<td>(n=300)</td>
<td>(n=50)</td>
</tr>
<tr>
<td>R. rubrum_ABG6X_dl</td>
<td>-</td>
<td>0.09 ± 0.07</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(n=3)</td>
<td>(n=3)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>R. rubrum_ABG6X_feo</td>
<td>0.8 ± 0.1</td>
<td>0.28 ± 0.07</td>
<td>37 ± 10</td>
<td>10 ± 4</td>
</tr>
<tr>
<td></td>
<td>(n=3)</td>
<td>(n=3)</td>
<td>(n=300)</td>
<td>(n=52)</td>
</tr>
</tbody>
</table>

*The slightly increased C\textsubscript{mag} is likely due to effects of the genuine cell division protein FtsZm on cell morphology, as no difference in iron content and crystal size or number per cell was detectable.

**64% of mutant cells (n=32) harbored clustered magnetosomes, whereas 36% still showed a chain-like alignment of magnetosomes (n=18).
Table S2: Magnetosome proteins identified in the MM of strain *R. rubrum_ABG6X* by nano-electrospray ionization-LC tandem MS (ESI-LC-MS/MS). Spectra were analyzed via Mascot™ software using the NCBI nr Protein Database and a database from *M. gryphiswaldense* (asterisks). Proteins are listed in the order of their exponentially modified protein abundance index (emPAI). The data have been deposited to ProteomeXchange with identifier PXD000348 (DOI 10.6019/PXD000348).

<table>
<thead>
<tr>
<th>Protein</th>
<th>Accession number</th>
<th>Coverage (%)</th>
<th>No. of spectrum matches</th>
<th>No. of sequence peptides</th>
<th>Molecular weight (kDa)</th>
<th>Calculated pI</th>
<th>emPAI</th>
<th>Putative function</th>
</tr>
</thead>
</table>
| MamK    | MGR_4093         | 57           | 9                       | 9                       | 39.6                   | 5.4          | 1.51  | Magnetosome chain assembly/positioning 
<sup>2,3</sup> |
| MamC    | MGR_4078         | 32           | 4                       | 3                       | 12.4                   | 5.1          | 1.01  | Crystal size and shape control <sup>4</sup> |
| MamJ    | MGR_4092         | 32           | 10                      | 6                       | 48.6                   | 4.0          | 0.76  | Magnetosome chain assembly <sup>5</sup> |
| MamA    | MGR_4099         | 37           | 1                       | 1                       | 23.9                   | 5.7          | 0.65  | TPR-like protein associated with the magnetosome membrane <sup>6,7</sup> |
| MamF    | MGR_4076         | 17           | 1                       | 1                       | 12.4                   | 9.1          | 0.60  | Magnetosome size and shape control <sup>8</sup> |
| Mms6    | MGR_4073         | 19           | 1                       | 1                       | 12.7                   | 9.5          | 0.58  | Magnetosome crystallization <sup>8,9</sup> |
| MamD    | MGR_4077         | 20           | 3                       | 3                       | 30.2                   | 9.8          | 0.49  | Crystal size and shape control <sup>9</sup> |
| MamM*   | MGR_4095         | 15           | 3                       | 3                       | 34.7                   | 5.8          | 0.42  | Iron transport/MM assembly <sup>10</sup> |
| MmsF*   | MGR_4072         | 8            | 2                       | 1                       | 13.9                   | 9.3          | 0.23  | Crystal size and shape control <sup>11</sup> |
| MamB*   | MGR_4102         | 7            | 1                       | 1                       | 32.1                   | 5.4          | 0.21  | Iron transport/MM assembly <sup>10</sup> |
| MamY*   | MGR_4150         | 18           | 2                       | 2                       | 40.9                   | 4.8          | 0.16  | Tubulation and magnetosome membrane formation <sup>12</sup> |
| MamO*   | MGR_4097         | 6            | 3                       | 3                       | 65.3                   | 6.5          | 0.15  | Magnetosome crystallization <sup>13,14</sup> |
| MamE    | MGR_4091         | 4            | 2                       | 2                       | 78.3                   | 8.1          | 0.08  | Magnetosome crystallization <sup>13,14</sup> |
Table S3: Strains and plasmids used in this study. Km\(^R\) = kanamycin resistance, Tc\(^R\) = tetracycline resistance, Ap\(^R\) = ampicillin resistance, BSD\(^R\) = blasticidin S resistance, Cm\(^R\) = chloramphenicol resistance, Gm\(^R\) = gentamicin resistance, Spec\(^R\) = spectinomycin resistance.

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Characteristics</th>
<th>Reference(s) or source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magnetospirillum gryphiswaldense strains</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>M. gryphiswaldense</em> MSR-1</td>
<td>Wild-type (wt)</td>
<td>DSM-6361,(^15)</td>
</tr>
<tr>
<td><em>M. gryphiswaldense</em> MSR-1B</td>
<td>spontaneous unmagnetic mutant lacking parts of the MAI</td>
<td>(^16)</td>
</tr>
<tr>
<td><em>M. gryphiswaldense ∆mamAB</em></td>
<td>mamAB deletion mutant</td>
<td>(^17)</td>
</tr>
<tr>
<td><em>M. gryphiswaldense ∆mamAB _AB</em></td>
<td>Km(^R), transposon mutant with inserted mamAB operon</td>
<td>This study</td>
</tr>
<tr>
<td><em>M. gryphiswaldense</em> MSR-1B _AB</td>
<td>Km(^R), transposon mutant with inserted mamAB operon</td>
<td>This study</td>
</tr>
<tr>
<td><em>M. gryphiswaldense</em> MSR-1B _ABG</td>
<td>Km(^R), Spec(^R), transposon mutant with inserted mamAB and mamGFDC operon</td>
<td>This study</td>
</tr>
<tr>
<td><em>M. gryphiswaldense</em> MSR-1B _ABG6</td>
<td>Km(^R), Cm(^R), transposon mutant with inserted mamAB, mamGFDC and mms6 operon</td>
<td>This study</td>
</tr>
<tr>
<td>Rhodospirillum rubrum strains</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>R. rubrum</em> ATCC 11170</td>
<td>wt</td>
<td>(^18) (kindly provided by H. Grammel, Magdeburg)</td>
</tr>
<tr>
<td><em>R. rubrum</em> _AB*</td>
<td>Km(^R), transposon mutant with inserted mamAB operon</td>
<td>This study</td>
</tr>
<tr>
<td><em>R. rubrum</em> _ABG*</td>
<td>Km(^R), Spec(^R), transposon mutant with inserted mamAB and mamGFDC operon</td>
<td>This study</td>
</tr>
<tr>
<td><em>R. rubrum</em> _ABG6*</td>
<td>Km(^R), Cm(^R), transposon mutant with inserted mamAB, mamGFDC and mms6 operon</td>
<td>This study</td>
</tr>
<tr>
<td><em>R. rubrum</em> _ABG6X*</td>
<td>Km(^R), Cm(^R), Gm(^R), transposon mutant with inserted mamAB, mamGFDC, mms6 and mamXY operon (without ftsZm)</td>
<td>This study</td>
</tr>
<tr>
<td><em>R. rubrum</em> _ABG6X_dJ*</td>
<td>Km(^R), Cm(^R), Gm(^R), Ap(^R), transposon mutant with inserted mamAB (mamJ deletion), mamGFDC, mms6 and mamXY operon (without ftsZm)</td>
<td>This study</td>
</tr>
<tr>
<td><em>R. rubrum</em> _ABG6X_dl*</td>
<td>Km(^R), Cm(^R), Gm(^R), Ap(^R), transposon mutant with inserted mamAB (mamI deletion), mamGFDC, mms6 and mamXY operon (without ftsZm)</td>
<td>This study</td>
</tr>
<tr>
<td><em>R. rubrum</em> _ABG6X_ftsZm*</td>
<td>Km(^R), Cm(^R), Gm(^R), Tc(^R), transposon mutant with inserted mamAB, mamGFDC, mms6 and mamXY operon (without ftsZm) and ftsZm</td>
<td>This study</td>
</tr>
<tr>
<td>Strain</td>
<td>Description</td>
<td>Source</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>--------</td>
</tr>
<tr>
<td>R. rubrum_ABG6X_feo</td>
<td>transposon mutant with inserted mamAB, mamGFDC, mms6, mamXY and feoAB1 operon</td>
<td>This study</td>
</tr>
<tr>
<td>R. rubrum_GFDC-EGFP</td>
<td>transposon mutant with inserted mamGFDC-EGFP</td>
<td>This study</td>
</tr>
<tr>
<td>R. rubrum_ABG6X_GFDC-EGFP</td>
<td>transposon mutant with inserted mamAB, mamGFDC, mms6 and mamXY operon (without ftsZm) and mamGFDC-EGFP</td>
<td>This study</td>
</tr>
<tr>
<td>R. rubrum_J-EGFP</td>
<td>transposon mutant with inserted mamGFDC-EGFP</td>
<td>This study</td>
</tr>
<tr>
<td>R. rubrum_ABG6X_J-EGFP</td>
<td>transposon mutant with inserted mamAB, mamGFDC, mms6 and mamXY operon (without ftsZm) and mamJ-EGFP</td>
<td>This study</td>
</tr>
</tbody>
</table>

### Escherichia coli strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH10b</td>
<td>F− mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara leu) 7697 galU galK rpsL nupG λ−</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>BW29427</td>
<td>dap auxotroph derivative of E. coli strain B2155</td>
<td>K. Datsenko and B. L. Wanner, unpublished</td>
</tr>
<tr>
<td>WM3064</td>
<td>thrB1004 pro thi rpsL hsdS lacZΔM15 RP4-1360 Δ(araBAD)567 ΔdapA1341::[erm pir]</td>
<td>W. Metcalf, kindly provided by J. Gescher, KIT Karlsruhe</td>
</tr>
</tbody>
</table>

