Versatile enzyme expression and characterization system for Aspergillus

characterization of the polyketide synthase from the mycophenolic acid gene cluster from Penicillium brevicompactum as a case study

Hansen, Bjarne Gram; Salomonsen, Bo; Hansen, Bjørn; Nielsen, Morten Thrane; Nielsen, Jakob Blæsbjerg; Regueira, Torsten Bak; Nielsen, Kristian Fog; Patil, Kiran Raosaheb; Mortensen, Uffe Hasbro

Published in:
Applied and Environmental Microbiology

DOI (link to publication from Publisher):
10.1128/AEM.01768-10

Publication date:
2011

Document Version
Early version, also known as pre-print

AEM Accepts, published online ahead of print on 11 March 2011
Appl. Environ. Microbiol. doi:10.1128/AEM.01768-10

Running title: Versatile gene expression and characterization system for *Aspergillus*

Title:

A versatile gene expression and characterization system for *Aspergillus*: heterologous expression of the gene encoding the polyketide synthase from the mycophenolic acid gene cluster from *Penicillium brevicompactum* as a case study

Authors:

Bjarne G. Hansen[1], Bo Salomonsen[1], Morten T. Nielsen[1], Jakob B. Nielsen[1], Niels B. Hansen[1], Kristian F. Nielsen[1], Torsten B. Regueira[1,3], Jens Nielsen[2], Kiran R. Patil[1,4]* and Uffe H. Mortensen[1]*

Affiliation:

[1]Technical University of Denmark, Department of Systems Biology, Center for Microbial Biotechnology, 2800 Kgs Lyngby, Denmark.

[2]Department of Chemical and Biological Engineering, Systems Biology Group, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden.

[3]Current address: Novozymes A/S, Krogshøjvej 36, 2880 Bagsværd.

[4]Current address: Structural and Computational Biology Unit, EMBL-Heidelberg, Meyerhofstrasse 1, 69117 Heidelberg, Germany.

Corresponding authors:

Uffe H. Mortensen and Kiran R. Patil

Technical University of Denmark

Department of Systems Biology

Center for Microbial Biotechnology

Building 223

2800 Kgs Lyngby

Denmark

Email: um@bio.dtu.dk and kiran.patil@embl.de

Phone: +45 45252701

Fax: +45 45884148

Abstract

Assigning functions to newly discovered genes constitutes one of the major challenges en route to fully exploit the data becoming available from the genome sequencing initiatives. Heterologous expression in an appropriate host is central in functional genomics studies. In this context, filamentous fungi offer many advantages over bacterial and yeast systems. To facilitate the use of filamentous fungi in functional genomics, we present a versatile cloning system that allows a gene of interest to be expressed from a defined genomic location of *A. nidulans*. By a single USER cloning step, genes are easily inserted into a combined targeting-expression cassette ready for rapid integration and analysis. The system comprises a vector set that allows genes to be expressed either from the constitutive PgpdA promoter or from the inducible PalcA promoter. Moreover, by using the vector set, protein variants can easily be made and expressed from the same locus, which is mandatory for proper comparative analyses. Lastly, all individual elements of the vectors can easily be substituted for other similar elements ensuring the flexibility of the system. We have demonstrated the potential of the system by transferring the 7745 bp large *mpaC* gene from *P. brevicompactum* to *A. nidulans.* In parallel, we produced defined mutant derivatives of *mpaC,* and the combined analysis of *A. nidulans* strain expressing *mpaC* or mutated *mpaC* genes unequivocally demonstrated that *mpaC* indeed encodes a polyketide synthase that produces the first intermediate in the production of the medically important immunosuppressant mycophenolic acid.

Introduction

Filamentous fungi have the ability to produce a plethora of bioactive metabolites and enzymes enabling them to thrive in competitive environments. Amongst the metabolites are not only mycotoxins, but also compounds that are used as drugs e.g. the antibiotic penicillin and the immunosuppressant mycophenolic acid (MPA). Since many of the desirable products are naturally secreted in large amounts, fungi possess considerable potential as expression hosts for the production of small molecules as well as proteins. The wide interest in fungi has led to the sequencing of an increasing number of fungal genomes and this number is expected to increase dramatically in the coming years (29). This resource constitutes a tremendous potential for future advances in the basic understanding and industrial exploitation of fungal biology. For example, the number of gene clusters predicted to produce secondary metabolites such as polyketides and non-ribosomal peptides constantly increases as new fungal genome sequences are released. However, at present, it is only few of these gene clusters where the compound produced by the enzymes encoded by the gene cluster has been identified. Since for most organisms no or inefficient gene targeting technology exist, it is often difficult to assign secondary metabolites to specific genes in the natural host. This problem can be solved by transferring genes of interest to a suitable heterologous host that provides a wide range of genetic tools for gene characterizations. Considering the large number of genes to be analyzed, it is important to develop high throughput methods that facilitate this process.

Preferentially, the foreign gene should be expressed from a defined well characterized location. As compared to random integration this provides several advantages. Firstly, integration at a random site may cause mutation or alter expression of neighbouring genes causing unpredictable pleiotropic effects (3). Secondly, since integrated genes are differentially expressed depending

on the genomic context at the site of integration it eliminates undesired positioning effects by ensuring that the novel genes are integrated at a location that accommodates a high expression level (28). Thirdly, it allows for comparative studies where the phenotypes of strains expressing wild-type and mutated alleles can be reliably compared as the gene variants are expressed from the same locus in the different strains.

A. nidulans serves as a widely used model for filamentous fungi and has been extensively used for basic genetic research. Many genetic tools are therefore available including efficient gene targeting in strains where the non-homologous end-joining pathway for DNA integration has been eliminated (14, 16, 17). However, compared to the yeast, *Saccharomyces cerevisiae*, (where only 20 – 50 bp are needed) gene-targeting substrates need to contain large, > 1500 bp, homologous sequences to ensure integration at the selected locus. Accordingly, a gene targeting substrate that contains an expression cassette and a selectable marker is constructed from six pieces of DNA and often exceeds a total size of 10 kb complicating its construction (Figure 1). Constructing gene-targeting substrates therefore constitute a potential bottleneck in a high throughput gene analysis process.

Here we present a vector set based on the USER (uracil-specific excision reagent) cloning technique, which allows rapid and easy generation of constructs for targeted integration and heterologous expression of a gene of interest in *A. nidulans.* As a proof of concept, the vector set was used to express the heterologous *mpaC* gene from *P. brevicompactum* (see accompanying manuscript by Regueira et al.). *MpaC* is located in a gene cluster predicted to encode the biosynthetic enzymes for the production of the immunosuppressant mycophenolic acid. By analyzing *A. nidulans* strains that expresses *mpaC* under the control of both constitutive and inducible promoters, as well as constitutive expression of a point-mutated *mpaC,* we

conclusively show that MpaC catalyzes the production of 5-methyl orsellinic acid the first intermediate in mycophenolic acid production.

