

Investigation of the Biosynthetic Potential of Endophytes in Traditional Chinese Anticancer Herbs

Kristin I. Miller^{1,2}, Chen Qing³, Daniel Man Yuen Sze^{1,4}, Brett A. Neilan^{2,5*}

1 Faculty of Pharmacy, The University of Sydney, Sydney, New South Wales, Australia, **2** School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Sydney, New South Wales, Australia, **3** Yunnan Key Laboratory of Pharmacology for Natural Products, School of Pharmaceutical Science, Kunming Medical University, Kunming, China, **4** Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hung Hom Kowloon, Hong Kong, **5** Australian Centre for Astrobiology, The University of New South Wales, Sydney, New South Wales, Australia

Abstract

Traditional Chinese medicine encompasses a rich empirical knowledge of the use of plants for the treatment of disease. In addition, the microorganisms associated with medicinal plants are also of interest as the producers of the compounds responsible for the observed plant bioactivity. The present study has pioneered the use of genetic screening to assess the potential of endophytes to synthesize bioactive compounds, as indicated by the presence of non-ribosomal peptide synthetase (NRPS) and polyketide synthase (PKS) genes. The total DNA extracts of 30 traditional Chinese herbs, were screened for functional genes involved in the biosynthesis of bioactive compounds. The four PCR screens were successful in targeting four bacterial PKS, six bacterial NRPS, ten fungal PKS and three fungal NRPS gene fragments. Analysis of the detected endophyte gene fragments afforded consideration of the possible bioactivity of the natural products produced by endophytes in medicinal herbs. This investigation describes a rapid method for the initial screening of medicinal herbs and has highlighted a subset of those plants that host endophytes with biosynthetic potential. These selected plants can be the focus of more comprehensive endophyte isolation and natural product studies.

Citation: Miller KI, Qing C, Sze DMY, Neilan BA (2012) Investigation of the Biosynthetic Potential of Endophytes in Traditional Chinese Anticancer Herbs. PLoS ONE 7(5): e35953. doi:10.1371/journal.pone.0035953

Editor: William C. S. Cho, Queen Elizabeth Hospital, Hong Kong

Received: September 2, 2011; **Accepted:** March 26, 2012; **Published:** May 22, 2012

Copyright: © 2012 Miller et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: KIM and DMYS were supported by the Australia-China Special Fund for S&T Cooperation (CH070067) <https://grants.innovation.gov.au/ISL/Pages/Doc.aspx?name=ChinaFund.htm>. CQ was financially supported by the National Natural Science Foundation of China (No.30811120009) and the International Cooperation Foundation of Yunnan Province of China (No.30811120009). BAN was supported by a grant and fellowship from the Australian Research Council (FF0883440) <http://www.arc.gov.au/>. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: b.neilan@unsw.edu.au

Introduction

There is an ongoing need for novel drugs that are highly effective in the treatment of cancer, drug resistant bacteria, fungal infections, emerging viruses and parasitic protozoan infections. Historically, natural products have provided the basis for the majority of new drugs and the bioactive properties of a wide variety of flora is reflected in their continued roles in the traditional healthcare systems of many cultures [1]. The successful use of plants in traditional medicine and modern natural products research has meant that there is a renewed interest in exploiting various aspects of the underlying bioactivities.

Still widely practiced in the modern era, Traditional Chinese Medicine (TCM) can trace its origins back thousands of years, to the dawn of civilisation in China. TCM theory is based on experiences of the effects of the medicines through documented trials to establish knowledge of the use of plants (approximately 5000 species) for the treatment of many diseases [2]. This TCM background has provided the basis for the discovery of several therapeutic agents, including the anticancer agents indirubin [3], camptothecin [4] and harringtonine [5]; artemisinin (antimalarial) [6] and ephedra (central nervous system stimulant) [7].

Compounds isolated from plant preparations are largely the products of plant metabolism, however, microorganisms living in symbiosis with plants also produce bioactive compounds. Taxol

and camptothecin are examples of anticancer compounds synthesized by both plants and endophytes [8,9]. In their natural environment the internal tissues of plants are frequently colonized by numerous different microorganisms, termed endophytes. Many endophytes produce secondary metabolites which confer major ecological benefits to their host plants including plant growth promotion [10], enhanced resistance to various predators and pathogens [11,12], and increased drought resistance [13]. Consequently, endophyte derived metabolites may also cause the observed bioactivity and associated beneficial health claims of the TCM host plants [14].

Low molecular weight endophytic secondary metabolites demonstrate a large degree of structural diversity with the largest and most important groups of compounds including the polyketides, amino acid derived compounds, and terpenes [15]. To date genetic methods have been used to screen for biosynthetic pathways involved in secondary metabolism. Genetic screening for microbial natural product genes has largely focused on the detection of the polyketide and non-ribosomal peptide synthesis pathways [16,17].

Polyketides are produced by many fungi, bacteria, plants and marine organisms [15]. This large family of structurally diverse natural products have already found widespread application as pharmaceuticals including rapamycin (immunosuppressant), erythromycin (antibiotic), lovastatin (anticholesterol drug), and

epothilone B (anticancer drug). Polyketides are biosynthesized by large multimodular enzyme complexes termed polyketide synthases (PKSs) which catalyse the polymerisation of acyl-CoA thioester building blocks [18]. PKSs are typically categorized based on their number of subunits (single or multiple) and their mode of synthesis (iterative or modular). The best characterized group of PKSs are the type I PKSs with a single enzyme complex responsible for the biosynthesis of the polyketide backbone. They can be either modular as found in bacteria, or iterative as found in fungi [19]. Extension of the polyketide backbone requires three core PKS domains: an acyltransferase (AT) domain, acyl carrier proteins (ACP) domain and a ketosynthase (KS) domain. The modular structure of type I PKSs and their assembly of polyketides, resembles the organisation of non-ribosomal peptide synthetases (NRPSs) and the production of non-ribosomal peptides.

Complex oligopeptides are formed by NRPSs through the linear condensation of proteinogenic and non-proteinogenic amino acids. NRPSs are involved in the biosynthesis of compounds including penicillin and cephalosporin (antibiotics), cyclosporine A (anti-inflammatory and immunosuppressive activities, and the endophyte-produced fusaricidin (antibiotic). NRPSs are composed of a series of modules, each containing the catalytic units required for the recognition, activation and peptide bond formation of the growing peptide chain by adenylation (A), thiolation (T), and condensation (C) domains, respectively [20]. NRPSs are found in bacteria and fungi feature the three core domains arranged in the order C-A-T [20].

While mechanistically distinct, the highly conserved regions within the KS and A domains are logical targets for the detection of PKS and NRPS biosynthesis genes, respectively. This facilitates the design of degenerate oligonucleotide primers, suitable for genetic screening of secondary metabolism pathways [21,22,23].

The location of host plants can affect endophyte populations [24,25] and for drug discovery potential it is imperative to investigate TCM plants sampled from their natural environment or those cultivated for medicinal purposes. Considering this, plant collections from Yunnan province were selected for two main reasons. Firstly, Yunnan is recognized for its large botanical diversity [26] and secondly, it is a source of many pharmaceutical herbs. TCM herbs with reported anticancer activity were chosen in an effort to increase the likelihood of discovering endophyte produced anticancer compounds [2].

The aim of this study was to assess the use of genetic screening for the identification of plant leads for endophyte natural product investigations. The biosynthetic potential of endophytes was considered via the detection of endophyte PKS and NRPS biosynthesis genes. Total DNA from thirty selected anticancer TCM herbs was genetically screened using four PCR primers sets. This study has identified a subset of TCM herbs found to harbor endophytes with bioactive potential, and these plants can be targeted in future investigations. The use of this genetic screen increases the likelihood of isolating endophytes which biosynthesize novel pharmaceuticals independently of the host plant.

Methods

Selection, collection and storage of TCM herbs

Yunnan province is situated in the southwestern region of the People's Republic of China. There is a strong temperature gradient across the province and it contains areas of permanent snow as well as rainforests with temperatures perennially above 10°C [26]. Thirty plant tissue samples (Table 1) were collected during spring in April, 2007 from sites within and surrounding Kunming City. No specific permits were required, however, the

collections were acknowledged and supported by the Yunnan Provincial government. Potential anticancer TCM herbs were selected based on scientific and traditional literature, their maturity for use as TCMs and their availability for collection at the time of collection. Following collection, samples were placed in sterile bags and stored in the dark at 4°C until processed. Processing was completed as soon as possible following receipt of samples, usually within 24 h.

Surface sterilisation of plant samples

Sections of each plant were carefully washed in running water to remove external soil and debris. Samples were examined to exclude those that showed symptoms of disease or superficial damage. In a method adapted from Cankar et al. [27], plant material was surface sterilized by immersion in 30% hydrogen peroxide for 30 min, followed by rinsing in sterile Milli-Q filtered water (Millipore) and washing with 70% ethanol for 1 min [27]. To confirm that the surface sterilisation process was successful, a section of each sterilized plant was placed into an aliquot of heart infusion broth and incubated at 30°C for three days before examining the media for any growth of plant surface-associated contaminating microorganisms.

DNA extraction

The different tissue types collected from the one plant were processed together for the total DNA extract. Total genomic DNA was extracted from the surface sterilized plant samples using the Power Plant DNA isolation kit (MoBio Laboratories). Extracted DNA consisted of host plant nuclear and plastid DNA and any endophyte microbial DNA. Extracted DNA was run on an agarose gel and NanoDrop (Thermo Scientific) readings taken to assess concentration and purity. Working stocks of DNA extracts were diluted to 1–10 ng/μL for PCR assays.