### Plasmids

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSC101-BAD-gbaA</td>
<td>TcR, replicative plasmid containing redα/redβ recombinases under the control of a L-Arabinose inducible promoter, temperature sensitive origin of replication</td>
<td>19</td>
</tr>
<tr>
<td>p15A-Tps-oriT-Km</td>
<td>KmR, BSDR, oriT, p15A origin of replication, mariner tps, cloning cassette</td>
<td>20</td>
</tr>
<tr>
<td>pSSK18 (BAC_mamAB)</td>
<td>BAC containing the mamAB operon from M. gryphiswaldense</td>
<td>16</td>
</tr>
<tr>
<td>pTps_AB</td>
<td>KmR, BSDR, mariner tps vector containing mamAB operon</td>
<td>This study</td>
</tr>
<tr>
<td>pTps_ABG</td>
<td>SpecR, KmR, BSDR, mariner tps vector with mamAB and mamGFDC operon</td>
<td>This study</td>
</tr>
<tr>
<td>pTps_ABG6</td>
<td>CmR, KmR, BSDR, mariner tps vector with mamAB, mamGFDC, and mms6 operon</td>
<td>This study</td>
</tr>
<tr>
<td>pTps_XYZ</td>
<td>GmR, BSDR, mariner Tps vector with mamY, mamX and mamZ</td>
<td>This study</td>
</tr>
<tr>
<td>Vector Name</td>
<td>Description</td>
<td>Source</td>
</tr>
<tr>
<td>------------</td>
<td>-------------</td>
<td>--------</td>
</tr>
<tr>
<td>pTps_ABG6_dJ</td>
<td>Cm&lt;sup&gt;+&lt;/sup&gt;, Km&lt;sup&gt;+&lt;/sup&gt;, BSD&lt;sup&gt;+&lt;/sup&gt;, Ap&lt;sup&gt;+&lt;/sup&gt;, mariner tps vector with mamAB, mamGFDC, and mms6 operon, (mamJ deletion)</td>
<td>This study</td>
</tr>
<tr>
<td>pTps_ABG6_dl</td>
<td>Cm&lt;sup&gt;+&lt;/sup&gt;, Km&lt;sup&gt;+&lt;/sup&gt;, BSD&lt;sup&gt;+&lt;/sup&gt;, Ap&lt;sup&gt;+&lt;/sup&gt;, mariner tps vector with mamAB, mamGFDC, and mms6 operon, (mamI deletion)</td>
<td>This study</td>
</tr>
<tr>
<td>pBAM1</td>
<td>Km&lt;sup&gt;+&lt;/sup&gt;, Ap&lt;sup&gt;+&lt;/sup&gt;, γR6K origin of replication, oriT, Tn5 vector</td>
<td></td>
</tr>
<tr>
<td>Tet-pBAM1</td>
<td>Tc&lt;sup&gt;+&lt;/sup&gt;, Ap&lt;sup&gt;+&lt;/sup&gt;, γR6K origin of replication, oriT, Tn5 vector</td>
<td>This study</td>
</tr>
<tr>
<td>Tet-pBam_mamGFDC-EGFP</td>
<td>Tc&lt;sup&gt;+&lt;/sup&gt;, Ap&lt;sup&gt;+&lt;/sup&gt;, mamGFDC operon under control of P&lt;sub&gt;mamDC&lt;/sub&gt; with a C-terminal EGFP fusion, Tn5 vector</td>
<td>This study</td>
</tr>
<tr>
<td>Tet-pBam_MamJ-EGFP</td>
<td>Tc&lt;sup&gt;+&lt;/sup&gt;, Ap&lt;sup&gt;+&lt;/sup&gt;, mamJ under control of P&lt;sub&gt;mamDC&lt;/sub&gt; with a C-terminal EGFP fusion, Tn5 vector</td>
<td>This study</td>
</tr>
<tr>
<td>pRU-1feoAB</td>
<td>Km&lt;sup&gt;+&lt;/sup&gt;, broad host range pBBRMCS2, feoAB&lt;sub&gt;1&lt;/sub&gt; operon under the control of P&lt;sub&gt;mamJ&lt;/sub&gt;</td>
<td>R. Uebe, unpublished</td>
</tr>
<tr>
<td>Tet-pBam_feoAB1</td>
<td>Tc&lt;sup&gt;+&lt;/sup&gt;, Ap&lt;sup&gt;+&lt;/sup&gt;, feoAB&lt;sub&gt;1&lt;/sub&gt; operon under the control of P&lt;sub&gt;mamH&lt;/sub&gt;, Tn5 vector</td>
<td>This study</td>
</tr>
<tr>
<td>Tet-pBam_ftsZm_mCherry</td>
<td>Tc&lt;sup&gt;+&lt;/sup&gt;, Ap&lt;sup&gt;+&lt;/sup&gt;, ftsZm, lacI with a C-terminal mCherry fusion under control of inducible P&lt;sub&gt;lac&lt;/sub&gt;, Tn5 vector</td>
<td>This study</td>
</tr>
<tr>
<td>pFM211</td>
<td>Km&lt;sup&gt;+&lt;/sup&gt;, broad host range pBBRMCS2, lacI, ftsZm with C-terminal mCherry fusion, mamK with N-terminal EGFP fusion</td>
<td>F. Müller, unpublished</td>
</tr>
<tr>
<td>Primer</td>
<td>Nucleotide sequence (5'-3')</td>
<td>Product</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Mam-tps5</td>
<td>AATTTCGACAGGAGTATAGCAATCGAGGTGTTGAC AAGCCATATAATCGAAAACTCTGCTAGAAGCG</td>
<td>p15A-Tps-oriT-Km, ET-recombination with BAC_mamAB, pTps_AB</td>
</tr>
<tr>
<td>Mam-tps3</td>
<td>GAGGCAAGGAGTATAGCAATCGAGGTGTTGAC AAGCCATATAATCGAAAACTCTGCTAGAAGCG</td>
<td>mamGFDC operon, ET-recombination with pTps_AB</td>
</tr>
<tr>
<td>mamGF3</td>
<td>CGCTAGCTGCGGGTATTTACCAATTTCC</td>
<td></td>
</tr>
<tr>
<td>spectMam3</td>
<td>TCAAAACCGGCGAGGGATAAACGAAATACCCGCAGCT AGTCGTCATATTCTGTTATT</td>
<td>Spectinomycin resistance cassette, ET-recombination with pTps_AB</td>
</tr>
<tr>
<td>spectMam5</td>
<td>TGATCCGCTGTGTGTAAGCGCATCATGTCCGGATCCCATGG CGTTCCGCTCGTAACGTGACTGGCAAGAGATATT</td>
<td>mamGFDC operon, ET-recombination with pTps_AB</td>
</tr>
<tr>
<td>mms2cm5</td>
<td>TACTGCGATGAGTGGCAGGGCGGGGCGTAAGCTTACAATT TCCATGCGCCATTC</td>
<td></td>
</tr>
<tr>
<td>mms2mam3</td>
<td>GTGCTTCGCTGTGCACAAGAACC</td>
<td>mms6 operon, ET-recombination with pTps_AB</td>
</tr>
<tr>
<td>cm-mms6-3</td>
<td>TGGCCAATGGAATTGTGAAAGGTACCTGCCCCGCTGCCAC TC</td>
<td>Chloramphenicol resistance cassette, ET-recombination with pTps_AB</td>
</tr>
<tr>
<td>cm-mms6-5</td>
<td>TGATCCGCTATGTTAAGGCGCATATGCTCGGATCCCCAGTCG CTGGCCAGCTGCTCTGGTGTCCCTGATACC</td>
<td>mamGFDC operon in pBAM1</td>
</tr>
<tr>
<td>IK097</td>
<td>TCTAGAGGGCCCCAACCCTTTTCTGTACTAGCTCTTAATT CTCAAAAATATCGGCCAGGGAAT</td>
<td>P_mamOC in pBAM1</td>
</tr>
<tr>
<td>IK098</td>
<td>CATATGCTGATCTCCGGCAATGTAACGGATTCCTCCTTCGTAATCGGCAAGGGAAT</td>
<td>mamGFDC operon in pBAM1</td>
</tr>
<tr>
<td>IK101</td>
<td>GGTACCAGGGCCATTTCTTCTCAGA</td>
<td>egfp in pBAM1</td>
</tr>
<tr>
<td>IK102</td>
<td>GTGACCAGGGCCAGGGGAGGCT</td>
<td>mamXY operon (without ftsZm), ET-recombination with p15A-Tps-oriT-Gm</td>
</tr>
<tr>
<td>IK103</td>
<td>GAATCTTACGTCTGTACAGG</td>
<td>p15A-Tps-oriT-Km</td>
</tr>
<tr>
<td>IK104</td>
<td>GAATCTTACGTCTGTACAGG</td>
<td>Gentamicin resistance cassette, ET-Recombination with p15A-Tps-oriT-Gm</td>
</tr>
<tr>
<td>IK105</td>
<td>CCCGCTACCAGCTTTTGTCTGTGGAAGGTTAATGGGG CCGCTGTGTCATCTTTCTCAGTATACGAC</td>
<td>pTps_AB</td>
</tr>
<tr>
<td>IK106</td>
<td>GGGCTTACCCACTTAATCAGGTGACATCCACATCTCCCTTT TCAGCCGCTGCTCTGCTTGGTAACGTAATTTG</td>
<td>mamGFDC operon in pBAM1</td>
</tr>
<tr>
<td>IK107</td>
<td>CATATGATCAGGCGCATCGG</td>
<td>mamGFDC operon in pBAM1</td>
</tr>
<tr>
<td>IK108</td>
<td>GACGTCGAGCCACGGG</td>
<td>tetragen in pBAM1</td>
</tr>
<tr>
<td>IK109</td>
<td>GACGTCGAGCCACGGG</td>
<td>Ampicillin resistance cassette insertion into mamJ of pTps_AB</td>
</tr>
<tr>
<td>IK110</td>
<td>GGCGGTACCTTTATTCTTACTTCTGACAGCATAC</td>
<td>Tet-pBAM1, ET-recombination with</td>
</tr>
<tr>
<td>IK252</td>
<td>CAGCACATCCCCCTTTCCGCCAGCTGGGTAAAATAGCGAAGA</td>
<td>recombination with pFM211</td>
</tr>
<tr>
<td>-------</td>
<td>---------------------------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td></td>
<td>GCCCGCACCAGGATTGTGAGACACAGACGTC</td>
<td></td>
</tr>
</tbody>
</table>
References

Publication III

Biosynthesis of magnetic nanostructures in a foreign organism by transfer of bacterial magnetosome gene clusters.

Isabel Kolinko, Anna Lohse, Sarah Borg, Oliver Raschdorf, Christian Jogler, Qiang Tu, Mihály Pósfai, Éva Tompa, Jürgen M. Plitzko, Andreas Brachmann, Gerhard Wanner, Rolf Müller, Youming Zhang and Dirk Schüler

Author contributions

I.K., D.S., Y.Z., Q.T., C.J. and R.M. planned and performed cloning experiments. I.K. and A.L. performed genetic transfers and cultivation experiments. G.W. prepared cryo- and chemically fixed cells. S.B., O.R. and G.W. performed TEM and I.K. analysed the data. J.P. and O.R. performed cryo-electron tomography experiments. E.T. and M.P. took highresolution TEM micrographs and analysed the data. I.K. and A.L. took fluorescence micrographs and performed phenotypization experiments. I.K. performed western blot experiments and analysed proteomic data. A.B. performed Illumina genome sequencing and I.K. analysed the data. I.K. and D.S. designed the study and wrote the paper. All authors discussed the results and commented on the manuscript.

