Phytochemistry 152 (2018) 113–124

Contents lists available at ScienceDirect

Phytochemistry

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

Probing the transcriptome of *Aconitum carmichaelii* reveals the candidate genes associated with the biosynthesis of the toxic aconitine-type C_{19}-diterpenoid alkaloids

Dake Zhao [a, b, c, 1], Yong Shen [d, 1], Yana Shi [e, 1], Xingqiao Shi [f], Qin Qiao [f], Shuhui Zi [f], Erqiang Zhao [f], Diqiu Yu [c, **], Edward J. Kennelly [g, h, *]

[a] *Laboratory of Ecology and Evolutionary Biology, State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan University, Kunming, China*
[b] *Yunnan Institute of Materia Medica, Yunnan Baiyao Group Company Limited, Kunming, China*
[c] *Key Laboratory of Tropical Plant Resources and Sustainable Use, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Kunming, China*
[d] *College of Agriculture and Biotechnology, Yunnan Agricultural University, Kunming, China*
[e] *Institute of Medicinal Plants, Yunnan Academy of Agricultural Sciences, Kunming, China*
[f] *School of Agriculture, Yunnan University, Kunming, China*
[g] *Department of Biological Sciences, Lehman College, City University of New York, United States*
[h] *Ph.D. Programs in Biochemistry, Biology, and Chemistry, The Graduate Center, City University of New York, United States*

ARTICLE INFO

Article history:
Received 14 September 2017
Received in revised form
25 April 2018
Accepted 29 April 2018
Available online 21 May 2018

Keywords:
Aconitum carmichaelii (Ranunculaceae)
Fuzi
Transcriptome
Biosynthesis
Aconitine-type C_{19}—diterpenoid alkaloids
Acute toxicity
Traditional Chinese medicine (TCM)

ABSTRACT

Aconitum carmichaelii has long been used as a traditional Chinese medicine, and its processed lateral roots are known commonly as fuzi. Aconitine-type C_{19}-diterpenoid alkaloids accumulating in the lateral roots are some of the main toxicants of this species, yet their biosynthesis remains largely unresolved. As a first step towards understanding the biosynthesis of aconitine-type C_{19}-diterpenoid alkaloids, we performed *de novo* transcriptome assembly and analysis of rootstocks and leaf tissues of *Aconitum carmichaelii* by next-generation sequencing. A total of 525 unigene candidates were identified as involved in the formation of C_{19}-diterpenoid alkaloids, including those encoding enzymes in the early steps of diterpenoid alkaloids scaffold biosynthetic pathway, such as *ent*-copalyl diphosphate synthases, *ent*-kaurene synthases, kaurene oxidases, cyclases, and key aminotransferases. Furthermore, candidates responsible for decorating of diterpenoid alkaloid skeletons were discovered from transcriptome sequencing of fuzi, such as monooxygenases, methyltransferase, and BAHD acyltransferases. In addition, 645 differentially expressed genes encoding transcription factors potentially related to diterpenoid alkaloids accumulation underground were documented. Subsequent modular domain structure phylogenetics and differential expression analysis led to the identification of BAHD acyltransferases possibly involved in the formation of acetyl and benzoyl esters of diterpenoid alkaloids, associated with the acute toxicity of fuzi. The transcriptome data provide the foundation for future research into the molecular basis for aconitine-type C_{19}-diterpenoid alkaloids biosynthesis in *A. carmichaelii*.

1. Introduction

Aconitum, a large genus of the Ranunculaceae family, consists of approximately 400 species distributed in the temperate regions of the northern hemisphere, with 211 of these species in China (Li and Kadota, 2001; Ma et al., 2015). As one widely used traditional Chinese medicines (TCM), the tubers of *Aconitum* are commonly used for the treatment of various conditions like rheumatic fever, painful joints, gastroenteritis, diarrhoea, oedema, bronchial asthma, tumours, and some endocrinal disorders like irregular menstruation, and some of these uses have been validated by modern pharmacology (reviewed by Nyirimigabo et al. (2015)). In addition, our previous studies have investigated the biological activity of this