Polymerase Chain Reaction

DNA extracts from TCM plants were screened using PCR primers that targeted various plant, bacterial and fungal genes. The PCR targets and primers employed in this investigation (NRPS, PKS, and the plant RUBISCO genes) are listed in Table 2. Typical gene-specific PCR reactions were carried out in 20 μL reaction volumes that contained 1×reaction buffer (Bioline), 2.5 mM MgCl₂, 0.2 mM dNTPs (Bioline), 10 pmol each of forward and reverse primers (Sigma-Aldrich, MO, USA), 0.2 U *Taq* DNA polymerase (Bioline) and 1–10 ng of DNA template. Thermal cycling was performed in a MyCycler Thermal Cycler (BioRad), with an initial denaturation at 94°C for 5 min, followed by 30 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min, followed by a final extension step at 72°C for 7 min. Degenerate PCRs were also performed in 20 μL reaction volumes containing the same reaction mix as for the non-degenerate primers except that 25 pmol of forward and reverse primers were used and the number of amplification cycles was increased to 35. PCR amplicons were separated by agarose gel electrophoresis in TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 7.8), visualized by ethidium bromide staining (0.5 μg/ml) and the GelDoc UV transilluminator (BioRad).

A spiked PCR targeting the conserved plant RUBISCO gene was used to ensure the absence of PCR inhibitors in the DNA samples. This procedure involved two PCRs per DNA sample, the first containing both sample DNA and an inhibitor free control DNA and the second reaction contained only sample DNA. Lack of amplification in control-spiked reactions was likely to be due to PCR inhibitors in the sample DNA. In this event, DNA was diluted to an extent where the PCR reaction was not inhibited.

Table 1. Traditional Chinese medicinal herbs collected for genetic screening.

| Host Plant Species | Location | Plant Sample | Pharmacological or TCM Activity | Reference |
|--|-----------------------|-----------------|---|-----------|
| <i>Aconitum carmichaeli</i> Debx. | Countryside | Leaf and root | Sedative, analgesic, antiinflammatory, anticancer, toxic | [65] |
| <i>Actinidia chinensis</i> Planch. | Medicinal herb farm | Root and stem | "Heat clearing", antimutagenic, anticancer | [66] |
| <i>Agrimonia pilosa</i> Ledeb. var. <i>japonica</i> (miq.) Nakai | Medicinal herb farm | Above ground | Haemostatic, antiinflammatory, antiinfective | [2] |
| <i>Artemisia capillaris</i> Thunb. | Medicinal herb garden | Leaf and stem | "Heat clearing", antihepatotoxic, antibacterial, antiinflammatory | [2] |
| <i>Astilbe rivularis</i> Buch.-Ham. | Medicinal herb garden | Whole plant | Antiviral | [67] |
| <i>Belamcanda chinensis</i> (L.) DC. | Medical herb farm | Root | Antifungal, antiviral, antiinflammatory, diuretic | [2] |
| <i>Bletilla striata</i> (Thunb.) Reichp. | Medical herb farm | Root/bulb | Haemostatic, anticancer | [2] |
| <i>Cephalotaxus fortunei</i> Hook. F. | Countryside | Leaf and stem | Anticancer | [68] |
| <i>Clematis hexapetala</i> Pall. | Medicinal herb farm | Root | Antibacterial, antifungal | [2] |
| <i>Coptis chinensis</i> Franch. | Medicinal herb farm | Root | Antibacterial, antiinflammatory, affects central nervous system | [2] |
| <i>Dendrobium nobile</i> , Lindl. | Countryside | Stem | Anticancer, antimutagenic | [69] |
| <i>Digitalis purpurea</i> Ehrh. | Medicinal herb garden | Root | Cardiotonic, anticancer | [70] |
| <i>Eleutherococcus senticosus</i> (Rupr. & Maxim) Maxim | Medicinal herb farm | Root | Anticancer, expectorant | [71] |
| <i>Leonurus heterophyllus</i> Sweet | Medicinal herb farm | Above ground | Anticancer, "invigorates blood" | [68] |
| <i>Lithospermum erythrorhizon</i> Sieb.et Zucc. | Medicinal herb farm | Root | Anticancer, antibacterial, antiinflammatory | [2] |
| <i>Lonicera japonica</i> Thunb. | Countryside | Leaf | Antiinfective | [2] |
| <i>Lycium chinense</i> Mill | Medicinal herb garden | Branch and leaf | Anticancer, immune stimulating | [68] |
| <i>Pinellia ternata</i> (Thunb.) Breit. | Countryside | Root | Anti-emetic, mucolytic, anticancer (cervical), discutient | [65] |
| <i>Pinellia pedatisecta</i> Schott | Medicinal herb farm | Root | Toxic, anticancer | [66] |
| <i>Paris polyphylla</i> Smith var. <i>chinensis</i> (Franch.) Hara | Medicinal herb farm | Root and leaf | Antiinflammatory, detoxicant, anticancer | [66] |
| <i>Rheum palmatum</i> L. | Medicinal herb farm | Root | Anticancer, detoxicant, antiinflammatory, laxative | [2] |
| <i>Rubia cordifolia</i> L. | Countryside | Leaf and root | Antiinflammatory, antiviral | [72] |
| <i>Salvia przewalskii</i> Maxim | Medicinal herb farm | Root | Anticancer, immune stimulating, antioxidant | [73] |
| <i>Salvia yunnanensis</i> C.H. Wright | Medicinal herb farm | Root | Anticancer, immune stimulating | [73] |
| <i>Senecio scandens</i> Buch.-Ham. Ex D. | Countryside | Leaf and root | Antiinflammatory, anticancer, detoxicant, antibacterial | [68] |
| <i>Solanum nigrum</i> L. | Countryside | Leaf and root | Antimicrobial | [68] |
| <i>Sophora davidii</i> (Franch) Kom. Ex Pavol. | Medicinal herb garden | Leaf and root | Anticancer, antiinflammatory | [74] |
| <i>Sophora japonica</i> L. | Countryside | Leaf and stem | Hemostatic | [68] |
| <i>Taxus baccata</i> L. | Countryside | Leaf and stem | Anticancer | [75] |
| <i>Thalictrum foliolosum</i> DC | Medicinal herb garden | Leaf and root | Anticancer, antiinfective, antihypertensive | [74] |

doi:10.1371/journal.pone.0035953.t001

The second PCR reaction, containing only sample DNA, was used to ensure that DNA was not diluted beyond the PCR detection limits.

DNA sequencing

PCR amplicons, of the correct size, were purified directly by ethanol precipitation or gel purification using the UltraClean GelSpin DNA Extraction Kit (MoBio Laboratories) in accordance with the manufacturer's instructions.

Amplicons were sequenced using the PRISM BigDye™ cycle sequencing system v3.1 on an ABI PRISM 373 DNA Sequencer (Life Technologies, Carlsbad, USA). PCR products that contained multiple sequences, as observed by multiple peaks in the sequencing chromatograms, were cloned using the TOPO cloning vector (Invitrogen Life Technologies, Carlsbad, USA), trans-

formed into *Escherichia coli* DH5 α and a clone library established according to the manufacturer's instructions. Thereafter, colony PCR screening and DNA sequencing of cloned inserts was performed to characterize single amplicons. Sequences were deposited into GenBank.

Phylogenetic analysis of KS and NRPS fragments

Sequence data were compared to the NCBI dataset on GenBank with the BLASTX algorithm (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The amino acid substrates recognized by the A domain binding pockets were predicted using the NRPSpredictor analysis tool available at (<http://www-ab.informatik.uni-tuebingen.de/toolbox/index.php?view=domainpred>) [28]. Multiple sequence alignments of KS and A domain protein sequences of PKS or NRPS pathways, respectively, were generated using

Table 2. PCR primers and reaction conditions used in the screening of total DNA extracts from TCM plants.

| Target | Primer | Primer Sequence (5'→3') | Product size (bp) | Altered* reaction conditions | Altered* thermocycling conditions | Reference |
|--|---------|----------------------------|-------------------|---|--|-----------|
| Plant RUBISCO enzyme | rbcl-F | tgtcaccacaacagarackaa | ~1400 | | | [76] |
| | rbcl-R | caaaattaatmsgatctcttccatac | | | | |
| Adenylation domain, Bacterial NRPSs | MTF2 | gcnggygygycntaygtnc | ~1000 | | 94°C for 10 s, 52°C for 30 s, 72°C for 30 s | [77] |
| | MTR2 | ccncgdaytnacytg | | | | |
| Ketosynthase domain, Bacterial PKSs (Type 1) | DKF | gtgccggtncrrtngngytc | ~650–700 | | | [41] |
| | DKR | gcgatggaycncarcarmg | | | | |
| Adenylation domain, Fungal NRPSs | RJ016-R | arrtncncngtyttrta | ~300 | | 94°C for 30 s, 50°C for 30 s, 60°C for 1 min; final extension: 60°C for 10 min | [78] |
| | RJ016-F | tayggnccnacnga | | | | |
| Ketosynthase domain, Fungal PKSs (Type 1) | LC1 | gayccimgittytyaayatg | ~700 | Primers (pmol): LC1, 50; LC2c, 8; LC3, 24; LC5c, 16 | 94°C for 30 s, 56°C for 10 s, 72°C for 30 s | [79] |
| | LC2c | gticcigticrctgcatytc | | | | |
| | LC3 | gcigarcaratggayccica | ~700 | | | |
| | LC5c | gtigaigticrtgigcytc | | | | |
| Adenylation domain, Streptomyces NRPSs | A3F | gcstacsysatstacacstcsgg | ~700 | 10 pmol each primer, 0.25 mM dNTPs, 10% DMSO | 94°C for 30 s, 58°C for 1 min, 72°C for 2 min | [33] |
| | A7R | sasgtcvccsgtscggtas | | | | |
| Ketosynthase domain, Streptomyces PKS (Type 1) | K1F | tsaagtcsaacatcggbca | ~1200–1400 | 20 pmol each primer, 0.25 mM dNTPs, 10% DMSO | 94°C for 30s, 55°C for 1 min, 72°C for 3 min | [33] |
| | M6R | cgcaggttscsgtaccagta | | | | |

