Reconstructing fungal natural product biosynthetic pathways

Journal:	*Natural Product Reports*
Manuscript ID:	NP-HIG-06-2014-000084.R1
Article Type:	Highlight
Date Submitted by the Author:	05-Aug-2014

Complete List of Authors:	Lazarus, Colin; University of Bristol, School of Biological Sciences Bailey, Andrew; University of Bristol, School of Biological Sciences Williams, Katherine; University of Bristol, School of Chemistry

Journal Name

RSCPublishing

ARTICLE

Reconstructing Fungal Natural Product Biosynthetic Pathways

Cite this: DOI: 10.1039/x0xx00000x

C.M. Lazarus,[a] K. Williams[b] and A. M. Bailey[a]

Received ooth January 2012,
Accepted ooth January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

Large scale fungal genome sequencing has revealed a multitude of potential natural product biosynthetic pathways that remain uncharted. Here we describe some of the methods that have been used to explore them *via* heterologous gene expression. We focus on filamentous fungal hosts and discuss the technological challenges and successes behind the reconstruction of fungal natural product pathways. Optimised, efficient heterologous expression of reconstructed biosynthetic pathways promises progress in the discovery of novel compounds that could be utilised by the pharmaceutical and agrochemical industries.

1. Introduction

The past couple of decades have witnessed the steady encroachment of molecular biology into the field of natural product research, traditionally the near exclusive preserve of chemistry. The isolation and analysis of a bewildering array of plant, fungal and bacterial secondary metabolites, many of which display significant biological activity, provided the basis of the modern pharmaceutical and agrochemical industries.[1] Conventional strain improvement by mutagenesis and screening of the producing organism for increased titres of product has sometimes been effective as a means of generating commercial titres of some fungal natural products, such as the penicillins.[2] However this approach is limited to those compounds already produced in reasonable titres and only in species which are amenable to large-scale growth in submerged culture. Many fungal products do not meet these requirements and in these cases sophisticated synthetic chemistry procedures offer one approach to the bulk production of high value natural products only available from nature in tiny amounts, sometimes from unidentified producer organisms. Unfortunately the complex structures of some natural products makes a synthetic approach rather complicated, requiring multiple steps and so costly to perform.

Genome sequences of many fungi have now been obtained and these have highlighted a higher number of potential secondary metabolism pathways than expected. Even in fungi which have had a long history of investigation, many of the gene clusters had unknown function. This shows that there is a large potentially untapped reserve of natural products awaiting discovery and exploitation; however the genome sequences alone are not sufficient to elucidate the likely structure of the products concerned. To exploit this diversity an approach, currently applied mainly in the research lab rather than the pharmaceutical factory, is to isolate and heterologously express the genes that specify the enzymes of the biosynthetic pathway for the compound of interest. From initial identification of such genes, for example by gene-knockout techniques, to the reconstruction of the biosynthetic pathway in a foreign host, molecular biology and analytical chemistry now go hand in glove as complementary tools of natural product research.

Heterologous gene expression has played a major role in identifying or confirming the importance of individual enzymes in specific metabolite production. In fungi, where the major secondary metabolite classes include polyketides, nonribosomal

peptides and terpenes, this may involve expression of the enzymes at the head of a biosynthetic pathway – a polyketide synthase (PKS), nonribosomal peptide synthetase (NRPS) or a terpene synthase (TS). Analysis of the product(s) may indicate a precursor relationship to the target compound. However in most cases such a product will only be an intermediate, with the final product resulting from the additional activities of a number of downstream enzymes. Uncovering the steps of the biosynthetic pathway may require the heterologous expression of several genes both individually and in combination with others. Production of the final compound requires the simultaneous expression of all the relevant genes, and in heterologous systems this means identifying and reconstructing the biosynthetic pathway in a foreign host.

A convenient feature of fungal secondary metabolism is that the genes encoding the various enzyme activities for a given biosynthetic pathway tend be clustered at a single genetic locus (Figure 1).[3] In contrast to prokaryotes, where this is the norm for all biosynthetic pathways, fungal gene clustering does not imply the existence of operons. All fungal genes are transcribed separately from their own promoters, with regulation of co-expression brought about by the binding of common transcription factors. In some cases, but by no means all, regulatory transcription factors are encoded within the gene clusters that they regulate. Over-expression of these transcriptional regulators has often been shown to induce expression of the cluster and hence synthesis of the corresponding natural product. In a pioneering example of this approach Bergmann *et al*.[4] activated a silent gene cluster in *Aspergillus nidulans* and identified aspyridone, a previously unknown 2-pyridone compound. In a similar approach, modifying the expression pattern of global transcription regulators, such as LaeA, PacC and CreA can also be an effective means to induce expression of some gene clusters.[5 6] Such approaches however are limited to fungi which are already amenable to genetic manipulation, and even when successful, often only result in low titres of product. Whilst this is sufficient for identification of the product and initial screening for biological activity, such titres may not be adequate for commercial exploitation. However it is clear that gene clustering and co-regulation greatly aid the identification of the genetic components of a biosynthetic pathway.

2. Heterologous hosts

The choice of a heterologous host is usually based on convenience, with a hierarchy of standard alternatives starting with *Escherichia coli* and increasing in complexity. While the prokaryotic nature of *E. coli* has recommended it for work on bacterial, particularly actinomycete, natural products (see, for example, Antosch *et al*.[7]), its role in elucidating fungal natural product biosynthetic pathways has tended to be limited to the characterisation of individual genes and usually only where the resulting proteins or metabolites show no antibacterial activity. In basidiomycete fungi, terpenes tend to be the predominant class of secondary metabolite. Agger *et al*.[8] expressed cDNA of various sesquiterpene synthetases from the basidiomycete fungus *Coprinopsis cinerea* in *E. coli* and using GC-MS, identified the resulting products for 5 of the 6 genes tested. This approach was then also used to analyse the 11 sesquiterpene synthetases from the Jack O'Lantern fungus *Omphalotus olearus*.[9] In both of these cases, the resulting metabolite could be analysed directly from the recombinant *E. coli* culture. Using a slightly different approach, Davison *et al*.,[10] elucidated early steps of tropolone biosynthesis by *in vitro* analysis of proteins purified from *E. coli* cultures expressing genes from the stipitatic acid gene cluster of *Talaromyces stipitatus*. Similarly, enzymes AzaH and AzaC from a silent azaphilone gene cluster in *Aspergillus niger* were characterised *in vitro* following overexpression in *E. coli*.[11]

