

2018

**MORPHOLOGICAL, GENETIC AND BIOCHEMICAL
CHARACTERIZATION OF THE ANTI-MALARIAL HERB, *Artemisia
annua* GERMPLASM COLLECTION AT WEST VIRGINIA
UNIVERSITY**

Delini K. Samarasinghe
dksamarasinghe@mix.wvu.edu

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>



Part of the [Biochemistry Commons](#), [Plant Biology Commons](#), and the [Plant Breeding and Genetics Commons](#)

Recommended Citation

Samarasinghe, Delini K., "MORPHOLOGICAL, GENETIC AND BIOCHEMICAL CHARACTERIZATION OF THE ANTI-MALARIAL HERB, *Artemisia annua* GERMPLASM COLLECTION AT WEST VIRGINIA UNIVERSITY" (2018). *Graduate Theses, Dissertations, and Problem Reports*. 3691.
<https://researchrepository.wvu.edu/etd/3691>

This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

MORPHOLOGICAL, GENETIC AND BIOCHEMICAL CHARACTERIZATION OF THE ANTI-MALARIAL HERB,
Artemisia annua GERMPLASM COLLECTION AT WEST VIRGINIA UNIVERSITY

Delini Kanchana Samarasinghe

Thesis submitted to the Davis College of Agriculture, Natural Resources and Design at West Virginia
University in partial fulfillment of the requirements for the degree of

Master of Science

in

Genetics and Developmental Biology

Vagner A. Benedito, Ph.D., Chair

Nicole Waterland, Ph.D.

Kang Mo Ku, Ph.D.

Division of Plant and Soil Sciences

Morgantown, West Virginia

2018

Keywords: Malaria, *Artemisia annua*, Artemisinin, Breeding, Morphology, GC-MS, RT-qPCR

Copyright 2018 Delini Kanchana Samarasinghe

ABSTRACT

MORPHOLOGICAL, GENETIC AND BIOCHEMICAL CHARACTERIZATION OF THE ANTI-MALARIAL HERB, *Artemisia annua* GERMPLASM COLLECTION AT WEST VIRGINIA UNIVERSITY

Delini Kanchana Samarasinghe

Malaria is one of the deadliest diseases in human history. Nearly half of the world's population, is at the risk in 106 countries. Only in 2016, this disease killed about 445,000 people, 72% of them being children under age five. It also accounts for US \$12 billion dollars of direct costs in Africa alone. Five different species of *Plasmodium* cause malaria but *P. falciparum* is the most detrimental one, causing 50% of all malaria cases and is considered as the deadliest parasite in humans. Artemisinin (ART), a 15 C sesquiterpenoid is currently the only precursor to the most effective anti-malarial drugs. The World Health Organization (WHO) recommends the use of ART-derived drugs in combination with other anti-malarial drugs (collectively called as Artemisinin-based Combination Therapies, ACT), to prevent the development of ART resistance by the parasite. The natural source of ART is *Artemisia annua*, an Asteraceae plant native to East Asia. ART is synthesized in 10-cell glandular trichomes and accumulates extra-cellularly in the sub cuticular space. ART is synthesized in *A. annua* in low quantities, about 0.01-3% DW and, due to its complex molecular structure, it cannot be synthesized chemically in an economically viable way. Even though the semi-synthesis of ART has been successful in bioreactors, ART extraction from *A. annua* is much more economically viable. Therefore, it is important to understand the *A. annua* genetic contribution to ART synthesis to improve the ART yield of the plants. We characterized a collection of 12 different genotypes of *Artemisia annua* at West Virginia University in terms of developmental traits, biochemical profile, and gene expression to better understand the phenotypical differences of the germplasm in the collection. Plants were analyzed for height, biomass, leaf shape and leaf size, adaxial and abaxial trichome densities for morphological characterization. Freeze dried leaf material was extracted with hexane and an internal standard and the extractions were analyzed with GC-MS. Flash frozen leaf material were used to extract total RNA and analyzed for the gene expression of 15 genes with qRT-PCR. Based on the results, it was revealed that there is high phenotypical and metabolic variation among the genotypes. Monoterpenes made up for a large component of the hexane-extracted metabolome. Camphor and endo-borneol contributed significantly to differentiate the genotypes. Out of the all the genes analyzed, gene expression of three genes (FPS, CPR, and DBR2) were significantly different among the genotypes. Based on data, genotype J has the highest accumulation of ART and DHAA with anatomical characteristics supporting towards high ART and DHAA yield. Next, I and G genotypes also showed anatomical features supporting towards high ART and DHAA accumulation. The genotypes can be used to further investigate ideal genetic variations and allelic combinations for crop improvement in future.

ACKNOWLEDGEMENTS

I would like to thank Dr. Vagner Benedito, my advisor for the opportunity given and for all his advices, guidance and all the help during my master's degree program. More than anything, I am grateful to Dr. Benedito for the freedom given in the lab that let me to grow as a researcher. I would like to thank Dr. Kang Mo Ku and Dr. Nicole Waterland for their guidance and advices given not only to make this research a success but also to improve myself in conducting a given research. I would like to thank Dr. Youyoun Moon for his initial guidance during this research project and spending his time to give me a solid understanding about essentials and basics of scientific research.

