Quantitative and qualitative transcriptome analysis of four industrial strains of Claviceps purpurea with respect to ergot alkaloid production

ARTICLE  in  NEW BIOTECHNOLOGY · JANUARY 2016
Impact Factor: 2.9 · DOI: 10.1016/j.nbt.2016.01.006

4 AUTHORS, INCLUDING:

Maria Majeska Cudejkova
Palacký University of Olomouc
9 PUBLICATIONS  24 CITATIONS

Josef Valík
Teva Pharmaceutical Industries Ltd.
5 PUBLICATIONS  9 CITATIONS

All in-text references underlined in blue are linked to publications on ResearchGate, letting you access and read them immediately.
Quantitative and qualitative transcriptome analysis of four industrial strains of Claviceps purpurea with respect to ergot alkaloid production

Mária Majeská Čudejková¹, Petr Vojta¹,², Josef Valík³ and Petr Galuszka¹

¹ Department of Molecular Biology, Centre of the Region Haná for Biotechnological and Agricultural Research, Faculty of Science, Palacký University in Olomouc, Šlechtitelů 27, 783 71 Olomouc, Czech Republic
² Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacký University Olomouc, Hněvotínská 1333/5, 779 00 Olomouc, Czech Republic
³ Teva Czech Industries s.r.o., Ostravská 305/29, 747 70 Opava-Komárov, Czech Republic

The fungus Claviceps purpurea is a biotrophic phytopathogen widely used in the pharmaceutical industry for its ability to produce ergot alkaloids (EAs). The fungus attacks unfertilized ovaries of grasses and forms sclerotia, which represent the only type of tissue where the synthesis of EAs occurs. The biosynthetic pathway of EAs has been extensively studied; however, little is known concerning its regulation. Here, we present the quantitative transcriptome analysis of the sclerotial and mycelial tissues providing a comprehensive view of transcriptional differences between the tissues that produce EAs and those that do not produce EAs and the pathogenic and non-pathogenic lifestyle. The results indicate metabolic changes coupled with sclerotial differentiation, which are likely needed as initiation factors for EA biosynthesis. One of the promising factors seems to be oxidative stress. Here, we focus on the identification of putative transcription factors and regulators involved in sclerotial differentiation, which might be involved in EA biosynthesis. To shed more light on the regulation of EA composition, whole transcriptome analysis of four industrial strains differing in their alkaloid spectra was performed. The results support the hypothesis proposing the composition of the amino acid pool in sclerotia to be an important factor regulating the final structure of the ergopeptines produced by Claviceps purpurea.

Introduction

The ascomycete fungus Claviceps purpurea (Cp) is a biotrophic plant pathogen that produces toxic alkaloids causing poisoning referred to as ergotism. Ergot alkaloids (EAs) are pharmacologically important mycotoxins that have a wide range of therapeutic applications, for example, anti-migraine drugs (ergotamine), anti-Parkinson’s agents (bromocriptine – a semisynthetic derivative of ergocryptine), and symptomatic therapy of dementia (ergotoxine – a mixture of ergocryptine, ergocristine, and ergocornine) etc. [1]. The fungus attacks the ovaries of young grasses exclusively, follows the route of the pollen tube during pollination, and forms blackened sclerotia that are substituted for grains in the floret [2]. Under natural conditions, EAs are produced only in the sclerotia that are formed after successful infection and not in axenic culture. However, several mutated strains that produce alkaloids in culture were described [3]. In both conditions, synthesis of EAs is correlated with two processes: cell differentiation into sclerotal cells and lipid synthesis. Sclerotinized cells accumulate lipids up to 50% of their weight, of which the principal fatty acid is ricinoleate [4, 5]. Presence of ricinoleate is specific to sclerotinized cells, and therefore it is used as a marker for ergot impurities in rye and rye products [6, 7]. The factor triggering the initiation of sclerotia development is unknown; however, morphological and metabolic changes linked with developmental transition from sphecidal to sclerotal phase were studied in detail [8–10]. When the cells accomplish the transition to the sclerotal phase, the synthesis of EAs begins.

EAs are nitrogenous secondary metabolites derived from L-tryptophan. They can be divided into three major structural groups:
TABLE 1

Composition of amino acids bound to the d-lysergic acid and resulting ergopeptines. At the third position is proline. Adopted from Haarmann et al. [11]

<table>
<thead>
<tr>
<th>Amino acid at the position I</th>
<th>Amino acid at the position II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phenylalanine</td>
<td>Leucine</td>
</tr>
<tr>
<td>Alanine</td>
<td>Ergotamin</td>
</tr>
<tr>
<td>Valine</td>
<td>α-Ergosine</td>
</tr>
<tr>
<td>α-Amino butyric acid</td>
<td>Ergostine</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Amino acid at the position I</th>
<th>Amino acid at the position II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isoleucine</td>
<td>β-Ergosine</td>
</tr>
<tr>
<td>Valine</td>
<td>Ergovamine</td>
</tr>
<tr>
<td>α-Amino butyric acid</td>
<td>Ergocornine</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Amino acid at the position I</th>
<th>Amino acid at the position II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valine</td>
<td>Ergocristin</td>
</tr>
<tr>
<td>α-Amino butyric acid</td>
<td>Ergocystine</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Amino acid at the position I</th>
<th>Amino acid at the position II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valine</td>
<td>β-Ergocryptine</td>
</tr>
<tr>
<td>α-Amino butyric acid</td>
<td>Ergocristine</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Amino acid at the position I</th>
<th>Amino acid at the position II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valine</td>
<td>Ergocornine</td>
</tr>
<tr>
<td>α-Amino butyric acid</td>
<td>Ergobutidine</td>
</tr>
</tbody>
</table>