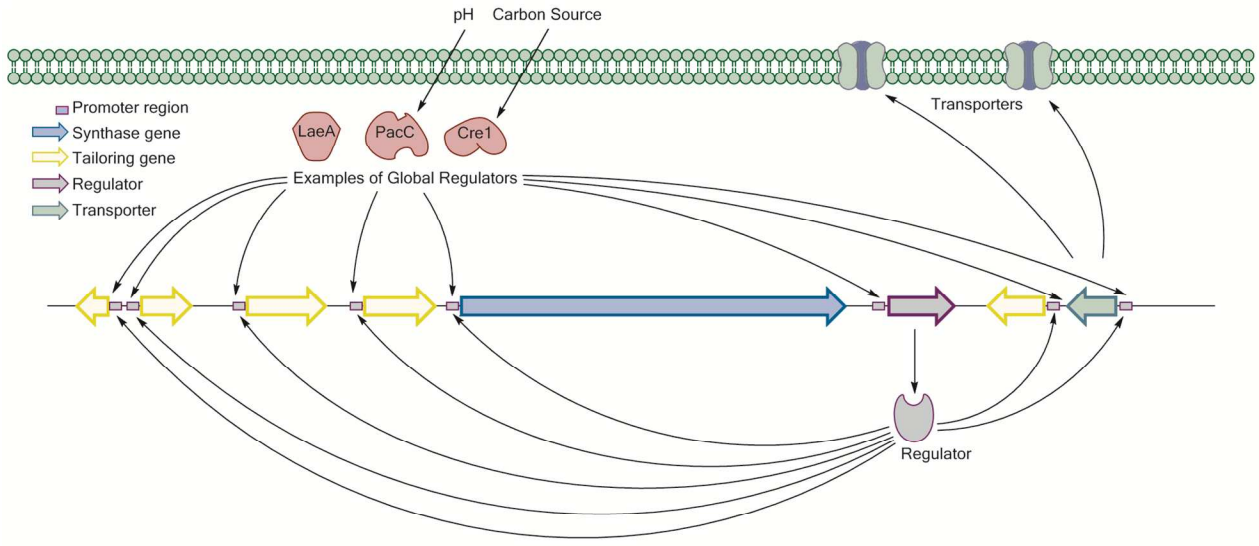

Figure 1: Depiction of a generalised secondary metabolite gene cluster, showing the main synthase gene, tailoring genes, and potential regulatory and transport elements

As a single-celled, eukaryotic microbe, *Saccharomyces cerevisiae* ("yeast") is used widely as an expression host. Yeast clearly has considerable opportunity for exploitation as a host system as demonstrated by the recent production of the plant metabolite artemisinin by heterologous expression of the six genes needed to convert farnesyl diphosphate into the mature sesquiterpene,[12] but again has found limited use in elucidating fungal natural product biosynthesis. Notable exceptions to this are provided by the work of Tang and colleagues, who have been able to exploit the full potential of yeast to study polyketide biosynthesis. For example, authentic activity of the lovastatin nonaketide synthase was obtained when the *lovB* gene from *Aspergillus terreus* was coexpressed in yeast with *lovC*, which encodes its partner enoyl reductase.[13] More recently Wang *et al.*[14] used expression of cryptic non-reducing (NR)-PKS and NRPS-like genes from *A. terreus* in yeast to discover a novel mechanism governing aryl-aldehyde formation. As is usual when using yeast as an expression host for eukaryotic genes this work required removal of fungal introns from the target genes, and the yeast strain used was engineered to produce the *A. nidulans* 4'-phosphopantetheine transferase (PPTase).[15] This is necessary to produce an active enzyme by transfer of a phosphopantetheine group to either the acyl carrier protein (ACP) domain of a PKS or the peptidyl carrier protein (PCP) domain of an NRPS.[16] Ishiuchi *et al*.[17] have also had considerable success with yeast-based expression systems for fungal metabolite pathways. Their similar approach has been successful for characterisation of both PKSs and NRPSs from several different fungi, showing the broad utility of this system for characterisation of individual genes. They further modified the system by expression of a *Rhizobium* gene for malonyl CoA synthetase, increasing titres of PKS-derived metabolites.[18] In the above mentioned examples only megasynthases were characterised, but there are also examples where various tailoring enzymes have been co-expressed to generate a mature secondary metabolite. Rugbjerg *et al*.[19] expressed the *Fusarium graminearum* PKS12 alongside an *A. fumigatus* PPTase and two tailoring genes, *aur*J and *aur*Z to obtain nor-rubofusarin. The utility of a different yeast system, the methylotrophic yeast *Hansenula polymorpha,* was demonstrated by successful expression of the genes for penicillin biosynthesis.[20] As with the *S. cerevisiae* examples, the host first required modifying to express a suitable PPTase before a functional NRPS could be generated. Here the entire genetic pathway was expressed giving the mature metabolite rather than an early intermediate, and the need for a functioning peroxisome system for efficient penicillin synthesis was also demonstrated.

The requirement for a functioning PPTase and sometimes also an efficient peroxisome system[21] help explain why most fungal natural product work now uses (other) filamentous fungi as expression hosts, and this is particularly the case where the objective is to reconstruct whole biosynthetic pathways rather than to express a megasynthase (PKS or NRPS) gene with or without a small number of ancillary enzymes. Filamentous fungal hosts are convenient in that they do not require co-expression of a foreign PPTase gene to provide the prosthetic group for the ACP and PCP domains of PKSs and NRPSs.[22] Their genetic systems are also generally considered to be compatible, obviating the need to express cDNAs or engineer intron removal from genomic DNA.[22, 23] However the reliability of this last point is open to question, as demonstrated in the examples below.

3. Fungal hosts for heterologous natural product synthesis

A further theoretical advantage of using fungal hosts for reconstruction of fungal natural product pathways is that gene clusters can be moved *en masse* for expression of the biosynthetic genes from their own promoters. Smith *et al.*[24] pioneered this approach by cloning the penicillin biosynthetic gene cluster from *Penicillium chrysogenum* on a cosmid vector, transferring it to *Neurospora crassa* and *A. niger* and demonstrating penicillin production. More recently Sakai *et al.*[25] used the same approach to produce citrinin, a mycotoxin of *Monascus purpureus*, in *Aspergillus oryzae* although with rather low titres. These cases demonstrate that while direct transfer of an entire gene cluster may be an easy means of characterisation, it does rely on the various promoters retaining function in the new host and the mRNA transcripts all being correctly processed. The resulting low titres are manageable provided that all that is required is proof of function of the gene cluster and that suitably sensitive detection facilities are available for metabolite analysis.

