Metabolic Engineering 14 (2012) 104–111

Contents lists available at SciVerse ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/ymben

Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform

Michael Dalgaard Mikkelsen [a,1,2], Line Due Buron [b,1], Bo Salomonsen [a,1], Carl Erik Olsen [c], Bjarne Gram Hansen [b,3], Uffe Hasbro Mortensen [b], Barbara Ann Halkier [a,*]

[a] University of Copenhagen, Faculty of Life Sciences, Department of Plant Biology and Biotechnology, Molecular Plant Biology, VKR Research Centre for Pro-Active Plants, 40 Thorvaldsensvej, DK-1871 Frederiksberg C, Copenhagen, Denmark
[b] Technical University of Denmark (DTU), Department of Systems Biology, Center for Microbial Biotechnology, Søltofts Plads, Building 223, DK-2800 Kgs. Lyngby, Denmark
[c] University of Copenhagen, Faculty of Life Sciences, Department of Basic Sciences and Environment/Bioorganic Chemistry, 40 Thorvaldsensvej, DK-1871 Frederiksberg C, Copenhagen, Denmark

ARTICLE INFO

Article history:
Received 5 October 2011
Received in revised form
17 January 2012
Accepted 26 January 2012
Available online 4 February 2012

Keywords:
Glucosinolates
Yeast expression platform
Pathway engineering
Natural products
Biosynthetic pathway

ABSTRACT

Epidemiological studies have shown that consumption of cruciferous vegetables, such as, broccoli and cabbages, is associated with a reduced risk of developing cancer. This phenomenon has been attributed to specific glucosinolates among the ~30 glucosinolates that are typically present as natural products characteristic of cruciferous plants. Accordingly, there has been a strong interest to produce these compounds in microbial cell factories as it will allow production of selected beneficial glucosinolates. We have developed a versatile platform for stable expression of multi-gene pathways in the yeast, *Saccharomyces cerevisiae*. Introduction of the seven-step pathway of indolylglucosinolate from *Arabidopsis thaliana* to yeast resulted in the first successful production of glucosinolates in a microbial host. The production of indolylglucosinolate was further optimized by substituting supporting endogenous yeast activities with plant-derived enzymes. Production of indolylglucosinolate serves as a *proof-of-concept* for our expression platform, and provides a basis for large-scale microbial production of specific glucosinolates for the benefit of human health.

1. Introduction

Plants are capable of producing a wide variety of secondary metabolites of which many have value as pharma- and nutraceuticals. These compounds are typically produced in small amounts and at specific time points during development in their host plants that in addition may be difficult to propagate in a large scale. In the past two decades it was demonstrated that it is possible to transfer the plant biosynthetic pathways to a microorganism for production, thereby bypassing these limitations. Some noteworthy examples are the antimalarial artemisinic acid (Ro et al., 2006), the antioxidant resveratrol (Becker et al., 2003; Lim et al., 2011; Wang et al., 2011), the flavor component vanillin (Hansen et al., 2009), the universal sesquiterpene precursor, farnesyldiphosphate (Farhi et al., 2011) and taxadiene, precursor for the anticancer agent Taxol (Engels et al., 2008). Such successes have furthered the already high interest of using microbial cell factories as production platforms for high-value plant compounds.

Epidemiological studies have shown that consumption of cruciferous vegetables, such as, broccoli and cabbages, is associated with a number of benefits to human health including prevention of cardiovascular diseases and reduced risk of developing cancer (for review see Hayes et al., 2008; McGrath and Spigelman, 2008; Traka and Mithen, 2009). The health-promoting effects have been attributed to glucosinolates that are secondary metabolites characteristic of cruciferous plants (Halkier and Gershenzon, 2006). The sulfur-rich, amino acid-derived glucosinolates are divided into three groups according to their amino acid precursor: aliphatic, aromatic and indolic glucosinolates. Their biosynthesis proceeds through three independent stages, (i) elongation of side chain (relevant for only methionine and phenylalanine), (ii) formation of core glucosinolate functional group and (iii) secondary modifications of the amino acid side chain. The combination of chain-elongation and side chain modifications is responsible for the high diversity of glucosinolate structures consisting of >120 structures (Fahey et al., 2001).

It is important to stress that the health-promoting effects have been attributed to specific glucosinolates (or rather their hydrolysis

Abbreviations: IG, indol-3-methyl glucosinolate; YPD, rich yeast extract based medium containing glucose; YPgal, rich yeast extract based medium containing galactose; GSH, glutathione; YFG, your favorite gene; 5-FOA, 5-fluoroorotic acid

* Corresponding author. Fax: +45 35333300, +45 4588 4148.
E-mail addresses: um@bio.dtu.dk (U.H. Mortensen), bah@life.ku.dk (B.A. Halkier).

[1] M.D.M., L.D.B. and B.S. contributed equally to this work.
[2] Present address: Evolva A/S, Lersø Park Allé 42-44, DK-2100 Copenhagen Ø, Denmark.
[3] Present address: Novozymes A/S, Krogshøjvej 36, DK-2880, Denmark.

1096-7176/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymben.2012.01.006

products) among the ~30 glucosinolates typically present in a single cruciferous vegetable (Verkerk et al., 2009). As an example, tryptophan-derived indolylglucosinolates (IGs) have received substantial attention due to a potential cancer-preventive role in prostate and breast cancer (Brew et al., 2006; Rahman et al., 2003; Wang et al., 2012), although a role in liver tumorogenesis has also been proposed (Kim et al., 1997; Shimamoto et al., 2011a, 2011b). To enable intake of well-defined amounts of the beneficial glucosinolates, there is a strong interest in developing microbial cell factories to generate a rich, reliable and cost-effective source of specific glucosinolates. Using heterologous expression in tobacco, we have previously succeeded in producing phenylalanine-derived benzylglucosinolates (Geu-Flores et al., 2009) and tryptophan-derived indolylglucosinolates (IGs) (Pfalz et al., 2011) as well as the aliphatic glucosinolate 4-methylsulfinylbutylglucosinolate derived from dihomomethionine, a chain-elongated derivative of methionine (Mikkelsen et al., 2010). In the previous studies, production of the IG (Pfalz et al., 2011) and 4-methylsulfinylbutylglucosinolate (Mikkelsen et al., 2010) in tobacco was obtained by engineering six and 11 biosynthetic genes, respectively. Importantly, through engineering of glucosinolate biosynthetic pathways into tobacco, we identified two metabolic bottlenecks associated with the two S atoms incorporated into the glucosinolate functional group. This includes incorporation of the oxidized sulfur atom that is derived from the sulfotransferase co-substrate PAPS (3′-phosphoadenosine 5′-phosphosulfate) into the sulfate group (Møldrup et al., 2011), and incorporation of the reduced sulfur atom derived from glutathione (GSH) into the thioglucose moiety (Geu-Flores et al., 2009) (Fig. 1). These findings underpin the unique sulfur chemistry of cruciferous plants and provide important clues to the portfolio of plant genes that may be required for transfer of glucosinolate production into a non-plant host. In the tobacco host, glucosinolate biosynthesis was supported by an endogenous NADPH-cytochrome P450-reductase (which donates electrons to cytochromes P450) and an endogenous glutathione-*S*-transferase (conjugating GSH to the product of CYP83s). It remains to be shown to what extent endogenous yeast genes will support the glucosinolate pathway.