* Corresponding author. Department of Biological Sciences, Lehman College, City University of New York, New York, USA.
** Corresponding author.
E-mail addresses: ydq@xtbg.ac.cn (D. Yu), Edward.kennelly@lehman.cuny.edu (E.J. Kennelly).
[1] Dake Zhao, Yong Shen, and Yana Shi contributed equally to this work.

https://doi.org/10.1016/j.phytochem.2018.04.022
0031-9422/

important TCM ingredient (Zhao et al., 2013, 2017; Guo et al., 2017). Of the 211 recognized *Aconitum* species in China, *A. carmichaelii* Debx., known in Chinese as 附子 (fuzi, which translates to "lateral root"), has been used in TCM for over 1000 years (Yu et al., 2016), and it is one of two species officially recorded as *Radix Aconite* in the *Chinese Pharmacopoeia* (Chinese Pharmacopoeia Commission, 2015). Ethnobotanical investigations have shown that fuzi is used specifically against certain diseases and is perceived to be effective as well (Kang et al., 2012; Yu et al., 2016).

The diterpenoid alkaloids (DAs) and their derivatives, such as the C_{18}, C_{19}, and C_{20} classes, are the principal constituents responsible for the biological activity of *Aconitum* (Nyirimigabo et al., 2015; Singhuber et al., 2009; Wang and Chen, 2010; Weber et al., 2015). To date, 75 C_{19}-DAs and 18 C_{20}-DAs, have been identified from the root of *A. carmichaelii* (Wang and Liang, 2002; Zhou et al., 2015). Furthermore, Yan et al. (2008) found that the DAs are mainly distributed in fuzi tubers, while their content is very low in the leaves, clearly demonstrating the enriched accumulation in roots. However, the content of toxic DAs in the tubers makes this important TCM particularly dangerous. The cardio- and neurotoxicities of these DAs are potentially lethal, and the improper use of *Aconitum* in China, India, Japan, and elsewhere still results in cases of human poisonings (Nyirimigabo et al., 2015). The toxicity of *Aconitum* DAs arises from their effect on voltage-gated sodium channels, the release of neurotransmitters and changes in receptors, the promotion of lipid peroxidation and cell apoptosis in the heart, liver and other tissues (Zhou et al., 2015). The principal toxic and pharmacological ingredients of fuzi are aconitine-type C_{19}-DAs (Yu et al., 2016; Zhou et al., 2015). Ester hydrolysis of fuzi can transform diester-DAs (DDA) into monoester-DAs (MDA) or ester-free amine-DAs (ADA), which have markedly lower toxicities. For example, the toxicity of MDA in rats was at least 64–180 times lower than that of DDA after hydrolysis (Wen et al., 2013), while their pharmacological effects did not weaken (Tong et al., 2013; Zhou et al., 2015). The C_{19}-DAs in the processed roots of *Aconitum* after hydrolysis are possibly the more beneficial bioactive constituents of fuzi (Singhuber et al., 2009).

DAs, so-called pseudoalkaloids, are characterized by complicated structural features with a large number of stereocenters. A variety of diterpenoids are produced in higher plants and fungi (Toyomasu, 2008), while DAs only emerge in a limited number of taxa, including *Aconitum* (Devkota and Sewald, 2013), indicating the specialized biosynthetic pathways that evolved in these species to produce DAs. Though a variety of DA skeletons have been isolated and their chemical synthesis characterized, their biosynthesis, such as aconitine-type C_{19}-DAs in fuzi, is still sparse (Devkota and Sewald, 2013; Dewick, 2009; Rai et al., 2017). Based upon the previous metabolic research and chemical syntheses, the biosynthesis of DAs has been initiated by the production of diterpenes from the linear primary metabolite (*E,E,E*)-geranylgeranyl diphosphate (GGPP) via enzymatic-mediated cyclization and Wagner–Meerwein rearrangements. The alkaloid skeleton is produced at a late stage after the formation of diterpenes through enzymatic transamination from L-serine, thus converting diterpenes to the C_{20}-diterpenoid alkaloid scaffold and then forming the various C_{19}- and C_{18}-alkaloids by sequential Wagner–Meerwein rearrangements or carbocationic cyclization followed by subsequent chemical modifications such as oxidation, and in some cases, further rearrangement, methylation, and acylation (Xiao et al., 2006; Rai et al., 2017; Chen and Baran, 2009; Devkota and Sewald, 2013; Hong and Tantillo, 2010; Ichinohe, 1978; Kodama et al., 1975; Wang and Chen, 2010; Weber et al., 2015; Zhao et al., 2009).