In the last example, the small amount of citrinin produced was enhanced >400-fold by additional expression of *ctnA*, an activator gene from the citrinin gene cluster, under the control of the *A. nidulans trpC* promoter. Despite the over-expression of the transcriptional regulator, the level of citrinin remained rather low at maximally 1.5 mg per litre. More recently Yin *et al.*[26] produced neosartoricin B at 10 mg per litre in *A. nidulans* transformed with a 5-gene cluster from *Trichophyton tonsurans*; the 4 biosynthetic genes of the cluster were expressed from their own promoters, whereas the transcription factor promoter was replaced with the strong *A. nidulans gpdA* promoter. The monacolin K and terrequinone A gene clusters from *Monascus pilosus* and *Aspergillus nidulans*, respectively, do not include any apparent pathway-specific transcription factor genes, so to express these clusters in *A. oryzae* Sakai *et al.*[27] modified the host to over-express one of the (native) global regulators of fungal secondary metabolism, LaeA. Over-expression of LaeA has been shown to up-regulate several secondary metabolite pathways in *Aspergillus spp*. and other fungi.[28, 29]

The downside of using natural promoters to drive expression of biosynthetic genes, either in the native or a foreign host, is that they tend to be relatively weak and are sometimes only

expressed under very specific conditions, so that production levels remain low even on massive enhancement of expression by multiple copies of a regulator gene. The more laborious alternative, therefore, is to replace the natural promoters of all the biosynthetic genes by strong promoters that will function in the heterologous host.

A potential drawback of using filamentous fungi as hosts for heterologous expression of fungal natural product pathways is that they will themselves have the genetic capability of producing numerous secondary metabolites. The fungi that are most commonly used for heterologous expression are those which are familiar to the biotech industry such as *Aspergillus* or *Trichoderma spp*. Strain development has provided fungi which are stable and easily cultured in a lab environment, usually amenable to genetic modification, grow well in submerged culture so are suitable for large-scale fermentation; importantly, they are usually free of contaminating secondary metabolites. The potential for confusion between native and heterologously produced secondary metabolites can be circumvented by deleting whole gene clusters from the host organism prior to introduction of foreign genes. Chiang *et al.*[30] deleted the asperfuranone gene cluster from the *A. nidulans* genome in their investigation of a cryptic gene cluster from *A. terreus*. This was one element of the development of an efficient system for expressing secondary metabolism genes in *A. nidulans*, while another member of the Aspergillaceae, *A. oryzae*, has been used extensively as a host because strains have been selected which naturally produce no secondary metabolites under normal culture conditions. In common with the domesticated forms of *S. cerevisiae* they are "Generally Regarded As Safe" for food and pharmaceutical production.[31] The first use of this system for heterologous expression of a fungal PKS gene, encoding the naphthopyrone synthase gene (*wA*) from *A. nidulans*,[32, 33] involved transformation of the *A. oryzae* arginine auxotrophic strain M-2-3[34] with the expression vector pTAex3,[35] which carries the complementing *argB* gene from *A. nidulans* and the starch-inducible *amyB* expression cassette from *A. oryzae*. Improvement of *A. oryzae* as a host for expressing whole or partial gene clusters has resulted from combining a total of four nutritional requirements (for adenine, arginine, methionine and nitrate) in the quadruply auxotrophic strain NSAR1, which enables multiple rounds of transformation.[36]

4. Assembling genes, cassettes and pathways

Most fungal natural product biosynthetic pathway reconstruction has been achieved in *Aspergillus* hosts, and that is the focus for the remainder of this article. A first point to consider is the sheer size of some of the genes concerned; this makes it necessary to employ non-conventional techniques, meaning more than simple restriction cutting and ligation, for their manipulation into vectors. Among the first fungal PKS (*Penicillium patulum* MSAS), NRPS (*A. nidulans* ACV synthetase) and hybrid PKS-NRPS (*Fusarium moniliforme* fusarin C synthase; *Fusarium heterosporum* equisitin synthase) genes to be isolated, coding regions were of the order of 5.3, 11.3 11.9 and 12.3 kb, respectively.[37-40] Nowadays it is generally considered most convenient to reconstruct large coding regions from genomic or cDNA by amplifying smaller overlapping fragments by PCR and joining them together *in vitro* or *in vivo*. Joining *in vitro* is usually achieved by overlap extension PCR, but other sequence-specific recombinase enzymes can also be used. Hansen *et al*.[41] used a USER recombinase both to insert the *Penicillium. brevicompactum mpa*C gene into an expression cassette and to recombine partial PKS fragments in order to perform site-directed mutagenesis of the phosphopantetheine attachment site of the ACP. In this example, the PKS was then transformed into a specific locus in *A. nidulans* by targeted transformation, resulting in the production of 5-methylorsellinic acid. In contrast, the *in vivo* approach exploits homologous recombination in an intermediate or final host. The intermediate host is yeast, and the process has been given specific names by some researchers; Shao *et al.*[42, 43] called the method "DNA Assembler", while Tsunematsu *et al.*[18] combined it with overlap extension PCR in a process called "ExRec" (Figure 2). In this last example yeast was also the expression host for fungal PKS genes. Homologous recombination to reassemble a very long coding region may be performed directly in an expression cassette in a shuttle vector, as in the case of the 14.1 kb ferrirhodin synthetase NRPS gene from *Fusarium sacchari*,[44] or in an intermediate vector for subsequent transfer to an expression cassette in the final vector. The "Yeast Assembly" method discussed by Pahirulzaman *et al.*[45] (Figure 2), is an example of the latter approach. While this requires additional steps of pathway reconstruction it allows for extra flexibility in that, for example, a megasynthase can be easily expressed with different combinations of auxiliary enzymes.[46] Module- or domain-swapping experiments to produce novel activities are also easily performed using homologous recombination in yeast.[47, 48] However, homologous recombination in the final expression host can also be used to combine chimaeric genes or gene fragments produced by fusion (overlap-extension) PCR at a chosen expression site in the host genome.[30] In a series of transformations, the same site within the *Aspergillus nidulans* genome was targeted and selectable markers recycled to build the 6-gene asperfuranone cluster from *A. terreus*; this involved the prior deletion of the host asperfuranone gene cluster in a KU70 mutant of *A. nidulans* that has a reduced frequency of non-homologous recombination.