Materials and methods

Strains and media. The *A. nidulans* strains IBT28738 (*argB2, veA1, pyrG89, nkuA-trS*::AF*pyrG*) and IBT29539 (*argB2, veA1, pyrG89,* Δ*nkuA*) (17) were used for strain constructions when *argB* or *pyrG* was used as selection marker, respectively. The IBT30750 (*veA1*, *pyrG89*), NID127, was used as reference strain in growth experiments. Full list of *A. nidulans* strains are given in Table 1. All plasmids were propagated in *Escherichia coli* strain DH5α. Minimal medium (MM) contained 1 % glucose, 10 mM $NaNO_3$, 1x salt solution (5), and 2 % agar for solid media. MM was supplemented with 10 mM uridine (Uri), 10 mM uracil (Ura), and 4 mM L-arginine (Arg) when necessary. Solid plates containing 5-fluoroorotic acid (5-FOA) were made as MM+Uri+Ura medium supplemented with filter sterilized 5-FOA (Sigma-Aldrich) to a final concentration of 1.3 mg/ml. YES medium (yeast extract sucrose) was made as previously described (8), and supplemented with 10 mM Uri, 10 mM Ura, and 4 mM Arg when necessary. PalcA induction medium consisted of 100 mM L-threonine, 100 mM glycerol, 10 mM $NaNO_3$, Mineral Mix (1x), 2 g agar/l. PgpdA*::lacZ* activity was determined on MM medium supplemented with 0.122 mM 5-bromo-4-chloro-3-indolyl- beta-D-galactopyranoside (X-gal), 10 mM Uracile and 10 mM Uridine.

PCR and USER cloning. Amplification of DNA by PCR to produce DNA fragments suitable for USER cloning was performed in 30 PCR cycles using proof-reading PfuTurbo® Cx Hotstart polymerase (Stratagene) or PfuX7 (21) in 50 µl according to the manufacturer's instructions.

USER cloning was performed as previously described (22) with minor modifications. The USER vectors were digested for 6 h with AsiSI for the AsiSI/Nb.BsmI and AsiSI/Nb.BtsI USER cassettes or PacI for the PacI/Nt.BbvCI USER cassettes A and B followed by digestion with the appropriate nicking endonuclease for 1 h. 0.1 pmol purified digested vector was mixed with 1 pmol purified PCR products amplified with primers that were extended by the appropriate tails for USER cloning into a designated USER cassette (see, Supplementary figure 1,2 3, 4 and 5). When more than one PCR product was cloned simultaneously the combined concentration of PCR product was kept at 1 pmol and same concentration of each PCR product was used. The DNA mix was adjusted to 8 µl by adding Milli-Q purified water followed by addition of 1 µL of 10x TE buffer [100 mM Tris-HCl, 1 mM EDTA (pH 8.0)] and 1 U of USER™ enzyme mix (New England Biolabs). The reaction mixture was incubated for 20 min at 37 °C, followed by 20 min at 25 °C. Next, the 10 µL reaction mix was used directly to transform chemically competent *E. coli* cells.

Construction of USER vectors. Four different USER cassettes were designed to allow for the construction of the USER vector set: the two PacI/Nt.BbvCI USER cassettes A and B (Supplementary figure 3)(9), one AsiSI/Nb.BsmI USER cassette (Supplementary figure 1), and an AsiSI/Nb.BtsI USER cassette (Supplementary figure 2). The founder vector used to construct the USER vector set, pU0002, was custom made by DNA2.0 (Menlo Park, USA). pU0002 is based on pJ204 (www.dna20.com) and contains a USER linker containing two successive PacI/Nt.BbvCI USER cassettes, A and B. The USER linker is flanked by SwaI and NotI restriction sites on both sides. The cloning strategies and cloning into the individual USER cassettes and vectors are outlined in Supplementary figures 1,2,3,4 and 5. Sequence and

description of primers used towards creating this vector set is found in supplementary table 1 and 2.

Nomenclature for USER vector set. Nomenclature for our USER vector set follows the general system pU QXYZ-IS. Q describes the marker present in the plasmid for use after *A. nidulans* transformation. 0 = no marker, 1 = *argB*, 2 = AF*pyrG* (flanked by direct repeats to allow marker excision), 3 = *Ble* (resistance to the genotoxin bleomycin). X denotes which promoter is present in the plasmid. 0 = no promoter, 1 = PgpdA, 2 = PalcA. Y denotes which terminator is present in the plasmid. 0 = no terminator, 1 = TtrpC. Z defined the USER cassettes that are present in the plasmid. 0 = no USER cassette, 1 = AsiSI/Nb.BtsI casette, 2 = PacI/Nt.BbvCI USER cassettes A and B, 3 = AsiSI/Nb.BtsI cassette and PacI/Nt.BbvCI cassettes A and B, 4 = AsiSI/Nb.BsmI cassette and PacI/Nt.BbvCI cassettes A and B. IS is present in the name of the plasmid if PCR fragment for targeting has been inserted into the PacI/Nt.BbvCI USER cassettes A and B. 1 = targeting-regions for homologous recombination into "Insertion Site 1" is present.

USER cloning of *lacZ*, *RFP* and *mpaC* into USER vectors. *E. coli lacZ* was amplified from pWJ1042 (7) using the primers BGHA503 and BGHA504. RFP was amplified from pSK800 (27) using the primers BGHA564 and BGHA565. The *P. brevicompactum mpaC* coding sequence including introns was amplified from BAC1E13 (Regueira et al, accompanying manuscript) using the primers BGHA296 and BGHA297. Purified PCR products were USER cloned into the AsiSI/Nb.BtsI USER cassette in pU2111-1, pU2211-1, pU1111-1 and pU1211-1.

PCR generated sections of plasmids were sequenced (StarSeq, Germany) to confirm that no mutations were introduced by PCR errors.

***A. nidulans* strain construction.** Protoplasting and gene-targeting procedures were performed as described previously (11, 20). 5 µg of plasmid were digested either with NotI (*argB* and *Ble* containing constructs) or with SwaI (*pyrG* containing constructs) to liberate the gene targeting substrate, which was used for transformation of IBT28738 (using *argB* as selection marker) or IBT29539 (using *pyrG* as selection marker). Streak purified transformants were grown on 5-FOA medium to select for recombinants where the *nkuA* locus was restored to wild type as described previously (17). All gene-targeting events were verified by analytical PCR using *Taq*-polymerase (Sigma-Aldrich) and genomic DNA obtained from individual transformants. List of primers for verifying transformants in IS1 can be found in Supplementary Table 3. The upstream region of the integration events was tested using primer BGHA163 that anneals upstream of the TS1 sequence and a primer that anneals to the inserted promoter (BGHA502 for PgpdA and BGHA267 for PalcA), see Supplementary figure 6. Similarly, the downstream region of the integration event was tested using a primer BGHA162 that anneals downstream of the TS2 sequence and one that anneals to the selectable marker (BGHA98 for *argB* and BGHA182 for AF*pyrG*).