TABLE 2


<table>
<thead>
<tr>
<th>Working name of strain</th>
<th>Strain ID</th>
<th>Patent</th>
<th>Average content of EA in sclerotia (% of dry mass)</th>
<th>Limit contents of EA in sclerotia (% of dry mass)</th>
<th>Spectrum of EA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cp20.1</td>
<td>Reference [27]</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>TAM – main alkaloid, KRY – minor alkaloid</td>
</tr>
<tr>
<td>Gal012</td>
<td>CCM8178</td>
<td>CZ279533</td>
<td>1.3%</td>
<td>0.78–1.7%</td>
<td>90% α-KRY, 9% MET, 1% valinamid</td>
</tr>
<tr>
<td>Gal130</td>
<td>CCM8058</td>
<td>CZ279877</td>
<td>1.26%</td>
<td>0.73–2.66%</td>
<td>82.8% CRI, rest α-KRY, MET and clavins</td>
</tr>
<tr>
<td>Gal310</td>
<td>CCM8057</td>
<td>CS276359</td>
<td>0.91%</td>
<td>0.52–1.25%</td>
<td>93.5% KORα-KRY, β-KRY = 4.2:1:6:1</td>
</tr>
<tr>
<td>Gal404</td>
<td>CCM8243</td>
<td>CZ287130</td>
<td>1.56%</td>
<td>0.4–2.2%</td>
<td>93.5% TAM, 1.1% MET, 0.5% STIN, 0.8% TOX</td>
</tr>
</tbody>
</table>

clavines, lysergic acid amides, and peptides (designated as ergopeptides or ergopeptines). The genes involved in the biosynthesis of EAs form a cluster, the ergot alkaloid synthesis (EAS) cluster, consisting of 14 genes spanning approximately 68.5 kb of the genome [11]. The first step in biosynthesis is catalyzed by the gene dmaW [12], that prenylates L-tryptophan in the presence of dimethyllall pyrophosphate (DMAPP) to form dimethyll all tryptophan (DMAT) [13]. Subsequent steps leading to chanoclavine-I synthesis involve methylation (product of easF) [14] and subsequent oxidation (products of easC and easE) [15]. Oxidation of chanoclavine-I (product of easD) [16] forms chanoclavine-I aldehyde, which is cyclized by the product of the easA gene [17] and/or agroclavine synthase (encoded by easG) to form agroclavine [18]. P450 monoxygenase catalyzes the oxidation of agroclavine and formation of elymoclavine, which is subsequently oxidized to psalpolic acid by clavine oxidase (encoded by cloA) [19]. Psalpolic acid is spontaneously converted to 3-lyseric acid [20]. 3-Lyseric acid is captured and activated by the monomodular non-ribosomal peptide synthetase (NRPS) LPS2 (encoded by the lpsB gene) [21], and transferred to the trimodular LPS1 or LPS4 NRPS (encoded by the genes lpsA1 or lpsA2) [11]. Then, it is condensed with three bound amino acids, resulting in the formation of tripeptide lactam [22], which is finally converted to ergopeptide in a reaction catalyzed by Fe3+/2-ketoglutarate-dependent dioxygenase (encoded by easH) [23]. LPS1 and LPS4 enzymes contain adenylation domains responsible for the activation of amino acid to adenylate, and acyl carrier domains referred to as peptide synthetase domains serve as carriers for three amino acids [24]. The structure of various ergopeptides found in nature depends on the first two amino acids bound to 3-lyseric acid, the third one always being a proline (Table 1) [25]. The last NRPS encoded in the EAS is LPS3 (encoded by lpsC), an ergometrine synthetase [26]. The variability of alkaloid production spectra appears to be determined by the substrate specificity of the NRPS modules [11,19]. However, the composition of the amino acid pool in the sclerotal cells can also contribute to the final structure of ergopeptines [24].

In this study, we used massively parallel sequencing to sequence the whole transcriptomes (RNA-seq) of the reference strain Cp20.1 [27] and four Cp industrial strains, which differ in their spectra of produced alkaloids. Differential gene expression analysis between the different stages of fungus development was performed to obtain a comprehensive view of the processes involved in sclerotic differentiation. Moreover, quantitative and qualitative analysis of the transcriptomes of strains producing different spectra of alkaloids was performed to shed more light on the regulation of the final composition of EAs.

Materials and methods

Sample characterization and preparation

For transcriptome sequencing, sclerotia of four industrial strains of Cp (Gal404, Gal012, Gal130, and Gal310) and the reference strain Cp20.1 were used. All industrial strains were derived from the wild type strains found on rye as an original host, which underwent a series of chemical, UV, and/or X-ray mutagenesis to generate the industrial strains. The strain Gal012 was derived from Gal130, while the other strains have different origins. The characterization of the strains is summarized in Table 2. To produce sclerotia, field-grown male sterile hybrid rye (Secale cereale) cultivar L25P × L130N bred at the Plant Breeding and Acclimatization Institute (Radzików, Poland) was inoculated using the conidial suspension of analyzed strains. Inoculation was performed by piercing ears with a needle and syringe containing the conidial suspension. The conidial suspension was prepared by homogenizing 4 g of dried mycelia using a disperser in 1 L of 0.1% corn starch solution buffered to pH 5.3 with 0.1 M K+ -phosphate buffer. Approximately one month after

2 Please cite this article in press as: Majeska Cudjakowska, M. et al., Quantitative and qualitative transcriptome analysis of four industrial strains of Claviceps purpurea with respect to ergot alkaloid production, New Biotechnol. (2016), http://dx.doi.org/10.1016/j.nbt.2016.01.006
infection, when the solid purple/brown sclerotia were formed, the material was harvested, frozen in liquid nitrogen, and homogenized for RNA extraction (several sclerotia per sample) using mortar and pestle. Sclerotia of the strain Gal404 were harvested during two seasons in 2013 and 2014, but sclerotia of other strains were harvested only in one season: either 2013 (Gal012, Gal130, Gal310) or 2014 (Cp20.1). For the strain Gal404, two independent samples of mycelia growing on Mantle agar [10] were prepared for transcriptome sequencing, one in 2013 and another in 2014.

RNA extraction, transcriptome sequencing and alignment
RNAqueous® Total RNA Isolation Kit (Thermo Scientific, Waltham, MA, USA) was used for total RNA extraction according to the manufacturer’s recommendations, and the samples were treated with DNase I (TURBO DNA-free kit, Thermo Scientific, Waltham, MA, USA). Library construction and sequencing of samples harvested in 2013 were performed by GATC Biotech AG (Cologne, Germany) on the Illumina HiSeq 2000 platform in the 50 bases paired end mode. For samples harvested in 2014, 2.5 μg of total RNA was used for cDNA library preparation using Illumina® TruSeq® Stranded mRNA Sample Preparation Kit (Illumina, San Diego, CA, USA). Libraries were pooled (three samples per run) and sequenced using the MiSeq Reagent Kit v3 (2 × 75 cycles) (Illumina, San Diego, CA, USA) on a MiSeq Desktop Sequencer (Illumina, San Diego, CA, USA). The reads generated by sequencing were aligned to the reference genome of Cp strain Cp20.1 [27] using the splice-read mapper TopHat v.2.0.12 [28], allowing maximally five mismatches/INDELS per read to be accepted.