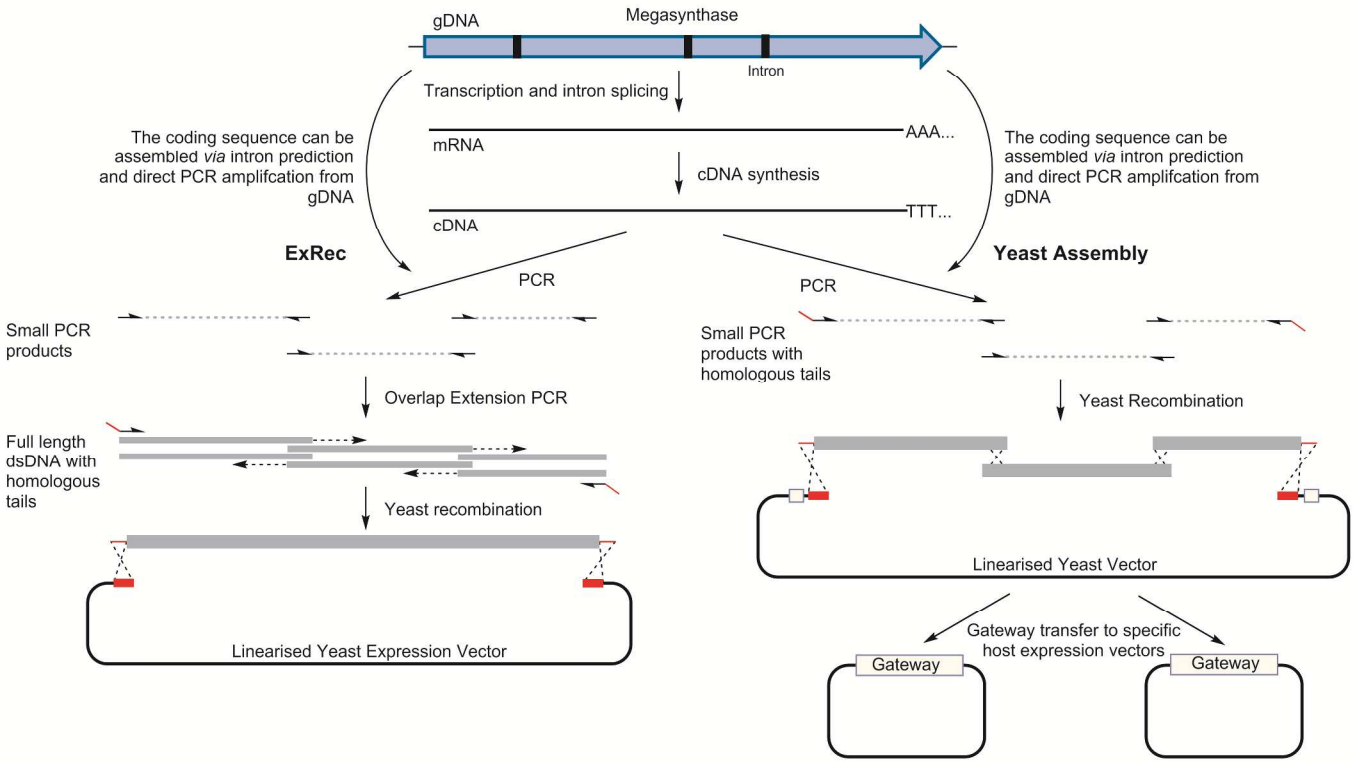

Figure 2: Comparison of two methods for assembling large genes using PCR and homologous recombination in yeast

Transformation of filamentous fungi is usually integrative, in contrast to the use of autonomous replicating plasmids commonly in yeast. This generates a population of transformants, each of which can have the plasmid inserted at a different chromosomal location. This means that some transformants have the cassette at genomic loci where the chromatin structure is such that expression may be rather low, whilst other integration events place the transgenes in loci more amenable to high-level expression.[49] In addition the transforming DNA may be present in differing copy numbers, single copy in some transformants and multiple copies in others.[50] Screening of several independent transformants is usually performed because the titre of the desired metabolites can differ enormously from transformant to transformant. Whilst it is possible to target insertion of the DNA to a specific locus, as in the asperfuranone example above, this is usually more time-consuming and delivers few advantages.

A major consideration in heterologous gene expression is the choice of promoter and the method of insertion of coding regions into expression cassettes. In terms of promoter selection, some workers have chosen constitutively expressed promoters while others have preferred regulated expression. Constitutive promoters mean that the desired product should be made irrespective of the growth conditions, which can help when screening transformants, while inducible promoters may have the advantage that fungal biomass can be generated before inducing the expression of what may be deleterious products. Similarly the strength of expression has rarely been addressed in a rational manner. Overexpression may increase titre, but may also trigger gene silencing pathways which cause degradation of the mRNA and hence no production.

In the asperfuranone example, Chiang *et al.*[30] linked the inducible *A. nidulans* alcohol dehydrogenase promoter to each of the coding regions to be expressed by fusion PCR; polyadenylation of transcripts was ensured by amplifying each coding region with its 3'-UTR. Repeated use of the strong, starch-inducible *amyB* expression cassette is a feature of many pathway reconstruction exercises in *A. oryzae*. Heneghan *et al.*[51] employed a multivector approach to reconstitute the 4-gene tenellin pathway from *Beauveria bassiana* in strain M-2-3 (using one auxotrophic and two dominant markers), with each cluster gene under the control of the *amyB* promoter. Fujii *et al.*[52] placed the 4 *Phoma betae* genes required for aphidicolin synthesis in *amyB* expression cassettes on separate plasmids to utilise all four auxotrophic markers available in strain NSAR1. By adopting a co-transformation approach termed "tandem transformation" Tagami *et al.*[53] reconstructed the 6-gene pathway for paxilline biosynthesis from *Penicillium paxilli* in *A. oryzae* using only 4 plasmid vectors; 2 vectors were used twice each, and the other 2 once each. Tandem transformation involves simultaneous introduction of two plasmids with the same selectable marker, resulting in the generation of transformants expressing one or other transgene or both by co-transformation. Even more extensive use of the *amyB*

expression cassette underpinned elucidation of the andrastin A biosynthetic pathway,[54] although the 9 genes involved were distributed among two transformants. Andrastin E resulted from the expression of 5 genes introduced on 4 plasmids, one of which carried 2 expression constructs. A second strain, expressing 4 genes introduced on 4 plasmids, converted andrastin E added to the culture medium to andrastin A.[54] It is worth noting that this *tour de force* of pathway reconstruction was achieved using only the conventional techniques of plasmid assembly (PCR, restriction and ligation) followed by sequential transformation of the hosts.