Signatures:
Isabel Kolinko:  
Anna Lohse:  
Sarah Borg:  
Oliver Raschdorf:  
Christian Jogler:  
Qiang Tu:  
Mihály Pósfai:  
Éva Tompa:  
Jürgen M. Plitzko:  
Andreas Brachmann:  
Gerhard Wanner:  
Rolf Müller:  
Youming Zhang:  
Dirk Schüler:  
Direct cloning and heterologous expression of the salinomycin biosynthetic gene cluster from *Streptomyces albus* DSM41398 in *S. coelicolor*A3(2).

Jia Yin1,2,3,4, Michael Hoffmann3, Xiaoying Bian1, Qiang Tu1,3, Fu Yan3, Liqiu Xia4, Xuezhi Ding4, A. Francis Stewart2, Rolf Müller3*, Jun Fu1,2* & Youming Zhang1*

1Shandong University–Helmholtz Institute of Biotechnology, State Key Laboratory of Microbial Technology, School of Life Science, Shandong University, Shanda Nanlu 27, Jinan, 250100, People’s Republic of China.

2Department of Genomics, Dresden University of Technology, BioInnovations-Zentrum, Tatzberg 47-51, Dresden, 01307, Germany.

3Helmholtz Institute for Pharmaceutical Research, Helmholtz Centre for Infection Research and Department of Pharmaceutical Biotechnology, Saarland University, PO Box 151150, Saarbrücken, 66041, Germany.

4Hunan Provincial Key Laboratory for Microbial Molecular Biology-State Key Laboratory Breeding Base of Microbial Molecular Biology, College of Life Science, Hunan Normal University, Changsha, 410081, People’s Republic of China.

*Correspondence and requests for materials should be addressed to R.M. (email: Rolf.Mueller@helmholtzhzi.de) or J.F. (email: fujun@sdu.edu.cn) or Y.M.Z. (email: zhangyouming@sdu.edu.cn)
Direct cloning and heterologous expression of the salinomycin biosynthetic gene cluster from *Streptomyces albus* DSM41398 in *Streptomyces coelicolor* A3(2)

Jia Yin1,2,3,4, Michael Hoffmann3, Xiaoying Bian1, Qiang Tu1,3, Fu Yan3, Liqiu Xia4, Xuezhi Ding1, A. Francis Stewart2, Rolf Müller3, Jun Fu1,2 & Youming Zhang1

Linear plus linear homologous recombination-mediated recombineering (LLHR) is ideal for obtaining natural product biosynthetic gene clusters from pre-digested bacterial genomic DNA in one or two steps of recombineering. The natural product salinomycin has a potent and selective activity against cancer stem cells and is therefore a potential anti-cancer drug. Herein, we separately isolated three fragments of the salinomycin gene cluster (salO-orf18) from *Streptomyces albus* (*S. albus*) DSM41398 using LLHR and assembled them into intact gene cluster (106 kb) by Red/ET and expressed it in the heterologous host *Streptomyces coelicolor* (*S. coelicolor*) A3(2). We are the first to report a large genomic region from a Gram-positive strain has been cloned using LLHR. The successful reconstitution and heterologous expression of the salinomycin gene cluster offer an attractive system for studying the function of the individual genes and identifying novel and potential analogues of complex natural products in the recipient strain.

Red/ET recombineering in *E. coli*1,2, is a powerful technique for the genetic engineering of natural product biosynthetic pathways, especially for large polyketide synthetase (PKS) as well as nonribosomal peptide-synthetase (NRPS)3–6. Recently, this technique was used to clone large biosynthetic gene clusters from a complex DNA source into a vector by linear plus linear homologous recombination (LLHR)7. LLHR is mediated by the full-length Rac prophage protein RecE, an exonuclease, its partner RecT, a single-strand DNA-binding protein, and Redγ, an inhibitor of the major exonuclease. RecA, a repair protein, is also included8. Fu et al., 2012 cloned ten hidden biosynthetic pathways from digested genomic DNA of Gram-negative *P. luminescens* using LLHR, and two of these have been successfully expressed in *E. coli*7,9. Many gene clusters have also been cloned by this method, including the syringolin, glidobactin, and colibactin gene clusters10–12, and all are from Gram-negative strains.

1Shandong University–Helmholtz Institute of Biotechnology, State Key Laboratory of Microbial Technology, School of Life Science, Shandong University, Shanda Nanlu 27, Jinan, 250100, People’s Republic of China. 2Department of Genomics, Dresden University of Technology, BioInnovations-Zentrum, Tatzberg 47-51, Dresden, 01307, Germany. 3Helmholtz Institute for Pharmaceutical Research, Helmholtz Centre for Infection Research and Department of Pharmaceutical Biotechnology, Saarland University, PO Box 151150, Saarbrücken, 66041, Germany. 4Hunan Provincial Key Laboratory for Microbial Molecular Biology-State Key Laboratory Breeding Base of Microbial Molecular Biology, College of Life Science, Hunan Normal University, Changsha, 410081, People’s Republic of China. Correspondence and requests for materials should be addressed to R.M. (email: Rolf.Mueller@helmholtz-hzi.de) or J.F. (email: fujun@sdu.edu.cn) or Y.M.Z. (email: zhangyouming@sdu.edu.cn)
An emerging idea in cancer biology is that tumors harbor a group of cells, known as cancer stem cells (CSCs), which have the unique ability to regenerate cancers\textsuperscript{13,14}. In addition to promoting tumor growth, growing evidence indicates that CSCs may be responsible for cancer recurrence, resistance to conventional treatments and metastasis\textsuperscript{15–18}. Recently, Lander et al., 2009 showed that salinomycin can selectively kill breast CSCs after screening 16,000 compounds\textsuperscript{19}. Further studies revealed that salinomycin has potent and selective activity against other cancer cell lines\textsuperscript{20,21}. In vitro data revealed that salinomycin pre-treatment reduced the tumor-seeding ability of cancer cell lines greater than 100-fold over the chemotherapy drug paclitaxel. Furthermore, salinomycin reduced mammary tumor size in mice to a greater extent than paclitaxel\textsuperscript{19}.

Salinomycin is produced by \textit{Streptomyces albus}\textsuperscript{22} and has been used to prevent \textit{Coccidioidomycosis} in poultry and alter gut flora to improve nutrient absorption in ruminants\textsuperscript{23}. The compound interferes with potassium transport across mitochondrial membranes, thus reducing intracellular energy production. It may also disrupt Na\textsuperscript{+}/Ca\textsuperscript{2+} exchange in skeletal and, in some cases, cardiac muscle, allowing a fatal accumulation of intracellular calcium\textsuperscript{24}.

Earlier results revealed that the polyketide chain of salinomycin is synthesized by an assembly line of nine PKS multienzymes (\textit{salAI–IX}). The nine PKS genes are collinearly arranged in the cluster. Four of these multienzymes (\textit{salAI}, \textit{salAVI}, \textit{salAVII}, and \textit{salAIX}) each catalyze a single extension module, while the other five (\textit{salAI}, \textit{salAV}, \textit{salAVIII}, \textit{salAV}, and \textit{salAIV}) encode two extension modules. In addition to the nine PKS genes, some other genes play vital roles in salinomycin biosynthesis\textsuperscript{25,26}. Upstream of the PKS genes, the adjacent \textit{orf1}, \textit{orf2}, and \textit{orf3} do not belong to the salinomycin cluster, but \textit{salN} and \textit{salO} encode putative regulatory proteins. \textit{SalP} and \textit{SalQ} are involved in the formation of the butyrate extender unit for salinomycin biosynthesis, and inactivation of \textit{salP} and \textit{salQ} reduced the yields of salinomycin by 10% and 36%, respectively when compared to wild-type\textsuperscript{26}. Downstream of the PKS genes, \textit{orf18} is predicted to encode a peptidyl carrier protein, and targeted inactivation of \textit{orf18} results in a 50–60% reduction in salinomycin production compared to wild-type\textsuperscript{25}.

Herein, we report the cloning of the 106-kb salinomycin gene cluster (\textit{salO–orf18}) from the genomic DNA of \textit{Streptomyces albus} DSM41398 by three rounds of direct cloning followed by assembling. All of the genes are oriented in the same direction and under the original promoters. The gene cluster was introduced into \textit{S. coelicolor} A3(2) for successful heterologous production of salinomycin.

**Results**

**Constructing a BAC vector for direct cloning of the salinomycin gene cluster by quadruple recombineering.** In order to construct a vector for direct cloning of the salinomycin gene cluster, the four fragments (backbone of \textit{pBeloBAC11}, \textit{amp-ccdB}, \textit{salO}, and \textit{orf18}) each had a 50-bp overlapping sequence, as illustrated in Fig. 1, and were co-electroporated into GB05dir-\textit{gyrA}\textsubscript{405}, a CcdB-resistant \textit{E. coli}.
coli strain containing the mutation GyrA R462M and LLHR-proficient recombinase (RecET, RecD, and RecA), to form the BAC vector by quadruple recombineering.

The BAC vector contained a homology arm to salO (292 bp) and orf18 (238 bp) and a cassette of the counterselection marker CcdB, which can be used to delete the background from the original BAC vector for direct cloning. A CcdB function test was performed as described previously.

Direct cloning of the salinomycin gene cluster. As mentioned above, salO encodes putative regulatory protein and orf18 is an essential factor for salinomycin production. Additionally, the restriction site (EcoRV), which can be utilized for direct cloning, is located in salO and orf18. Thus, we attempted to directly clone the 106-kb fragment (salO-orf18) using one and two-step recombination reactions with the BAC vector but were unsuccessful.

Hence, we divided the gene cluster into three fragments for direct cloning (Fig. 2a). We directly cloned the fragments of salO-salAIV (F1) and salAIX-orf18 (F3) using one step of LLHR with an efficiency of 4/24 and 1/24, respectively (Fig. S1a,c). We directly cloned the fragment of salAIV-salAVIII (F2) by a two-step recombination with an efficiency of 8/24 (Fig. S1b). Due to the repeated sequence in salAIV-salAVIII (Fig. S2), we were unable to directly clone this fragment by one step of LLHR. Therefore, this fragment was isolated using a neomycin selection marker flanked by lox71-lox66, which could be utilized to delete the selection marker conveniently in the assembling procedure. The three desired fragments were inserted in plasmids p15A-amp-F1, p15A-amp-F2-lox71-neo-lox66, and p15A-amp-F3, respectively.