Rai et al. (2017) also conducted a transcriptome analysis of *A. carmichaelii*, but mainly focused on the candidate genes in the mevalonate (MVA) and methylerythritol (MEP) biosynthetic pathways involved with GGPP, the precursor of DAs; however, few unigenes after GGPP were revealed. In this paper, we analysed both the aerial parts and rootstocks of *A. carmichaelii* using RNA-Seq. This study seeks to discover genes that encode for enzymes involved in aconitine-type C_{19}-DAs biosynthesis via comparison of genes expressed in rootstocks as compared to leaves. By a transcriptomic analysis utilizing next-generation sequencing (NGS), we identified the novel *ent*-CPP synthases (CPS), *ent*-kaurene synthases (KS), kaurene oxidases (KOX), cyclases, aminotransferases, monooxygenases, methyltransferase, and BAHD acyltransferases that are likely to be involved in the biosynthesis of C_{19}-DAs of fuzi, together with transcription factors (TFs) potentially associated with the accumulation of DAs. With the development of synthetic biology, this information could be used to regulate the DA biosynthetic pathway, and potentially produce DAs in higher quantities or with more bioactivity. This data in turn may be helpful for the molecular breeding of *A. carmichaelii*.

2. Results

2.1. De novo assembly and functional annotation

Using NGS applied to RNA extracts of aerial parts and rootstocks (see methods), six libraries of non-redundant unigenes were assembled, consisting of 74,122 and 92,947 genes, respectively, and a total of 95,812 transcripts (SRA accession number in NCBI: PRJNA378328). All unigenes ranged from 201 bp to 15,893 bp, with an average of 595 bp. Given a set of contiguous sequences, each with its own length, the N50 length is defined as the shortest sequence length at 50% of the genome. N50 statistic defines assembly quality in terms of contiguity, and thus N50 with 844 bp indicated a good sequencing result in the study. The unigenes were annotated based on similarities to the sequences in public databases, like the NCBI non-redundant protein (Nr) database, the Protein family database (Pfam), the Swiss-Prot protein database (Swiss-Prot), the KEGG Orthology database (KO), and the Clusters of Orthologous Groups of proteins database (KOG/COG), and Gene ontology (GO) with an E-value of less than 10^{-5}. GO analysis provides functional classification of genes, which defines the properties of genes and their products. A total of 19,703 transcripts were annotated using the GO database, and 3 GO categories (biological process, cellular component, and molecular function) including 24 sub-categories were identified (Fig. 1). Six GO categories regarding 'biological process' (namely metabolic processes, cellular processes, responses to stimuli, localization, establishment of localization, and biological regulation) each accounted for over 10% of the total annotated transcripts. Five GO categories regarding 'molecular function' (namely catalytic activity, binding, transporter activity, structural molecule activity and translation regulation activity) each accounted for over 10% of the total annotated transcripts. All the assembled unigenes were also subjected to the COG database to further evaluate the functional predictions and classifications and were assigned to 25 COG functional category classifications (Fig. 2). The category of general function prediction was the largest group, followed by transcription, post-translational modification, protein turnover and chaperones, replication, recombination and repair. In particular, categories associated with DA biosynthesis were observed such as specialised metabolite biosynthesis, transport and catabolism, and amino acid transport and metabolism.

2.2. Differential expression analysis and enriched specialised metabolic pathways of identified DEGs

The expression of read-mapped genes was analysed based on