The high number of genes required for glucosinolate production excluded plasmid-based expression in yeast as it would require a large number of genetic markers to allow selection for the many plasmids. Such a system would also be prone to plasmid loss and be vulnerable to gene re-arrangement and gene loss due to recombination between vectors. To facilitate construction of a microbial cell factory for glucosinolate production, we established a versatile expression platform that allows stable integration and engineering of complex multi-gene pathways into the well characterized cell factory *Saccharomyces cerevisiae*.

As a *proof-of-concept* of our expression platform, we successfully produced IG by stably and iteratively integrating into the genome of *S. cerevisiae* eight plant genes, including seven biosynthetic genes and one supporting gene required for the catalysis of the seven-step biosynthetic pathway. The ability to produce specific glucosinolates in microbial cell factories has not previously been reported, and has great application potential within both pharmaceutical and nutraceutical industries.

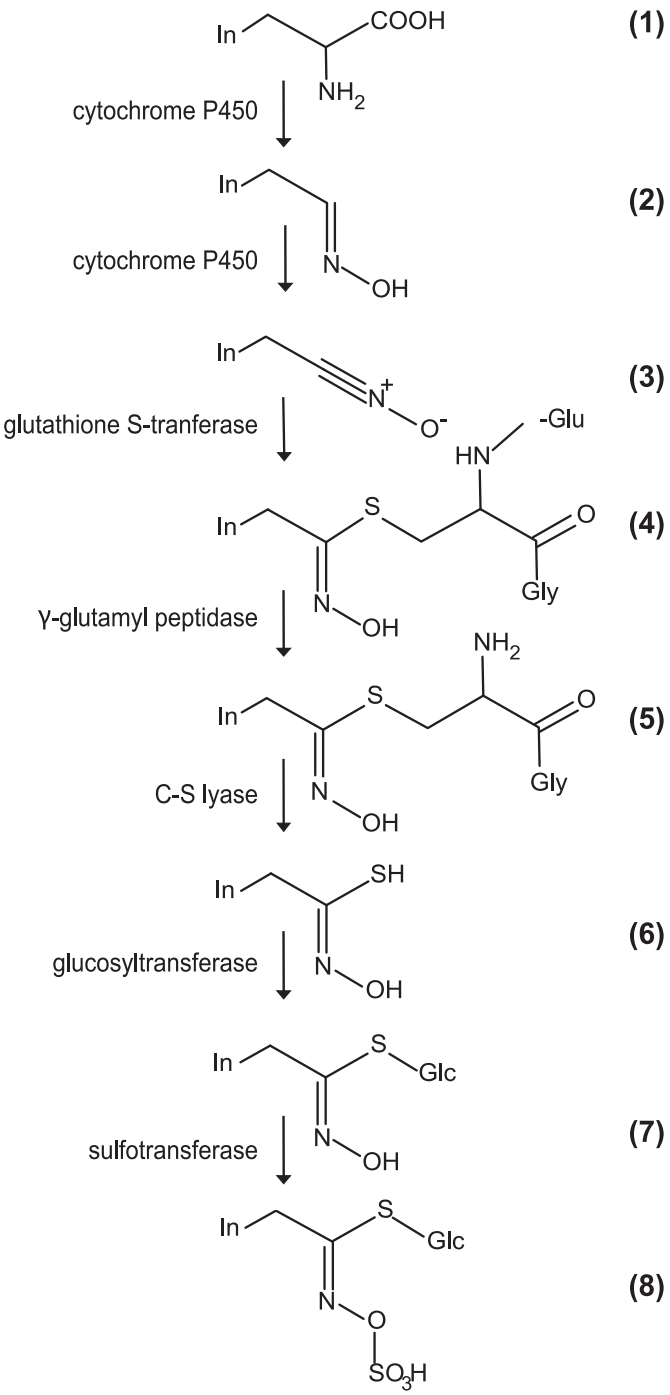

Fig. 1. The biosynthetic pathway of indolylglucosinolate. All genes in the pathway have been identified and characterized in the model plant *Arabidopsis thaliana* (for review see Sønderby et al., 2010), except for the glutathione-*S*-transferase that has only been indicated based on co-expression analysis. The two cytochromes P450 are encoded by *CYP79B2* and *CYP83B1*, and the glutathione-*S*-transferase, γ-glutamyl peptidase, C-S lyase, glucosyltransferase and the sulfotransferase are encoded by *GSTF9*, *GGP1*, *SUR1*, *UGT74B1* and *ST5a*, respectively. Compounds are as follows: (1) tryptophan, (2) indolylacetaldoxime, (3) indolylacetonitrile oxide, (4) S-[(Z)-indolylacetohydroximoyl]-L-glutathione (GSH conjugate), (5) Cys-Gly conjugate, (6) indolylacetothiohydroximic acid, (7) desulfoindolylglucosinolate, and (8) indolylglucosinolate.

2. Materials and methods

2.1. Generation of constructs

Vectors designed for insertion of "your favorite genes" (YFGs) into the selected integration sites were generated in four steps from pU0002 (Hansen et al., 2011) (Fig. S1). All fragments inserted in the vectors were amplified by PCR with either PfuTurbo Cx Hotstart polymerase (Stratagene) or PfuX7 polymerase (Nørholm, 2010) using the primers listed in Table S1. All USER clonings were carried out as previously described (Nour-Eldin et al., 2006) and all plasmids were verified by sequencing (Starseq). In brief, 400–600 bp regions upstream and downstream

of the selected integration sites were amplified by PCR using genomic DNA originating from CEN.PK113-11C as template. USER cloning of the resulting UP and DOWN targeting fragments into pYU-URA3-3 (Fig. S1) generated the vectors X-1 through X-4, XI-1 through XI-5 and XII-1 through XII-5. Vectors harboring the genes of the IG pathway, i.e., *CYP79B2*, *CYP83B1*, *GSTF9*, *SUR1*, *GGP1*, *UGT74B1* and *ST5a*, and the cytochrome P450 electron-donating support gene *ATR1* were amplified by PCR using cDNA of *Arabidopsis thaliana* Col-0 as template. The cDNAs and a bidirectional *GAL1*/*GAL10*-promoter cassette were USER cloned into vectors designed to integrate on chromosome XII. A total of four constructs were produced: Vector XII-1 harboring *CYP79B2* and *CYP83B1*, vector XII-5 harboring *GSTF9* and *ATR1*, vector XII-2 harboring *SUR1* and *GGP1* and vector XII-4 harboring *UGT74B1* and *ST5a* (see Table S2 for a list of all plasmids used or constructed).

2.2. Media

All media for genetic manipulations of yeast were prepared as previously described (Sherman et al., 1986) with the modification that the synthetic medium contained twice the amount of leucine (60 mg/l). All yeast transformants were streak purified on synthetic complete media lacking uracil (Sherman et al., 1986). Subsequently, strains with the *URA3* marker excised through direct repeat recombination were selected on synthetic complete media containing 30 mg/l uracil and 740 mg/l 5-fluoroorotic acid (5-FOA). In shake flask experiments, strains were grown in rich media composed of 10 g/l Bacto yeast extract (BD Bionutrients), 20 g/l Bacto Peptone (BD Bionutrients) and either 20 g/l glucose (media name: YPD) or 20 g/l galactose (media name: YPgal) as the carbon source.