Assembling of the salinomycin gene cluster and engineering for conjugation and integration. Figure 2b show the assembling procedure to reconstitute the entire cluster. F2 and F3 were ligated using the original restriction site of AsiSI/EcoRV in the gene cluster, which did not cause any open reading frame shift. The neomycin selection marker was deleted by Cre from the plasmid
p15A-amp-F2-lox71-neo-lox66 to produce p15A-amp-F2. Modifications were made to p15A-amp-F3 with two steps of recombineering. The neomycin selection marker flanked by lox71-lox66 was inserted into the non-coding sequence of F3 in the first recombineering step. The second recombineering step replaced the ampicillin selection marker with the hyg-ccdB cassette to produce p15A-hyg-ccdB-F3-lox71-neo-lox66. F3 was excised by AsIS/EcoRV and inserted into the AsIS/EcoRV site in p15A-amp-F2 by ligation to produce p15A-amp-F2 & F3-lox71-neo-lox66.

The ampicillin selection marker of the previous ligation product was replaced by the hyg-ccdB cassette to produce the plasmid p15A-hyg-ccdB-F2 & F3-lox71-neo-lox66. The plasmid p15A-amp-F1 was digested by EcoRV to release the fragment F1, and p15A-hyg-ccdB-F2 & F3-lox71-neo-lox66 was digested by EcoRV/MseI to excise F2 & F3-lox71-neo-lox66. The two fragments overlapped by 592 bp, and each fragment had a homologous arm with previously constructed BAC vector. The BAC vector was transformed into GB05 cells harboring the plasmid pSC101-cdad-gbaA. As CcdB is toxic, we induced CcdA, that inactivates the CcdB toxin, by rhamnose in the liquid medium or culture plates. The two previous linearize fragments were co-transformed into GB05 cells containing the BAC vector and the expression plasmid (pSC101-cdad-gbaA) to produce pBeloBAC11-sal-lox71-neo-lox66. We verified pBeloBAC11-sal-lox71-neo-lox66 using three restriction enzymes, the results (Fig. S3) showed that the pBeloBAC11-sal-lox71-neo-lox66 was correct.

To introduce the gene cluster into a heterologous expression host, few necessary elements were engineered before conjugation. The two step engineering procedure for conjugation and integration is diagrammed in Fig. 2c. Finally, the gene cluster was introduced into S. coelicolor A3(2) by conjugation and integrated into its chromosome.

**Heterologous production of salinomycin in S. coelicolor A3(2).** The genetic organization and promoters of the obtained salinomycin gene cluster are identical to those of the original producer S. albus DSM41398. After conjugation, the exconjugant colonies were confirmed by PCR and subsequently analyzed for heterologous salinomycin production. The salinomycin gene cluster was successfully inserted into the attB site of S. coelicolor A3(2) (Fig. S4).

The metabolite profiles of the wild-type S. coelicolor and the mutant strains S. coelicolor::sal were analyzed by HPLC-MS and compared with the salinomycin standard (Fig. 3a (Ref)). Thus, we were able to identify Salinomycin in extracts of the mutant strains S. coelicolor::sal via HPLC-MS (Fig. 3a, b) and
heterologous expression could be unambiguously confirmed by comparing MS² fragmentation pattern (Fig. 3c).

Discussion
Over the past several decades, numerous multifunctional megasynthases have been identified, cloned, sequenced, engineered, and heterologously expressed in suitable hosts. Traditionally, natural product biosynthetic gene clusters were retrieved from a single cosmid or reconstructed from several cosmids within a genomic library of the natural producer stain, which was time consuming due to subsequent cloning steps following the screening process from a genomic library²⁹.

LLHR-mediated recombineering was ideal for direct cloning of the salinomycin gene cluster from pre-digested genomic DNA after one or two steps of recombineering. Red/ET recombineering has traditionally been applied for heterologous expression of biosynthetic pathways to modify the biosynthetic pathways³⁰.

The failure to directly clone the 106-kb fragment with the BAC vector may have resulted from several considerations. First, the recombineering efficiency is very low for large fragments. Although the developed method of direct cloning is efficient for cloning up to ~52-kb fragments from a bacterial genome⁷, it is limited by inefficient co-transformation of two linear molecules, especially for long fragments (106kb). Moreover, the gene cluster contains GC-rich sequences. We studied the impact of the GC content on the recombineering efficiency and found that it was decreased for sequences with high GC content (data not shown). Second, enrichment of the target DNA is difficult after extracting the genomic DNA. Genomic DNA is susceptible to shearing forces associated with mechanical destruction and degradation by nuclease activity. Therefore, it is difficult to obtain the intact salinomycin biosynthesis gene cluster, especially for S. albus DSM 41398, the gram-positive strain. Third, previous data revealed that the Red/ET monomer anneals ~11 bp of DNA, and the smallest stable annealing intermediate requires only 20 bp of DNA and two Red/ET monomers³¹. In this study, we found that most of the colonies resulted from self-circularization of the vector used for direct cloning after recombineering although there were no obvious homologous regions in the backbone of the vector. As a result, it is difficult to screen the correct clone from thousands of self-circularized vectors.

In parallel to our LLHR-mediated direct cloning, the other DNA cloning methods for bioprospecting have their distinct merits. LLHR-mediated RecET direct cloning was not accessible to metagenomic DNA. Bioprospecting of metagenomics needs DNA synthesis and assembly method. Streptomyces phage φBT1 integrase-mediated in vitro site-specific recombination could assemble the 56kb epothilone biosynthetic gene cluster using modules as units. The authors didn’t prove that the complete gene cluster with att site scars could be expressed in a heterologous host³². The incorrect linker between modules might affect the biosynthesis³³. An intact DNA sequence can be obtained by the Gibson assembly³⁴,³⁵, which is the most efficient ‘chew back and anneal’ method³⁶-³⁸. The Gibson assembly was also proved to be capable of direct cloning of a 41-kb conglabatin biosynthetic gene cluster³⁹. Much larger DNA fragment can be directly cloned by transformation-associated recombination (TAR) in yeast Saccharomyces cerevisiae⁴⁰,⁴¹. However unrelated yeast homologous recombining might cause rearrangement of repetitive PKS/NRPS biosynthetic DNA sequences. The oriT-directed cloning for Gram-negative bacteria relies on available genetic tools to insert conjugation elements on the genome by two elaborated vectors. Although it is not straight forward, it has a capacity of cloning regions up to 140 kb from the genome of Burkholderia pseudomallei⁴². The phage φBT1 integrase-mediated direct cloning was developed for Gram-positive bacteria Streptomyces. It has the similar logic to oriT-directed cloning, which starts with integration of a capture vector by genome engineering at two spots, but both excision and circularization happen in the original bacteria⁴³. If Bacillus subtilis is justified as a suitable heterologous host for a biosynthetic gene cluster, its genome can be used as a vector for direct cloning of giant DNA, which has the potential to overcome the capacity limit of the BAC vector⁴⁴.

Compare to above methods our LLHR-mediated direct cloning has a significant feature. It is a genetic tool in E. coli, which is simple, convenient and cost-effective. The important improvement in this study is to combine RecET mediated direct cloning and lambda Red mediated plasmids stitching to hierarchically clone the intact 106kb salinomycin gene cluster. The reliability of the cloning method has been proved by subsequently successful heterologous expression in S. coelicolor A3(2). Our results represent a potent approach to mine the function of the individual genes and identify novel and potentially useful analogues of the complex natural products through module exchange in the recipient.

Methods
Strains, plasmids and culture conditions. The bacterial strains and plasmids used in this study are shown in Table S1. All primers were synthesized by Sigma-Genosys (Germany) (Table S2). All restriction enzymes, Taq polymerase, and DNA markers were purchased from New England Biolabs (UK).

E. coli cells were cultured in Luria-Bertani (LB) liquid media or on LB agar (1.2% agar). Ampicillin (amp, 100μg mL⁻¹), kanamycin (km, 15μg mL⁻¹), chloramphenicol (cm, 15μg mL⁻¹), hygromycin (hyg, 30μg mL⁻¹), apramycin (apr, 15μg mL⁻¹), and tetracycline (tet, 5μg mL⁻¹) were added to the media as required.

For sporulation and conjugation, S. coelicolor A3(2) was grown on mannitol salt (MS) agar plates for 10 days. If necessary for conjugation, apr (50μg mL⁻¹) and nalidixic acid (NA, 50μg mL⁻¹) were added.
S. albus DSM41398, S. coelicolor A3(2), and mutant strains were cultivated in M1 medium (10 g L\(^{-1}\) starch, 4 g L\(^{-1}\) yeast extract, 2 g L\(^{-1}\) peptone) at 30°C with constant agitation at 180 rpm.

**Bacterial genomic DNA isolation.** S. albus DSM41398 was cultured in 30 mL medium (4 g L\(^{-1}\) glucose, 4 g L\(^{-1}\) yeast extract, 10 g L\(^{-1}\) malt extract, pH 7.2) at 30°C for two days. After centrifugation, the cells were resuspended in 5 mL SET buffer (75 mM NaCl, 25 mM EDTA, 20 mM Tris, pH 7.5). After adding lysozyme to a final concentration of 1 mg mL\(^{-1}\) and incubating at 37°C for 0.5–1 h, 500 μL 10% SDS and 125 μL 20 mg mL\(^{-1}\) proteinase K were added, and the mixture was incubated at 55°C with occasional inversion for 2 h until the solution became clear. The solution was combined with 2 mL 5 M NaCl and 8 mL phenol:chloroform:isoamyl alcohol (25:24:1) and incubated at room temperature for 0.5 h with frequent inversion. After centrifuging at 4500 × g for 15 min, the aqueous phase was transferred to a new tube using a blunt-ended pipette tip, and the DNA was precipitated by adding one volume of isopropanol and gently inverting the tube. DNA was transferred to a microfuge tube, rinsed with 75% ethanol, dried under vacuum, and dissolved in ddH\(_2\)O.

**Preparation of electrocompetent cells for recombineering.** Recombineering and direct cloning were performed as described previously\(^4\) with several small modifications. The linear cloning vector p15A-amp, flanked with homology arms to target genes, was amplified by PCR using p15A-amp-ccdB\(^+\) as a template. Digested genomic DNA (10 μg) was mixed with 2 μg linear cloning vector and co-transformed into competent cells by electroporation.

**Conjugation.** Conjugation between E. coli and S. coelicolor A3(2) was performed as described previously\(^4\), with minor modifications.\(^45\). The plasmid containing the salinomycin gene cluster and elements was linearized with EcoRI as a template. Digested genomic DNA (10 μg) were performed as described previously\(^7\) with several small modifications. The linear cloning vector pUZ8002 was used as a template. The plasmid containing the salinomycin gene cluster and elements were digested with EcoRI and ligated into the linear cloning vector pUZ8002. The linear cloning vector and co-transformed into competent cells by electroporation.