*Altered from standard PCR procedures described.
doi:10.1371/journal.pone.0035953.t002

ClustalX [29] (Figure S1). Gaps and positions with ambiguities were excluded from the phylogenetic analysis. The ProtTest program was used to determine the most appropriate substitution model for the dataset [30]. Phylogenetic analysis was performed using the maximum likelihood methods [31]. Node support values were determined by the approximate Likelihood-Ratio Test (aLRT).

Results

Identification of genes encoding bioactivity in TCM plants

The total DNA extracted from thirty TCM plants was expected to contain a mixture of plant and endophyte genes. The DNAs were free from PCR inhibitors, as evidenced by successful PCR amplification of the plant RUBISCO gene. Degenerate PCRs were used to detect putative NRPS and PKS gene sequences originating from bacterial and fungal endophytes in the DNA extracts. Specifically, the PCR screens were successful in targeting endophyte sequences with four streptomyces KS, six streptomyces A, ten fungal KS and three fungal A domain genes detected (Table 3). Some of the amplified PKS and NRPS PCR products

could be directly sequenced, however, many contained a mixture of amplicons that needed to be cloned into a TOPO cloning vector. In such cases, at least ten transformants were randomly selected for direct colony PCR and the products sequenced. Amplification of bacterial and fungal A domains and KS domains was confirmed via sequencing and BLASTX (translated) analysis (Table 3). BLASTN (nucleotide) analysis of the PCR products did not reveal any significant matches to the nucleotide database, indicating that the sequences were novel at this level. Endophyte-derived genes were amplified from DNA extracted from different types of plant tissues including the roots, leaves, stems and bark (Table 1).

Ten unique fungal KS domain sequences (~700 bp) were amplified from total DNA extracts of *Bl. striata*, *C. fortunei*, *L. heterophyllus*, *Lo. japonica*, *P. temata*, *S. scandens*, and *T. baccata* (Table 3). The amplified sequences possessed between 76% and 100% amino acid identity to known fungal PKSs. Specifically, sequence LH2 shared 100% sequence identity with an uncharacterized PKS from *Botryotinia fuckeliana*. Other fungal PKS amplicons were similar to genes involved in the production of polyketides, such as melanin and the toxic pigment elsinochrome.

Table 3. PKS and NRPS genes identified with degenerate PCR primers.

| Gene | Plant | Sequence identifier | BLASTX Match | Identity (%) | Accession number | Predicted binding pocket (amino acid substrate) ^{1,2} |
|----------------|-------------------------|---------------------|---|----------------|------------------|--|
| Fungal PKS | <i>Bl. striata</i> | BS1 | Elsinochrome PKS, <i>Elsinoe fawcettii</i> | 124/161 (77%) | ABU63483 | ND |
| Fungal PKS | <i>C. fortunei</i> | CF1 | PKS involved in melanin production, <i>Bipolaris oryzae</i> | 189/210 (90%) | BAD22832 | ND |
| Fungal PKS | <i>L. heterophyllum</i> | LH1 | non-reduced type PKS for melanin pigment, <i>Ascochyta rabiei</i> | 232/236 (98%) | ACS74449 | ND |
| Fungal PKS | <i>L. heterophyllum</i> | LH2 | Uncharacterized PKS, <i>Botryotinia fuckeliana</i> | 221/221 (100%) | AAR90249 | ND |
| Fungal PKS | <i>Lo. japonica</i> | LJ1 | Putative non-reduced type PKS, <i>Penicillium</i> sp. | 193/207 (93%) | ABQ85550 | ND |
| Fungal PKS | <i>P. ternata</i> | PT1 | Conidial pigment PKS, <i>Verticillium albo-atrum</i> | 197/211 (93%) | EEY14472 | ND |
| Fungal PKS | <i>S. scandens</i> | SS1 | PKS 1, <i>Glarea lozoyensis</i> | 198/220 (90%) | AAN59953 | ND |
| Fungal PKS | <i>S. scandens</i> | SS2 | Putative melanin PKS, <i>Penicillium</i> sp. | 205/243 (98%) | ACJ13039 | ND |
| Fungal PKS | <i>S. scandens</i> | SS3 | PKS 1, <i>Glarea lozoyensis</i> | 205/243 (84%) | AAN59953 | ND |
| Fungal PKS | <i>T. baccata</i> | TB1 | Elsinochrome PKS, <i>Elsinoe fawcettii</i> | 205/243 (84%) | ABU63483 | ND |
| Fungal NRPS | <i>Pi. pedatisecta</i> | PPe1 | MicC synthetase, <i>Planktothrix rubescens</i> | 20/34 (58%) | CAQ48260 | No prediction |
| Fungal NRPS | <i>Pi. pedatisecta</i> | PPe2 | AerB synthetase, <i>Planktothrix rubescens</i> | 56/95 (58%) | CAQ48266 | No prediction |
| Fungal NRPS | <i>Pi. pedatisecta</i> | PPe3 | Arthrofactin synthetase, <i>Burkholderia ghumae</i> | 65/95 (68%) | YP_002908545 | No prediction |
| Bacterial PKS | <i>D. purpurea</i> | DL1 | Nystatin synthase (NysC) <i>Streptomyces halstedii</i> | 103/174 (59%) | AAF71776 | ND |
| Bacterial PKS | <i>Pa. polyphylla</i> | PPol1 | NapC synthase <i>Streptomyces hygrosopicus</i> | 157/218 (72%) | ABB86421 | ND |
| Bacterial PKS | <i>Pa. polyphylla</i> | PPol2 | NapC synthase <i>Streptomyces hygrosopicus</i> | 166/235 (70%) | ABB86421 | ND |
| Bacterial PKS | <i>Pi. pedatisecta</i> | PPe4 | Arthrofactin synthetase, <i>Pseudomonas</i> sp. MIS38 | 47/63 (74%) | BAC67535 | ND |
| Bacterial NRPS | <i>D. purpurea</i> | DL2 | Peptide synthetase PhsB, <i>Streptomyces viridochromogenes</i> | 83/130 (63%) | CAJ14037 | DVEHLSLID- (pro) |
| Bacterial NRPS | <i>D. purpurea</i> | DL3 | Ansamitocin PKS, <i>Actinosynnema pretiosum</i> | 121/231 (52%) | AAM54075 | -AFALACGM- (val/leu/ile/abu/iva) |
| Bacterial NRPS | <i>L. heterophyllum</i> | LH3 | Amino acid adenylation domain <i>Acidovorax avenae</i> | 18/19 (94%) | YP_972054 | DVWNIGLI(thr) |
| Bacterial NRPS | <i>L. heterophyllum</i> | LH4 | Pyoverdine synthetase, <i>Pseudomonas fluorescens</i> | 154/181 (85%) | AAF40219 | CVWHFGRI (glu) |
| Bacterial NRPS | <i>B. chinensis</i> | BC1 | Amino acid adenylation domain, <i>Rhodococcus erythropolis</i> | 192/195 (98%) | ZP_04387369 | DATFAGGI (leu/ile/val) |
| Bacterial NRPS | <i>B. chinensis</i> | BC2 | Amino acid adenylation domain, <i>Rhodococcus erythropolis</i> | 224/228 (98%) | ZP_04387369 | DATFAGGI (leu/ile/val) |

¹In silico prediction of the amino acid substrate recognized by putative NRPS fragments [28].

²ND= not done.

doi:10.1371/journal.pone.0035953.t003

Screening of total TCM plant DNA successfully amplified ~300 bp fragments of the fungal NRPS A domain. Of all the DNAs screened only DNA from *Pi. pedatisecta* produced NRPS gene fragments (PPe1, PPe2, PPe3). BLASTX analysis revealed

that the predicted translated DNA sequences were homologous to bacterial NRPSs, with sequence similarity between 58% and 68%. Prediction of the eight amino acids that form the A domain active site, and the amino acids that bind to the active residues, was not

successful possibly indicating that the algorithm is not sufficiently robust to predict fungal NRPS module specificity [32].