2.3. Strain construction

The genotype and the source of strains are provided in Table S3. CEN.PK113-5D (*MAT*a *MAL2-8^c SUC2 ura3-52*) and CEN.PK113-11C (*MATa MAL2-8C SUC2 his3Δ ura3-52*) were kindly donated by Dr. Peter Kötter, Institut für Mikrobiologie, der Johan Wolfgang Goethe-Universität, Frankfurt am Main, Germany. All constructed strains were made using the lithium acetate/polyethylene glycol/single carrier DNA transformation method (Gietz and Schiestl, 2007) with either a bipartite substrate or full length substrate designed to specifically target the locus of interest.

To construct a yeast strain containing the seven genes of the IG pathway and the supporting *ATR1* gene, the following vectors XII-1, XII-5, XII-2 and XII-4 harboring the following gene pairs (*CYP79B2* and *CYP83B1*), (*SUR1* and *GPP1*), (*UGT74B1* and *ST5b*) and (*GSTF9* and *ATR1*), respectively, were digested with NotI and the individual fragments were transformed iteratively into the yeast strain CEN.PK113-11C in 4 consecutive transformations. The *URA3* marker was eliminated by direct repeat recombination and 5-FOA selection allowing the *URA3* marker to be recycled.

To insert *lacZ* at all integration sites (see Table S4 for information on the chromosomal coordinates of the different integration sites), bipartite substrates designed to target the 14 different integration sites were constructed by fusion PCR (Fig. S2). First, a fragment containing *lacZ* under the control of the *TEF1* promoter was amplified together with 2/3 of the upstream part of the *URA3* marker using pWJ1042-TZC (Flagfeldt et al., 2009) as template. Secondly, the 2/3 downstream part of the *URA3* marker was amplified using the same plasmid as template. Finally, the two *URA3* containing fragments were fused by PCR with UP and DOWN targeting sequences specifically designed to target each of the 14 integration sites. The UP and DOWN fragments were amplified by PCR using CEN.PK113-5D as template and the primers listed in Table S1. The constructed bipartite substrates were transformed into CEN.PK113-5D to yield the strains CEN.bgal.*sitename* (e.g., a strain with *lacZ* integrated at site X-1 is annotated as CEN.bgal.X-1). For all strains correct integration of substrates was verified by colony PCR with one outlying and one substrate-specific primer.

2.4. Shake flask experiments

To determine the growth rate and prepare extracts for metabolite analysis and determination of β-galactosidase activity, strains were grown in 500 ml baffled Erlenmeyer flasks containing either 100 ml glucose-based media (YPD) or 100 ml galactose-based media (YPgal). Flasks were incubated in a 30 °C shaker running at 150 rpm. Growth of the cultures was followed by measuring A_{600} every 1–2 h during the exponential phase.

2.5. β-Galactosidase activity

To determine the β-galactosidase activity of strains with one copy of *lacZ* integrated at the 14 different integration sites tested in this study, yeast cells were grown to A_{600}=1–1.2 in Erlenmeyer flasks containing 100 ml YPgal. Subsequently, β-galactosidase activity was determined according to standard methods. The specific assay used was a modified version of the one published by Miller et al. (1977). For a detailed description of the assay see Fig. S3.

2.6. Metabolite extraction, purification

Total extracts of intracellular yeast metabolites were produced by harvesting cells from a 50 ml culture at OD_{600}=30 by centrifugation. The cells were washed twice in 0.9% (w/v) of NaCl, prior to addition of 500 μl MeOH. This solution was transferred to 2 ml FastPrep tubes containing 0.2 ml of acid-washed glass beads (0.45–0.55 mm). The FastPrep tubes were processed 3 times 20 s in a FastPrep FP120 Instrument (Savant Instruments, New York, USA). The methanol was evaporated and the remaining metabolites resuspended in 100 μl H_2O and analyzed by LC-MS as described previously (Mikkelsen et al., 2010). To detect extracellular metabolites, 1 ml of media was concentrated *in vacuo*, resuspended in 100 μl H_2O and analyzed as described above. Analysis of glucosinolates was performed as previously described (Hansen et al., 2007). In brief, the total methanol yeast extract was applied on a Sephadex DEAE A25 column, washed with 85% methanol, water, desulfonated with a Helix Pomatia sulfatase and eluted with water. The resulting desulfoglucosinolates were analyzed by LC-MS. For media samples, 0.4–1 ml of media was applied directly to the Sephadex column and processed as described above.

3. Results and discussion

3.1. Validation of integration sites for expression of heterologous genes in S. cerevisiae

Establishment of stable, reliable microbial production of glucosinolates requires insertion of several plant genes into the cell factory. To meet this challenge we decided to set up a versatile expression platform that supports stable integration of multiple genes into the *S. cerevisiae* genome at well-defined positions. The platform consists of well-characterized integration sites and matching plasmid-borne tools that allow individual genes to be equipped with yeast promoters and terminators as well as the sequences necessary for specific integration at defined loci using a recyclable marker system (Fig. 2). The integration sites in the platform are grouped in clusters to facilitate that strains containing multi-gene pathways can be constructed and/or improved by

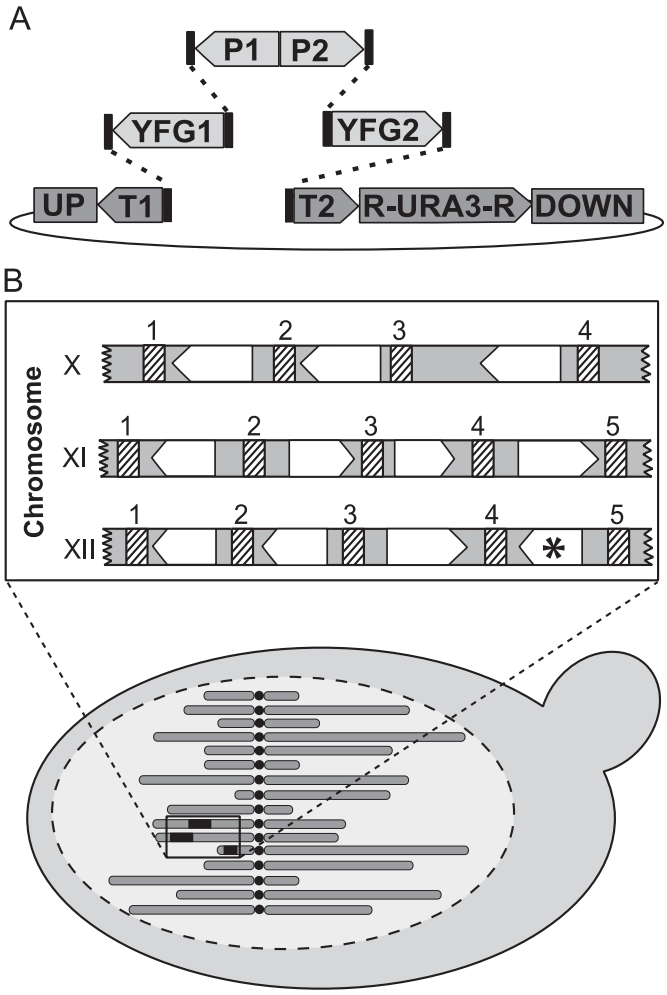

Fig. 2. Overview of the yeast expression platform. (A) Vectors were designed to facilitate stable, pairwise insertion of "your favorite genes" (YFGs) together with promoters (P) and terminators (T) into the genome at selected integration sites. Each vector contains 'UP' stream and 'DOWN' stream yeast genomic sequences of approximately 500 bp designed to target a specific integration site. *ADH1*(T1) and *CYC1*(T2) were used as terminators, and *URA3* of *Kluyveromyces lactis* was used as selection marker. To allow marker recycling *URA3* was flanked by direct repeats (R). Genes of interest along with a bidirectional promoter of choice are to be inserted into vectors through a USER-fusion strategy (Fig. S1). (B) The 14 integration sites were organized in clusters on chromosomes X, XI and XII. All integration sites (hatched boxes) were separated by either genetic elements which are essential for growth (white arrows) or by genes essential for maintaining wild type growth rates (white arrow with star).