**Conjugation.** Conjugation between E. coli and S. coelicolor A3(2) was performed as described previously\(^4\), with minor modifications. The plasmid containing the salinomycin gene cluster and elements was linearized with EcoRI as a template. Digested genomic DNA (10 μg) were performed as described previously\(^7\) with several small modifications. The linear cloning vector pUZ8002 was used as a template. The plasmid containing the salinomycin gene cluster and elements were digested with EcoRI and ligated into the linear cloning vector pUZ8002. The linear cloning vector and co-transformed into competent cells by electroporation.

**Conjugation.** Conjugation between E. coli and S. coelicolor A3(2) was performed as described previously\(^4\), with minor modifications. The plasmid containing the salinomycin gene cluster and elements was linearized with EcoRI as a template. Digested genomic DNA (10 μg) were performed as described previously\(^7\) with several small modifications. The linear cloning vector pUZ8002 was used as a template. The plasmid containing the salinomycin gene cluster and elements were digested with EcoRI and ligated into the linear cloning vector pUZ8002. The linear cloning vector and co-transformed into competent cells by electroporation.

**Conjugation.** Conjugation between E. coli and S. coelicolor A3(2) was performed as described previously\(^4\), with minor modifications. The plasmid containing the salinomycin gene cluster and elements was linearized with EcoRI as a template. Digested genomic DNA (10 μg) were performed as described previously\(^7\) with several small modifications. The linear cloning vector pUZ8002 was used as a template. The plasmid containing the salinomycin gene cluster and elements were digested with EcoRI and ligated into the linear cloning vector pUZ8002. The linear cloning vector and co-transformed into competent cells by electroporation.

**Conjugation.** Conjugation between E. coli and S. coelicolor A3(2) was performed as described previously\(^4\), with minor modifications. The plasmid containing the salinomycin gene cluster and elements was linearized with EcoRI as a template. Digested genomic DNA (10 μg) were performed as described previously\(^7\) with several small modifications. The linear cloning vector pUZ8002 was used as a template. The plasmid containing the salinomycin gene cluster and elements were digested with EcoRI and ligated into the linear cloning vector pUZ8002. The linear cloning vector and co-transformed into competent cells by electroporation.
Scientific Reports | 5:15081 | DOI: 10.1038/srep15081


Acknowledgments
This work was supported by funding to Y.Z. from the Recruitment Program of Global Experts, funding to J.Y. from China/ Shandong University International Postdoctoral Exchange Program, funding to J.F. from the International S&T Cooperation Program of China (ISTCP 2015DFE32850), funding to A.F. S. from the TUD Elite University Support the Best program. The authors acknowledge Vinothkannan Ravichandran's help in proofreading this manuscript.

Author Contributions
J.Y. participated in the design of this study, performed data collection analysis, and drafted the manuscript; H.M., X.Y., T.Q., F.Y., L.Q. and X.Z. participated in interpretation data; A.F.S. and R.M.
helped with the revision of the final manuscript. J.F. and Y.Z. designed and oversaw the study, performed data interpretation and drafted the manuscript.

**Additional Information**

**Supplementary information** accompanies this paper at http://www.nature.com/srep

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Yin, J. et al. Direct cloning and heterologous expression of the salinomycin biosynthetic gene cluster from *Streptomyces albus* DSM41398 in *Streptomyces coelicolor* A3(2). *Sci. Rep.* 5, 15081; doi: 10.1038/srep15081 (2015).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/
Direct cloning and heterologous expression of the salinomycin biosynthetic gene cluster from *Streptomyces albus* DSM41398 in *S. coelicolor* A3(2)

Jia Yin, Michael Hoffmann, Xiaoying Bian, Qiang Tu, Fu Yan, Liqiu Xia, Xuezhi Ding, A. Francis Stewart, Rolf Müller, Jun Fu and Youming Zhang
<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Characteristics</th>
<th>References or sources</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB05</td>
<td>$\Delta$(mrr-hsdRMS-mcrBC) $\varphi$B80lacZΔM15 ΔlacX74 recA1 endA1 araD139 $\Delta$(ara, leu)7697 galU galK $\lambda$ rpsL nupGfhuA::IS2 recET redα, phage T1-resistant</td>
<td>1</td>
</tr>
<tr>
<td>GB06</td>
<td>endA1 glnV44 thi-1 relA1 gyrA96 recA1 mcrB+ $\Delta$(lac-proAB) e14- [F' traD36 proAB+ lacIq lacZΔM15] hsdR17(rK-mK+) rpsL recET phage T1-resistant</td>
<td>1</td>
</tr>
<tr>
<td>GB05-dir</td>
<td>GB2005, araC-BAD-ETγA</td>
<td>2</td>
</tr>
<tr>
<td>GB05-red</td>
<td>GB2005, araC-BAD-γ</td>
<td>A</td>
</tr>
<tr>
<td>GB05dir-gyrA462</td>
<td>GB05dir, gyrA462 (ccdB resistant)</td>
<td>3</td>
</tr>
<tr>
<td>GB05red-gyrA462</td>
<td>GB05red, gyrA462 (ccdB resistant)</td>
<td>3</td>
</tr>
<tr>
<td>ET12567(pUZ8002)</td>
<td>recF, dam-, dcm-, CmR, KmR</td>
<td>4</td>
</tr>
<tr>
<td><strong>Streptomyces</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S. albus DSM41398</td>
<td>Salinomycin-producing wild-type strain</td>
<td>DSMZ</td>
</tr>
<tr>
<td>S. coelicolor A3(2)</td>
<td>Recipient strain of gene cluster</td>
<td>5</td>
</tr>
<tr>
<td>S. coelicolor::sal</td>
<td>Salinomycin gene cluster in the chromosome, aprR</td>
<td>This study</td>
</tr>
<tr>
<td><strong>Plasmid</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pSC101-BAD-Cre-tet</td>
<td>Cre under BAD promoter, tetR</td>
<td>6</td>
</tr>
<tr>
<td>pSC101-rha-ETgA-tet</td>
<td>RecET, redy and RecA under rha promoter, tetR</td>
<td>This study</td>
</tr>
<tr>
<td>pBeloBAC11-amp-ccdB-4Sal</td>
<td>pBeloBAC11, the vector for direct cloing, ampR</td>
<td>This study</td>
</tr>
<tr>
<td>p15A-amp-F1</td>
<td>p15A replicon, contained the fragment of salO-salAIV, ampR</td>
<td>This study</td>
</tr>
<tr>
<td>p15A-amp-F2-lox71-neo-lox66</td>
<td>p15A replicon, contained the fragment of salAIV-salAVIII and the cassette of lox71-neo-lox66, ampR, KmR</td>
<td>This study</td>
</tr>
<tr>
<td>p15A-amp-F3</td>
<td>p15A replicon, contained the fragment of salAIX-orf18, ampR</td>
<td>This study</td>
</tr>
<tr>
<td>Plasmid</td>
<td>Description</td>
<td>Sources</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>p15A-amp-F2</td>
<td>p15A replicon, the neo selection marker was deleted by Cre from the p15A-amp-F2-lox71-neo-lox66, ampR</td>
<td>This study</td>
</tr>
<tr>
<td>p15A-hyg-ccdB-F3-lox71-neo-lox66</td>
<td>p15A replicon, two cassettes of hyg-ccdB and lox71-neo-lox66 was integrated into the p15A-amp-salAI-ori18, hygR, kmR</td>
<td>This study</td>
</tr>
<tr>
<td>p15A-amp-F2&amp;3-lox71-neo-lox66</td>
<td>p15A replicon, contained the fragment of salAI-ori18 and the cassette of lox71-neo-lox66, ampR, kmR</td>
<td>This study</td>
</tr>
<tr>
<td>pBeloBAC11-sal-lox71-neo-lox66</td>
<td>pBeloBAC11, contained the whole salinomycin PKS and the cassette of lox71-neo-lox66, kmR</td>
<td>This study</td>
</tr>
<tr>
<td>pBeloBAC-sal-int-attP-oriT-apr</td>
<td>pBeloBAC11, contained the whole salinomycin PKS, an origin of transfer(oriT), attP sequence and integrase gene, aprR</td>
<td>This study</td>
</tr>
<tr>
<td>Gene</td>
<td>Primer name</td>
<td>Primer sequence 5'-3'</td>
</tr>
<tr>
<td>------------</td>
<td>-------------</td>
<td>-------------------------------------------------------------</td>
</tr>
<tr>
<td>salO</td>
<td>salO-5</td>
<td>AGTGAATTGTATACGACTCCTACTATAGGGGAATCCGTGAGCTCGGTACCCCGGTGGACGAGACCCGACCGCG</td>
</tr>
<tr>
<td>salO-3</td>
<td></td>
<td>GGCACCGCGCGGTCCATCA</td>
</tr>
<tr>
<td>orf18</td>
<td>Orf18-5</td>
<td>CTATAGGAAGGGGGCTTTTTTTAGAACAACAAACATTTATATCTTGAGGGGCTGAGATCt</td>
</tr>
<tr>
<td>Orf18-3</td>
<td></td>
<td>GTGACACTATGATAATAATCTGACGGTACGGAGCTGAAGAgATCACCAGGCCCCCCTGG</td>
</tr>
<tr>
<td>Amp-ccdB</td>
<td>Amp-ccdB-5</td>
<td>ATGCAGAGCAGACTGCCACCCTGAGTGGACCCTGGGCGGCGCGGCGGGAGATTAGGGAGGTCTCTTTTTTAAA</td>
</tr>
<tr>
<td>salO-salAIV</td>
<td>salO-salAIV-3</td>
<td>GCGTTCGCGCGGCGGCGGCGGGTTCGCGCTCCCAGCCGACCATGTTGAGGATCTGCAGGAGGATCGGACCCGCG</td>
</tr>
<tr>
<td>salO-salAIV-5</td>
<td></td>
<td>GTATACCTTTACAAATATTATTGACGTGAGG</td>
</tr>
<tr>
<td>salAIV-salAVIII</td>
<td>salAIV-salAV-3</td>
<td>TCCGAGACGAGCGAGCGGCGGCGGTTCGCGACCCAGCTCCCAGCTCCAGCCTCACCAGCCTCCAGCGACG</td>
</tr>
<tr>
<td>salAIV-salAVIII</td>
<td>salAIV-salAV-5</td>
<td>GCGTATTACAGCAATATTATTGACGTGAGG</td>
</tr>
<tr>
<td>salAIX-orf18</td>
<td>salAIX-orf18-3</td>
<td>TCGGCTCCAGCGGCGGCGGTTCGAAGTGATCGTCTGGTTKGAGGAGGAGCGAGGCGCGGCGA</td>
</tr>
<tr>
<td>salAIX-orf18</td>
<td>salAIX-orf18-5</td>
<td>CATCTGGCAGAGCGGCGGCGGTTCGAGCTCCCTCCCTGAGCCTCCGCTGCCGCGGCGGA</td>
</tr>
<tr>
<td>lox71-neo-lox66</td>
<td>sal-km-5</td>
<td>CTCGACTCGAGAAGCTCGGACCGACGAGAATACCTAGCAGACGACGACGAGGAGCGGCGGCGGTTCGCTGATATGAAGTCTATACGAAAGTAGTTACGAGAATCGCAAAGCTC</td>
</tr>
<tr>
<td>lox71-neo-lox66</td>
<td>sal-km-3</td>
<td>AAATCTGATCCACCGGGAAAGGGAAGGGAAGGGCGGCGGTACCGTGTAGCGTCTTAATACGCTTATACGAAAGTAGTTACGAGAATCGCAAAGCTC</td>
</tr>
<tr>
<td>Hyg-ccdB</td>
<td>hygccdb-5</td>
<td>ATGCAGAGCAGACTGCCACCCTGAGTGGACCCTGGGCGGCGCGGCGGGAGATTAGGGAGGTCTCTTTTTTAAA</td>
</tr>
<tr>
<td>hygccdb-3</td>
<td></td>
<td>ATGTAATCCGCTACGAGAAGAATAATCCTCTTGCAGTATATATTTGAGGCTAGGGTATGGAAAAGCCCTGAAA</td>
</tr>
<tr>
<td>kan-3</td>
<td>GAAGTTATGCTTGAGTTGGGCTTACAT</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td>GACGAGGAGGATTTGCTGGCGGACCTCAGAACTCCTCGGAGATCAGCATACCTGTGTAATATGTATTGCTTAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAAGTTATGCTTGAGTTGGGCTTACAT</td>
<td>GACGAGGAGGATTTGCTGGCGGACCTCAGAACTCCTCGGAGATCAGCATACCTGTGTAATATGTATTGCTTAC</td>
<td></td>
</tr>
<tr>
<td>Integrase-attP-o</td>
<td>TAGAGCGCCACCATCAGATGCGGCTGGAACTCCTCGGAGGGAACATTAGATCAGCTTTCCCGGGTGTCC</td>
<td></td>
</tr>
<tr>
<td>rT-Arapmycin</td>
<td>TAGAGCGCCACCATCAGATGCGGCTGGAACTCCTCGGAGGGAACATTAGATCAGCTTTCCCGGGTGTCC</td>
<td></td>
</tr>
<tr>
<td>intApSal5</td>
<td>TAGAGCGCCACCATCAGATGCGGCTGGAACTCCTCGGAGGGAACATTAGATCAGCTTTCCCGGGTGTCC</td>
<td></td>
</tr>
<tr>
<td>Engineering</td>
<td>TAGAGCGCCACCATCAGATGCGGCTGGAACTCCTCGGAGGGAACATTAGATCAGCTTTCCCGGGTGTCC</td>
<td></td>
</tr>
<tr>
<td>intApSal3</td>
<td>TAGAGCGCCACCATCAGATGCGGCTGGAACTCCTCGGAGGGAACATTAGATCAGCTTTCCCGGGTGTCC</td>
<td></td>
</tr>
<tr>
<td>attB-L</td>
<td>TAGAGCGCCACCATCAGATGCGGCTGGAACTCCTCGGAGGGAACATTAGATCAGCTTTCCCGGGTGTCC</td>
<td></td>
</tr>
<tr>
<td>attB-R</td>
<td>TAGAGCGCCACCATCAGATGCGGCTGGAACTCCTCGGAGGGAACATTAGATCAGCTTTCCCGGGTGTCC</td>
<td></td>
</tr>
<tr>
<td>attP-dn</td>
<td>TAGAGCGCCACCATCAGATGCGGCTGGAACTCCTCGGAGGGAACATTAGATCAGCTTTCCCGGGTGTCC</td>
<td></td>
</tr>
<tr>
<td>attP-up</td>
<td>TAGAGCGCCACCATCAGATGCGGCTGGAACTCCTCGGAGGGAACATTAGATCAGCTTTCCCGGGTGTCC</td>
<td></td>
</tr>
</tbody>
</table>