Bacterial directed primers were also used to screen for PKS and NRPS genes [33]. Degenerate primers targeting the genes coding for the KS domain of actinobacterial PKSs amplified a total of four unique amplicons from the plant DNA extracts, including *D. purpurea* (one), *P. polyphylla* (two), and *P. pedatisecta* (one). Predicted translation of the gene fragments was afforded through the BLAST program and subsequent peptide sequence analysis established that three of the fragments were similar to sequences of PKSs involved in the biosynthesis of the antifungal compound nystatin (DL1, 59%) and the anticancer and antibiotic compound geldanamycin (PpO1, 72% and PpO2, 70%). The sequence of PpE4 was shorter (~235 bp) and had 74% similarity to arthrofactin synthetase from *Pseudomonas* sp. MIS38 (Table 3).

Screening of the TCM plant DNA extracts for bacterial NRPS A domain gene sequences revealed six unique amplicons from the extracts of *D. purpurea*, *L. heterophyllus* and *B. chinense*. The translated sequences of five of these amplicons showed high sequence similarity, 57–98%, to bacterial NRPS synthetases (Table 3). Importantly, amplicons BC1 and BC2 showed significant similarity (98%) to an adenylation domain of an uncharacterized *Rhodococcus erythropolis* NRPS. Of relevance to drug discovery, the product amplified from *D. purpurea* (DL3) had 65% sequence similarity to the ansamitocin synthase from *Actinosynnema pretiosum*. Unlike the PKS amplicons, which were all similar to published *Streptomyces* sp. sequences, the NRPS PCR amplicons were similar to sequences of diverse bacteria, including the Proteobacteria and different genera of the Actinobacteria group (Table 3).

The deduced translated A domains sequences were analysed to determine if they contained an eight residue binding pocket, and the amino acid that was likely to be bound by the pocket. Predictions were successful for the bacterial NRPS sequences including fragments BC1 and BC2 that are likely to possess the same amino acid binding pocket, and could bind either a Leu, Ile or Val residue (Table 3). Analysis of the two fragments detected from the endophytes of *L. heterophyllus*, LH3 and LH4, were shown to possess different binding pockets and thus would be involved in the addition of different amino acids to the growing peptide chain produced by the NRPS (Table 3).

Phylogenetic distribution of PKS and NRPS genes from TCM endophytes

In an attempt to investigate the homology and novelty of amplicons from the genetic screen, the deduced amino acid sequence of the KS and A domain gene fragments were used in multiple sequence alignments and the phylogenies reconstructed. The unrooted phylogenetic tree of fungal PKSs (Figure 1) revealed the presence of two main clades that correspond to the two largest structural classes of fungal polyketides: reduced and unreduced. All fungal KS domain sequences amplified in this study belonged to the clade of PKSs which synthesize unreduced PKs. The proteins were closely grouped to PKSs involved in the biosynthesis pathways of various pigments including melanin (LH1, PT1), other pigments from *Penicillium* sp. and *Elsinoe fawcetti* (CF1, LJ1, SS2), or putative PKSs derived from endophytes and other fungal symbionts (BS1, LH2, SS1, SS3, PT1). The gene products of SS1 and SS3 formed a highly supported clade and are possible homologs of the same biosynthetic pathway.

Phylogenetic analysis was performed on the inferred amino acid sequences of NRPS A domains (Figure 2). The three NRPS A domains sequences identified in this study were amplified from the extract of *Pi. pedatisecta* and although each sequence was unique, the sequences (PpE1, PpE2 and PpE3) were clustered in the

phylogenetic tree indicating that they most probably have a common phylogenetic ancestor and/or similar function. This cluster of fragments also grouped with cyanobacterial NRPSs. Related NRPSs included those involved in the synthesis of a range of bioactive compounds; MicC in the synthesis of the pentapeptide microginin in *Planktothrix rubescens* [34], AerB in the synthesis of the aeruginosin protease inhibitors in *Microcystis viridis* [35], and OciA in the synthesis of cyanopeptolin [36]. Despite the fact that fungal specific PCR primers were used, these A domain sequences had greater similarity to NRPSs of bacterial origin with bacterial and fungal NRPSs clustered separately in the phylogenetic tree (Figure 2).

Phylogenetic analysis of inferred bacterial KS domains revealed that the endophyte sequences clustered with KS domains of Actinobacteria (Figure 3). Sequences PpO1 and PpO2 were closely related to each other and also clustered with PKS fragments involved in the biosynthesis of the anticancer antibiotic geldanamycin, from *Streptomyces hygroscopicus* [37]. The DL1 sequence clustered within a clade containing KS domains, involved in the tailoring of geldanamycin and herbimycin in *S. hygroscopicus*, and an uncharacterized KS domain from *Pseudonocardia autotrophica*. The sequence PpE4 was not included in the bacterial PKS multiple sequence alignment and phylogenetic tree, because it was significantly shorter than the other fragments (~80 amino acids).

The phylogenetic tree containing bacterial NRPS A domain sequences demonstrated two main clusters from the Actinobacteria and Proteobacteria (Figure 4). Sequences BC1 and BC2 were closely related and clustered with an uncharacterized NRPS biosynthesis pathway from *Rhodococcus erythropolis*. Although the function of the *R. erythropolis* NRPS is unknown, it also grouped with A domain sequences from other Actinobacteria. Sequences LH3 and LH4 were related to NRPS sequences from the Proteobacteria, specifically *Pseudomonas fluorescens* and *Acidovorax avenae*, respectively. The sequences DL3 grouped with the clade of bacterial A domain sequences, however, had low homology to known NRPSs.

Discussion

Natural product drug discovery investigations are supported by a vast number of candidate systems and environments available for study, presenting the necessity to optimise the targets chosen for investigation. Ethnopharmacological approaches have been affirmed by the discovery of novel bioactive compounds from traditional medicinal plants [3,6,38]. Additionally endophytes are capable of producing a wide variety of natural products, contributing to the overall bioactivity of the host plant. Thus, they are considered to be a valuable source of novel bioactive compounds and an avenue to increase the breadth of natural product studies [39]. The present study took advantage of the immense knowledge of TCM plants to focus on plants of pharmaceutical interest. This preliminary study screened the metabolic potential of endophytes using PKS and NRPS genes as a proxy for endophytic production of bioactive compounds. The use of these genes reflects the noted involvement of PKSs and NRPSs in the production of many natural products [21,40,41]. This screen has the potential to be applied to the endophytes of other TCM plant species or growth climates.

PKS and NRPS genes are appropriate targets for detection of small molecule biosynthesis systems [21,40,41]. Genetic screening results demonstrated the presence of at least one NRPS/PKS biosynthetic pathway in 36% (11/30) of the plant DNA extracts. Furthermore, PKS genes were detected more frequently in the total DNA samples (30%, 9/30) than NRPS genes (13%, 4/30).

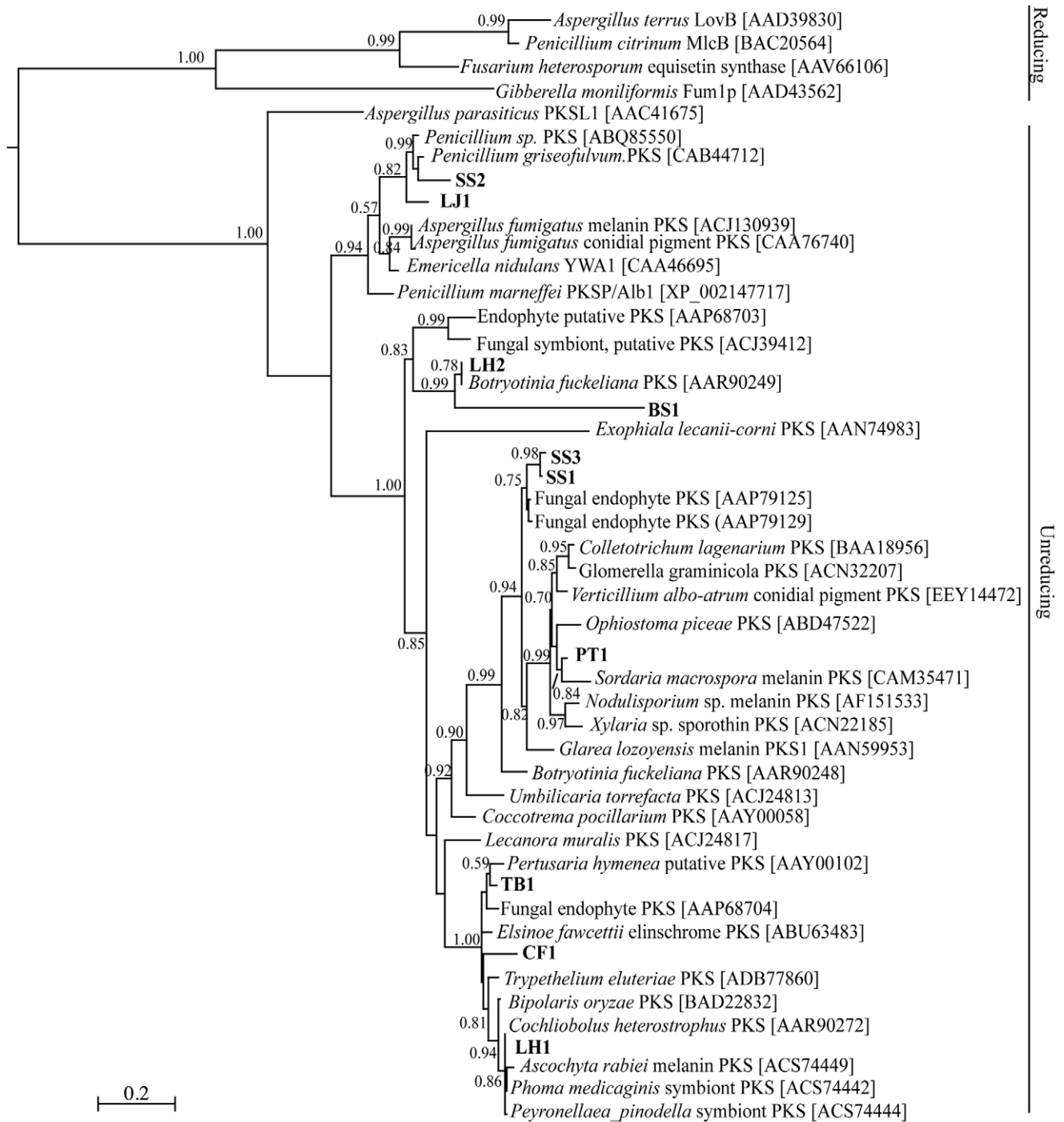


Figure 1. Phylogenetic analysis of putative fungal type I PKSs in TCM herbs. Evolutionary relationships were determined by maximum likelihood analysis using the LG substitution model. Branch lengths indicate inferred divergence of amino acid sequences. Numbers adjacent to the nodes indicate aLRT support, with support values >50% considered to be significant. Accession numbers for sequences obtained from GenBank are indicated. The scale bar represents 0.2 amino acid changes.
doi:10.1371/journal.pone.0035953.g001