strategies that involve sexual crossings. Here we exploit that genes integrated in the same cluster are genetically closely linked, hence, increasing the chance that they all co-segregate during meiosis. This will enable that a cell factory expressing a large multi-gene pathway can be constructed using less consecutive transformations steps by integrating sections of a pathway in parallel in different strains, which can then be combined by crossing into one strain that contains the entire pathway. The design and construction of the platform is described in more detail below.

We investigated 14 locations in the yeast genome for gene integration suitability by testing whether a given site supported high gene expression and accepted fitness-neutral DNA without any significant impact on the growth rate. The selected integration sites are located in regions where gene expression is high as judged by inspecting the yeast GFP fusion localization database (Howson et al., 2005). Secondly, each site is situated centrally in large intergenic regions of minimum 750 bp, to reduce the chance that integration influences the fitness of the strain by adversely affecting neighboring genes. The target sites are clustered at three locations on chromosomes X, XI and XII. The sites were named X-1 through X-4, XI-1 through XI-5, and XII-1 through XII-5 reflecting the chromosomal position and order of the targeting site (Fig. 2). A clustered gene arrangement can potentially introduce genetic instability if the foreign genes are equipped with identical promoters and terminators, which is often the case. Each integration site in a cluster is therefore separated by genetic elements that are essential for sustaining wild type growth rates (e.g., essential genes) to prevent evolution of strains, which have lost genes due to direct repeat recombination (see Fig. S4 for principle of gene loss through direct repeat recombination).

For the platform, we generated 14 plasmids to facilitate construction of the gene targeting substrates directed towards the 14 integration sites. Each vector contained the necessary sequences to target a specific site as well as the selectable/counter selectable *Kluyveromyces lactis*-derived *URA3* marker gene, which was flanked by directly repeated sequences to allow for iterative gene targeting cycles (Siewers et al., 2010). In addition, each vector contained a USER cloning cassette (Nour-Eldin et al., 2006) flanked by two commonly used *S. cerevisiae* terminators derived from *ADH1* and *CYC1* (see Figs. 2 and S2 for vector layout). By applying the efficient USER cloning technique (Nour-Eldin et al., 2010), one or two genes of interest and a fragment containing a uni- or bidirectional promoter of choice can be simultaneously integrated into the vectors by USER-fusion cloning (Geu-Flores et al., 2007).

Validation of the level of expression at the different integration sites in the genome was performed by determination of the relative *lacZ* gene expression levels at each site. Specifically, a gene targeting cassette containing the *E. coli lacZ* gene under the control of the strong constitutive *TEF1* promoter was integrated into each of the 14 sites. The successful integrations resulted in 14 strains that were named CEN.bgal followed by the designated site name (e.g., a strain with *lacZ* integrated at site X-1 is annotated as CEN.bgal.X-1). After growth on YPgal medium, β-galactosidase activity was determined for all 14 strains expressing *lacZ* (Table S5). We found that all sites supported high *lacZ* expression from the *TEF1* promoter and only small expression differences were observed. The average specific β-galactosidase activity for all sites was 1168 ± 293 nmol/min/mg protein and the difference between the strain with the highest (CEN.bgal.XI-5) and the lowest activity level (CEN.bgal.XII-1) was 2.5 fold. Only four strains displayed activity levels that were significantly different from the average expression level of the combined set of strains, one with higher (CEN.bgal.XI-5) and three with lower activity (CEN.bgal.X-3, CEN.bgal.XII-1 and CEN.bgal.XII-4) (Table S5). The effect of the individual integrations on fitness was evaluated by measuring growth rates in rich media containing either glucose (YPD) or galactose (YPgal) as carbon sources. 11 out of the 14 strains grew at a rate comparable to the reference strain or slightly faster (Table S5). Of the remaining three strains, one, CEN.bgal.X-1, grew to $58 \pm 2\%$ of the reference strain in YPD, one, CEN.bgal-XI-5, grew to $81 \pm 3\%$ of reference strain in YPgal medium, and one, CEN.bgal.XI-4, grew to, respectively, $46 \pm 11\%$ in YDP and $52 \pm 3\%$ YPgal medium (Table S5). In summary, 11 of the 14 integration sites (three on chromosomes X and XI, and five on chromosome XII) were suitable for integration and expression of heterologous genes. This enables engineering of biosynthetic pathways depending on up to 22 genes.

3.2. Pathway engineering

Towards our goal of producing glucosinolates in yeast and as a *proof-of-concept* of our technology platform, we set out to engineer IG production in yeast. The biosynthetic pathway of IG

consists of seven biosynthetic genes, i.e., two cytochromes P450 *CYP79B2*, *CYP83B1*, a glutathione-*S*-transferase *GSTF9*, a γ-glutamyl peptidase *GGP1*, a C-S lyase *SUR1*, a glucosyltransferase *UGT74B1* and a sulfotransferase *ST5a* (Sønderby et al., 2010) (Fig. 1). We engineered the pathway by building it up from the entry point by pairwise and stepwise integration of two genes into the genome, as this allowed us to test for functional expression of the inserted genes directly on the yeast extract as the first committed step catalyzed by CYP79B2 uses tryptophan as substrate. In subsequent introductions of genes, functionality was followed by either accumulation or reduction of the respective intermediate pools.