Underlined sequences indicate homology arms.

Bold letters represent restriction enzyme sites.
Figure S1 Digestion of the direct cloning products. (a) Products of fragment F1 were digested by NcoI. M, 1-kb NEB ladder. 1–24, clones obtained from direct cloning. (b) Products of fragments F2 and (c) F3 digested by PvuII. M, 1-kb NEB ladder. 1–24, clones obtained from direct cloning.
Figure S2 Alignment analysis of the F2 fragment and itself. The repetitive sequences were marked as same numbers.
Figure S3 Digestion of pBeloBAC11-sal-lox71-neo-lox66. The three fragments were stitched into pBeloBAC11, and the correct clone was confirmed by three restriction enzymes: AscI (1), ApaLI (2), and BgII (3). M: Marker (NEB 1-kb DNA ladder).
Figure S4 PCR verification of salinomycin gene cluster integration into the \textit{attB} site in \textit{S. coelicolor} A3(2). (a, b) Diagram of the PCR procedure in the wild-type (a) and mutant (b) strains using the indicated primers. Primers \textit{attB-L} and \textit{attB-R} were dependent on the sequence of \textit{S. coelicolor} A3(2). Primers \textit{attP-down} and \textit{attP-up} were based on the BAC sequence. (c) PCR products and primer pairs. (d) Agarose gel containing PCR-amplified DNA fragments using the primer pairs in (c). M, Marker; \textit{N}, wild-type \textit{S. coelicolor} A3(2); 1, exconjugant colonies.
References

Publication IV

Direct cloning and heterologous expression of the salinomycin biosynthetic gene cluster from *Streptomyces albus* DSM41398 in *S. coelicolor* A3(2).

Jia Yin, Michael Hoffmann, Xiaoying Bian, Qiang Tu, Fu Yan, Liqiu Xia, Xuezhi Ding, A. Francis Stewart, Rolf Müller, Jun Fu & Youming Zhang

Author Contributions

J.Y. participated in the design of this study, performed data collection analysis, and drafted the manuscript; H.M., X.Y., T.Q., F.Y., L.Q. and X.Z. participated in interpretation data; A.F.S. and R.M. helped with the revision of the final manuscript. J.F. and Y.Z. designed and oversaw the study, performed data interpretation and drafted the manuscript.

Signatures:

Jia Yin: 

Michael Hoffmann: 

Xiaoying Bian:  

Qiang Tu:  

Fu Yan: 

Liqiu Xia:  

Xuezhi Ding:  

A. Francis Stewart:  

Rolf Müller:  

Jun Fu:  

Youming Zhang:  

C. Final discussion

1. Development of a novel improved Red/ET recombination method

Introduction of foreign DNA into *E. coli* (transformation) is a basic and essential step for mutagenesis and genetic engineering of microorganisms. In order to enhance transformation efficacy, several methods have been reported in the literature to introduce exogenous DNA into the cells, including chemical treatment (using divalent cations like Ca$^{2+}$), electroporation, polyethylene glycol, utilization of a biolistic gun, ultrasound, hydrogel, and microwave techniques. Among these methods, electroporation has been proven to be most useful and convenient to genetically transform a number of microorganisms used for genetic studies.

The phage-derived homologous recombination systems, well known as recombineering, have been developed into very helpful DNA engineering technologies performing in electrocompetent cells, and it is very important to increase the transformation efficiency which means the number of cells transformed out of one microgram DNA.

The novel transformation of electrocompetent cells presented here, prepared at room temperature (named as room temperature competent cells), showed almost 10 times higher transformation efficiency than the conventional ice-cold electrocompetent cells in our research (Figure 3-1). As is well known, the electrocompetent cells were always prepared on ice in the conventional electroporation transformation, including other supplies, such as cold pre-chilled cuvette, buffer, and centrifuge. At such low temperatures, cell membrane topology would be modified for electrocompetent cells, and the fatty acid tails of the phospholipids in the cell membrane become more rigid, which impacts membrane fluidity and permeability, and in turn, cell survival.

In addition, a remarkable raise occurred in the permeability of cell’s plasma membrane, which was used to introduce exogenous DNA into bacterial cell during electroporation process due to externally applied electric field. These structural perturbations were associated with characteristic disturbances of function such as loss
C. Final discussion

of selective permeability. Afterwards, according to the electroporation theory, hydrophobic pores in the cell membrane were formed spontaneously by lateral thermal fluctuations of the lipid molecules, which suggested that hydrophobic pores formation would be enhanced at higher temperatures. Moreover, previous research found that the cell membranes underwent major morphological changes with respect to constantly changing environmental conditions, including temperature.

![Effect of temperature shift](image)

**Figure 3-1. Transformation efficiency of competent cells with temperature shift.** *E. coli* GB2005 cells transformed by plasmid pGB-amp-Ptet-plu1880 (27.8 kb) were plated on Amp plates. 1, the normal ice-cold method for preparing electrocompetent cells; 2, every step was done at RT; 5, no plasmid DNA. Error bars, s.d.; n = 3.

Recombineering is now a central technology for recombinant DNA engineering, as well as the method of choice for bacterial genome engineering. In order to improve recombination efficiency, many parameters had been described previously. For our Red/ET recombineering, whether LLHR by RecET or LCHR by Redαβ, the recombineering efficiency tests were also performed at room temperature. Interestingly, the situation in these two types of recombineering were not as good as we expected and the result was different. The reason was attributed to the instability of the Red recombinases (Red alpha and beta), which are related to LCHR recombineering while preparing the competent cells at room temperature. In a word, no one protocol was best for all the DNA manipulation.
C. Final discussion

Using improved Red/ET recombineering with room temperature competent cells, the secondary metabolite pathways can be more easily and rapidly performed in *E. coli*. We can take the appropriate procedure in the complex different cases, whether using room temperature competent cells in transformation, LLHR or cold temperature competent cells in LCHR. Additionally, in some cases, it is more deliberate in the construction of metabolite biosynthetic gene clusters because of knowing the minimum homology arms for recombineering. DNA transformation and recombineering are very critical steps in the study of synthetic biology in the recent revolution for discovering thousands of biosynthetic gene clusters and novel natural compounds. The transformation of exogenous DNA by room temperature competent cells in several other Gram-negative bacteria (*Agrobacterium, Burkholderia, Photorhabdus* and *Xenorhabdus*) have also been accomplished by our group, which demonstrates the powerful capability and wide adaptability of room temperature competent cells for potential bioprospecting. The improved approach will be a new guide and standard for our Red/ET recombineering in the future.