Differential gene expression analysis
The reads aligned to the transcripts annotated in the reference genome and were quantified using FeatureCounts [29] with respect to the stranded pair-end library. The tests for the differential gene expression were performed using the DESeq [30] (for tests without biological replicates, which include a comparison of the sclerotial samples) and DESeq2 packages [31] (for tests with biological replicates, which include a comparison of the mycelial and sclerotial samples) implemented in R (R Development Core Team 2008).

Annotation enrichment and GO annotation
The proteome of the reference strain Cp20.1 was extracted from the reference genome based on its annotation, using the gffread program included in the Cufflinks package [32]. The proteome was subsequently used for annotation enrichment performed using blastp included in the program ncbi-blast+ v.2.2.28 [33]. Protein blast of the Cp proteome was performed against the database of all fungal proteins downloaded from UniProtKB (http://www.uniprot.org/, 2015_07), the database of all fungal transcriptional factors and regulators downloaded from UniProtKB (http://www.uniprot.org/, 2014_06), the database of all fungal proteins associated with amino acid metabolism (http://www.uniprot.org/, 2015_06), and the COG database (ftp://ftp.ncbi.nih.gov/pub/COG/COG2014/data/) [34] accepting the first best blast hit with e-value < 10⁻³. The search for conserved domains in the proteins was performed using InterProScan (InterProScan 5.3-46.0) [35]. Gene Ontology (GO) annotation of the reference genome was performed using the program Blast2GO v.3.0 [36] and the nt database (b2g_2015_01). Single nucleotide polymorphism discovery
The data obtained from the commercial sequencing at GATC Biotech AG (Cologne, Germany) were used for the discovery of single nucleotide polymorphisms (SNP) in the transcriptomes of the four industrial strains. Reads were aligned to the reference genome as described above. Alignments were then processed using the SamTools and Bcftools packages [37]. Bias reads originated by enrichment (PCR duplicates) were removed using the Genome Analysis Tool Kit [38]. Variant extraction and annotation were accomplished using in-house pipelines in comparison with the reference genome and GTF (gene transfer format) annotation of the strain Cp20.1 [27]. qRT-PCR
Isolation of total RNA and DNase I treatment were performed as described above. Samples for qRT-PCR analysis were grown during the two seasons: 2014 (representing first biological replicate, the same as for transcriptome sequencing) and 2015 (representing second biological replicate). 2 μg of RNA were used for the synthesis of cDNA using the RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Scientific, Waltham, MA, USA). Real-time PCR was set up with the cDNA as template in a total volume of 5 μL containing gb 9G PCR Master mix with 100 nM reference dye, ROX (Generi Biotech, Hradec Kralove, Czech Republic), and 500 nM of each primer. Reactions were run on the Vial™ Real-Time PCR System using a default program (Thermo Scientific, Waltham, MA, USA). Primers were designed using the Primer3 (web version 4.0.0) software (Supplemental Table 1). A melting curve analysis was performed at the end of the PCR reaction to confirm product quality. All data were normalized with respect to the β-tubulin amplicon. The efficiency corrected ΔΔCT method [39], and the differences in the cycle number during the linear amplification phase between samples were used for the determination of relative gene expression. Efficiency of qPCR was estimated for each primer pair using a method based on the linear regression slope of a dilution series [40]. Gene expression in the sclerotia was expressed relative to that estimated in the mycelia.

Quantification of ergot alkaloids
Sclerotia, homogenized as described above, were dissolved in 50 mL of extraction buffer (26% ammonium hydroxide and 90% acetone, 1:100, v/v) on a shaker for 2 h. A 10 mL aliquot was filtered through glass wool and dried using a vacuum evaporator at 50°C. The residue was dissolved in 2 mL of 90% methanol, filtered through a 0.22 mm nylon filter and analyzed for EA content using ultra high performance liquid chromatography on a Nexera system (Shimadzu, Prague, Czech Republic) equipped with a C18 reverse-phase column (Zorbax RRHD Eclipse Plus, 1.8 mm, 2.1 mm ID × 50 mm, Agilent, Santa Clara, CA, USA). An EA reference mixture was prepared in 90% methanol and contained 0.001% (w/v) of each of the following alkaloids: ergamine, ergostine, 8-hydroxyergotamine, ergocornine, α-ergokryptine, β-ergokryptine, ergocristine, ergogaline, ergotamine, ergostinine, ergocorninine, α-ergokryptinine, β-ergokryptinine, ergogaline, and ergocristinine (all from Teva Czech Industries, Opava-Komárov, Czech Republic). The analytes were eluted with solvents A (36 mM triethylamine/phosphate, pH 4.4, and acetonitrile, 4:1, v/v) and B (H₂O and acetonitrile, 1:4, v/v) using the following gradient: 0 min, 9% B; 10–12 min, 9–13% B; 12–14 min,
13–29% B; 14–16 min, 29–44% B; 18–20 min, 44–59% B; 26–28 min, 59–9% B, at flow rate 0.4 mL/min and column temperature of 30°C. Monitoring was performed at 317 nm.

**Genomic DNA extraction and PCR amplification**

Genomic DNA was extracted from mycelia grown on Mantle agar [10] using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany). Total 50 ng of DNA was used for the amplification of the region of interest in the *ipsA1* gene using the forward primer: 5’-CCGC-AAGAAGTCACAGG-3’ and reverse primer: 5’-GGAGC-CTTTCTCGAGTTTCA-3’. The reaction mixture contained 0.2 μM of each primer, 0.2 mM dNTPs, 1× PCR reaction buffer A (containing 1.5 mM MgCl₂ in the final volume) and 1 U of Kapa Taq polymerase (Kapa Biosystems, Wilmington, MA, USA). Amplifications were carried out as follows: initial denaturation at 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 60°C for 30 s, 72°C for 1 min and final extension for 10 min. PCR products were purified using the GenElute™ PCR Clean-Up Kit (Sigma–Aldrich, St. Louis, Missouri, USA), and sequencing was performed at SEQme s.r.o. (Dobřš, Czech Republic).

**Results and discussion**

A combination of two sequencing platforms was used to obtain enough sequencing data for comparative transcriptomic study. Samples sequenced on the HiSeq and MiSeq platforms differed in the total number of generated reads. With the HiSeq platform, approximately 90 million reads per sample were obtained, whereas with the MiSeq platform, approximately 20 million reads were generated per sample. From 75.1% up to 90.5% of these reads were mapped to the reference genome and covered the transcriptome 365 to 92 times. An average of 72% of mapped reads were successfully assigned to the transcripts annotated in the reference genome [27] and quantified for differential gene expression (DEG) analysis (Table 3). Several genes were analyzed by qRT-PCR to validate the results of transcriptome sequencing.