First, a strain named I(A) expressing gene pair A, i.e. *CYP79B2* and *CYP83B1*, under the control of the strong bidirectional *GAL1/GAL10* promoter cassette was constructed. We chose to drive expression of all genes by this inducible promoter pair, which are almost completely repressed by glucose (Johnston et al., 1994). Accordingly, strain construction and maintenance could be performed in the repressive state minimizing the risk that undesired mutations develop in strains where the inserted genes result in production of toxic intermediates or products. When extracts of the intermediary strain I(A) grown under inducing conditions were analyzed by LC-MS, a major peak at 7 min appeared (Fig. 3A). The *m/z* ratio of this compound was 480 $[M+Na]^+$, corresponding to the expected mass of the tryptophan-derived GSH conjugate intermediate (compound 4 in Fig. 1) of the IG pathway (Geu-Flores et al., 2011) (Fig. 4). This was further supported by the MS2 spectrum, which showed a dominant ion of *m/z* ratio 318, a characteristic fragment ion of this GSH conjugate (Pfalz et al., 2011) (Fig. 3B). A minor peak with *m/z* and a fragmentation pattern identical to the dominant peak was observed. This is likely to be the stereoisomer of the GSH conjugate intermediate. The data showed that the two cytochromes P450 were functionally expressed, and that their catalytic activity was supported by endogenous yeast NADPH-cytochrome P450 reductase. Furthermore, it showed that an endogenous yeast glutathione-*S*-transferase was able to conjugate GSH to the product of *CYP83B1*. In the natural host *A. thaliana*, the NADPH-cytochrome P450 reductase encoded by *ATR1* supports *CYP79B2* and *CYP83B1* (Urban et al., 1997), and the glutathione-*S*-transferase, *GSTF9*, is the most likely glutathione-*S*-transferase candidate as evidenced by co-expression analysis. Although the two plant genes, *ATR1* and *GSTF9* were not required for production of the GSH conjugate in *S. cerevisiae*, their expression could potentially result in more efficient IG production compared to when only the corresponding native yeast enzymes were present. In an effort to optimize the production of the GSH conjugate of the IG pathway, we constructed strains that in addition to *CYP79B2* and *CYP83B1* also expressed *ATR1* (I(A)+ATR1), *GSTF9* (I(A)+GSTF9), or both from the *GAL1/GAL10* promoter cassette. Analysis of the I(A) strain co-expressing *ATR1*, *GSTF9* or both resulted in, respectively, 1.29 ± 0.07-fold, 1.25 ± 0.08-fold and 2.62 ± 0.16-fold increases in levels of the GSH conjugate intermediate when intracellular extracts were prepared from cells harvested in the late exponential phase, i.e., after 24 h (Fig. 3C). This demonstrated that the plant-derived enzymes *ATR1* and *GSTF9* were more efficient in supporting the pathway than the

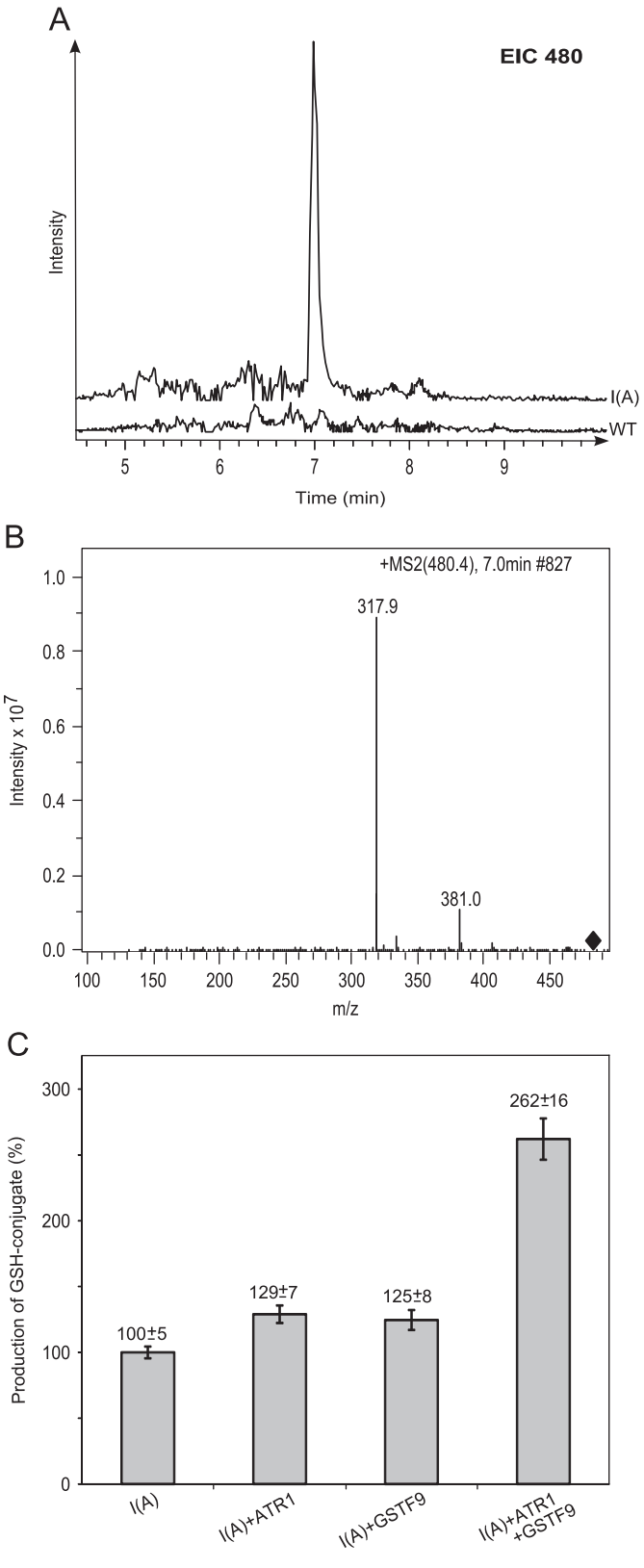

Fig. 3. Comparison of the activity of NADPH cytochrome P450 reductase and glutathione-*S*-transferase from yeast and plant. (A) Extracted ion chromatogram (EIC) (*m/z* 480) corresponding to $[M+H^+]$ of the stable GSH conjugate on intracellular extracts of wild type yeast (WT) and the strain I(A) encompassing the two cytochromes P450 responsible for the first two steps of indolylglucosinolate biosynthesis. (B) MS2 spectrum of the peak detected in extracts from I(A) strain revealed a dominant ion of *m/z* ratio 318, corresponding to a fragment ion of the cyclized glutathione conjugate intermediate (for details see Geu-Flores et al., 2009). (C) Accumulation of GSH conjugate in extracts from yeast strains expressing the two cytochromes P450 alone or together with either plant *GSTF9*, *ATR1*, or both, and grown for 24 h in rich media under inducing conditions. The GSH conjugate intermediate was extracted from yeast pellet and detected by LC-MS. Each column represents production of GSH conjugate relative to the I(A) strain, which is normalized to 100%. Values are based on product peak areas from three biological replicates each harvested in four technical replicates, error bars represent standard error (SEM, $N=12$).

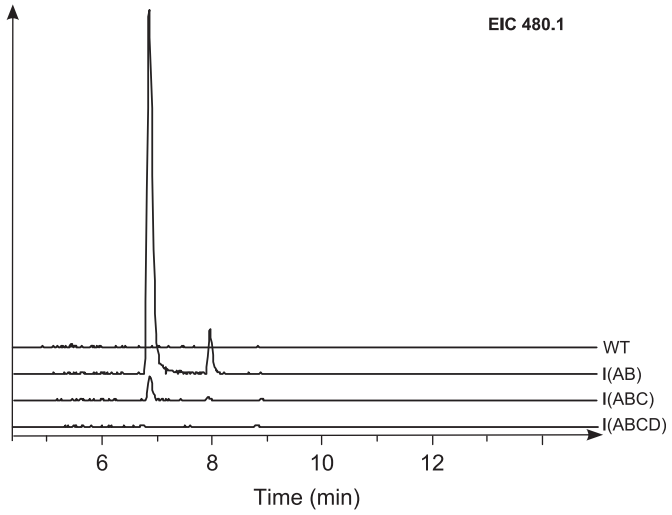

Fig. 4. Accumulation of the tryptophan-derived glutathione-conjugate intermediate in different engineered yeast strains. Extracted ion chromatogram (EIC) of extracts from the four yeast strains (WT, I(AB), I(ABC), and I(ABCD)) grown in rich media under inducible conditions. *m/z* 480 corresponds to the $[M+H]^+$ ion of the glutathione conjugate intermediate.

corresponding endogenous yeast enzymes that possess the same enzymatic activities. Accordingly, we used the I(AB) strain harboring gene pair A (*CYP79B2*, *CYP83B1*), and gene pair B (*GSTF9* and *ATR1*) as the basis for the continued pathway engineering.