Creation of a point-mutation in *mpaC*. Point-mutations in the DSL motif of the Acyl-Carrier-Protein (ACP) domain in *mpaC* were created using a variation of the USER fusion method previously described (10). The DSL to ASL mutation was obtained by USER fusing two PCR

fragments generated by primer pairs (BGHA456 and BGHA297) and (BGHA457 and BGHA296), respectively, using *mpaC* as template. Simultaneously, the fusion fragment was USER cloned into the pU1111-1 vector fragment, which was included in the same reaction, see Results and Discussion for further details. The DSL to DAL mutation was made in the same way except that the two PCR fragments were generated by primer pairs (BGHA458 and BGHA297) and (BGHA459 and BGHA296). The two mutated *mpaC* genes were inserted into IS1 using the method described above.

Batch fermentation. Batch fermentations of *A. nidulans* were performed in 2-liter Braun fermentors with a working volume of 1.6 liters and equipped with two Rushton four-blade disk turbines. The bioreactor was sparged with air, and the concentrations of carbon dioxide in the exhaust gas were measured in a gas analyzer. Temperature was maintained at 30 °C and pH was kept constant at 5.5 controlled by automatic addition of either 2 M NaOH or 2 M HCl. Agitation and aeration were controlled throughout the cultivations. For inoculation of the bioreactor and germination of spores the stirring rate was set to 200 rpm, and aeration to 0.2 l/min. The bioreactors were inoculated with spores suspended in 0.9 % NaCl to a concentration of $1.5x10^9$/l. 12 hours after inoculation, the stirring rate was increased to 350 rpm and the air flow to 1.5 l/min and kept at that level for the remainder of the process. Statistical analysis of growth rates were performed using simple linear regression for multi-replicate data on natural logarithmic transformed data and comparison are based on a Student's t-test adopted for linear regression data.

Determination of β-galactosidase activity. Assays for β-galactosidase activity were performed using mycelia harvested from batch fermentations in the stationary phase (45 hours after inoculation) containing five grams of biomass (dry weight)/liter. Quantification was performed using the o-nitrophenyl-β-galactopyranoside method (15) with the following modifications. Mycelia from 1 mL fermentation broth was transferred to a 2 mL screwcap tube containing 250 µL glass beads (0.5 mm diameter), 250 µL ice cold Z-buffer (25) and 12.5 µL of a 100 mM 4-(2-amino-ethyl)- benzenesulfonyl fluoride hydrochloride solution. The samples were homogenized by shaking 30 seconds at maximum speed in a Fastprep FP120 (Bio 101 Savant). The extraction volumes were increased by addition of 250 µL ice-cold Z buffer and cleared for cell debris by centrifugation at 10.000 g for 30 min at 4°C. Total protein content in the cleared extract was determined using the Quant-It™ Protein Assay kit (Invitrogen) with the Quibit® fluorometer (Invitrogen) according to the instructions of the manufacturer.

Fluorescence microscopy. Microscopy was essentially performed as previously described (23). Images were captured with a cooled Evolution QEi monochrome digital camera (Media Cybernetics Inc.) mounted on a Nikon Eclipse E1000 camera (Nikon).

RNA isolation and qRT-PCR. *A. nidulans* total RNA, from four strains grown at identical conditions (NID127, NID210, NID211 and NID257), were isolated with Qiagen Plant RNAeasy kit. 10 µg of RNA was DNAse I (Qiagen) treated prior to cDNA amplification of 1 µg of DNAse I treated RNA samples by Phusion RT-PCR Kit (Finnzymes) according to protocol. The subsequent qRT-PCR was performed in a Chromo 4™ Detector/PTC-200 (MJ Research) using the SYBR® Green JumpStart Taq ReadyMix (Sigma). The *A. nidulans* actin gene, *actA* AN6542, was the internal standard for normalization of expression levels. Primer combination AN6638-F/AN6638-R was used for AN6638, AN6639-F/AN6639-R for AN6639, AN6640-F/AN6640-R

for AN6640, AN6636-F/AN6636-R for AN6636, ANactA-F/ANactA-R for *actA* and AN10837-F/AN10837-R for AN10837. All primers are listed in Supplementary table 3. Two types of control samples were included for the qPCR; DNAse treated RNA sample, and a template free reaction to test for primer-dimer influence on overall fluorescence. The individual cDNA samples were run both as concentrated samples (1/10 of the cDNA prep) and 10 times diluted samples (1/100x). Samples were run in triplicates. The program was 94 °C for 2 min and cycling conditions 40 times; 94 °C for 10 s, 60 °C for 15 s, 72 °C for 30 s. A melting curve from 65 °C to 95 °C with reads every 0.2 min was ending the program to evaluate the purity of the reaction products. The fluorescence threshold values, C(t), was determined through the OpticonMonitor 3.1 software (MJ Research). The relative expression levels was approximated by $2^{-\Delta\Delta C(t)}$ as $\Delta\Delta C(t) = \Delta C(t)_{normalized} - \Delta C(t)_{calibrator}$. $\Delta C(t)_{normalized} = \Delta C(t)_{target\ gene} - \Delta C(t)_{actA}$.The calibrator C(t) values are the values from the reference strain, NID127.

Chemical characterization of mutants. Three 6 mm diameter plugs were taken from each strain grown as three-point inoculations in the dark at 25 °C for 7 days on YES media (26). The plugs were transferred to a 2 mL vial and 1 mL of acetonitrile (ACN) was added. The plugs were placed in an ultrasonication bath for 60 min. The ACN was filtered and transferred to a new 2 mL clean vial, in which the organic phase was evaporated to dryness by applying nitrogen airflow at 30 °C. The residues were re-dissolved ultrasonically for 10 min in 150 µL ACN/H_2O (1:1, v/v) mixture. Samples of 1-5 µL were analysed by HPLC-UV/VIS using 15 to100 % ACN in 20 min on a Luna C_{18} column (19). HPLC-UV/VIS-high resolution mass spectrometry (LC-

HRMS) analysis was performed with an Agilent 1100 system (Waldbronn, Germany) equipped with a diode array detector and coupled to a Micromass LCT (Micromass, Manchester, U.K.) equipped with an electrospray (ESI) (18, 19). Separations of 1-5 µL sample was performed on a 100 × 2 mm inner diameter, 2.6 µm Kinetex C_{18}- column (Phenomenex, Torrance, CA) using a linear water-ACN gradient at a flow of 0.400 ml/min from 10 to 65 % ACN within 14 min, then to 100% ACN in 3 min, followed by a plateau at 100 % ACN for 3 min. Both solvents contained 20 mM formic acid. Samples were analyzed both in ESI^- and ESI^+ mode. For compound identification, each peak was matched against an internal reference standard database (~800 compounds) (18). 3-methyl orsellinic acid (Ambinter, Paris, France) and orsellenic acid (Apin chemicals, Oxon, UK) were co-analysed. Other peaks were tentatively identified by matching data from previous studies in our lab and searching the accurate mass in the ~ 13.000 fungal metabolites reported in Antibase 2009 (12). Here UV–VIS data, fragmentations, ionization efficiency in ESI^- versus ESI^+ and the retention time were used.