PKS and NRPS biosynthesis genes have previously been detected in fungal and bacterial isolates from ecological niches similar to those inhabited by endophytes. Fungal and bacterial isolates of sponges (18% and 13.8%, respectively) and marine (90%) and freshwater (65%) cyanobacteria were positive for PKS and NRPS genetic screens [23].

Due to the fact that plants are host to diverse endophyte populations, it was speculated that many of the total DNA extracts would contain PKS and NRPS genes. However, genetic screening for endophyte biosynthesis genes has not previously been attempted using total plant DNA extracts. In this study, DNA samples often contained only one PKS/NRPS sequence, inferring that there may be few biosynthesis genes present in the endophytes

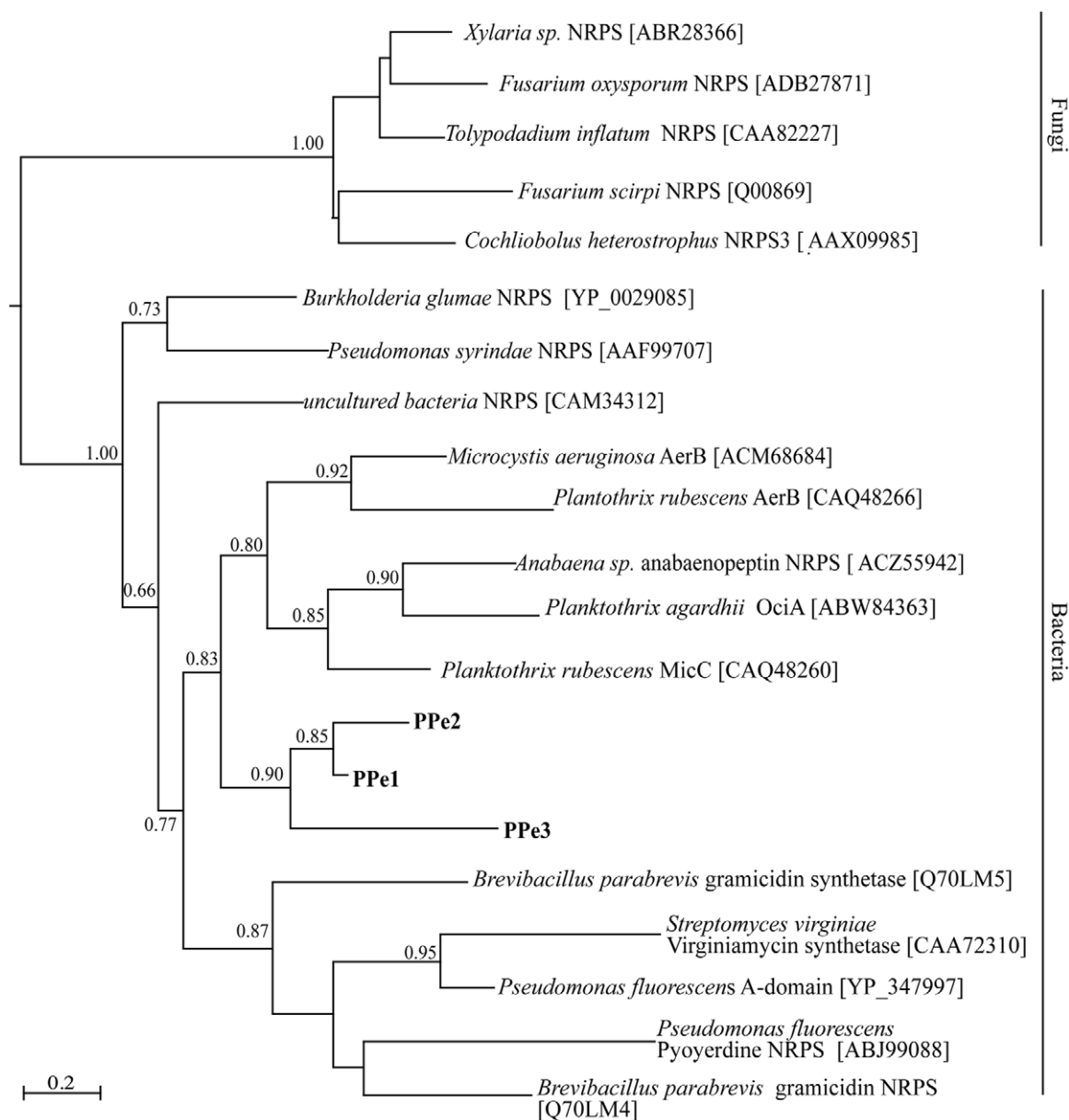


Figure 2. Phylogeny of fungal NRPSs. Relationships inferred by maximum likelihood analysis of fungal NRPS adenylation domain fragments using the WAG substitution model. Branch lengths indicate inferred divergence of amino acid sequences. Numbers adjacent to nodes indicate aLRT support, with support values >50% considered significant. Bacterial and fungal reference sequences are included in the analysis and the accession numbers for these sequences obtained from GenBank are indicated. The scale bar represents 0.2 amino acid changes. doi:10.1371/journal.pone.0035953.g002

of the sampled plants. It is probable that only a few endophytes biosynthesize compounds that increase plant viability, and these endophytes may dominate the population, and any PCR results. Additionally, many samples failed to reveal any endophyte PKS or NRPS genes. This indicates that either endophytes within the samples did not contain PKS or NRPS biosynthesis genes or that the biosynthesis gene sequences were too divergent to be amplified by the reaction described. Alternatively, endophytes involved in the biosynthesis of bioactive compounds in these plants may have been absent from the tissue type used for the DNA extractions [42]. Instead endophytes could inhabit an alternate host tissue and export bioactive compounds for storage in the tissue used medicinally [43].

Actinobacteria, in particular *Streptomyces* sp., are significant secondary metabolite producers, synthesising a functionally diverse class of bioactive compounds, often through PKS and NRPS pathways [44,45]. Although Actinobacteria are only a subset of the total bacterial population, sequences related to non-actinobacteria NRPS genes were also detected using the actinobacterial primers (Figure 4). This is possible, due to the conserved gene regions within the KS and A domains [46]. However, the type I PKS and NRPS primers used in this study cannot amplify all PKS or NRPS genes, suggesting that these primers underestimate endophytes' KS and NRPS gene diversity within the plants. This limitation is in part due to the absence of divergent sequences in the sequence databases and primers which are not designed to detect atypical PKS or NRPS genes. Consequently, some interesting pathways