The next two genes in the pathway, i.e. *GGP1* and *SUR1* (gene pair C), were integrated in site XII-2, resulting in generation of the intermediary strain I(ABC). Assuming functional expression of both *GGP1* and *SUR1*, we would expect that the GSH conjugate intermediate was converted to the corresponding thiohydroximic acid (Pfalz et al., 2011), which notoriously is hard to detect (Walter and Schaumann, 1971). Upon LC-MS analysis of methanolic extracts from I(ABC) grown under inducible conditions, only trace amounts of the tryptophan-derived GSH conjugate were detected (Fig. 4), which implied that insertion of at least one of these genes resulted in metabolism of the GSH conjugate. If GGP1 was functionally expressed, it would metabolize this intermediate to produce the corresponding Cys-Gly conjugate intermediate. In the absence of a functional SUR1 such a product would spontaneously cyclize (Geu-Flores et al., 2009). However, no such intermediate was observed when we inspected ion traces of intracellular extracts of the strain or medium samples. This suggested that both GGP1 and SUR1 could be functionally expressed and that the corresponding thiohydroximic acid had evaded detection. As a final step towards engineering the IG pathway into yeast, we integrated *UGT74B1* and *ST5a* (gene pair D) into site XII-4, resulting in generation of the strain harboring the entire pathway I(ABCD). LC-MS analysis of methanolic extracts of the I(ABCD) strain grown under inducing conditions showed that the strain was able to produce IG as demonstrated by two lines of evidence (Fig. 5). First, a peak co-eluting with the authentic IG standard contained a dominant ion mass of *m/z* 391 $[M+Na^+]$, corresponding to the sodium adduct of desulfoIG (glucosinolates are determined as desulfoglucosinolates after treatment with sulfatase, see Section 2) (Fig. 5A and B). Secondly, *m/z* 185 and 219, that are characteristic of glucosinolates and representing the glucose and thioglucose fragment ions (Geu-Flores et al., 2009), were observed on MS2 spectra (Fig. 5C). Time-course experiments showed that detectable amounts of IG accumulated after approximately 12 h (Fig. 6). In the stationary phase that was reached after 32 h, a maximum concentration of 1.07 ± 0.38 mg IG/l was obtained. This value

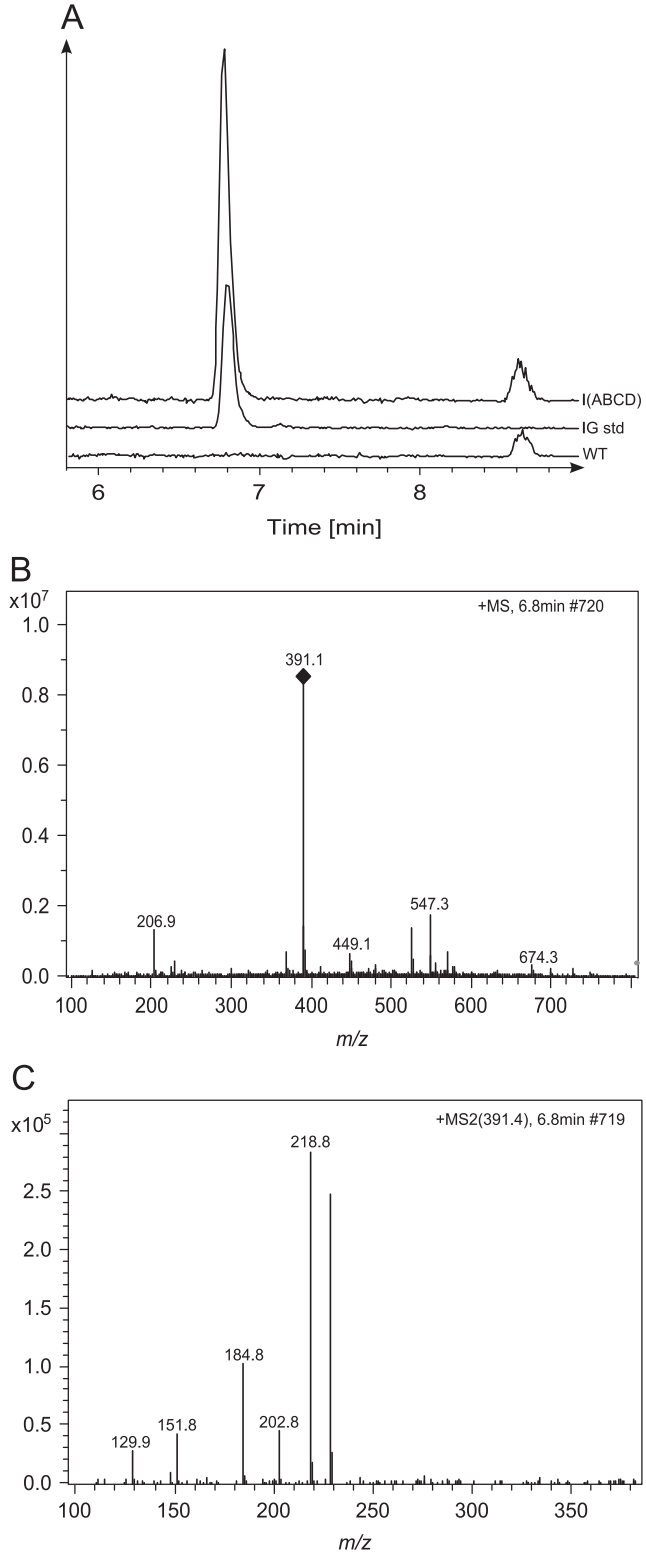

Fig. 5. Production of indolylglucosinolate in *Saccharomyces cerevisiae*. (A) LC-MS analysis of medium samples from cultures of the I(ABCD) strain. The peak at 6.9 min co-migrated with an authentic indolylglucosinolate standard. (B) The dominant ion with an *m/z* ratio of 391.1 corresponded to desulfoindolylglucosinolate (as glucosinolates are treated with sulfatase prior to LC-MS analysis, see Section 2). (C) The main fragment ions of *m/z* 185 and 219 correspond to the glucose and thioglucose moiety, characteristic of desulfoglucosinolates.

remained constant for at least 16 h. The produced IG was recovered from the medium and only minute quantities were found in the intracellular fraction. This suggests that yeast has