**Differential gene expression in mycelial and sclerotial tissues**

Global gene expression was compared between the mycelia and sclerotia of the industrial strain Gal404. The analysis was performed using two biological replicates of sclerotia obtained during two different seasons, and two biological replicates of axenically cultivated mycelia. A total of 846 genes were found to be significantly differentially expressed at the p-adjusted value < 0.01. The genes whose expression differed <4-fold (Log2FoldChange < -2 and >2) were filtered out, which reduced the number of DEG to 627, with 334 being downregulated and 293 being upregulated in sclerotia. Approximately half of the genes could be assigned to a gene ontology (GO) number. GO distribution by level 5 was analyzed in aspects of ‘Biological process’ and ‘Molecular function’ and compared between the downregulated and upregulated genes. The terms which appeared in one group two or more times than in the other one were considered as up/downregulated in sclerotia (Tables 4 and 5). Although the GO annotation is not complete, an expected trend connected with sclerotia formation could be observed. Considering ‘Biological process’ (Table 4), the most significantly downregulated genes in the sclerotia were involved in transport, protein metabolism (translation, un/folding, signal peptide processing, repair, and degradation), and protein modification (involving pre-translational, co-translational, and post-translational modifications). Because the sclerotia serve as resting structures, downregulation of these processes was expected. Organic acid biosynthesis and its degradation was also notably downregulated in the sclerotia. These processes include amino acid metabolism (e.g., CPUR_00803.1 – branched-chain amino acid transaminase, or CPUR_03287.1 – delta-1-pyrroline-5-decarboxylase dehydrogenase involved in proline and arginine metabolism, etc.; Supplemental Table 2); the downregulation of these genes in sclerotia was already observed by Oeser et al. [41] in an analysis of ESTs generated from pathogenic cultivation of *Cp* compared with axenic culture. A similar observation was previously made at the protein and amino acid level [4]. Moreover, changes in the amino acid pool composition were also reported by the same group, which correspond with our results showing several upregulated genes involved in amino acid biosynthesis and transport (e.g., CPUR_05321.1 – aspartate-tRNA ligase, CPUR_01235.1 – LEU4-2-isopropylmalate synthase, CPUR_04007.1 – phenylalanyl-tRNA synthetase alpha subunit (*podG*), and CPUR_04624.1 – proline-specific permease) (Supplemental Table 2). Expression of three of these genes (CPUR_01235.1, CPUR_04007.1, CPUR_04624.1) was verified using qRT-PCR, showing upregulation in sclerotia in the samples from the two seasons 2014 and 2015 (CPUR_05321.1 was not tested) (Supplemental Figure 1). Another expected event was the upregulation of alkaloid biosynthesis and indole alkaloid metabolism in sclerotia, which represent the only ergot alkaloid
TABLE 4
GO terms assigned to the differentially expressed genes (adjusted p-value < 0.01) at the level 5, GO domain – biological process. The distribution of terms was compared between the down/upregulated genes (DGs/UGs). The terms, which appeared in one group of the genes two or more times than in other one, were considered as down/upregulated (D/U) in sclerotia.

<table>
<thead>
<tr>
<th>GO-id</th>
<th>GO-term</th>
<th>#UGs</th>
<th>#DGs</th>
<th>D/U</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0006778</td>
<td>Porphyrin-containing compound metabolic process</td>
<td>0</td>
<td>2</td>
<td>D</td>
</tr>
<tr>
<td>GO:006811</td>
<td>Ion transport</td>
<td>3</td>
<td>13</td>
<td>D</td>
</tr>
<tr>
<td>GO:0016053</td>
<td>Organic acid biosynthetic process</td>
<td>2</td>
<td>6</td>
<td>D</td>
</tr>
<tr>
<td>GO:0016054</td>
<td>Organic acid catabolic process</td>
<td>0</td>
<td>2</td>
<td>D</td>
</tr>
<tr>
<td>GO:0016310</td>
<td>Phosphorylation</td>
<td>4</td>
<td>19</td>
<td>D</td>
</tr>
<tr>
<td>GO:0036211</td>
<td>Protein modification process</td>
<td>4</td>
<td>20</td>
<td>D</td>
</tr>
<tr>
<td>GO:0044267</td>
<td>Cellular protein metabolic process</td>
<td>5</td>
<td>20</td>
<td>D</td>
</tr>
<tr>
<td>GO:0046434</td>
<td>Organophosphate catabolic process</td>
<td>0</td>
<td>2</td>
<td>D</td>
</tr>
<tr>
<td>GO:0055085</td>
<td>Transmembrane transport</td>
<td>14</td>
<td>22</td>
<td>D</td>
</tr>
<tr>
<td>GO:006023</td>
<td>Aminoglycan biosynthetic process</td>
<td>2</td>
<td>0</td>
<td>U</td>
</tr>
<tr>
<td>GO:006733</td>
<td>Oxidoreduction coenzyme metabolic process</td>
<td>3</td>
<td>1</td>
<td>U</td>
</tr>
<tr>
<td>GO:007166</td>
<td>Cell surface receptor signaling pathway</td>
<td>2</td>
<td>0</td>
<td>U</td>
</tr>
<tr>
<td>GO:009311</td>
<td>Oligosaccharide metabolic process</td>
<td>4</td>
<td>1</td>
<td>U</td>
</tr>
<tr>
<td>GO:009821</td>
<td>Alkaloid biosynthetic process</td>
<td>3</td>
<td>0</td>
<td>U</td>
</tr>
<tr>
<td>GO:0016051</td>
<td>Carbohydrate biosynthetic process</td>
<td>6</td>
<td>4</td>
<td>U</td>
</tr>
<tr>
<td>GO:003203</td>
<td>Glycosaminoglycan metabolic process</td>
<td>2</td>
<td>0</td>
<td>U</td>
</tr>
<tr>
<td>GO:0034645</td>
<td>Cellular macromolecule biosynthetic process</td>
<td>14</td>
<td>12</td>
<td>U</td>
</tr>
<tr>
<td>GO:0034654</td>
<td>Nucleobase-containing compound biosynthetic process</td>
<td>9</td>
<td>7</td>
<td>U</td>
</tr>
<tr>
<td>GO:0035834</td>
<td>Indole alkaloid metabolic process</td>
<td>3</td>
<td>0</td>
<td>U</td>
</tr>
<tr>
<td>GO:0044042</td>
<td>Glucan metabolic process</td>
<td>3</td>
<td>1</td>
<td>U</td>
</tr>
<tr>
<td>GO:0044106</td>
<td>Cellular amine metabolic process</td>
<td>3</td>
<td>0</td>
<td>U</td>
</tr>
<tr>
<td>GO:0044724</td>
<td>Single-organism carbohydrate catabolic process</td>
<td>3</td>
<td>1</td>
<td>U</td>
</tr>
<tr>
<td>GO:0055086</td>
<td>Nucleobase-containing small molecule metabolic process</td>
<td>4</td>
<td>2</td>
<td>U</td>
</tr>
<tr>
<td>GO:0090304</td>
<td>Nucleic acid metabolic process</td>
<td>13</td>
<td>10</td>
<td>U</td>
</tr>
</tbody>
</table>