Results and discussion

Identification of a genomic insertion site for heterologous gene expression

An ideal all-round genomic integration site for characterization of heterologous genes and their products should accommodate the new genes without interfering with the fitness of the strain and allow for high and stable expression levels in a tissue unspecific manner. To identify such a site, we exploited the existing transcriptome microarray data of *A. nidulans* (1) obtained at

exponential growth on different carbon sources to identify genomic regions that support high expression at all the different growth conditions investigated. One possible site of integration, IS1, which fulfilled these criteria, is situated 202 bp downstream of AN6638 and 245 bp upstream of AN6639, was selected for further characterization. To evaluate the usefulness of IS1 as a site for heterologous expression we inserted *lacZ* and *RFP* into this locus (NID192 and NID257, respectively) by taking advantage of the vector set described below. Expression of *lacZ* from IS1 was first visualized by growing NID192 on X-gal containing plates. As expected NID192 colonies were bright blue due to β- galactosidase production whereas colonies from a reference strain (NID127), which does not contain *lacZ*, were pale, Figure 2A. Importantly, expression of *lacZ* appears stable over time as the blue appearance of colonies do not change over time as judged from visual inspection of colonies obtained from three successive re-stabbings (data not shown). To examine the level of protein production NID192 was grown in MM medium in well controlled bioreactors using a batch fermentation setup. Crude protein extracts prepared from NID192 mycelia, which were harvested in the stationary phase of the fermentation, converted o-nitrophenyl-β-D-galactoside at a rate of 5.3 µmol/min/mg total protein. This level of activity is similar to what was obtained by Lubertozzi and Keasling who inserted PgpdA::*lacZ* into three different loci, *argB*, *trpC*, or *niaD,* of *A. nidulans* (13). The reference strain, NID127, was analyzed in parallel and produced no detectable β-galactosidase activity (Supplementary Figure 7). Importantly we note that given these experimental conditions the growth rates of NID192 and NID127 were not significantly different ($p>0.2$) indicating that insertion of foreign DNA into IS1 does not impair fitness of the strain in MM medium. Inspection of NID257, which expresses RFP, by fluorescent microscopy reveals easily detectable RFP distributed uniformly throughout the mycelia (Figure 2B). Similarly, RFP was also present

in the spores (Figure 2B) indicating that RFP is expressed from IS1 in a tissue independent manner. We also investigated whether inserting an expression cassette into IS1 influences expression of the flanking genes AN6638 and AN6639. The expression levels of the two genes were determined in a reference strain as well as in three different strains containing the *PgpdA/TtrpC* expression cassette and the *argB* selection marker. In the three strains, the cassette contained either, nothing, *RFP* or *mpaC*. Of the two flanking genes, only AN6638, which is located just next to the constitutive promoter PgpdA, is significantly affected (~2.5 fold increase) in these three strains. To investigate whether other genes close to this locus were affected by the presence of an expression cassette in IS1, the same analysis was performed for AN6636, AN10837and AN6639. None of these genes were expressed differently in the three strains containing an expression cassette in IS1 as compared to the reference strain. The modest effect of gene expression in the IS1 region after integration of an expression cassette is in agreement with the finding that the growth rate of a strain containing an *lacZ* expression cassette in IS1 does not influence the growth rate. Based on the combined results of the experiments described above, we conclude that the integration site IS1 is useful for integrating novel genes for their further characterization.

Construction of a flexible USER vector set for gene analysis in *Aspergillus nidulans*

To facilitate exploiting IS1 as a convenient expression platform for foreign genes, we have constructed a flexible vector set, which allows for easy construction of gene targeting substrates for integration of Your Favorite Gene (YFG) into IS1 by taking advantage of the DNA ligase free improved USER cloning system (22). So far the set comprises six vectors, see Table 2, allowing for integration of YFG into IS1 under the control of either inducible PalcA or

constitutive PgpdA using either the selectable marker *argB*, *Ble* or *pyrG*. The latter marker can be recycled after transformation and used in subsequent experiments since it is flanked by direct repeats that allow *pyrG* to be eliminated by pop-out recombination (20). In all vectors, the terminator TtrpC is present downstream of the *YFG* integration site.

Using one (or more) of the vectors in our vector set, YFG can rapidly be introduced into IS1 in *A. nidulans* by performing four simple steps, see Figure 3. Step 1) the gene of interest is PCR amplified with primers containing the appropriate tails for USER cloning into the AsiSI/Nb.BtsI USER cassette; Step 2) PCR fragment is USER cloned into the appropriate vector; Step 3) the completed gene targeting substrate containing YFG is released from this vector by restriction enzyme digest or by PCR; Step 4), the gene targeting substrate is used for transformation of *A. nidulans* protoplasts. In our hands, this can be done in less than 48 hours and *A. nidulans* transformants are obtained a few days later.

The vector set is designed for optimal flexibility. Hence, in case other markers, promoters, terminators or other integration sites are preferred, the present repertoire of vectors can easily be expanded, since all parts can be replaced in simple USER cloning based reactions. As a guideline for future vector construction, we present the strategy for building pU2111-1 as a model; see Supplementary figure 5 (and details in Supplementary figures 1, 2, 3 and 4).

Cloning of *mpaC* from *Penicillium brevicompactum* and transformation into *A. nidulans*

To demonstrate the potential of the presented USER vector set for heterologus expression and characterization of *YFG* in *A. nidulans* we decided to investigate a recently discovered gene cluster from *Penicillium brevicompactum*, which is proposed to encode the enzymes required for

MPA production, see accompanying manuscript by Reguira et al. The *mpaC* gene in the cluster consists of 7745 bp, including introns, and is predicted to encode a polyketide synthase that catalyses production of the polyketide 5-methyl orsellinic acid (5-MOA), which is believed to be the first intermediate in MPA biosynthesis (2). The large size of *mpaC* makes it well suited to test the robustness of our USER cloning based system for cloning and inserting *YFG* into IS1. The large *mpaC* gene was PCR amplified as three fragments that were subsequently fused by taking advantage of the USER fusion technique (10), which allows several PCR fragments to be merged in a single cloning step. Using this principle, the entire *mpaC*, was readily inserted into the *argB* containing vectors pU1111-1 and pU1211-1, hence, equipping *mpaC* with the constitutive *PgpdA* and inducible *PalcA* promoters, respectively. In both cases more than 50% of the colonies had *mpaC* inserted between the promoter and terminator demonstrating the efficiency of the construction part of the system. The entire *mpaC* was sequenced for both constructs and no PCR generated errors were observed. Using the *argB* marker for selection, the two constructs were used for transformation of *A. nidulans* IBT28738 protoplasts. This strain allows for efficient integration of gene targeting substrates by homologous recombination since the competing pathway for genomic DNA integration, non-homologous end-joining, has been transiently eliminated due to a *pyrG* insertion in the *nkuA* locus (17). In agreement with this, a PCR test demonstrated that the *mpaC* expression cassette was integrated into IS1 in all transformants analyzed (data not shown). Finally, selected transformants were grown on medium containing 5-FOA to reconstitute *nkuA* to avoid any influence of a defective *nkuA* gene on further strain characterization, see reference 17 for details.