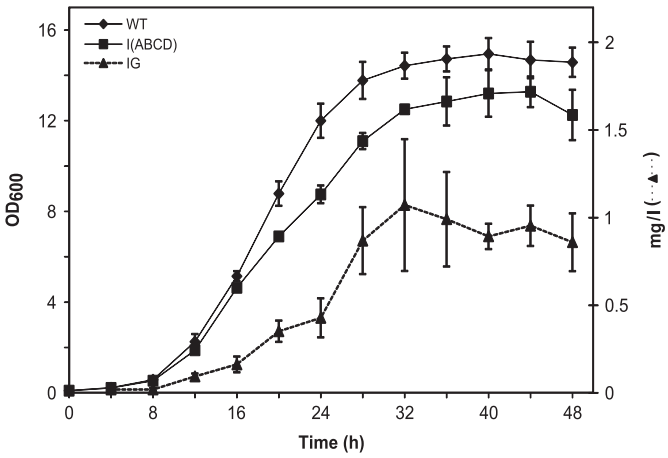

Fig. 6. Time-course experiment of indolylglucosinolate accumulation. Growth of WT(◆) and the indolylglucosinolate-producing I(ABCD)(■) from shake flask cultures grown in rich inducible medium was compared. Samples from four biological replicates were taken every 4 h for a 48 h period for OD_{600} measurements and indolylglucosinolate quantification (IG(Δ)) by LC-MS. The error bars indicate standard deviation (SD, $N=4$).

secreted the compound into the medium, a phenomenon that has previously been observed when yeast was engineered to produce bioactive plant compounds (Asadollahi et al., 2008; Hansen et al., 2009; Ro et al., 2006). No intermediates or other compounds were detected, neither in intracellular extracts of the IG-producing strain nor in the medium (Figs. 4 and 5). In the exponential phase, there was no significant difference in growth rates between I(ABCD) and wild-type strain ($P=0.052$) (Fig. 6). However, the final biomass accumulation was lower for the I(ABCD) strain, which implies that production of IG affects strain fitness.

4. Concluding remarks

In summary, we have established a technology platform designed for transfer of multi-gene pathways (up to 22 genes) into *S. cerevisiae*. The platform is robust and ensures stable expression, as all genes are genomically integrated in sites separated by genes essential for sustaining normal growth. All sites have been validated by analysis of growth rates and *lacZ* expression. The platform includes 11 vectors designed for integration into the validated sites in the *S. cerevisiae* genome. Hence, our platform supports production of compounds that require the expression of up to 22 genes. In our system, all markers in the host strain can be eliminated by further transformation or by sexual crossing. This allows for production in both minimal and rich media. Such large pathways will be difficult to establish using the otherwise popular plasmid-borne expression systems (De Jong et al., 2006; Yan et al., 2005) as the number of genes in the pathway will be restricted by the number of available markers in the host strain. Moreover, the Yan et al. (2005) stability of such expression systems will likely decrease as the number of plasmids is increased. Integrated multi-gene pathways have also been successfully puzzled together *in vivo* by homologous recombination (Shao et al., 2009). However, using this system it will be difficult to integrate fragments that contain repetitive elements, e.g., repetitive use of promoters or terminators, as these sequences will also have the potential to recombine during integration to complicate strain construction. Our system has a built-in safety system that prevents strains to propagate if they lose genes through direct repeat recombination. This allows expression of multiple copies of the same gene and/or use of identical promoters or terminators for the expression of several genes without compromising strain stability.

As a *proof-of-concept* for our technology platform, the ability to support expression of complex, multi-gene pathways was demonstrated with IG production. Further optimization is required to reach commercially viable levels. Transcript profiling and analysis of yeast fitness may provide important clues for optimization of the production. Our *S. cerevisiae* strain capable of synthesizing IG represents the first example of microbial production of glucosinolates. In addition to providing a stable, reliable source of glucosinolates, engineering of the specific biosynthetic pathway of a given glucosinolate into a microorganism enables purification of that glucosinolate without contamination from the other >30 glucosinolates typically present in cruciferous vegetables. Accordingly, production of glucosinolates from a microbial source has great potential for industrial-scale production of specific, health-promoting glucosinolates for the benefit of human health.

Acknowledgments

We thank the Danish Research Council for Technology and Production for financial support (J.no. 274-08-0221). We thank Tomas Strucko for inspiration to Fig. 2 and Jette Mortensen for technical assistance.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ymben.2012.01.006.

References

Asadollahi, M.A., Maury, J., Møller, K., Nielsen, K.F., Schalk, M., Clark, A., Nielsen, J., 2008. Production of plant sesquiterpenes in *Saccharomyces cerevisiae*: effect of *ERG9* repression on sesquiterpene biosynthesis. Biotechnol. Bioeng. 99, 666–677.

Becker, J.V.W., Armstrong, G., Merwe, M., Lambrechts, M., Vivier, M., Pretorius, I., 2003. Metabolic engineering of *Saccharomyces cerevisiae* for the synthesis of the wine-related antioxidant resveratrol. FEMS Yeast Res. 4, 79–85.

Brew, C., Aronchik, I., Hsu, J., Sheen, J., Dickson, R., Bjeldanes, L., Firestone, G., 2006. Indole-3-carbinol activates the ATM signaling pathway independent of DNA damage to stabilize p53 and induce G1 arrest of human mammary epithelial cells. Int. J. Cancer 118, 857–868.

De Jong, J., Liu, Y., Bollon, A., Long, R., Jennewein, S., Williams, D., Croteau, R., 2006. Genetic engineering of taxol biosynthetic genes in *Saccharomyces cerevisiae*. Biotechnol. Bioeng. 93.

Engels, B., Dahm, P., Jennewein, S., 2008. Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards Taxol (Paclitaxel) production. Metab. Eng. 10, 201–206.

Fahey, J., Zalcmann, A., Talalay, P., 2001. The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. Phytochemistry 56, 5–51.

Farhi, M., Marhevka, E., Masci, T., Marcos, E., Eyal, Y., Ovadis, M., Abeliovich, H., Vainstein, A., 2011. Harnessing yeast subcellular compartments for the production of plant terpenoids. Metab. Eng. 13, 474–481.

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

Geu-Flores, F., Møldrup, M., Böttcher, C., Olsen, C., Scheel, D., Halkier, B., 2011. Cytosolic gamma-glutamyl peptidases process glutathione conjugates in the biosynthesis of glucosinolates and camalexin in *Arabidopsis*. Plant Cell 23, 2456–2469.

Geu-Flores, F., Nielsen, M., Nafisi, M., Møldrup, M., Olsen, C., Motawia, M., Halkier, B., 2009. Glucosinolate engineering identifies a gamma-glutamyl peptidase. Nat. Chem. Biol. 5, 575–577.

Geu-Flores, F., Nour-Eldin, H., Nielsen, M., Halkier, B., 2007. USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. Nucleic Acids Res. 35, e55. doi:10.1093/nar/gkm106.

Gietz, R., Schiestl, R., 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat. Protoc. 2, 31–34.

Halkier, B., Gershenzon, J., 2006. Biology and biochemistry of glucosinolates. Annu. Rev. Plant Biol. 57, 303–333.

Hansen, B., Kliebenstein, D., Halkier, B., 2007. Identification of a flavin-monooxygenase as the S-oxygenating enzyme in aliphatic glucosinolate biosynthesis in *Arabidopsis*. Plant J. 50, 902–910.