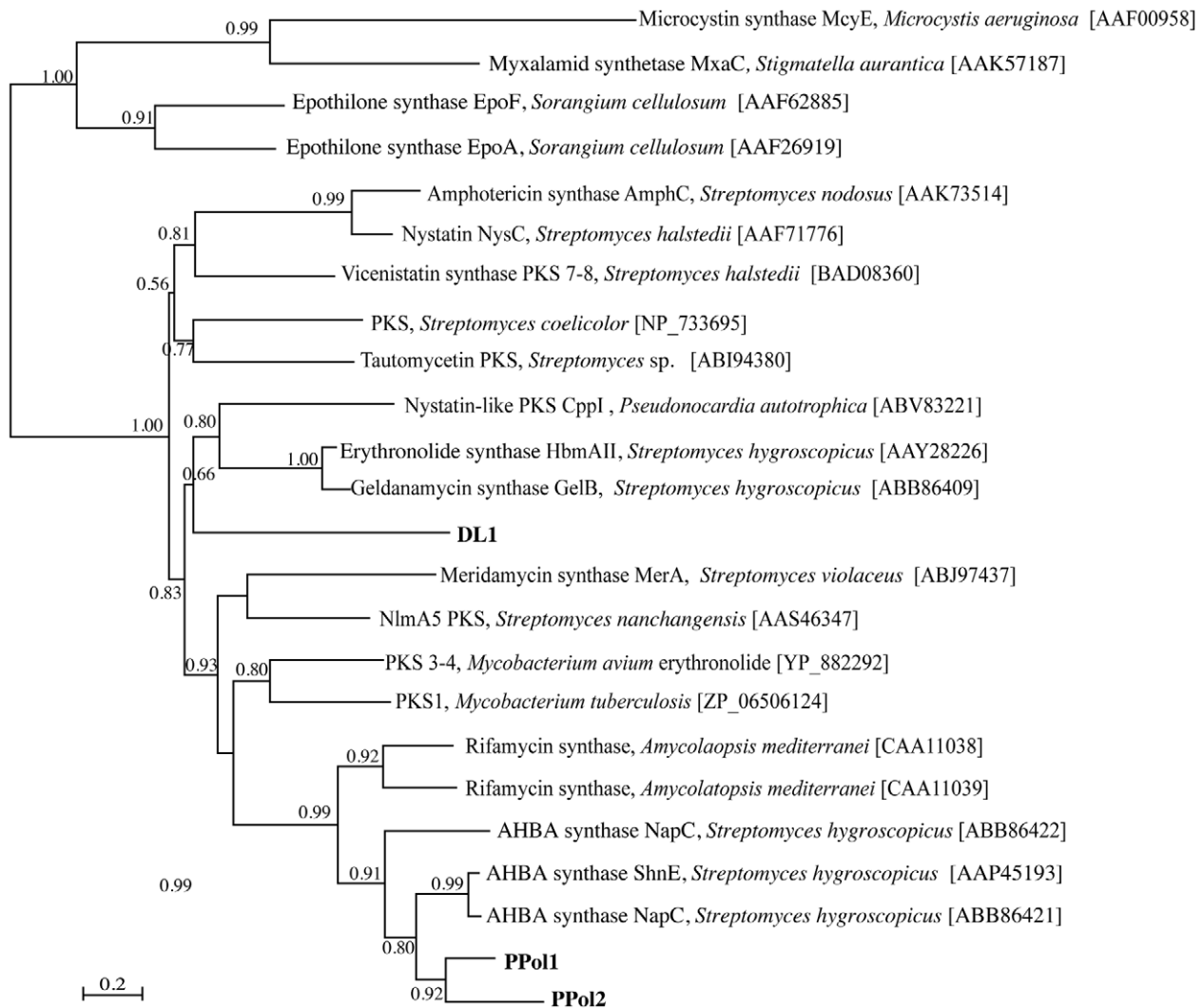


Figure 3. Phylogeny of bacterial type I PKSs. Relationships inferred by maximum likelihood analysis of bacterial PKS ketosynthase domain fragments, using the LG substitution model. Branch length indicates inferred divergence of amino acid sequences and numbers adjacent to the nodes indicate aLRT support, with support values >50% considered significant. The accession numbers for all sequences obtained from GenBank are indicated. The scale bar represents 0.2 amino acid changes.
doi:10.1371/journal.pone.0035953.g003

may be overlooked with this approach. Sequence-based approaches will improve from advances in our understanding of the molecular genetics of PKS and NRPS biosynthesis.

Phylogenetic analysis of putative fungal KS domains showed that sequences clustered into two main groups (Figure 1). PKSs which synthesize partially reduced polyketides, such as lovastatin and T-toxin, were clustered separately from those PKSs that synthesize unreduced polyketides, examples of which include aflatoxin and melanin. This division is supported by previous phylogenetic investigations of fungal PKSs [47,48]. All ten sequences from the TCM samples were within the cluster of PKSs that synthesize unreduced polyketides, specifically, pigments. Fungal pigments show a range of bioactivities such as elsinchrome, which possesses phytotoxic activities [49], and the naphopyrones, which possess antimicrobial activities [50,51]. It can be concluded that the detection of pigment-type biosynthesis genes indicates the biosynthetic potential of TCM endophytes. The putative fungal KS domain fragments identified; BS1, CF1, LH1, LH2, SS1, SS3 and TB1, showed close relationships with KS domains of

symbiotic fungi, including endophytes and mycobionts of lichens (Figure 1). These amplified KS domains may encode biosynthetic pathways unique to symbiotic fungi and thus be involved in the host-symbiont relationship.

Phylogenetic analysis of the fungal A domains demonstrated that the sequences were clustered with NRPSs from bacteria (Figure 2). Some sequences showed similarity to bacterial PKS sequences involved in the production of cyclic lipopeptides, such as arthrofactin synthetases (Table 3). Cyclic lipopeptide production by *Pseudomonas fluorescens* was shown to play a key role in biofilm formation, motility and the antifungal activity of strain SBW25. Thus the detected fungal A domains could be involved in the biosynthesis of novel fungal peptides. The amplified A domains were related to NRPS pathways involved in the synthesis of compounds with a range of bioactivities; microginin, an angiotensin-converting enzyme inhibitor [52], anabaenopeptin, a serine protease inhibitor [53], and aeruginosin, a thrombin inhibitor [54]. Protease inhibitor-type compounds are significant because of the biochemical processes they control. For example serine

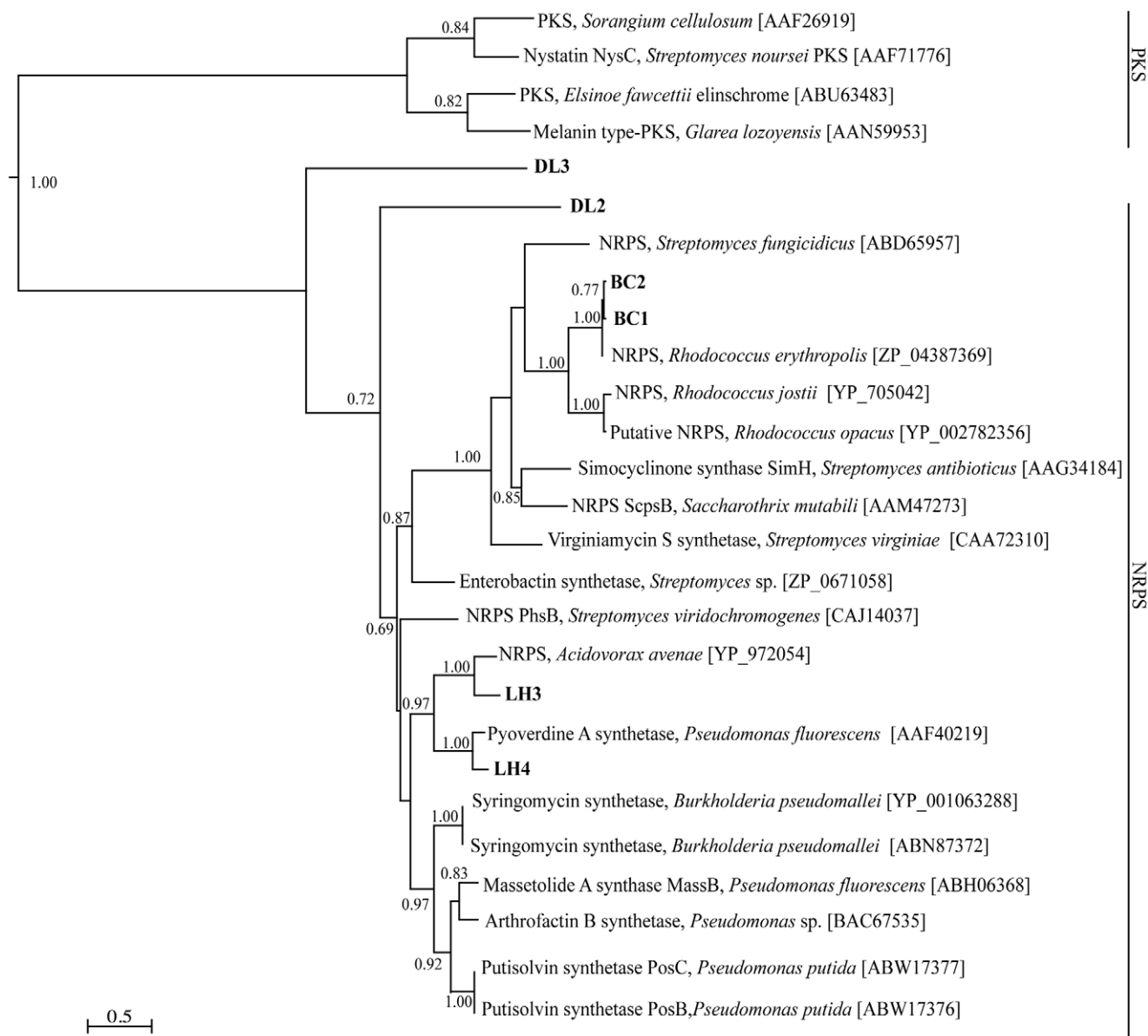


Figure 4. Phylogeny of bacterial NRPS sequences. Relationships inferred by maximum likelihood analysis of bacterial NRPS adenylation domain fragments using the LG substitution model. Branch lengths indicate inferred divergence of amino acid sequences. Numbers adjacent to the nodes indicate aLRT support, with support values >50% considered significant. The accession numbers for all sequences obtained from GenBank are indicated. The scale bar represents 0.5 amino acid changes. doi:10.1371/journal.pone.0035953.g004

protease inhibitors possess cell cytotoxic activities and are used for the treatment of a wide variety of human diseases [55].