Introduction of a point-mutation in the DSL motif in the ACP domain of *mpaC*

Detailed characterization of a gene product requires simple means to introduce genetic modifications like deletions and point mutations. Using the present commercially available techniques, introduction of such modifications in large vectors, like those containing an entire gene targeting substrate, constitutes a challenging and tedious task. By making a slight modification of the USER fusion technique described above, point-mutations and deletions can easily be introduced in YFG. Hence, if the primer tails used to merge individual segments of YFG contain the desired point mutation, defined sequence modifications can be inserted anywhere in a gene (Figure 4). Similarly, deletions can be introduced by designing matching USER primer tails, which at the fusion point, bridges two non-continuous, but successive, sections of YFG. Here we demonstrate the principle of introducing point mutations into *mpaC* by using USER cloning to independently introduce two alanine substitutions, D1622A and S1623A, in the conserved DSL motif in the ACP domain of MpaC. Both substitutions are predicted to impair polyketide synthase activity. It is known that the phosphopantetheine moiety of coenzyme A bind to the serine in the DSL domain of PKSs (6) and therefore a S1623A substitution in MpaC is likely to prevent the phosphopantetheine moiety of coenzyme A to bind to the serine in the DSL domain. Accordingly, MpaC cannot be converted from the inactive apo form to the active holo form in this mutant protein. From studies on other PKS it has been found that the negative charge aspartate in the DSL motif in the acyl-carrier protein domain creates a salt bridge to the acyl-transferase domain ensuring that these two domains interact properly (4). Therefore, the D1622A mutation in MpaC is predicted to disrupt MpaC activity. The creation of the two *mpaC* mutants was as fast and efficient as cloning of the wild-type *mpaC* described above. Both *mpaC* variants were verified by sequencing and no additional PCR generated errors

were observed. We note that the method is applicable of introducing several point-mutations in different regions of interest in a single round of cloning, simply by fusing additional PCR fragments. Specifically, simultaneous construction of two mutations require the fusion of three PCR fragments, three mutations require the fusion of four PCR fragments, and so on. To this end we note, that presently up to five fragments have been efficiently fused by USER fusion and that the upper limit has yet not been delineated (10).

Expression of *mpaC* in *A. nidulans* results in 5-methyl orsellinic acid production

To investigate for 5-MOA production in *A. nidulans*, we first analyzed a reference strain NID210 (AR1 that contains *argB* integrated at IS1) for its ability to produce 5-MOA. Since 5-MOA is not commercially available, a reference standard of 3-methyl orsellinic acid (3-MOA), which is expected to behave very similar to 5-MOA in LC-HRMS analyses were analysed. Since neither 3- or 5-MOA have previously been reported in *A. nidulans*, we surprisingly identified a compound eluting at 4.09 min with an elementary composition (Figure 5A) identical to that of both 3- and 5-MOA in the extract. However, since both the retention time and the UV-spectrum of this compound are identical to that of the 3-MOA standard, it was unambiguously assigned as 3-MOA. Next, a strain containing the *PgpdA::mpaC* expression cassette at IS1 (NID211) was analyzed by LC-HRMS. In contrast to the reference strain, NID211 produced a compound eluting as a prominent peak at 3.78 min with the mass expected for 5-MOA (Figure 5A). This peak contained a unique ion with m/z 181.050 corresponding to the $[M-H]^-$ ion of $C_9H_{10}O_4$. Moreover, it produces a UV spectrum (Figure 5B), which is identical to the previously published UV spectrum for 5-MOA (24). We note that, NID211, like the reference strains, also produced the compound eluting at 4.09 min supporting that this compound was 3-MOA.

Next we addressed whether the unique compound produced by NID211 is due to the *mpaC* gene product or whether it results from the expression of endogenous *A. nidulans* gene(s) that are accidentally activated by insertion of the PgpdA::*mpaC* expression cassette into IS1. To this end we first analyzed the two strains expressing PgpdA::*mpaC-D1622A* and PgpdA::*mpaC-S1623A* (NID189 and NID190, respectively). In both cases, no compound eluted at 3.78 min. Moreover, strain NID66, which harbors the expression cassette PalcA::*mpaC* inserted at IS1, produced a compound eluting at 3.78, but only when the strain was grown on media inducing expression from PalcA (Figure 5C). Together the data conclusively demonstrate that *mpaC* encodes a polyketide synthase that produces 5-MOA. In addition, since the *mpaC* gene expressed in *A. nidulans* contained its native *P. brevicompactum* introns, we conclude that these introns are efficiently removed by the splicing apparatus of *A. nidulans*.

Concluding remarks

In this report we have presented a simple USER cloning based system that allows genes to be transferred from organisms of interest into the well characterized fungal model *A. nidulans* for further characterization. As proof of concept we firmly demonstrate that *mpaC* from *P. brevicompactum* encodes the PKS responsible for production of 5-MOA, the first intermediate in MPA production. Importantly, since the vector set is constructed in a flexible manner, it can easily be modified to allow specific integration of YFG into other organisms that support efficient gene targeting if desirable. The strategy for gene characterization presented here is therefore widely applicable and should greatly facilitate assignment of gene functions in organisms where the genetic tool-box is poorly developed.

Acknowledgements

The authors would like to thank Dorte Marie Koefoed Holm for assisting in identification of IS1. The work was supported by grants number 09-064967 and 09-064240 from the The Danish Council for Independent Research, Technology and Production Sciences to KRP and UHM. JN was supported by the Chalmers Foundation. We thank Martin Engelhard Kornholt for valuable technical assistance in the laboratory.

Literature cited

1. **Andersen, M.R., Vongsangnak, W., Panagiotou, G., Salazar, M.P., Lehmann, L. and Nielsen, J.** 2008. A trispecies Aspergillus Microarray: comparative transcriptomics of three Aspergillus species. Proc. Natl. Acad. Sci. USA. **105**:4387-4392.
2. **Bentley, R.** 2000. Mycophenolic Acid: A one hundred year odyssey from antibiotic to immunosuppressant. Chem. Rev. **100**:3801-3825.
3. **Chiou, C., Miller, M., Wilson, D.L., Trail, F. and Linz, J.E.** 2002. Chromosome Location Plays a Role in Regulation of Aflatoxin Gene Expression in *Aspergillus parasiticus*. Appl. Environ. Microbiol. **68**:306-315.
4. **Cox, R.J. and Simpson, T.J.** 2009. Fungal type I polyketide synthases. Methods in Enzymol. **459**:49-78.
5. **Cove, D.J.** (1966) The induction and repression of nitrate reductase in the fungus Aspergillus nidulans. Biochim. Biophys. Acta. **113**: 51-56.