Hansen, B., Salomonsen, B., Nielsen, M., Nielsen, J., Hansen, N., Nielsen, K., Regueira, T., Nielsen, J., Patil, K., Mortensen, U., 2011. Versatile enzyme expression and characterization system for *Aspergillus nidulans*, with the *Penicillium brevicompactum* polyketide synthase gene from the mycophenolic acid gene cluster as a test case. Appl. Environ. Microbiol. 77, 3044–3051.

Hansen, E.H., Moller, B.L., Kock, G.R., Bunner, C.M., Kristensen, C., Jensen, O.R., Okkels, F.T., Olsen, C.E., Motawia, M.S., Hansen, J., 2009. De novo biosynthesis of vanillin in fission yeast (*Schizosaccharomyces pombe*) and baker's yeast (*Saccharomyces cerevisiae*). Appl. Environ. Microbiol. 75, 2765–2774.

Hayes, J., Kelleher, M., Eggleston, I., 2008. The cancer chemopreventive actions of phytochemicals derived from glucosinolates. Eur. J. Nutr. 47, 73–88.

Howson, R., Huh, W., Ghaemmaghami, S., Falvo, J., Bower, K., Belle, A., Dephoure, N., Wykoff, D., Weissman, J., O'Shea, E., 2005. Construction, verification and experimental use of two epitope-tagged collections of budding yeast strains. Comp. Funct. Genom. 6, 2–16.

Johnston, M., Flick, J., Pexton, T., 1994. Multiple mechanisms provide rapid and stringent glucose repression of *GAL* gene expression in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 14, 3834–3841.

Kim, D., Han, B., Ahn, B., Hasegawa, R., Shirai, T., Ito, N., Tsuda, H., 1997. Enhancement by indole-3-carbinol of liver and thyroid gland neoplastic development in a rat medium-term multiorgan carcinogenesis model. Carcinogenesis 18, 377–381.

Lim, C.G., Fowler, Z., Hueller, T., Schaffer, S., Koffas, M., 2011. High-yield resveratrol production in engineered *Escherichia coli*. Appl. Environ. Microbiol. 77, 3451–3460.

McGrath, D., Spigelman, A., 2008. Putative mechanisms of action for indole-3-carbinol in the prevention of colorectal cancer. Expert Opin. Ther. Targets 12, 729–738.

Mikkelsen, M., Olsen, C., Halkier, B., 2010. Production of the cancer-preventive glucoraphanin in tobacco. Mol. Plant 3, 751–759.

Miller, J.H., Platt, T., Müller-Hill, B., 1977. Experiments in molecular genetics. In: Assay of Betagalactosidase. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (Chapter VII).

Møldrup, M., Geu-Flores, F., Olsen, C., Halkier, B., 2011. Modulation of sulfur metabolism enables efficient glucosinolate engineering. BMC Biotechnol. 11, 12.

Nour-Eldin, H., Geu-Flores, F., Halkier, B., 2010. USER cloning and USER fusion: the ideal cloning techniques for small and big laboratories. Methods Mol. Biol. 643, 185–200.

Nour-Eldin, H., Hansen, B., Nørholm, M., Jensen, J., Halkier, B., 2006. Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. Nucleic Acids Res. 34, E122.

Nørholm, M.H.H., 2010. A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. BMC Biotechnol. 10, 21.

Pfalz, M., Mikkelsen, M., Bednarek, P., Olsen, C., Halkier, B., Kroymann, J., 2011. Metabolic engineering in *Nicotiana benthamiana* reveals key enzyme functions in *Arabidopsis* indole glucosinolate modification. Plant Cell 23, 716–729.

Rahman, K., Aranha, O., Sarkar, F., 2003. Indole-3-carbinol (I3C) induces apoptosis in tumorigenic but not in nontumorigenic breast epithelial cells. Nutr. Cancer 45, 101–112.

Ro, D.K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., Ho, K.A., Eachus, R.A., Ham, T.S., Kirby, J., Chang, M.C.Y., Withers, S., Shiba, Y., Sarpong, R., Keasling, J.D., 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 440, 940–943.

Shao, Z., Zhao, H., Zhao, H., 2009. DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. Nucleic Acids Res. 37, e16 (10 pp.).

Sherman, F., Fink, G.R., Hicks, J.B., 1986. Laboratory Course Manual for Methods in Yeast Genetics. Cold Spring Harbor Laboratory, New York.

Shimamoto, K., Dewa, Y., Ishii, Y., Kemmochi, S., Taniai, E., Hayashi, H., Imaoka, M., Morita, R., Kuwata, K., Suzuki, K., Shibutania, M., Mitsumori, K., 2011a. Indole-3-carbinol enhances oxidative stress responses resulting in the induction of preneoplastic liver cell lesions in partially hepatectomized rats initiated with diethylnitrosamine. Toxicology 283, 109–117.

Shimamoto, K., Dewa, Y., Kemmochi, S., Taniai, E., Hayashi, H., Imaoka, M., Shibutani, M., Mitsumori, K., 2011b. Relationship between CYP1A induction by indole-3-carbinol or flutamide and liver tumor-promoting potential in rats. Arch. Toxicol. 85, 1159–1166.

Siewers, V., Mortensen, U., Nielsen, J., 2010. Genetic engineering tools for *Saccharomyces cerevisiae*. In: Baltz, R., Davies, J.E., Demain, A.L. (Eds.), Manual of Industrial Microbiology and Biotechnology. ASM Press, Washington, DC, USA, pp. 287–301.

Sønderby, I., Geu-Flores, F., Halkier, B., 2010. Biosynthesis of glucosinolates—gene discovery and beyond. Trends Plant Sci. 15, 283–290.

Traka, M., Mithen, R., 2009. Glucosinolates, isothiocyanates and human health. Phytochem. Rev. 8, 269–282.

Urban, P., Mignotte, C., Kazmaier, M., Delorme, F., Pompon, D., 1997. Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 CYP73A5. J. Biol. Chem. 272, 19176–19186.

Verkerk, R., Schreiner, M., Krumbein, A., Ciska, E., Holst, B., Rowland, I., De Schrijver, R., Hansen, M., Gerhauser, C., Mithen, R., Dekker, M., 2009. Glucosinolates in Brassica vegetables: the influence of the food supply chain on intake, bioavailability and human health. Mol. Nutr. Food Res. 53, S219–S265.

Walter, W., Schaumann, E., 1971. The chemistry of thiohydroxamic acids. Int. J. Methods Synth. Org. Chem., 111–130.

Wang, T., Schoene, N., Milner, J., Kim, Y., 2012. Broccoli-derived phytochemicals indole-3-carbinol and 3,3′-diindolylmethane exerts concentration-dependent pleiotropic effects on prostate cancer cells: comparison with other cancer preventive phytochemicals. Mol. Carcinog. 51, 244–256.

Wang, Y., Halls, C., Zhang, J., Matsuno, M., Zhang, Y., Yu, O., 2011. Stepwise increase of resveratrol biosynthesis in yeast *Saccharomyces cerevisiae* by metabolic engineering. Metab. Eng. 11, 310–318.

Yan, Y., Kohli, A., Koffas, M., 2005. Biosynthesis of natural flavanones in *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. 71, 5610–5613.