TABLE 5
GO terms assigned to the differentially expressed genes (adjusted p-value < 0.01) at the level 5, GO domain – molecular function. The distribution of terms was compared between the down/upregulated genes (DGs/UGs). The terms, which appeared in one group of the genes two or more times than in other one, were considered as down/upregulated (D/U) in sclerotia.

<table>
<thead>
<tr>
<th>GO-id</th>
<th>GO-term</th>
<th>#UGs</th>
<th>#DGs</th>
<th>D/U</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0004672</td>
<td>Protein kinase activity</td>
<td>4</td>
<td>19</td>
<td>D</td>
</tr>
<tr>
<td>GO:0003513</td>
<td>Inorganic phosphate transmembrane transporter activity</td>
<td>0</td>
<td>3</td>
<td>U</td>
</tr>
<tr>
<td>GO:0008081</td>
<td>Phosphonic diester hydrolase activity</td>
<td>0</td>
<td>3</td>
<td>D</td>
</tr>
<tr>
<td>GO:008237</td>
<td>Metalloproteinase activity</td>
<td>2</td>
<td>5</td>
<td>D</td>
</tr>
<tr>
<td>GO:008324</td>
<td>Cation transmembrane transporter activity</td>
<td>1</td>
<td>8</td>
<td>D</td>
</tr>
<tr>
<td>GO:0015405</td>
<td>P-P-bond-hydrolysis-driven transmembrane transporter activity</td>
<td>1</td>
<td>6</td>
<td>D</td>
</tr>
<tr>
<td>GO:0016407</td>
<td>Acetyltransferase activity</td>
<td>0</td>
<td>2</td>
<td>D</td>
</tr>
<tr>
<td>GO:0016410</td>
<td>N-acetyltransferase activity</td>
<td>0</td>
<td>2</td>
<td>D</td>
</tr>
<tr>
<td>GO:0016462</td>
<td>Pyrophosphatase activity</td>
<td>2</td>
<td>7</td>
<td>D</td>
</tr>
<tr>
<td>GO:0016791</td>
<td>Phosphatase activity</td>
<td>2</td>
<td>4</td>
<td>D</td>
</tr>
<tr>
<td>GO:0017076</td>
<td>Purine nucleotide binding</td>
<td>9</td>
<td>28</td>
<td>D</td>
</tr>
<tr>
<td>GO:0032550</td>
<td>Purine ribonucleoside binding</td>
<td>9</td>
<td>28</td>
<td>D</td>
</tr>
<tr>
<td>GO:0032553</td>
<td>Ribonucleotide binding</td>
<td>11</td>
<td>29</td>
<td>D</td>
</tr>
<tr>
<td>GO:0050660</td>
<td>Flavin adenine dinucleotide binding</td>
<td>6</td>
<td>1</td>
<td>U</td>
</tr>
<tr>
<td>GO:0004096</td>
<td>Catalase activity</td>
<td>3</td>
<td>0</td>
<td>U</td>
</tr>
<tr>
<td>GO:0008762</td>
<td>UDP-N-acetyl muramate dehydrogenase activity</td>
<td>3</td>
<td>0</td>
<td>U</td>
</tr>
<tr>
<td>GO:0046914</td>
<td>Transition metal ion binding</td>
<td>23</td>
<td>11</td>
<td>U</td>
</tr>
</tbody>
</table>

Please cite this article in press as: Majeska Cudejko, M. et al., Quantitative and qualitative transcriptome analysis of four industrial strains of Chaetomium purpurea with respect to ergot alkaloid production, New Biotechnol. (2016), http://dx.doi.org/10.1016/j.nbt.2016.01.006
in 1832 by H. Wiggers [44], may play a variety of roles, such as a source of energy or an environmental stress protectant [45–47]. Moreover, a recent study demonstrated that in yeast, trehalose is effective against long-term desiccation; thus, it can be hypothesized that trehalose in Cp could contribute to the over-wintering function of sclerotia [48]. Another group of upregulated genes share the common property of being repressed by glucose; they are: CPUR_05946.1 – glycerol kinase (gut1), CPUR_05947.1 – glyceral uptake facilitator (both involved in the utilization of glycerol as a source of carbon and energy [49]), CPUR_04235.1 – fructose-1,6-bisphosphatase (a rate-limiting enzyme in gluconeogenesis), CPUR_01290.1 – suc2 invertase, and CPUR_01071.1 – glucose repressible protein (ggr1, expression increases during glucose starvation [50]). The results indicated a probable low content of glucose in sclerotia and the utilization of alternative carbon sources, which represents a crucial condition for alkaloid biosynthesis under saprophytic conditions, as glucose is known to inhibit alkaloid biosynthesis [5]. Among the 293 upregulated genes, at least two groups of genes were found located next to each other, which might form gene clusters, a typical organization of genes involved in secondary metabolism in fungi. One of the clusters was the EAS cluster. The other putative cluster/clusters are represented by the genes CPUR_05417.1–CPUR_05424.1 and CPUR_05426–CPUR_05437.1. CPUR_05437.1 encodes a polyketide synthase similar to bik1 (involved in bikaverin biosynthesis in Fusarium fujikuroi [51]) or pks12 (involved in the biosynthesis of autofusarin in Fusarium graminearum [52]). These polyketide synthases are responsible for pigment synthesis, thus genes located before CPUR_05437.1 are very likely involved in the biosynthesis of the typical purple pigment of ergot. A function of another putative cluster (CPUR_05407.1–CPUR_05421.1) might indicate the presence of the gene CPUR_05422.1 encoding a protein similar to the aflatoxin efflux pump (MFS family permease associated with secondary metabolites) and the gene CPUR_05423.1 probably encoding a maackian detoxification protein 1; both are known to be virulence factors involved in the detoxification of phytoalexins, plant antimicrobial compounds produced in response to pathogen attack or wounding [53–55]. Significant upregulation of the three tested genes (CPUR_05417.1, CPUR_05422.1, and CPUR_05423.1) from the putative cluster was confirmed using qRT-PCR. However, the intensity of expression appears to be regulated by environmental conditions (Supplemental Figure 1). Thus, we hypothesize a role for this cluster (or at least the two genes) in defense mechanisms against plant antimicrobial compounds. Among the other genes, which were already characterized as genes involved in pathogenicity and virulence in Cp, hydrophobin 1 (CPUR_00649.1) [50,56] was found to be approximately 30-fold upregulated in sclerotia (Supplemental Table 2). The genes cft1 (CPUR_07883.1) [57] and cmpk2 (CPUR_01700.1) [58] were found to be upregulated in sclerotia <4-fold and were filtered out, and other genes, such as cpcdt1 (CPUR_03237.1) [59], cpcdt4 (CPUR_04764.1) [60], cpdc42 (CPUR_06091.1) [61], cpdx1 (CPUR_02147.1) [62], and cpno2 (CPUR_03617.1) [63], were not found to be differentially expressed. These results can be explained by the fact that late stages of development of sclerotia were analyzed, while these genes may probably be active at earlier phases of pathogenesis, or they do not have to be differentially expressed to play a role in virulence and pathogenicity.