While the reconstruction of biosynthetic pathways by multiple sequential transformation steps has clearly produced excellent results, other workers have developed multigene expression systems that reduce the number of transformation steps required. One such approach involves linking all the genes of a biosynthetic pathway so that they are transcribed as a single polycistronic message, with viral 2A peptide co-translational cleavage signals separating the open reading frames.[22] To date this method has only been used in a proof-of-concept reconstruction of the 3-gene penicillin K pathway from *Penicillium chrysogenum* in *A. nidulans*, with *S. cerevisiae* used as an intermediate host for gene construction.[22] Homologous recombination in yeast is also a major feature of the one plasmid, multigene approach described by Pahirulzaman *et al.*,[45] in which up to 3 expression cassettes are filled in a single event. All cassettes on the vector need to be non-homologous to avoid undesirable crossovers resulting in deletion of segments of the plasmid. The selection of additional promoters was informed by an EST analysis that identified genes that are highly expressed in rich liquid cultures.[55] The fourth expression cassette, consisting of the *amyB* promoter and terminator, is modified to contain a Gateway destination fragment; this cassette is typically used to express a megasynthase gene assembled in a Gateway entry vector by homologous recombination in yeast, and inserted by site-specific recombination *in vitro* (Figure 3). Pahirulzaman *et al.*[45] proved the concept of the multigene expression system by reconstructing the 4-gene tenellin biosynthetic pathway on a single plasmid, and obtaining tenellin production in *A. oryzae* following a single transformation step. This contrasted with previous reconstruction of this pathway, in which all genes were expressed from *amyB* cassettes carried on 3 plasmids and sequentially transferred to *A. oryzae*.[51] While tenellin was successfully produced in the multigene system the most abundant expression product was pre-tenellin B, the immediate precursor of tenellin, indicating sub-optimal N-hydroxylase activity. While no explanation for this was immediately forthcoming RT-PCR ruled out low-level transcription of the *tenB* gene.

Further development of the multigene vector series is shown in Figure 4. All expression sites now have terminators, obviating the need to amplify coding regions with 3'-UTRs, and the dominant selectable markers previously used have been

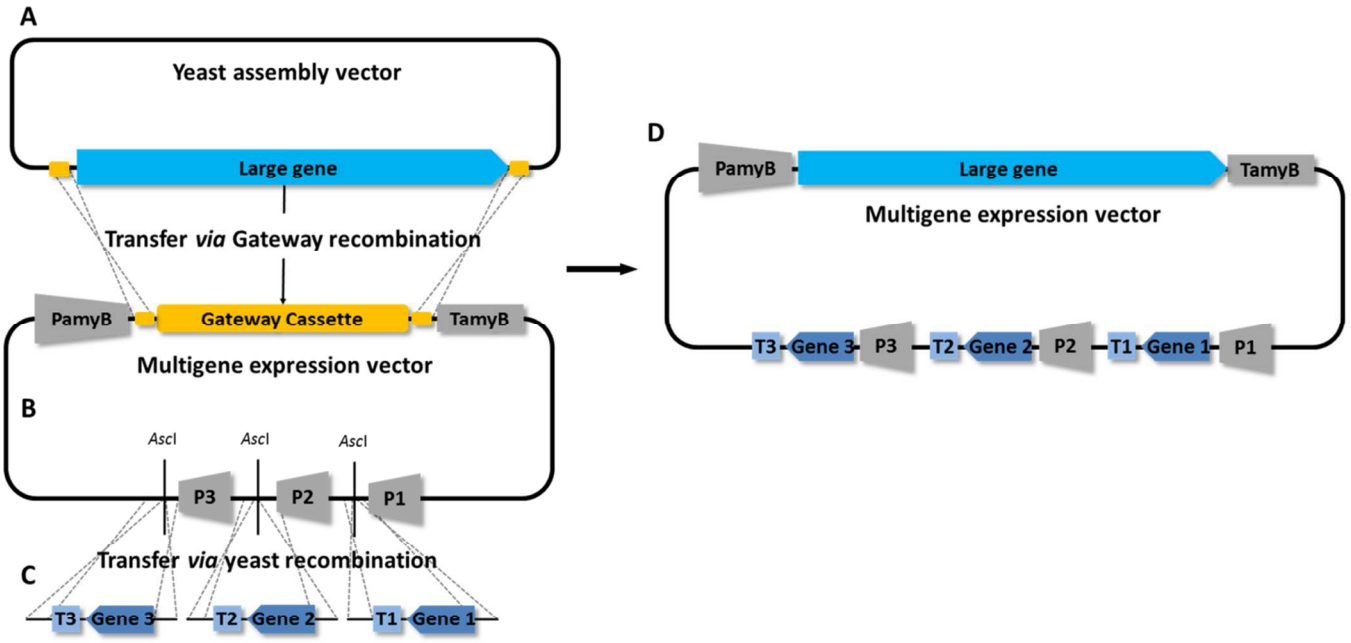

Figure 3: Illustration of multigene pathway reconstruction using homologous recombination in yeast and Gateway transfer **A)** The large gene in the yeast assembly vector can be transferred into the multigene expression vector via Gateway recombination. **B)** *Asc*I cuts the multigene expression vector downstream of each promoter (P1, P2 and P3). **C)** Genes encoding tailoring enzymes (plus terminators – T1, T2 and T3) can be amplified using primers that contain tails homologous to the cut ends of the vector. **D)** The final vector contains three genes encoding tailoring enzymes and the large synthase gene in expression cassettes tailored for high-level expression in *A. oryzae*

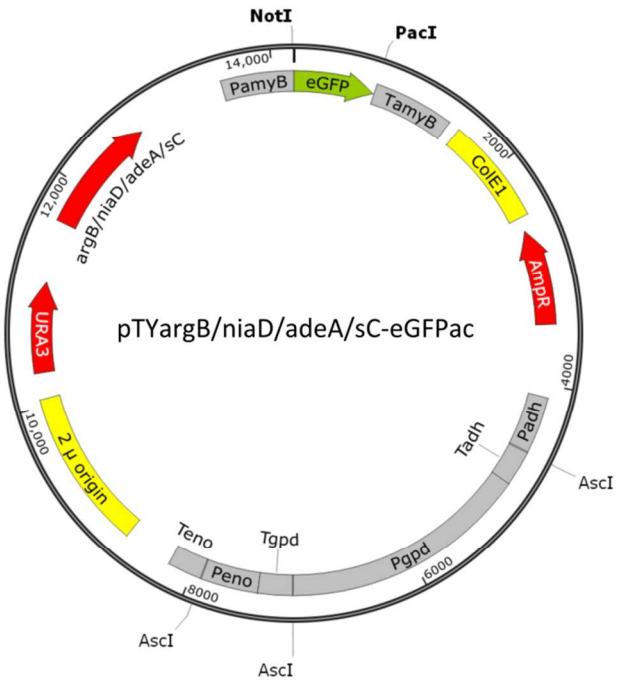

Figure 4: Map of the improved multigene expression vector. eGFP has replaced the Gateway cassette which can be used to create a gene fusion. Four versions of the vector exist, containing complementation markers for use with the *A. oryzae* strain NSAR1