6. **Evans, S.E., Williams, C., Arthur, C.J., Burston, S.G., Simpson, T.J., Crosby, J. and Crump, M.P**. 2008. An ACP Structural Switch: Conformational Differences between the Apo and Halo Forms of the Actinorhodin Polyketide Synthase Acyl Carrier Protein. ChemBioChem. **9**:2424-2432.

7. **Flagfeldt, D.B., Siewers, V., Huang, L. and Nielsen, J**. 2009. Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae. Yeast. **26**:545-551.

8. **Frisvad, J.C. and Thrane, U**. 1987. Standardised High-Performance Liquid Chromatography of 182 mycotoxins and other fungal metabolites based on alkylphenone retention indices and UV-VIS spectra (Diode Array Detection). J.Chromatogr. **404**:195-214.

9. **Frandsen, R.J.N., Andersson, J.A., Kristensen, M.B. and Giese, H**. 2008. Efficient four fragment cloning for the construction of vectors for targeted gene replacement in filamentous fungi. BMC Mol. Biol. **9**:70.

10. **Geu-Flores, F., Nour-Eldin, H.H., Nielsen, M.T. and Halkier, B.A**. 2007. USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. Nucleic Acids Res. **35**:e55.

11. **Johnstone, I.L., Hughes, S.G. and Clutterbuck, A.J**. 1985. Cloning an *Aspergillus nidulans* developmental gene by transformation. EMBO J. **4**:1307–1311.

12. **Laatsch, H**. 2009. AntiBase 2009. The Natural Compound Identifier. Wiley-VCH GmbH & Co. Weinheim, Germany.

13. **Lubertozzi, D. and Keasling, J.D**. 2006. Marker and promoter effects on heterologous expression in *Aspergillus nidulans*. Appl. Microbiol. Biotechnol. **72**:1014-1023.

14. **Meyer, V**. 2008. Genetic engineering of filamentous fungi - Progress, obstacles and future trends. Biotechnol. Adv. **26**:177-185.
15. **Miller, J.H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA.
16. **Nayak, T., Szewczyk E., Oakley, C.E., Osmani, A., Ukil, L., Murray, S.L., Hynes, M.J., Osmani, S.A. and Oakley, B.R**. 2006. A versatile and efficient gene-targeting system for *Aspergillus nidulans*. Genetics. 172:1557-1566.
17. **Nielsen, J.B., Nielsen, M.L. and Mortensen, U.H**. 2008. Transient disruption of non-homologous end-joining facilitates targeted genome manipulation in the filamentous fungus *Aspergillus nidulans*. Fungal Genet. Biol. **45**:165-170
18. **Nielsen, K.F. and Smedsgaard, J**. 2003. Fungal metabolite screening: database of 474 mycotoxins and fungal metabolites for de-replication by standardised liquid chromatography-UV-mass spectrometry methodology. J.Chromatogr.A **1002**:111-136.
19. **Nielsen, K.F., Mogensen, J.M., Johansen, M., Larsen, T.O. and Frisvad, J.C**. 2009. Review of secondary metabolites and mycotoxins from the Aspergillus niger group. Anal. Bioanal. Chem. **395**:1225-1242.
20. **Nielsen, M.L., Albertsen, L., Lettier., G., Nielsen, J.B. and Mortensen, U.H**. 2006. Efficient PCR-based gene targeting with a recyclable marker for *Aspergillus nidulans*. Fungal Genet. Biol. **43**:54-64.
21. **Norholm, M**.H.H. 2010. A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. BMC Biotechnology. **10**:21.

22. **Nour-Eldin, H.H., Hansen, B.G., Nørholm, M.H.H., Jensen, J.K. and Halkier, B.A**. 2006. Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. Nucleic Acids Res. **34**:e122.

23. **Plate, I., Hallwyl, S.C.L., Shi, I., Krejci, L., Muller, C., Albertsen, L., Sung, P. and Mortensen. U.H**. 2008. Interaction with RPA is necessary for Rad52 repair center formation and for its mediator activity. J Biol Chem. **283**:29077-29085.

24. **Puel, O., Tadrist, S., Galtier, P., Oswald, I.P. and Delaforge, M.** 2005. *Byssochlamys nivea* as a source of mycophenolic acid. Appl. Environ. Microbiol. **71**:550-553.

25. **Sambrook, J. and Russell, D.W.** 2001. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA.

26. **Smedsgaard, J**. 1997. Micro-scale extraction procedure for standardized screening of fungal metabolite production in cultures. J.Chromatogr. A. **760**:264-270.

27. **Toews, M.W., Warmbold, J., Konzack, S., Rischitor, P., Veith, D., Vienken, K., Vinuesa, C., Wei, H. and Fischer, R.** 2004. Establishment of mRFP1 as a fluorescent marker in Aspergillus nidulans and construction of expression vectors for high-throughput protein tagging using recombination in vitro (GATEWAY). Curr. Genet. **45**:383-389.

28. **Verdoes, J.C., Punt, P.J. and van den Hondel, C.A.M.J.J**. 1995. Molecular genetics strain improvement for the overproduction of fungal proteins by Filamentous fungi. Appl. Microbiol. Biotechnol. **43**:195-205.

29. **Wilson, R.A. and Talbot, N.J**. 2009. Fungal physiology – a future perspective. Microbiology. **155**:3810-3815.

FIG. 1. Integration of a general expression cassette into the integration site IS1 situated between AN6638 and AN6639 on chromosome I (represented as a fat black line) by homologous recombination. The gene targeting substrate contains six parts: YFG (your favourite gene), TSI and TSII (targeting sequence I ,1949 bp; and II, 1926 bp), prom, term and marker (promoter, terminator and marker of choice). The orientation of genes AN6638 and AN6639 are indicated by green arrows. Drawing is not made to scale.

FIG. 2. Expression of β-galactosidase and RFP from IS1 in *A. nidulans*. (A) To the left, diagram showing the positions of strains on the plate shown to the right. NID127 is a reference strain and NID192 contains PgpdA*::lacZ.* To the right, strains were stabbed on an X-gal containing plate and incubated for two days at 37^{o}C before it was subjected to photography. (B) NID257 containing PgpdA*::RFP* was examined by bright field (BF) microscopy and by fluorescent microscopy detecting RFP as indicated. Top panels show a full mycelium, middle panels show hyphal tips and bottom panels show conidia spores. A scale bar is included in all panels for comparison.