Except several uncharacterized genes, four genes from the EAS cluster (easC, easA, easG, and easG), genes which we propose to be responsible for pigment biosynthesis, and the gene encoding fatty acid hydroxylase cpfah (CPUR_04032.1), responsible for the biosynthesis of ricinoleic acid [64], were found among the twenty most upregulated genes in sclerotia. The most upregulated gene (1184x, not expressed in mycelia) was CPUR_03366.1, a homolog of monodehydroascorbate reductase (MDAR) known to be a key enzyme in the ascorbate H2O2 detoxification machinery in plants [65]. Ascorbate is also produced by fungi and its role as an antioxidant was previously proposed and confirmed [66–68]. Moreover, the reduction of oxidative stress via ascorbate suppresses sclerotal differentiation in Sclerotium rolfsii [68] and Rhizoctonia solani [69]. Because there is no evidence of an oxidative burst in Cp-infected rye tissues [70], it is very likely that Cp possesses a defense mechanism against reactive oxygen species (ROS) generated by the host. The expression of MDAR in sclerotia, but not in mycelia, may indicate its specific function in the detoxification of ROS produced by the plant, thus it is a promising candidate involved in the fungal defense response. In 1997, Georgiou proposed that oxidative stress, specifically H2O2 formation, is the triggering factor of sclerotial differentiation [71], and this hypothesis was subsequently supported by several studies on several sclerotia-forming phytopathogens [72–75] including Cp [63]. Moreover, oxidative stress induces aflatoxin B1 biosynthesis in Aspergillus flavus [76]; thus it seems to be required not only for sclerotia formation but also for the biosynthesis of secondary metabolites associated with sclerotia. This hypothesis might be supported by our results. Indeed, the elevated expression of antioxidant defense genes (e.g., CPUR_07394.1 – peroxidase/catalase and CPUR_02927.1 – Cu/ Zn superoxide dismutase), which were shown to be stress-induced in Neurospora crassa or Saccharomyces cerevisiae [77,78], might be indicative of stress conditions in mature sclerotia. Increased expression of both these genes in sclerotia was confirmed using qRT-PCR (Supplemental Figure 1).

Among the twenty most downregulated genes in sclerotia were two genes encoding copper transport proteins (CPUR_02520.1 and CPUR_05851.1), ferric/cupric reductase component 2 (CPUR_05852.1) – a metalloreductase reducing iron and copper [79], and probable dis1-suppressing protein kinase (dsk1, CPUR_06787.1), a protein kinase specific for serine/arginine-rich proteins, an important factor for alternative splicing [80]. Interestingly, seven and three other genes annotated as dsk1 were found among the downregulated and upregulated genes, respectively, indicating that alternative splicing may play an important role in tissue differentiation.

**Identification of differentially expressed genes encoding putative transcription factors and regulators in the mycelial and sclerotal tissues**

The identification of putative transcription factors and regulators (TFsTRs) was performed using a combination of two techniques: Blast (ncbi-blast+ v.2.2.28 [33]) and InterProScan (InterProScan 5.3.46.0 [35]). Blasting the whole Cp proteome against the database of all fungal transcription factors and regulators downloaded from UniprotKB resulted in 1176 putative TFsTRs. Filtering the InterProScan results according to PFAM and Superfamily ID, which have been proposed to occur in fungal species [81], identified
approximately 252 TFsTRs. Taken together, we could predict 1214 putative TFsTRs in the Cp genome (data not shown). Seventy four of these genes were found to be significantly differentially expressed in sclerotia when compared with mycelia: 45 genes being downregulated and 29 being upregulated in sclerotia (Supplemental Table 3). Among the downregulated TFsTRs in sclerotia, we could identify multiple genes having the same function, for example: four genes annotated as serine/threonine protein kinases psk1 (CPUR_04539.1, CPUR_04553.1, CPUR_04554.1, and CPUR_04555.1), three genes coding for proteins related to the mck1 protein kinase (CPUR_04441.1, CPUR_06787.1, and CPUR_06788.1), two putative mitogen-activated protein kinases hog1 (CPUR_04543.1 and CPUR_08159.1), and two putative transcriptional regulators ure2 (CPUR_00641.1 and CPUR_00643.1). The serine/threonine protein kinase psk1 belongs to the PAS domain containing kinases, which are a part of the signaling pathway that controls and connects the balance of carbohydrate consumption and its storage to protein synthesis in yeast [82]. According to the described function of psk1, a decreased level of psk1 should result in increased carbohydrate storage and decreased protein synthesis, which is in compliance with our results described above. The protein kinase mck1 fulfills several functions, for example, it is a positive regulator of meiosis and spore formation [83] and also a negative regulator of pyruvate kinase activity in yeast [84]. However, its role in Cp sclerotia formation is not clear, and we cannot exclude the possibility that genes annotated as mck1 protein kinases are involved in different signaling pathways. A mitogen-activated protein kinase hog1 was found to play a role in the osmostic stress response in yeast [85] and in vegetative and pathogenic development in Botrytis cinerea, where deletion of the hog1 homolog caused an increase in sclerotia formation [86]. The function of ure2 regulator might be associated with nitrogen metabolism [87]; however, it appears to be variable depending on species [88]. The list of upregulated TFsTRs included two genes (CPUR_00005.1 and CPUR_08625.1) coding for the secreted protein pry1 (Pathogenesis Related in Yeast 1), a sterol-binding protein involved in fungal virulence [89,90]. Other upregulated TFsTRs in sclerotia were genes probably implicated in sexual reproduction and spore formation, for example, CPUR_03624.1 – transcriptional regulatory protein moc3 [91], CPUR_07653.1 – flug protein [92], CPUR_04653.1 – transcriptional regulator crz1 [93], and CPUR_06568.1 – mating-type protein mat alpha 1 [94]. The transcription factor crz1 is important for calcium homeostasis [95], stress response [96], secondary metabolism [97], and virulence in fungi [98]. The transcriptional regulator vib-1 (CPUR_01075.1), a regulator of responses to nitrogen and carbon starvation in Neurospora crassa [99], and the zinc-responsive transcriptional regulator zap1 (CPUR_01124.1), a regulator of zinc homeostasis in yeast [100], were also found to be significantly upregulated in sclerotia (Supplemental Table 3).