Sequences identified from the bacterial NRPS screen were phylogenetically grouped with diverse bacterial NRPS biosynthesis pathways. The sequences LH3 and LH4 were grouped with proteobacterial NRPSs involved in the synthesis of the siderophores pyoverdine and enterobactin [56,57]; antibiotics, massetolide A and virginiamycin [58,59]; and the biosurfactant, arthrofactin [60]. It appears that these putative A domain genes could be involved in the synthesis of novel and biologically active peptides, which in turn add to the medicinal activity of the host plants. Gene fragments BC1 and BC2 were related to an A domain of an uncharacterized NRPS pathway of *Rhodococcus erythropolis*. While this specific NRPS gene is uncharacterized,

Rhodococcus spp. are known to contain a large set of NRPS and PKS genes with their biotechnological significance arising from their existence in harsh environmental niches [61,62]. It is proposed that the products of this NRPS may assist endophytes to survive inside the plant. The putative A domain sequence DL3 failed to cluster with known NRPSs and therefore may represent divergent NRPSs, for the production of unique peptide-derived compounds.

Of interest in the phylogenetic analysis of bacterial PKS fragments, is that fragments PPol1 and PPol2 grouped with various enzymes of *Streptomyces hygroscopicus* involved in geldanamycin biosynthesis. This suggests that the endophyte genes could encode pathways for the production of similar anticancer antibiotics [63]. The detection of endophyte associated biosynthe-

sis genes in *P. polyphylla* highlights the fact that this plant is a good candidate for further endophyte investigations involved in isolation endophyte and characterisation of their associated bioactive compounds. Endophytes offer sustainable production of plant associated natural products, and have undeniable application in the development of compounds in drug discovery [64].

In summary, TCM herbs are host to diverse bacteria and fungi with the potential to synthesize secondary metabolites which would contribute to the plant's chemical composition. In this particular study, genetic screening revealed endophyte-derive PKS and NRPS fragments with putative roles in the biosynthesis of secondary metabolites with a wide array of biological activities. This screening allows further investigations to focus on plants that host endophytes with the greatest potential to produce polyketide and nonribosomal peptide-based bioactive substances. To further investigate links between PKS/NRPS genes and bioactivity, elucidation of the full biosynthetic pathways is required from endophytes. Future studies of the plant candidates may lead to harnessing endophytes, their genetic resources, and associated compounds which are valuable leads for drug discovery investigations, as well as understanding how endophytes contribute to the bioactivity of medicinal plants. In addition, it remains to be determined whether the PKS and NRPS genes identified in this

study are functionally dedicated to specialized endophyte-based activities.

Supporting Information

Figure S1 Multiple sequence alignments. The multiple sequence alignments of KS domains protein sequences, from PKS pathways, and A domain protein sequences, from NRPS pathways, were generated using ClustalX. (DOC)

Acknowledgments

The authors would like to thank Dr. Troco K. Mihali (University of New South Wales), Yanli Zhang and Yajuan Chen (Kunming Medical University) for advice and helpful discussions. We would also like to acknowledge Bin Qiu (Yunnan Institute of Materia Medica) for assistance with collection of some of the plant samples.

Author Contributions

Conceived and designed the experiments: KIM CQ DMYS BAN. Performed the experiments: KIM CQ. Analyzed the data: KIM BAN. Contributed reagents/materials/analysis tools: CQ BAN. Wrote the paper: KIM BAN. Assisted in organising trip to China: KIM DMYS. Assisted in organising plant collection: CQ.

References

- Newman DJ, Cragg GM (2007) Natural products as sources of new drugs over the last 25 years. *J Nat Prod* 70: 461–477.
- Zhu Y-P (1998) Chinese materia medica: chemistry, pharmacology, and applications. USA: CRC Press.
- Hoessel R, Leclerc S, Endicott JA, Nobel MEM, Lawrie A, et al. (1999) Indirubin, the active constituent of a Chinese antileukaemia medicine, inhibits cyclin-dependent kinases. *Nat Cell Biol* 1: 60–67.
- Wall ME, Wani MC, Cook CE, Palmer KH, McPhail HT, et al. (1966) Plant antitumor agents. I. The isolation and structure of camptothecin, a novel alkaloidal leukemia and tumor inhibitor from *Camptotheca acuminata*. *J Am Chem Soc* 88: 3888–3890.
- Perdue RE, Spetzman LA, Powell RG (1970) Cephalotaxus- Source of harringtonine, a promising new anti-cancer alkaloid. *Amer Horticult Mag* 49: 19–22.
- Acton N, Klayman DL (1985) Artemisitene, a new sesquiterpene lactone endoperoxide from *Artemisia annua*. *Planta Med* 51: 441–442.
- Chen KK, Schmidt CF (1924) The action of ephedrine, the active principle of the Chinese drug ma huang. *J Pharmacol Exp Ther* 24: 339–357.
- Puri SC, Verma V, Amna T, Qazi GN, Spittler M (2005) An endophytic fungus from *Nothapodytes foetida* that produces camptothecin. *J Nat Prod* 68: 1717–1719.
- Stierle A, Strobel G, Stierle D (1993) Taxol and taxane production by *Taxomyces andreanae*, an endophytic fungus of Pacific yew. *Science* 260: 214–216.
- Compant S, Duffy B, Nowak J, Clement C, Barka EA (2005) Use of plant growth-promoting bacteria for biocontrol of plant diseases: principles, mechanisms of action, and future prospects. *Appl Environ Microbiol* 71: 4951–4959.
- Arnold AE, Mejia LC, Kyllo D, Rojas EI, Maynard Z, et al. (2003) Fungal endophytes limit pathogen damage in a tropical tree. *PNAS* 100: 15649–15654.
- Mejia LC, Rojas EI, Maynard Z, Van Bael S, Arnold AE, et al. (2008) Endophytic fungi as biocontrol agents of *Theobroma cacao* pathogens. *Biol Control* 46: 4–14.
- Kannadan S, Rudgers JA (2008) Endophyte symbiosis benefits a rare grass under low water availability. *Funct Ecol* 22: 706–713.
- Strobel GA (2003) Endophytes as sources of bioactive products. *Microbes Infect* 5: 535–544.
- Keller NP, Turner G, Bennett JW (2005) Fungal secondary metabolism – from biochemistry to genomics. *Nat Rev Microbiol* 3: 937–947.
- Cox RJ (2007) Polyketides, proteins and genes in fungi: programmed nanomachines begin to reveal their secrets. *Org Biomol Chem* 5: 2010–2026.
- Mankel DP, Neilan B (2000) Non-ribosomal peptide antibiotics. *Expert Opin Ther Pat* 10: 1583–1591.
- Fischbach MA, Walsh CT (2006) Assembly-Line Enzymology for Polyketide and Nonribosomal Peptide Antibiotics: Logic, Machinery, and Mechanisms. *Chem Rev* 106: 3468–3496.
- Staunton J, Wilkinson B (2001) Combinatorial biosynthesis of polyketides and nonribosomal peptides. *Curr Opin Chem Biol* 5: 159–164.
- Mootz HD, Schwarzer D, Marahiel MA (2002) Ways of assembling complex natural products on modular nonribosomal peptide synthetases. *ChemBiochem* 3: 490–504.
- Sauer M, Lu P, Sangari R, Kennedy S, Polishook J, et al. (2002) Estimating polyketide metabolic potential among non-sporulating fungal endophytes of *Vaccinium macrocarpon*. *Mycol Res* 106: 460–470.
- Burns BP, Seifert A, Goh F, Pomati F, Jungblut AD, et al. (2005) Genetic potential for secondary metabolite production in stromatolite communities. *FEMS Microbiol Lett* 243: 293–301.
- Ehrenreich IM, Waterbury JB, Webb EA (2005) Distribution and diversity of natural product genes in marine and freshwater cyanobacterial cultures and genomes. *Appl Environ Microbiol* 71: 7401–7413.
- Gore ME, Bucak C (2007) Geographical and seasonal influences on the distribution of fungal endophytes in *Laurus nobilis*. *Forest Pathol* 37: 281–288.
- Hoffman MT, Arnold AE (2008) Geographic locality and host identity shape fungal endophyte communities in cupressaceous trees. *Mycol Res* 112: 331–344.
- Xiwen L, Walker D (1986) The plant geography of Yunnan Province, southwest China. *Journal of Biogeography* 13: 367–397.
- Cankar K, Kraigher H, Ravnikar M, Rupnik M (2005) Bacterial endophytes from seeds of Norway spruce (*Picea abies* L. Karst). *FEMS Microbiol Lett* 244: 341–345.
- Rausch C, Weber T, Kohlbacher O, Wohlleben W, Huson DH (2005) Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucl Acids Res* 33: 5799–5808.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl Acids Res* 25: 4876–4882.
- Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* 21: 2104–2105.
- Guindon S, Gascuel O (2003) A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52: 690–704.
- Keller NP, Hohn TM (1997) Metabolic pathway gene clusters in filamentous fungi. *Fungal Genet Biol* 21: 17–29.
- Ayuso-Sacido A, Genilloud O (2005) New PCR primers for the screening of NRPS and PKS-I systems in actinomycetes: detection and distribution of these biosynthetic gene sequences in major taxonomic groups. *Microb Ecol* 49: 10–24.
- Rounge TB, Rohrlack T, Nederbragt AJ, Kristensen T, Jakobsen KS (2009) A genome-wide analysis of nonribosomal peptide synthetase gene clusters and their peptides in a *Planktothrix rubescens* strain. *BMC Genomics* 10: 396.
- Ishida K, Welker M, Christiansen G, Cadel-Six S, Bouchier C, et al. (2009) Plasticity and evolution of aeruginosin biosynthesis in cyanobacteria. *Appl Environ Microbiol* 75: 2017–2026.
- Rounge TB, Rohrlack T, Kristensen T, Jakobsen KS (2008) Recombination and selection forces in cyanopeptolin NRPS operons from highly similar, but geographically remote *Planktothrix* strains. *BMC Microbiol* 8: 141.
- Rascher A, Hu Z, Viswanathan N, Schirmer A, Reid R, et al. (2003) Cloning and characterization of a gene cluster for geldanamycin production in *Streptomyces hygroscopicus* NRRL 3602. *FEMS Microbiol Lett* 218: 223–230.