2. Several applications in mining microbial genomes via Red/ET technology, including disorazol, salinomycin and magnetosome product

Next, we used this developed transformation technique combined with direct cloning and other Red/ET methods for cloning and engineering of known disorazol gene cluster from *Sorangium cellulosum* So ce12 and salinomycin gene cluster from *S. albus* DSM41398, and for the challenging magnetosome gene clusters from *M. gryphiswaldense*.

2.1 Disorazol

Disorazol, initially isolated from the fermentation broth of the myxobacterium *Sorangium cellulosum* So ce12 in 1994,\textsuperscript{196} represents a class of antibiotics exhibiting inhibitory activity on cancer cell proliferation at low picomolar concentrations.\textsuperscript{197,198,209} Disorazols are the third group of myxobacterial secondary metabolites that interfere with tubulin polymerization following behind epothilones\textsuperscript{238}
C. Final discussion

and tubulysins. Because of this extraordinary potency, disorazols were fostered for development as peptide-conjugates for cancer therapy in drug discovery. Later on, the dis biosynthetic gene cluster was identified in 2005 by two independent groups through transposon mutagenesis. The whole disorazol gene cluster was larger than 60 kb with four main genes which were named disA-D. After comparison with databases, the anticipated disA-C genes encoded hybrid trans-AT type I PKS/NRPS megaenzymes, and the last gene disD encoded an additional acyl transferase protein. Our findings are in agreement with these results.

In this work, we reconstituted the 58-kb dis core gene cluster from BACs into a p15A replication origin (p15A ori) via Red/ET recombineering and expressed in M. xanthus DK1622. Gene clusters with moderate sequence complexity are usually stable in the p15A plasmid. Additionally, disorazol was highly toxic to the host cell due to its biological activity. All the E. coli strains harboring the core dis gene cluster with native promoter were found to carry mutations after recombineering. Therefore, to solve the issue, an inducible promoter Ptet was used to regulate gene expression because of its achievements in enabling several mixed PKS/NRPS natural products in many diverse hosts unrelated to the native producing organisms, including E. coli, M. xanthus and P. putida.

The dis gene cluster was randomly transposed into the chromosome of the heterologous host M. xanthus DK1622 by electroporation and several recombinants were verified by PCR to confirm that the dis gene cluster had been integrated into the M. xanthus chromosome in each case. After fermentation of mutants M. xanthus::p15A-dis, we demonstrated that the main compound disorazol A2 was produced at 5-fold higher levels than described in the native producer strain Soce12. The reason for this enhancement could be the effect of promoter Ptet which could raise transcription rate of the dis gene cluster.

Disorazol was characterized to be a typical trans-AT type PKSs. Based on the successful heterologous expression model, more research had been focused on the special AT domain (disD gene) of this trans-AT type PKS. According to the results of
overexpression the discrete \textit{disD} gene, the yield of disorazol increased approximately 2.5-fold when compared to the ones without overexpression. The event obviously indicated and confirmed the crucial position of the special AT domain in disorazol biosynthetic pathway. Enough supplementation of DisD could boost metabolite quantities to a modest extent. It was worth mentioning that the production of disorazol was currently limiting further progress of the compounds as neither synthetic nor fermentative methods had resulted in sufficient production strategies. Notably, abundant substrates consisting of sufficient ATs would facilitate the PKS module efficiency for polyketide chain extension of disorazol biosynthesis leading to elevated production.

Furthermore, successful heterologous expression of known full-length disorazol gene cluster in \textit{M. xanthus} DK1622 is another considerable mirror for Red/ET recombineering application in engineering metabolites biosynthetic gene clusters in the myxobacteria which own an ample source of natural products with biological activity. Disorazol, a novel macrocyclic compound with two oxazol rings, has almost highest inhibitory effects on mammalian cells ever recorded for a natural product. Clearly, it seems that disorazol has great potential development and market value. However, the first limitation of disorazol promotion is the drug production. The native natural product producer \textit{Soce12} grow poor and slowly even in optimized laboratory conditions. In this case, in order to increase the yields and industrialize development of disorazols, a surrogate heterologous host must be invested for the reconstitution and genetic manipulation of \textit{dis} biosynthetic pathways. A multitude of shared heterologous hosts are chosen to express such potential clinic compound because of theirs fast growth rate and the ease of achieving genetic manipulations, including \textit{M. xanthus} DK1622 here. All sorts of trials are provided towards a total synthesis of disorazol and to expand a natural derivative as an anticancer drug.

\textbf{2.2 Salinomycin}

Salinomycin, an antibiotic potassium ionophore with an unique tricyclic spiroketal ring system and an unsaturated six-membered ring isolated from \textit{Streptomyces}
C. Final discussion

*Streptomyces albus*,\(^{201}\) has been described to exhibit high antimicrobial activity against Gram-positive bacteria and murder breast cancer stem cells in mice at least 100 times higher effectively than paclitaxel.\(^{217-219}\) Previous studies have stated that the polyketide chain of salinomycin was synthesized by an assembly line of nine PKS multienzymes (named *salAI*, *salAII*, *salAIII*, *salAIV*, *salAV*, *salAVI*, *salAVII*, *salAVIII* and *salAIX*).\(^{220,221}\) Here, the 106 kb salinomycin gene cluster (*salO-orf18*) was cloned from the genomic DNA of *Streptomyces albus* DSM41398 by three rounds of direct cloning (LLHR) followed by assembly and successfully expression in a heterologous host *S. coelicolor* A3 (2), which resulted in the identification of salinomycin and its derivatives. All of the genes are oriented in the same direction and under the original promoters.

At the beginning, we attempted to directly clone the 106 kb *sal* biosynthetic gene cluster (*salO-orf18*) using one and two-step recombination reactions\(^{116}\) with the BAC vector but were unsuccessful. The failure to this case might have resulted from several issues. Firstly, the direct cloning efficiency, both in transformation and recombineering, was quite low for large DNA fragments, especially with high GC content. Co-transformation of two large linear molecules inside the cell was difficult as well as the recombineering efficiency would also drop for sequences with GC-rich content (data not shown). Secondly, it was arduous to enrich the target DNA after extracting the genomic DNA whose stability could be damaged under the mechanical destruction and degradation by nuclease activity. Consequently, it was challenging to gain the intact *sal* biosynthesis gene cluster from *S. albus* DSM 41398 straight.

*Streptomyces* is the largest genus of *Actinobacteria* who produce the large majority of characterized natural products.\(^{8}\) Over 500 species of *Streptomyces* bacteria have been described. With the purpose to analyze the known natural products biosynthetic gene clusters and discover the silent biosynthetic gene clusters in these precious resources, lots of genetic tools and approaches are developed for *Streptomyces* strains. Target secondary metabolite genes from *Streptomyces* were constructed and manipulated into bifunctional vectors in traditional *E. coli* species by the management of the λ Red
C. Final discussion

recombination system. Afterwards, the vectors were transformed into *Streptomyces* by conjugation and site-specific integrated into bacteriophage attachment (*attB*) sites for stable maintenance of the points. Consequently, target genes were expressed in model heterologous hosts for natural product production in modern biotechnology.

In summary, our work represents a powerful approach to mine the function of the individual genes and identify novel and potentially valuable analogues of the complex natural products through module exchange in the recipient.

2.3 Magnetosome products

Magnetotactic bacteria (MTB), well known to synthesize magnetic nanocrystals with uniform shapes and sizes at physiological conditions, serve as a source of multiple biological macromolecules taken for the biomimetic synthesis of a whole host of magnetic nanomaterials.76 Only a limited number of MTB have been isolated in pure culture so far, including our target *M. gryphiswaldense* strain.222 According to the 4.3 Mb genome draft sequence of *M. gryphiswaldense*,223 magnetosome-related mam and mms genes are organized as clusters in close proximity to each other.76,224 These clusters are organized as a larger unit, the so-called genomic magnetosome island (MAI), in some species. However, although the gene clusters within the magnetosome island that are required for biomineralization have already been identified in 2010,228 successfully transferring these specific gene clusters or indeed the whole island to a different bacterium has remained an elusive aim till now.

In our work, a minimal set of genes from the magnetotactic bacterium *M. gryphiswaldense* linked to magnetosome biosynthesis has been successfully expressed in the photosynthetic model organism *Rhodospirillum rubrum* (very close biological relationship to *M. gryphiwaldense*, whose 16S rRNA similarity to *M. gryphiwaldense* is 90%) using powerful Red/ET recombineering technology. These cassettes consist of four chief gene clusters on the magnetosome island that play either vital (*mamAB* gene, for magnetite biomineralization225) or regulatory (*mamGFDC* gene, for involving in controlling the size of magnetite crystals,226 mms6 gene, for the magnetosome membrane protein227 and *mamXY* gene) roles in magnetosome.
formation. These expression cassettes were transformed into *R. rubrum* and cultivated under the same conditions (iron ions in solution, microaerobic or anaerobic environment) as *M. gryphiswaldense*. The host cells *R. rubrum_ABG6* were found to be able to produce iron oxide particles of varying quality depending on which gene clusters were transferred. Consequently, the successful synthesis of magnetosomes within another organism expands the feasibility of commercialization of modified magnetosome production within bacteria which breaks the inexorable laws of nanomagnets under the strict control of magnetotactic bacterial genes. Similarly, depending on the use of Red/ET technology, a crucial following procedure would be integrated magnetosome biosynthetic gene clusters in *E. coli* and other easy-to-grow model bacterium. If fast-growing microbes could be utilized in the future to manufacture nanomagnets in a satisfied yield, higher value-added and greener industries for nanomagnets might emerge. Accompanied by a mass production of nanomagnets in a well-known model organism which is with the advantage of easy-control and stabilization, it could be widely used in development of targeted drugs, kits of early-diagnosis disease, heat treatment for tumor magnetic, molecular environment monitoring, and even the sealing materials for aviation helmet seals.

In a word, the Red/ET recombineering art mediated activation and expression of complex biosynthetic gene clusters in various heterologous hosts is exhibited to be a cogent tool for synthetic biology, not only for antibiotic in PKS/NRPS but also for other type natural products.