Differential gene expression in the sclerotia of the four industrial strains

A comparison of the transcriptomes of sclerotia from four industrial strains of Cp (Table 2) and the reference strain Cp20.1 was performed to better understand the mechanisms regulating the final spectrum of EAs produced by the different strains. An overview of the numbers of differentially expressed genes between the strains is summarized in Table 6, and the complete results are summarized in Supplemental Table 4. A relatively lower number of significantly differentially expressed genes was found between the reference strain Cp20.1 and the industrial strains Gal012 (47 genes), Gal130 (61 genes), and Gal310 (57 genes) than in Gal404 compared with the other Gal strains (Gal012 – 132 genes, Gal130 – 98 genes, Gal310 – 192 genes); this is likely due to the analysis setup. Indeed, except for the strain Gal404, the analysis was based on only one biological replicate, which may have resulted in the identification of a lower number of significantly differentially expressed genes. However, only 7 downregulated genes (and no upregulated genes!) in Gal404 compared with Cp20.1 clearly indicated that global gene expression in Gal404 was very similar to the reference strain in contrast to the other Gal strains (Table 6, Supplemental Table 4). Identification of DEG between all the strains revealed no differences in the expression of genes belonging to the EAS cluster, except between the strains Gal012 and Gal310. Six EAS genes (cloA, easH, lpsB, easC, easE, and lpsA2) were found to be significantly upregulated in Gal310 compared with Gal012. Higher expression of these genes in Gal310 might be connected with the increased accumulation of alkaloids (Table 7); however, the correlation between expression and

<table>
<thead>
<tr>
<th>Strain</th>
<th>SUM of EA (%) of dry mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gal04</td>
<td>0.743 ± 0.424*</td>
</tr>
<tr>
<td>Gal012</td>
<td>0.596</td>
</tr>
<tr>
<td>Gal130</td>
<td>0.732</td>
</tr>
<tr>
<td>Gal310</td>
<td>1.064</td>
</tr>
<tr>
<td>Cp20.1</td>
<td>0.597</td>
</tr>
</tbody>
</table>

TABLE 6

<table>
<thead>
<tr>
<th></th>
<th>Cp20.1</th>
<th>Gal04</th>
<th>Gal012</th>
<th>Gal130</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Up</td>
<td>Down</td>
<td>Up</td>
<td>Down</td>
</tr>
<tr>
<td>Gal012</td>
<td>9</td>
<td>38</td>
<td>40</td>
<td>92</td>
</tr>
<tr>
<td>Gal130</td>
<td>19</td>
<td>42</td>
<td>39</td>
<td>59</td>
</tr>
<tr>
<td>Gal310</td>
<td>28</td>
<td>29</td>
<td>86</td>
<td>106</td>
</tr>
<tr>
<td>Gal404</td>
<td>7</td>
<td>0</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

TABLE 7

<table>
<thead>
<tr>
<th>Strain</th>
<th>SUM of EA of dry mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gal04</td>
<td>0.743 ± 0.424*</td>
</tr>
<tr>
<td>Gal012</td>
<td>0.596</td>
</tr>
<tr>
<td>Gal130</td>
<td>0.732</td>
</tr>
<tr>
<td>Gal310</td>
<td>1.064</td>
</tr>
<tr>
<td>Cp20.1</td>
<td>0.597</td>
</tr>
</tbody>
</table>

Please cite this article in press as: Majeska Čudežková, M. et al., Quantitative and qualitative transcriptome analysis of four industrial strains of Claviceps purpurea with respect to ergot alkaloid production, New Biotechnol. (2016), http://dx.doi.org/10.1016/j.nbt.2016.01.006
accumulation is not clear because Cp20.1 accumulated similar amounts of alkaloids as Gal012, and no DEGs among the EAS were found when compared with Gal310 (Supplemental Table 4). Among the DEGs found between Gal404 and other Gal strains, we found several up/downregulated genes, which are very likely involved in amino acid metabolism and transport, for example, CPUR_05447.1 – L-amino acid oxidase (downregulated in Gal404, upregulated in all other Gal strains), CPUR_05450.1 – thiamine pyrophosphosphate enzyme (downregulated in Gal404, upregulated in Gal130 and Gal310), or CPUR_00573.1 – arginine metabolism regulation protein 4 (upregulated in Gal404, downregulated in Gal012, and Gal310). A similar trend may be observed in the DEGs found between the reference strain and Gal012, Gal130, and Gal310 (Supplemental Table 4). However, much lower numbers of amino acid metabolism and transport associated genes were found among the DEGs identified between the Gal strains except Gal404 (Supplemental Table 4). These results indicate changes in amino acid metabolism between the strains producing different spectra of alkaloids and support the hypothesis that the composition of the amino acid pool in sclerotic cells affects the final ergopeptine structure [24].