38. Castillo UF, Strobel GA, Ford EJ, Hess WM, Porter H, et al. (2002) Munumbicins, wide-spectrum antibiotics produced by *Streptomyces* NRRL 30562, endophytic on *Kennedia nigricans*. *Microbiol* 148: 2675–2685.
39. Strobel G, Daisy B, Castillo U, Harper J (2004) Natural products from endophytic microorganisms. *J Nat Prod* 67: 257–268.
40. Wawrik B, Kerkhof L, Zylstra GJ, Kukor JJ (2005) Identification of unique type II polyketide synthase genes in soil. *Appl Environ Microbiol* 71: 2232–2238.
41. Moffitt MC, Neilan BA (2001) On the presence of peptide synthetase and polyketide synthase genes in the cyanobacterial genus *Nodularia*. *FEMS Microbiol Lett* 196: 207–214.
42. Bayman P, Lebron LL, Tremblay RL, Lodge DJ (1997) Variation in endophytic fungi from roots and leaves of *Lepanthes* (Orchidaceae). *New Phytol* 135: 143–149.
43. Marler TE, Lee V, Shaw CA (2005) Spatial variation of steryl glucosides in *Cycas micronesica* plants: Within and among plant sampling procedures. *Hortscience* 40: 1607–1611.
44. Moore BS, Kalaitzis JA, Xiang LK (2005) Exploiting marine actinomycete biosynthetic pathways for drug discovery. *Antonie Van Leeuwenhoek* 87: 49–57.
45. Weber T, Welzel K, Pelzer S, Vente A, Wohlleben W (2003) Exploiting the genetic potential of polyketide producing streptomycetes. *J Biotechnol* 106: 221–232.
46. Jenke-Kodama H, Sandmann A, Muller R, Dittmann E (2005) Evolutionary implications of bacterial polyketide synthases. *Mol Biol Evol* 22: 2027–2039.
47. Kroken S, Glass NL, Taylor JW, Yoder OC, Turgeon GB (2003) Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proc Natl Acad Sci USA* 100: 15670–15675.
48. Nicholson TP, Rudd BAM, Dawson M, Lazarus CM, Simpson TJ, et al. (2001) Design and utility of oligonucleotide gene probes for fungal polyketide synthases. *Chem Bio* 8: 157–178.
49. Liao HL, Chung KR (2008) Genetic dissection defines the roles of elsinochrome Phytotoxin for fungal pathogenesis and conidiation of the citrus pathogen *Elsinoe fawcettii*. *Mol Plant Microbe Interact* 21: 469–479.
50. Brimble MA, Nairn MR, Duncalf IJ (1999) Pyranonaphthoquinone antibiotics—isolation, structure and biological activity. *Nat Prod Rep* 16: 267–281.
51. Bianchi C, Ceriotti G (1975) Chemical and pharmacological investigations of constituents of *Eleutherine bulbosa* (Miller) Urb. (Iridaceae). *J Pharm Sci* 64: 1305–1308.
52. Okino T, Matsuda H, Murakami M, Yamaguchi K (1993) Microginin, an angiotensin-converting enzyme inhibitor from the blue-green alga *Microcystis aeruginosa*. *Tetrahedron Lett* 34: 501–504.
53. Murakami M, Shin HJ, Matsuda H, Ishida K, Yamaguchi K (1997) A cyclic peptide, anabaenopeptin B, from the cyanobacterium *Oscillatoria agardhii*. *Phytochemistry* 44: 449–452.
54. Kodani S, Ishida K, Murakami M (1998) Aeruginosin 103-A, a thrombin inhibitor from the cyanobacterium *Microcystis viridis*. *J Nat Prod* 61: 1046–1048.
55. Leung D, Abbenante G, Fairlie DP (2000) Protease inhibitors: Current status and future prospects. *J Med Chem* 43: 305–341.
56. Visca P, Imperi F, Lamont IL (2007) Pyoverdine siderophores: from biogenesis to biosignificance. *Trends in Microbiology* 15: 22–30.
57. Liu J, Duncan K, Walsh CT (1989) Nucleotide sequence of a cluster of *Escherichia coli* enterobactin biosynthesis genes—Identification of entA and purification of its product 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase. *J Bacteriol* 171: 791–798.
58. Namwat W, Kamioka Y, Kinoshita H, Yamada Y, Nihira T (2002) Characterization of virginiamycin S biosynthetic genes from *Streptomyces virginiae*. *Gene* 286: 283–290.
59. de Bruijn I, de Kock MJD, de Waard P, van Beek TA, Raaijmakers JM (2008) Massetolide a biosynthesis in *Pseudomonas fluorescens*. *J Bacteriol* 190: 2777–2789.
60. Morikawa M, Daido H, Takao T, Murata S, Shimonishi Y, et al. (1993) A new lipopeptide biosurfactant produced by *Arthrobacter* sp. strain MIS38. *J Bacteriol* 175: 6459–6466.
61. McLeod MP, Warren RL, Hsiao WWL, Araki N, Myhre M, et al. (2006) The complete genome of *Rhodococcus* sp RHA1 provides insights into a catabolic powerhouse. *P Natl Acad Sci U S A* 103: 15582–15587.
62. de Carvalho CCCR, da Fonseca MMR (2005) The remarkable *Rhodococcus erythropolis*. *Appl Microbiol Biotech* 67: 715–726.
63. Stebbins CE, Russo AA, Schneider C, Rosen N, Hartl FU, et al. (1997) Crystal structure of an Hsp90-geldanamycin complex: targeting of a protein chaperone by an antitumor agent. *Cell* 89: 239–250.
64. Kusari S, Spiteller M (2011) Are we ready for industrial production of bioactive plant secondary metabolites utilizing endophytes? *Nat Prod Rep* 28: 1203–1207.
65. Li X, Wei W (2002) Chinese materia medica: combinations and applications. UK: Donica.
66. Lee H, Lin J-Y (1988) Antimutagenic activity of extracts from anticancer drugs in Chinese medicine. *Mutat Res-Genet Tox* 204: 229–234.
67. Rajbhandari M, Mentel R, Jha PK, Chaudhary RP, Bhattacharai S, et al. (2009) Antiviral activity of some plants used in Nepalese traditional medicine. Evidence-based Complementary and Alternative Medicine 6: 517–522.
68. Huang KC, Williams WM (1999) The pharmacology of Chinese herbs. Boca Ranton: CRC Press.
69. Miyazawa M, Shimamura H, Nakamura S-i, Sugiura W, Kosaka H, et al. (1999) Moscatilin from *Dendrobium nobile*, a naturally occurring bibenzyl compound with potential antimutagenic activity. *J Agr Food Chem* 47: 2163–2167.
70. Lopez-Lazaro M, de la Pena NP, Pastor N, Martin-Cordero C, Navarro E, et al. (2003) Anti-tumour activity of *Digitalis purpurea* L. subsp. *heycwoodii*. *Planta Med* 69: 701–704.
71. Davydov M, Krikorian AD (2000) (Rupr. & Maxim.) Maxim. (Araliaceae) as an adaptogen: a closer look. *J Ethnopharmacol* 72: 345–393.
72. Ho L-K, Don M-J, Chen H-C, Yeh S-F, Chen J-M (1996) Inhibition of Hepatitis B surface antigen secretion on human hepatoma cells. Components from *Rubia cordifolia*. *J Nat Prod* 59: 330–333.
73. Lu Y, Yeap Foo L (2002) Polyphenolics of Salvia—a review. *Phytochemistry* 59: 117–140.
74. NiceData (2004) Traditional Chinese medicines: natural sources and applications. 2.1 p.
75. Wani MC, Taylor HL, Wall ME, Coggon P, McPhail AT (1971) Plant antitumor agents. VI. The isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. *J Am Chem Soc* 93: 2325–2327.
76. van der Bank H, Wink M, Vorster P, Treutlein J, Brand L, et al. (2001) Allozyme and DNA sequence comparisons of nine species of *Encephalartos* (Zamiaceae). *Biochem Syst Ecol* 29: 241–266.
77. Neilan BA, Dittmann E, Rouhiainen L, Bass RA, Schaub V, et al. (1999) Nonribosomal peptide synthesis and toxigenicity of cyanobacteria. *J Bacteriol* 181: 4089–4097.
78. Johnson R, Voisey C, Johnson L, Pratt J, Fleetwood D, et al. (2007) Distribution of NRPS gene families within the *Neotyphodium*/*Epichloe* complex. *Fungal Genet Biol* 44: 1180–1190.
79. Bingle LE, Simpson TJ, Lazarus CM (1999) Ketosynthase domain probes identify two subclasses of fungal polyketide synthase genes. *Fungal Genet Biol* 26: 209–223.