augmented with additional complementation markers that can be used in NSAR1, allowing for further rounds of transformation necessary for clusters containing many genes. A further modification has been to replace the Gateway destination cassette with the eGFP coding region. This allows megasynthase genes to be reconstructed, optionally as eGFP fusions, by homologous recombination directly in the expression plasmid, either before or after filling the other cassettes. Tenellin pathway reconstruction again confirmed the unchanged efficacy of the system. The full-pathway plasmid was then subjected to homologous recombination in yeast to remove the single intron in the *tenB* gene. The effect, shown in Figure 5, was to decrease pre-tenellin B and elevate tenellin to the major product. Indeed the titres of tenellin produced were far in excess of those commonly observed in the native producer *B. bassiana.* This result indicates that correct intron splicing cannot be taken for granted when heterologously expressing fungal genes in other fungi, and that it is advisable to express cDNAs rather than genomic fragments when it is possible to do so. The sub-optimal splicing of the *B. bassiana tenB* intron in *A. oryzae* could be due to the relatively uncommon GC dinucleotide at the 5' splice site, but there are other examples where more conventional-looking introns may also be problematic. For example, the *Magnaporthe grisea* gene *ACE1*, encoding a hybrid PKS-NRPS,[56] contains 3 introns all with conventional GT-AG splice sites, only 2 of which are correctly spliced in *A. oryzae*.[57] A more complex situation was found when exploring expression of AsPKS1. Heterologous expression in *A. oryzae* yielded only one form of the transcript, giving a PKS which produced 3-methylorcinaldehyde. Analysis of the same gene in the native producer *Acremonium strictum*, showed that the second intron in AsPKS could be spliced in two differing ways: one, as in *A. oryzae,* gave a PKS with a reductive release mechanism yielding an aldehyde product, while the alternative splicing mechanism gave a truncated PKS with 3-methylorcellinic acid as its product.[58] This highlights how heterologous expression may not always give the expected or "correct" splicing pattern.

While heterologous expression has clearly had some significant successes, there is still much to be achieved. To date there has

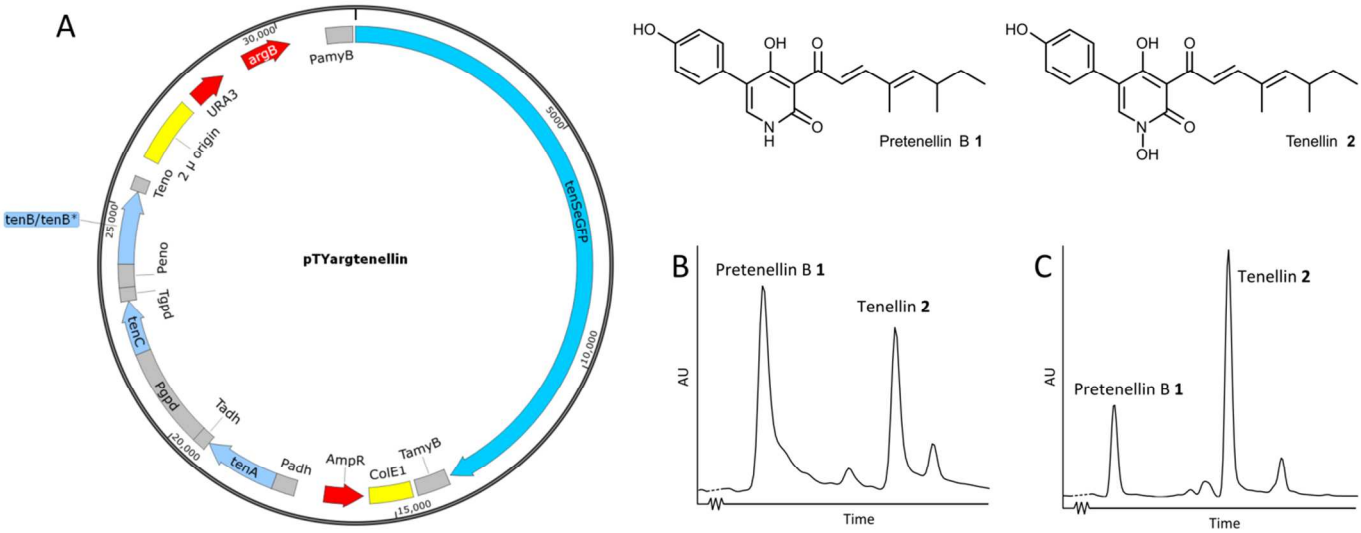

Figure 5: **A)** Map of the multigene expression plasmid containing the genes required to synthesise tenellin. The *tenB* gene used was either gDNA (*tenB*) or intronless (*tenB**). **B)** Diode array chromatogram of the crude extract from *A. oryzae* transformed with pTYargtenellin(tenB). The relative abundance of pretenellin B **1** compared to tenellin **2** using gDNA of tenB is 2.6:1. **C)** Diode array chromatogram of the crude extract from *A. oryzae* transformed with pTYargtenellin(tenB*).The relative abundance of pretenellin B **1** compared to tenellin **2** using the intronless tenB* is 0.6:1.

been little rationale in the choice of promoters used, in particular in assessing the comparative strength of expression needed for each step in a pathway. This may mean that flux through the pathway is not the same as in a native producer. Should this lead to accumulation of pathway intermediates, it is possible that one or more of the numerous oxidative enzymes present within the host genome may modify the products in an unexpected manner. Similarly, it is possible that careful selection of the appropriate promoter systems might enhance production. Thought may also need to be given to secretion of the metabolites. In the majority of cases to date, secretion has not been deliberately engineered, but rather has relied on the multitude of transporters already present in the heterologous host. Some secondary metabolite clusters include transporter genes, and it is entirely possible that increased heterologous expression of such transporters could lead to enhanced titre.