Common up/downregulated genes overlapping between the Gal strains were identified using Venn Diagrams (http://bioinformatics.psb.ugent.be/webtools/Venn/). The lists of genes found in the intersections are summarized in Supplemental Table 5. From the possible functions of these proteins based on the annotation or presence of conserved domains, their involvement in the regulation of alkaloid production was not apparent. The intersection of downregulated genes in Gal404 and Cp20.1 and upregulated in other Gal strains (Gal012, Gal130 and Gal310) consists of 14 and 12 common genes, respectively. Four of these genes are identical (CPUR_00025.1, CPUR_05443.1, CPUR_05445.1, and CPUR_08721.1) and are expressed at very low levels in ergotamine producers (Supplemental Table 5). The CPUR_00025.1 protein possesses three domains characteristic of helicases, for example, PIFI-like helicases (pfam05970), which are known to inhibit telomerase activity in yeast [101] and are also needed as accessory helicases for the stable maintenance of the genome [102]. The protein encoded by the gene CPUR_05445.1 contains a Class III extradiol dioxygenase domain, which might have a variety of functions from catabolism of aromatic compounds [103] to betalain biosynthesis [104]; however, betalain accumulation in Cp was not described so far. The CPUR_05443.1 protein seems to be a glutathione S-transferase, and CPUR_08721.1 is an uncharacterized protein containing no conserved domains.

Single nucleotide polymorphism discovery in the four industrial strains

In 2005, Haarmann and colleagues [11] proposed that the evolution of the ‘chemical races’ of Cep is determined by the evolution of the NRPS (encoded by genes lpsA1 and lpsA2) module specificity. To test this hypothesis, we performed a qualitative analysis (SNP discovery) of the transcriptomes of four industrial strains differing in their alkaloid spectra (Table 2). The sequencing data were aligned to the reference genome Cp20.1 using the same criteria as for differential gene expression analysis, which was 5 mismatches per 50-bp long read to be accepted into the alignment. Thus, this method is limited by this criterion, and does not cover more variable regions present in the aligned transcriptomes. Nevertheless, a relatively high percentage of successfully aligned reads (Table 3) indicated that low coverage regions were very likely rare. Moreover, a different setup of alignment allowing more mismatches either did not improve or decreased the percentage of successfully aligned reads (data not shown). Identification of SNPs present in the analyzed transcriptomes compared with the reference genome was performed using the SAMtools and BCFTools packages [37], complemented by in-house pipelines. The results are summarized in Tables 8 and 9; detailed results for EAS cluster are available in Supplemental Table 6 (complete results are available upon request).

The NRPS genes, lpsA1 (CPUR_04074.1) and lpsA2 (CPUR_04073.1), possess regions containing acyl carrier and peptide synthetase domains. These domains are organized into three modules, where each module carries one amino acid; the composition of the amino acids determines the final ergopeptine (Table 1) [11]. Based on the hypothesis that the production of specific ergopeptides is determined by NRPS module specificity, we should observe amino acid changes in the binding pocket of the first two modules (the third is specific for proline) of lpsA1 or/and lpsA2 genes in the strains Gal012, Gal130, and Gal310, but not in Gal404, which is an ergotamine producer similar to the reference strain. However, no nonsynonymous SNPs were found in the modules of these two genes in all analyzed strains (Supplemental Table 6; check column lpsA_module – 19th column). However, when we look at the alignment coverage of these regions visualized using the Integrative Genome Viewer (IGV) [105], we can observe no coverage of some parts of the sequences or even whole regions (Supplemental Figure 2). The least covered was the first module in the lpsA2 gene, where no reads coming from the transcriptome sequencing of Gal012, Gal130, and Gal310 were aligned. Thus, this module possesses the most variable sequence. Generally, concerning the lpsA1/lpsA2 modules, the strain Gal310 seem to
TABLE 9
Number of SNPs found within CDS of the genes belonging to the EAS cluster in industrial strains compared to reference strain Cp20.1. The results were filtered by quality ≥100, and only the non-reference homozygotes, and nonsynonymous SNPs were taken into consideration. SNP – single nucleotide polymorphism, EAS – ergot alkaloid synthesis

<table>
<thead>
<tr>
<th>Transcript ID</th>
<th>Gene name</th>
<th>Gal404</th>
<th>Gal012</th>
<th>Gal130</th>
<th>Gal310</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPUR_04073.1</td>
<td>lpsA2</td>
<td>42</td>
<td>50</td>
<td>22</td>
<td>30</td>
</tr>
<tr>
<td>CPUR_04074.1</td>
<td>lpsA1</td>
<td>89</td>
<td>99</td>
<td>63</td>
<td>28</td>
</tr>
<tr>
<td>CPUR_04075.1</td>
<td>easH</td>
<td>13</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>CPUR_04076.1</td>
<td>dmaW</td>
<td>16</td>
<td>14</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>CPUR_04077.1</td>
<td>easG</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>CPUR_04078.1</td>
<td>easF</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>CPUR_04079.1</td>
<td>easE</td>
<td>3</td>
<td>7</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>CPUR_04080.1</td>
<td>easD</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CPUR_04081.1</td>
<td>easC</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>CPUR_04082.1</td>
<td>cloA</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>CPUR_04083.1</td>
<td>lpsB</td>
<td>24</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>CPUR_04084.1</td>
<td>lpsA1</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>CPUR_04085.1</td>
<td>lpsC</td>
<td>42</td>
<td>45</td>
<td>47</td>
<td>42</td>
</tr>
</tbody>
</table>

FIGURE 1
Protein alignment of the 2nd module of LPS1 protein (encoded by lpsA1 gene). Sequences were obtained by sequencing of PCR products by Sanger method. Brown objects under the alignment annotate the binding pocket of the module.

have the most variable sequence when compared with the reference. As expected, the third module of both genes was the best covered module (Supplemental Figure 2). The problem with low coverage of the modules in the lpsA genes cannot be solved by increasing the mismatch criteria in the alignment or by de novo transcriptome assembly because these two genes show 89% mRNA sequence identity. Allowing more mismatches in the alignment would lead to an increased number of multi aligned reads, which may cause loss of specificity, and de novo assembly may result in the reconstruction of one ‘hybrid’ lpsA gene (data not shown). To resolve the sequence of the first two modules of both lpsA genes, gene and module specific primers were designed for conventional PCR followed by Sanger sequencing. Despite extensive optimization of the PCR, only the 2nd module of lpsA1 gene in more than one Gal strain (concretely Gal404, Gal012 and Gal130) was successfully amplified and sequenced. The alignment of the sequences showed no differences in the amino acid composition of this module between the strains. The sequence of the binding pocket was determined to be: DLVGMAAV (Fig. 1). The same sequence was already reported for Cp strains P1 and ECC93 [11]. The strain P1 is an ergotamine producer (similar to Gal404 and Cp20.1), and the strain ECC93 is an ergocristine producer.
References


Please cite this article in press as: Majeska Udovcikov, M. et al., Quantitative and qualitative transcriptome analysis of four industrial strains of Claviceps purpurea with respect to ergot alkaloid production, New Biotechnol. (2016), http://dx.doi.org/10.1016/j.nbt.2016.01.006