FIG. 3. Construction of a gene targeting substrate designed for integrating an expression cassette into IS1. The example shows construction of an expression cassette containing YFG flanked by the PgpdA promoter and the TtrpC terminator. A vector fragment with eight nucleotides ssDNA overhangs is generated by cutting pU2111-1 simultaneously with restriction endonuclease AsiSI and restriction nicking endonuclease Nb.BtsI. Note, the two overhangs contain three different nucleotides (presented by pale green and pale purple) to ensure directionality of subsequent fragment insertion. A PCR generated insert fragment containing YFG is generated by primers that are extended at the 5' ends by a uracil residue (highlighted in magenta) followed by seven non-priming nucleotides. Treatment of the PCR

fragment with USER enzyme removes uracil residues to generate ends with eight nucleotides ssDNA overhangs that match the ends of the vector fragment. Insert and vector fragments are mixed and co-transformed into *E. coli* where the two fragments are joined by *E. coli* mediated ligation. AsiSI and Nb.BtsI recognition sequences are marked in light blue and tan, respectively. Drawing is not made to scale.

FIG. 4. Method to introduce point mutations in YFG by USER cloning as exemplified by mutation in *mpaC*. (A) Graphical representation of *mpaC*. A section of *mpaC* encoding the region around and including the conserved DSL motif in the ACP domain is presented at the nucleotide level. The codon targeted for mutation is marked in bold. (B) Annealing positions of primers required for site directed mutagenesis of the DSL motif of MpaC are indicated by arrows. The 5' sections of the primers, which are relevant for USER fusion of fragments, are presented as nucleotides. Red nucleotides in primers BGHA458 and BGHA459 represent the desired mutation. (C) The entire *mpaC* gene is amplified as two fragments in two separate PCR reactions using the two primer pairs (BGHA296 and BGHA459) and (BGHA458 and BGHA297) and *mpaC* as template. The two fragments are subsequently combined and inserted into an expression vector in a single USER cloning step. The desired mutation is generated at the fusion point where the two fragments merge via the overlapping tails of the two mutagenic primers BGHA458 and BGHA459. Drawing is not made to scale.

FIG. 5. Expression of *mpaC* in *A. nidulans* results in 5-MOA production. (A) UV chromatograms from LC-UV/VIS-HRMS analyses of the standard 3-MOA and strains NID210, NID211 and NID190. Extracted ion chromatogram, m/z 181 corresponding to the $[M-H]^-$ ion of methyl orsellinic acids is inserted in each chromatogram. (B) UV spectra for

3-MOA (pure compound) and 5-MOA from strain NID211. (C) Overlay UV chromatogram from HPLC-UV/VIS for extract obtained from strain NID66 with *mpaC* under control of the inducible alcA promoter. Chromatograms representing inducing (+Thr) and non-inducing conditions (-Thr) are indicated by arrows.

Table 1. Name and description of the strains used in this work

Strain	Genotype
NID66	*argB2, pyrG89, veA1*, Δ*nkuA*, IS1::PalcA-*mpaC*-TtrpC-AF*pyrG*
NID127	*pyrG89*, *veA1*
NID189	*argB2*, *pyrG89*, *veA1*, IS1::PgpdA-*mpaCD1622A*-TtrpC-*argB*
NID190	*argB2*, *pyrG89*, *veA1*, IS1::PgpdA-*mpaCS1623A*-TtrpC-*argB*
NID192	*argB2*, *pyrG89*, *veA1*, IS1::PgpdA-*lacZ*-TtrpC-*argB*
NID210	*argB2*, *pyrG89*, *veA1*, IS1::PgpdA-TtrpC-*argB*
NID211	*argB2*, *pyrG89*, *veA1*, IS1::PgpdA-*mpaC*-TtrpC-*argB*
NID257	*argB2*, *pyrG89*, *veA1*, IS1::PgpdA-*mRFP*-TtrpC-*argB*

Table 2. Name and description of one step USER vectors for inserting YFG into IS1.

Vector	Marker	Promoter	Terminator	User Cloning Cassette	Integration site
pU1111-1	*argB*	PgpdA	TtrpC	AsiSI/Nb.BtsI	IS1
pU1211-1	*argB*	PalcA	TtrpC	AsiSI/Nb.BtsI	IS1
pU2111-1	*pyrG*	PgpdA	TtrpC	AsiSI/Nb.BtsI	IS1
pU2211-1	*pyrG*	PalcA	TtrpC	AsiSI/Nb.BtsI	IS1
pU3111-1	*Ble*[a]	PgpdA	TtrpC	AsiSI/Nb.BtsI	IS1
pU3211-1	*Ble*[a]	PalcA	TtrpC	AsiSI/Nb.BtsI	IS1