3. **New molecular technologies are required for construction and modification of large gene clusters in microbial genomes**

It is well known that the development of recombinant DNA technology in *E. coli* occurred in the early 1970s while natural products was playing an increasing important part in its progress as all vectors used for cloning depended on the use of genes that conferred resistance to antibiotics (such as penicillin, aureomycin, streptomycin, etc.). Subsequently, multifunctional plasmids, such as P1 vectors, BACs, harboring the specialized secondary metabolite genes were constructed and
engineered in *E. coli* under the use of λ Red recombination genetic engineering methods.\textsuperscript{229-231} Afterwards, the modified target constructs would be transformed and expressed in a host that was closely associated to its native producer. In addition, all analytical means were in a state of rapid advance with the developing of determining gene expression (transcriptomics, RNA-Seq), enzyme levels (proteomics), precursor supply (metabolomics) in *E. coli* and other well-characterized heterologous hosts.

Cloning, engineering and expression of long microbial genomic sequences in the chosen heterologous host is an effective approach in synthetic biology and genome engineering for natural products discovery and optimization, especially if the genomic sequences are derived from slow-growing bacteria or such one for which genetics are only poorly or not at all established. However, such large sequences (invariably more than 20 kb) are often arduous to acquire directly by traditional PCR or restriction enzyme digestion, and therefore the cloning of these sequences has remained a technical obstacle in molecular biology.

Much of the research in the last two decades has reported several synthetic biology tools which have already been developed to directly capture entire biosynthetic gene clusters from complex genomic DNA sources, circumventing the laborious construction and screening of genomic libraries.\textsuperscript{163} RecET-mediated linear–linear homologous recombination (LLHR),\textsuperscript{116} or transformation-associated recombination (TAR),\textsuperscript{175} CRISPR/Cas9-mediated TAR,\textsuperscript{233} Cas9-assisted targeting of chromosome segments (CATCH),\textsuperscript{234} and ϕBT1 integrase-mediated site-specific recombination\textsuperscript{235} are the present popular and powerful methods for direct capture of large DNA fragments from bacteria chromosome. Based on these advanced technologies, now gene cloning is not limiting genome mining anymore. Furthermore, with the continued technological and conceptual advances in bioinformatics, mass spectrometry, proteomics, transcriptomics, metabolomics and gene expression, the new field of microbial genome mining for applications in natural product discovery and development are gradually opening to scientists. Therefore, we predict a second “Golden Age for Antibiotics” and natural products discovery is coming.
D. References

### D. References

<table>
<thead>
<tr>
<th>Number</th>
<th>Author(s)</th>
<th>Title</th>
<th>Journal Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>38</td>
<td>Chan, Y. A., Boyne, M.T., Podevels, A.M., Klimowicz, A.K., Handelsman, J.,</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
References


D. References

3154 (2003).


D. References


83 Zhang Y., Muyrers J.P. and Stewart A.F. A new logic for DNA engineering using
D. References


101 Mullis, K. *et al.* Specific enzymatic amplification of DNA in vitro: the polymerase
References


van Kessel, J. C. & Hatfull, G. F. Recombinering in Mycobacterium tuberculosis.

46
D. References


154 Li, Y. et al. Directed natural product biosynthesis gene cluster capture and expression


170 Yin, J. *et al.* Direct cloning and heterologous expression of the salinomycin


186 Onaka, H., Taniguchi, S., Igarashi, Y. & Furumai, T. Cloning of the staurosporine biosynthetic gene cluster from *Streptomyces* sp. TP-A0274 and its heterologous
D. References


202 Singh, M., Yadav, A., Ma, X., & Amoah, E. Plasmid DNA Transformation in *Escherichia Coli* Effect of Heat Shock Temperature, Duration, and Cold Incubation of
D. References


D. References


233 Lee, N. C., Larionov, V. & Kouprina, N. Highly efficient CRISPR/Cas9-mediated TAR cloning of genes and chromosomal loci from complex genomes in yeast.
D. References


E. Appendix

Author’s efforts in publications presented in this work

Publication I:
Genetic engineering and heterologous expression of the disorazol biosynthetic gene cluster via Red/ET recombineering.
Q.T., S.H. and Y.Z. planned and performed cloning experiments. Q.T. and X.B. performed genetic transfers, cultivation experiments and data analysis. Q.T. and R.R. performed HPLC and compound isolation. R.R. performed NMR experiments and data analysis. J.H. performed biological functional studies. Q.T., Y.Z. and R.M. designed the study and wrote the paper. All authors discussed the results and commented on the manuscript.

Publication II:
Room temperature electrocompetent bacterial cells improve DNA transformation and recombineering efficiency.
Q.T. and J.Y. participated in the design of this study, performed data collection analysis, and drafted the manuscript; J.F, J.H, Y.L. and Y.Y. participated in interpretation data; A.F.S. and R.M. gave the advice for experimental design and discussed the data, also helped in the revision of the final manuscript. Y.Z. designed and oversaw the study, performed data interpretation and drafted the manuscript. All authors read and approved the final manuscript.

Publication III:
Biosynthesis of magnetic nanostructures in a foreign organism by transfer of bacterial magnetosome gene clusters.
I.K., D.S., Y.Z., Q.T., C.J. and R.M. planned and performed cloning experiments. I.K.
and A.L. performed genetic transfers and cultivation experiments. G.W. prepared cryo- and chemically fixed cells. S.B., O.R. and G.W. performed TEM and I.K. analyzed the data. J.P. and O.R. performed cryo-electron tomography experiments. E.T. and M.P. took high-resolution TEM micrographs and analyzed the data. I.K. and A.L. took fluorescence micrographs and performed phenotypization experiments. I.K. performed western blot experiments and analyzed proteomic data. A.B. performed Illumina genome sequencing and I.K. analyzed the data. I.K. and D.S. designed the study and wrote the paper. All authors discussed the results and commented on the manuscript.

Publication IV:

Direct cloning and heterologous expression of the salinomycin biosynthetic gene cluster from *Streptomyces albus* DSM41398 in *S. coelicolor* A3 (2).

J.Y. participated in the design of this study, performed data collection analysis, and drafted the manuscript; H.M., X.Y., T.Q., F.Y., L.Q. and X.Z. participated in interpretation data; A.F.S. and R.M. helped with the revision of the final manuscript. J.F. and Y.Z. designed and oversaw the study, performed data interpretation and drafted the manuscript.
Publication I

Genetic engineering and heterologous expression of the disorazol biosynthetic gene cluster via Red/ET recombineering.

Qiang Tu, Jennifer Herrmann, Shengbiao Hu, Ritesh Raju, Xiaoying Bian, Youming Zhang & Rolf Müller

Author Contributions

Q.T., S.H. and Y.Z. planned and performed cloning experiments. Q.T. and X.B. performed genetic transfers, cultivation experiments and data analysis. Q.T. and R.R. performed HPLC and compound isolation. R.R. performed NMR experiments and data analysis. J.H. performed biological functional studies. Q.T., Y.Z. and R.M. designed the study and wrote the paper. All authors discussed the results and commented on the manuscript.

Signatures:

Qiang Tu: 

Jennifer Herrmann: 

Shengbiao Hu: 

Ritesh Raju: 

Xiaoying Bian: 

Youming Zhang: 

Rolf Müller: 
Publication II

Room temperature electrocompetent bacterial cells improve DNA transformation and recombineering efficiency.

Qiang Tu*, Jia Yin*, Jun Fu, Jennifer Herrmann, Yuezhong Li, Yulong Yin, A. Francis Stewart†, Rolf Müller† and Youming Zhang†

Author Contributions

Q.T. and J.Y. participated in the design of this study, performed data collection analysis, and drafted the manuscript; J.F, J.H, Y.L. and Y.Y. participated in interpretation data; A.F.S. and R.M. gave the advice for experimental design and discussed the data, also helped in the revision of the final manuscript. Y.Z. designed and oversaw the study, performed data interpretation and drafted the manuscript. All authors read and approved the final manuscript.

Signatures:

Qiang Tu: Jia Yin:

Jun Fu: Jennifer Herrmann:

Yuezhong Li: Yulong Yin:

A. Francis Stewart: Rolf Müller:

Youming Zhang:
Publication III

Biosynthesis of magnetic nanostructures in a foreign organism by transfer of bacterial magnetosome gene clusters.

Isabel Kolinko, Anna Lohße, Sarah Borg, Oliver Raschdorf, Christian Jogler, Qiang Tu, Mihály Pósfai, Éva Tompa, Jürgen M. Plitzko, Andreas Brachmann, Gerhard Wanner, Rolf Müller, Youming Zhang¹ and Dirk Schüler¹

Author contributions
I.K., D.S., Y.Z., Q.T., C.J. and R.M. planned and performed cloning experiments. I.K. and A.L. performed genetic transfers and cultivation experiments. G.W. prepared cryo- and chemically fixed cells. S.B., O.R. and G.W. performed TEM and I.K. analysed the data. J.P. and O.R. performed cryo-electron tomography experiments. E.T. and M.P. took high resolution TEM micrographs and analysed the data. I.K. and A.L. took fluorescence micrographs and performed phenotypization experiments. I.K. performed western blot experiments and analysed proteomic data. A.B. performed Illumina genome sequencing and I.K. analysed the data. I.K. and D.S. designed the study and wrote the paper. All authors discussed the results and commented on the manuscript.

Signatures:
Isabel Kolinko: ¹Kolinko  Anna Lohße: ¹Lohße
Sarah Borg: ¹Sarah Borg  Oliver Raschdorf: ¹Raschdorf
Christian Jogler: ¹Jogler  Qiang Tu: ¹Tu
Mihály Pósfai: ¹Pósfai  Éva Tompa: ¹Tompa
Jürgen M. Plitzko: ¹Plitzko  Andreas Brachmann: ¹Brachmann
Gerhard Wanner: ¹Wanner  Rolf Müller: ¹Müller
Youming Zhang: ¹Zhang  Dirk Schüler: ¹Schüler
Publication IV

Direct cloning and heterologous expression of the salinomycin biosynthetic gene cluster from *Streptomyces albus* DSM41398 in *S. coelicolor A3(2).

Jia Yin, Michael Hoffmann, Xiaoying Bian, Qiang Tu, Fu Yan, Liqiu Xia, Xuezhi Ding, A. Francis Stewart, Rolf Müller, Jun Fu* & Youming Zhang

Author Contributions

J.Y. participated in the design of this study, performed data collection analysis, and drafted the manuscript; H.M., X.Y., T.Q., F.Y., L.Q. and X.Z. participated in interpretation data; A.F.S. and R.M. helped with the revision of the final manuscript. J.F. and Y.Z. designed and oversaw the study, performed data interpretation and drafted the manuscript.

Signatures:

Jia Yin:  
Michael Hoffmann:  
Xiaoying Bian:  
Qiang Tu:  
Fu Yan:  
Liqiu Xia:  
Xuezhi Ding:  
A. Francis Stewart:  
Rolf Müller:  
Jun Fu:  
Youming Zhang:  