5. Conclusions

Heterologous gene expression has long been used as a tool of fungal natural products research, but recent years have seen it move towards centre stage. Much has been learned by moving individual genes, partial- or indeed complete gene clusters encoding fungal secondary metabolite biosynthetic pathways into heterologous hosts. Whilst there has been some success in using *E. coli* or yeast as the host for this, the majority of successful examples come from the use of other filamentous fungi as host systems, primarily those that are suitable for industrial exploitation such as *A. oryzae*. These approaches have allowed the functional characterisation of previously silent gene clusters or provided a means by which optimisation and enhanced production of mature products or pathway intermediates could be achieved. Careful selection of appropriate host and vector systems is needed to maximise the chances of successful expression, and recent advances in the development of toolkits and technologies for biosynthetic pathway reconstruction, as described in the preceding sections, are helping to streamline the process. The current rapid increase in fungal genome sequencing projects is revealing ever more diversity of potential natural product biosynthetic pathways, and heterologous gene expression will doubtless play a major role in their elucidation. However, this approach depends on the quality of the bioinformatics used to identify and delimit gene clusters, as well as the ability to accurately predict gene starts and ends and particularly intron boundaries. A potential outcome of this knowledge and technology will be the generation of much-needed new lead compounds for future exploitation in the pharmaceutical or agrochemical arena. Looking further to the future the challenge will be to harness the knowledge gained on individual enzyme activities from a range of biosynthetic pathways to produce rationally designed novel molecules entirely by heterologous gene expression.

6. Acknowledgements

We gratefully acknowledge the contributions of our colleagues in the Bristol Polyketide Group in laying the foundations for this work. In particular we thank Professors Tom Simpson and Russell Cox, together with "generations" of hard-working postdocs and graduate students too numerous to name individually. KW is supported by a BBSRC IPA with Syngenta.

7. Notes and references

[a] School of Biological Sciences, University of Bristol, Bristol Life Sciences Building, 24 Tyndall Avenue, Bristol, BS8 1TQ.
[b] School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS.

1. D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2012, **75**, 311-335.
2. G. Ozcengiz and A. L. Demain, *Biotechnol. Adv.*, 2013, **31**, 287-311.
3. J. D. Walton, *Fungal Genet. Biol.*, 2000, **30**, 167-171.
4. S. Bergmann, J. Schumann, K. Scherlach, C. Lange, A. A. Brakhage and C. Hertweck, *Nat. Chem. Biol.*, 2007, **3**, 213-217.
5. S. Jain and N. Keller, *Fungal Biol. Rev.*, 2013, **27**, 51-59.
6. A. A. Brakhage, *Nat. Rev. Microbiol.*, 2013, **11**, 21-32.
7. J. Antosch, F. Schaefers and T. A. M. Gulder, *Angew. Chem. Int. Ed. Engl.*, 2014, **53**, 3011-3014.
8. S. Agger, F. Lopez-Gallego and C. Schmidt-Dannert, *Mol. Microbiol.*, 2009, **72**, 1307-1308.
9. G. T. Wawrzyn, M. B. Quin, S. Choudhary, F. Lopez-Gallego and C. Schmidt-Dannert, *Chem. Biol.*, 2012, **19**, 772-783.
10. J. Davison, A. al Fahad, M. H. Cai, Z. Song, S. Y. Yehia, C. M. Lazarus, A. M. Bailey, T. J. Simpson and R. J. Cox, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 7642-7647.
11. A. O. Zabala, W. Xu, Y.-H. Chooi and Y. Tang, *Chem. Biol.*, 2012, **19**, 1049-1059.
12. C. J. Paddon, P. J. Westfall, D. J. Pitera, K. Benjamin, K. Fisher, D. McPhee, M. D. Leavell, A. Tai, A. Main, D. Eng, D. R. Polichuk, K. H. Teoh, D. W. Reed, T. Treynor, J. Lenihan, M. Fleck, S. Bajad, G. Dang, D. Dengrove, D. Diola, G. Dorin, K. W. Ellens, S. Fickes, J. Galazzo, S. P. Gaucher, T. Geistlinger, R. Henry, M. Hepp, T. Horning, T. Iqbal, H. Jiang, L. Kizer, B. Lieu, D. Melis, N. Moss, R. Regentin, S. Secrest, H. Tsuruta, R. Vazquez, L. F. Westblade, L. Xu, M. Yu, Y. Zhang, L. Zhao, J. Lievense, P. S. Covello, J. D. Keasling, K. K. Reiling, N. S. Renninger and J. D. Newman, *Nature*, 2013, **496**, 528-+.
13. S. M. Ma, J. W. H. Li, J. W. Choi, H. Zhou, K. K. M. Lee, V. A. Moorthie, X. Xie, J. T. Kealey, N. A. Da Silva, J. C. Vederas and Y. Tang, *Science*, 2009, **326**, 589-592.
14. M. Wang, M. Beissner and H. Zhao, *Chem. Biol.*, 2014, **21**, 257-263.
15. S. Wattanachaisaereekul, A. E. Lantz, M. L. Nielsen, Ó. S. Andrésson and J. Nielsen, *Biotechnol. Bioeng.*, 2007, **97**, 893-900.
16. R. Lambalot, Gehring, A.M., Fluge, R.S., Zuber, P., and M. LaCelle, Marahiel, M. A., Reid, R., Khosia, C., Walsh, C. T., *Chem. Biol.*, 1996, **3**, 923-936.
17. K. i. Ishiuchi, T. Nakazawa, T. Ookuma, S. Sugimoto, M. Sato, Y. Tsunematsu, N. Ishikawa, H. Noguchi, K. Hotta, H. Moriya and K. Watanabe, *ChemBioChem*, 2012, **13**, 846-854.
18. Y. Tsunematsu, K. i. Ishiuchi, K. Hotta and K. Watanabe, *Nat. Prod. Rep.*, 2013, **30**, 1139-1149.
19. P. Rugbjerg, M. Naesby, U. H. Mortensen and R. J. N. Frandsen, *Microb. Cell Fact.*, 2013, **12**.
20. L. Gidijala, J. A. K. W. Kiel, R. D. Douma, R. M. Seifar, W. M. van Gulik, R. A. L. Bovenberg, M. Veenhuis and I. J. van der Klei, *PLoS ONE*, 2009, **4**, e8317.
21. W. H. Meijer, L. Gidijala, S. Fekken, J. A. K. W. Kiel, M. A. van den Berg, R. Lascaris, R. A. L. Bovenberg and I. J. van der Klei, *Appl. Environ. Microbiol.*, 2010, **76**, 5702-5709.
22. S. E. Unkles, V. Valiante, D. J. Mattern and A. A. Brakhage, *Chem. Biol.*, 2014, **21**, 502-508.
23. J. Yaegashi, B. R. Oakley and C. C. C. Wang, *J. Ind. Microbiol. Biotechnol.*, 2014, **41**, 433-442.
24. D. J. Smith, M. K. R. Burnham, J. Edwards, A. J. Earl and G. Turner, *Nat. Biotechnol.*, 1990, **8**, 39-41.