[a] *Ble* confers resistance to the genotoxin bleomycin

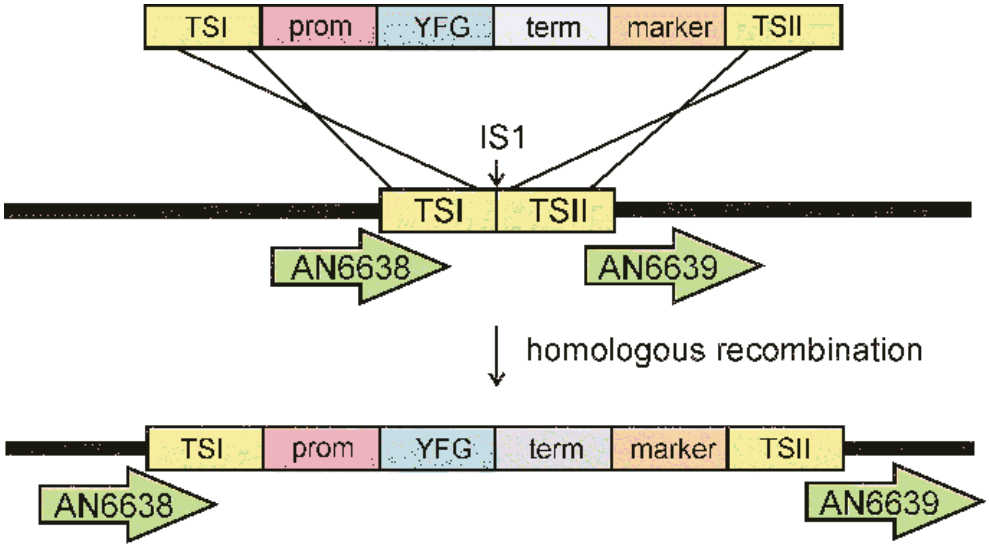

Figure 1

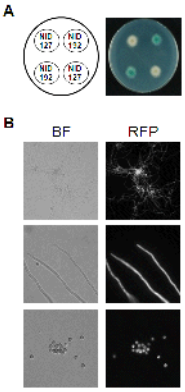

Figure 2

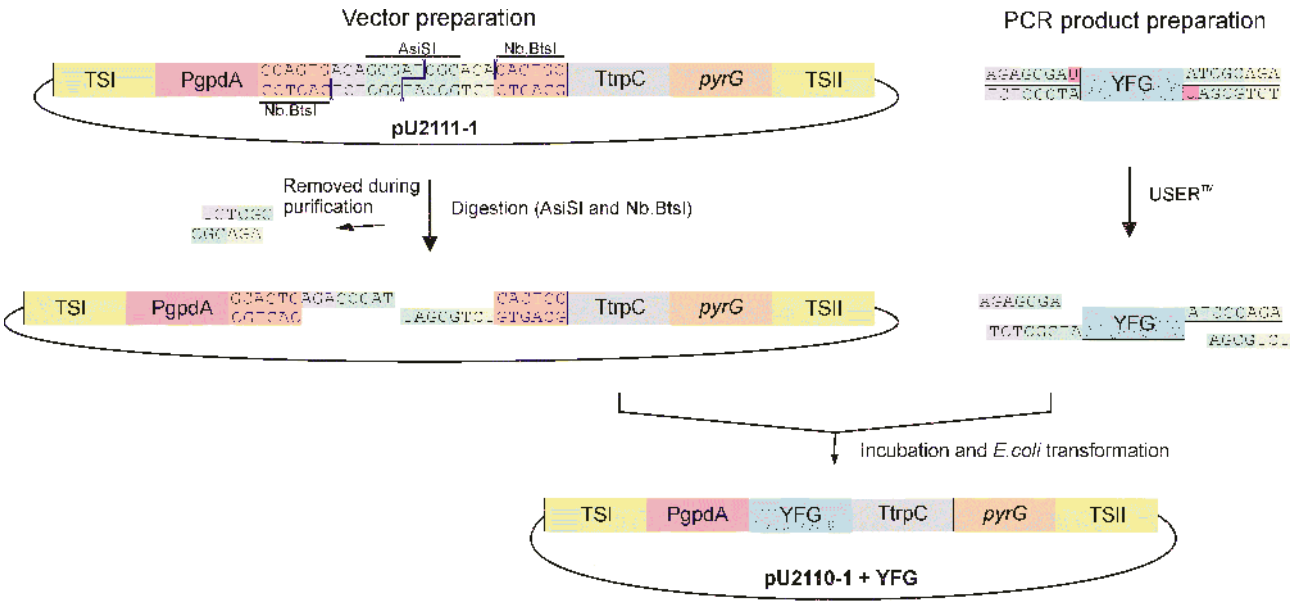

Figure 3

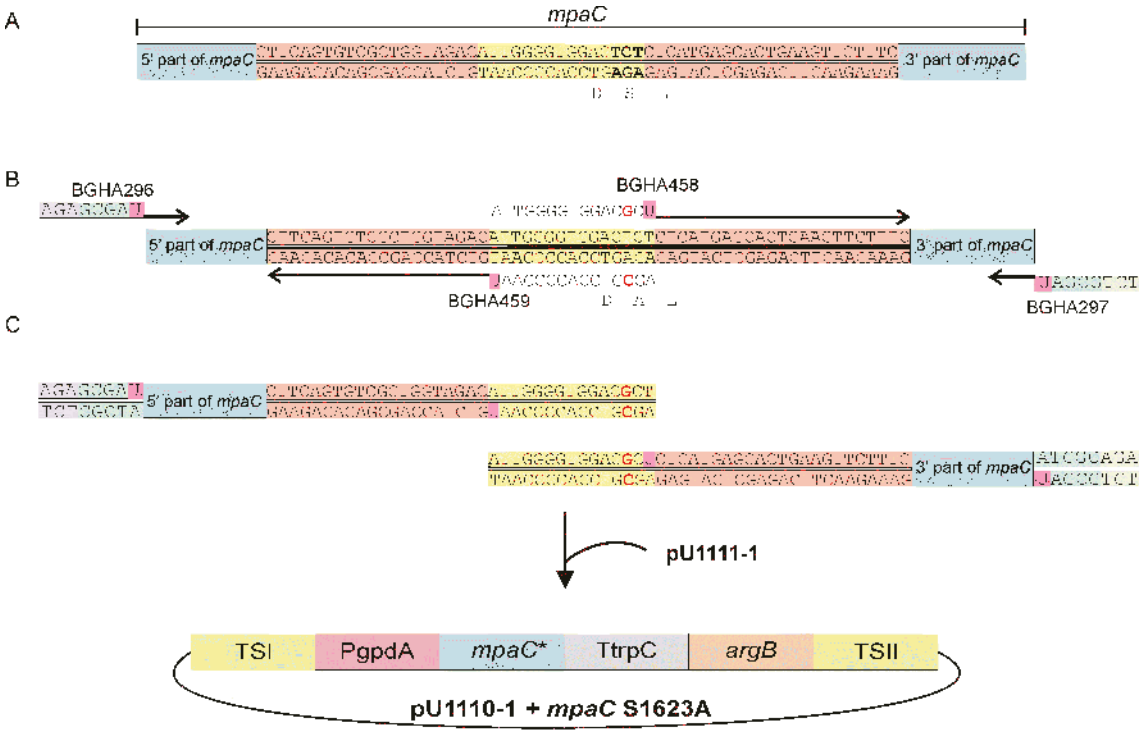

Figure 4

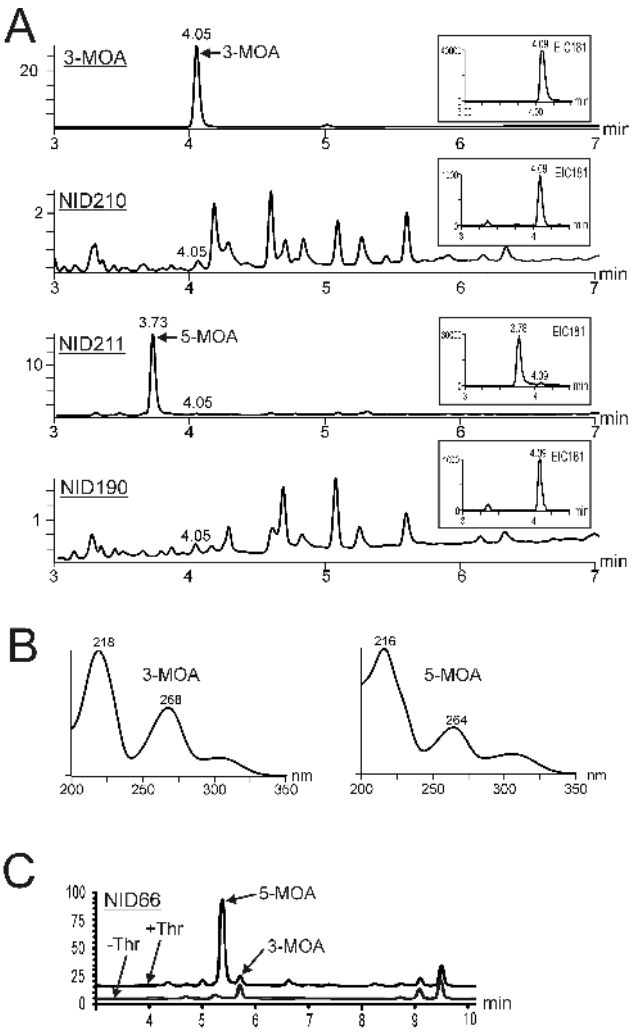

Figure 5