25. K. Sakai, H. Kinoshita, T. Shimizu and T. Nihira, *J. Biosci. Bioeng.*, 2008, **106**, 466-472.
26. W.-B. Yin, Y. H. Chooi, A. R. Smith, R. A. Cacho, Y. Hu, T. C. White and Y. Tang, *ACS Synth. Biol.*, 2013, **2**, 629-634.
27. K. Sakai, H. Kinoshita and T. Nihira, *Appl. Microbiol. Biotechnol.*, 2012, **93**, 2011-2022.
28. J. W. Bok and N. P. Keller, *Eukaryot. Cell*, 2004, **3**, 527-535.
29. S. S. Lee, J. H. Lee and I. Lee, *J. Microbiol. Biotechnol.*, 2013, **23**, 959-965.
30. Y.-M. Chiang, C. E. Oakley, M. Ahuja, R. Entwistle, A. Schultz, S.-L. Chang, C. T. Sung, C. C. C. Wang and B. R. Oakley, *J. Am. Chem. Soc.*, 2013, **135**, 7720-7731.
31. P. Barbesgaard, H. P. Heldthansen and B. Diderichsen, *Appl. Microbiol. Biotechnol.*, 1992, **36**, 569-572.
32. A. Watanabe, Y. Ono, I. Fujii, U. Sankawa, M. E. Mayorga, W. E. Timberlake and Y. Ebizuka, *Tetrahedron Lett.*, 1998, **39**, 7733-7736.
33. A. Watanabe, I. Fujii, U. Sankawa, M. E. Mayorga, W. E. Timberlake and Y. Ebizuka, *Tetrahedron Lett.*, 1999, **40**, 91-94.
34. K. Gomi, Y. Iimura and S. Hara, *Agric. Biol. Chem.*, 1987, **51**, 2549-2555.
35. T. Fujii, H. Yamaoka, K. Gomi, K. Kitamoto and C. Kumagai, *Biosci., Biotechnol., Biochem.*, 1995, **59**, 1869-1874.
36. F. H. Jin, J.-I. Maruyama, P. R. Juvvadi, M. Arioka and K. Kitamoto, *FEMS Microbiol. Lett.*, 2004, **239**, 79-85.
37. J. Beck, S. Ripka, A. Siegner, E. Schiltz and E. Schweizer, *Eur. J. Biochem.*, 1990, **192**, 487-498.
38. A. P. MacCabe, H. Vanliempt, H. Palissa, S. E. Unkles, M. B. R. Riach, E. Pfeifer, H. Vondohren and J. R. Kinghorn, *J. Biol. Chem.*, 1991, **266**, 12646-12654.
39. Z. Song, R. J. Cox, C. M. Lazarus and T. J. Simpson, *ChemBioChem*, 2004, **5**, 1196-1203.
40. J. W. Sims, J. P. Fillmore, D. D. Warner and E. W. Schmidt, *Chem. Commun.*, 2005, 186-188.
41. B. G. Hansen, B. Salomonsen, M. T. Nielsen, J. B. Nielsen, N. B. Hansen, K. F. Nielsen, T. B. Regueira, J. Nielsen, K. R. Patil and U. H. Mortensen, *Appl. Environ. Microbiol.*, 2011, **77**, 3044-3051.
42. Z. Shao, H. Zhao and H. Zhao, *Nucleic Acids Res.*, 2009, **37**.
43. Z. Shao, Y. Luo and H. Zhao, *Mol. Biosyst.*, 2011, **7**, 1056-1059.
44. A. Munawar, J. W. Marshall, R. J. Cox, A. M. Bailey and C. M. Lazarus, *ChemBioChem*, 2013, **14**, 388-394.
45. K. A K Pahirulzaman, K. Williams and C. M. Lazarus, *Methods Enzymol.*, 2012, **517**, 241-260.
46. Z. Wasil, K. A. K. Pahirulzaman, C. Butts, T. J. Simpson, C. M. Lazarus and R. J. Cox, *Chem. Sci.*, 2013, **4**, 3845-3856.
47. M. N. Heneghan, A. A. Yakasai, K. Williams, K. A. Kadir, Z. Wasil, W. Bakeer, K. M. Fisch, A. M. Bailey, T. J. Simpson, R. J. Cox and C. M. Lazarus, *Chem. Sci.*, 2011.
48. K. M. Fisch, W. Bakeer, A. A. Yakasai, Z. Song, J. Pedrick, Z. Wasil, A. M. Bailey, C. M. Lazarus, T. J. Simpson and R. J. Cox, *J. Am. Chem. Soc.*, 2011, **133**, 16635-16641.
49. J. M. Palmer and N. P. Keller, *Curr. Opin. Microbiol.*, 2010, **13**, 431-436.
50. R. J. Watson, S. Burchat and J. Bosley, *Fungal Genet. Biol.*, 2008, **45**, 1348-1363.
51. M. N. Heneghan, A. A. Yakasai, L. M. Halo, Z. Song, A. M. Bailey, T. J. Simpson, R. J. Cox and C. M. Lazarus, *ChemBioChem*, 2010, **11**, 1508-1512.
52. R. Fujii, A. Minami, T. Tsukagoshi, N. Sato, T. Sahara, S. Ohgiya, K. Gomi and H. Oikawa, *Biosci., Biotechnol., Biochem.*, 2011, **75**, 1813-1817.
53. K. Tagami, C. Liu, A. Minami, M. Noike, T. Isaka, S. Fueki, Y. Shichijo, H. Toshima, K. Gomi, T. Dairi and H. Oikawa, *J. Am. Chem. Soc.*, 2013, **135**, 1260-1263.
54. Y. Matsuda, T. Awakawa and I. Abe, *Tetrahedron*, 2013, **69**, 8199-8204.
55. T. Akao, M. Sano, O. Yamada, T. Akeno, K. Fujii, K. Goto, S. Ohasi-Kunihiro, K. Takase, M. Yasukawa-Watanabe, K. Yamaguchi, Y. Kurihara, J. I. Maruyama, P. R. Juvvadi, A. Tanaka, Y. Hata, Y. Koyama, S. Yamaguchi, N. Kitamoto, K. Gomi, K. Abe, M. Takeuchi, T. Kobayashi, H. Horiuchi, K. Kitamoto, Y. Kashiwagi, M. Machida and O. Akita, *DNA Res.*, 2007, **14**, 47-57.
56. H. U. Bohnert, I. Fudal, W. Dioh, D. Tharreau, J. L. Notteghem and M. H. Lebrun, *Plant Cell*, 2004, **16**, 2499-2513.
57. W. Bakeer, University of Bristol, 2011.
58. K. M. Fisch, E. Skellam, D. Ivison, R. J. Cox, A. M. Bailey, C. M. Lazarus and T. J. Simpson, *Chem. Commun.*, 2010, **46**, 5331-5333.