I would like to thank the Green house staff at WVU, Sue Myers, Gail Sikorsky and Silas Child for their help and care given to maintain this valuable collection of plants and during the experiments from taking care of plants to harvesting. I would like to thank Dr. Jianbo Yao and his lab members for generosity in sharing their resources during this research. I am very thankful to all my past and present colleagues and undergraduate students for all the big and small assistances I got from their own ways.

On a personal note, I would like to thank my husband for all his support and sacrifices made from the very first day of the classes to thesis submission to make this journey a smooth and a successful one. I would like to thank all my friends and siblings, especially for my sisters for their encouragements and for cheering me up when I needed it the most. I would like to thank my parents-in-law for their love and care I got during this time. And I am grateful for my parents for showing me what patience is and how not to give up by example. Finally I am very grateful to venerable Mankadawela Sudassana thero, for helping me to stay indifferent to both good and bad times and for teaching what the true happiness of life is which helped me to enjoy the current moment or in other words enjoy the journey without waiting for the destination.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	iv
OVERVIEW	vi
CHAPTER 1	1
The historical importance and current status of malaria in the world	1
Glandular secretory trichomes	6
The artemisinin biosynthesis pathway	7
Goal of this Research	12
CHAPTER 2	13
Morphological characterization of <i>Artemisia annua</i> germplasm collection at West Virginia University	13
INTRODUCTION	13
MATERIAL AND METHODS	15
RESULTS AND DISCUSSION	16
CONCLUSIONS	23
CHAPTER 3	25
Biochemical profiling of <i>Artemisia annua</i> genotypes in the germplasm collection at WVU	25
INTRODUCTION	25
MATERIAL AND METHODS	26
RESULTS AND DISCUSSION	26
CONCLUSION	40
CHAPTER 4	42
Transcriptional analysis of genes involved in the artemisinin pathway in the <i>Artemisia annua</i> germplasm collection at WVU	42
INTRODUCTION	42
MATERIAL AND METHODS	43
ABA receptor (AaPYL9)	44
RESULTS AND DISCUSSION	44
CONCLUSION	48
CHAPTER 5	50

Genetic improvement of <i>Artemisia annua</i>: honing domestication and boosting yields of artemisinin and dihydroartemisinic acid.....	50
INTRODUCTION.....	50
The general isoprenoid/terpenoid biosynthesis pathway.....	52
The committed ART biosynthesis pathway.....	53
Transmembrane transport - plastid to cytosol delivery of IPP and artemisinin efflux to the secretory reservoir.....	54
Breeding for high ART accumulation via targeted allelic selection of natural variants.....	54
Genome editing as a potential approach.....	58
Boosting ART and DHAA yields in <i>Artemisia annua</i> through crop management and post-harvest practices.....	60
Conclusions.....	60
SUMMARY AND CONCLUSION.....	61
REFERENCES.....	62
APPENDIX 1.....	75
Biomass.....	75
Height.....	75
Leaf Area.....	75
Adaxial and abaxial trichome density.....	76
APPENDIX 2.....	77
Anova table.....	77
PLS-DA loadings.....	87
PLSDA VIP values.....	88
Metabolite correlation coefficients.....	90
Pval correlation table.....	101
Metabolite yield.....	117
APPENDIX 3.....	119
mRNA Abundance.....	119
mRNA abundance graphs.....	123

OVERVIEW

Artemisinin (ART) is the sesquiterpenoid the precursor of the main component that the World Health Organization (WHO) recommends as treatment for *Plasmodium falciparum* malaria (WHO, 2006). Artemisinin is a secondary metabolite naturally synthesized in the annual herb *Artemisia annua* (Asteraceae). ART is produced only in *A. annua* in small quantities, 0.01-1.4% of dry weight and its supply does not meet the global demand, which results that ART-based treatments (ACTs) are costly and the prices fluctuate drastically. Even though the semi-synthesis of ART using yeast (*Saccharomyces cerevisiae*) as a heterologous system has been successfully achieved, the production costs have made it a pricier alternative (Pelpow, 2016). Therefore, *A. annua* remains the most economically viable source for the synthesis of ART.

The major goal of this research is to characterize phenotypically and genetically the twelve genotypes of *Artemisia annua* germplasm collection at West Virginia University regarding their accumulation of artemisinin (ART), dihydroartemisinic acid (DHAA) and artemisinic acid (AA), and compare the morphological and genetic factors that determine the levels of those metabolites within the genotypes. Understanding the factors that drive high ART and DHAA yield is important to breed for superior *A. annua* genotypes that can be used to fulfill the demand for ART-based active pharmaceuticals on time and at a reasonable and constant price.

The first chapter of this thesis is a brief introduction to the importance of *A. annua* in malaria treatments and factors related to ART biosynthesis pathway with the current available literature.

The second chapter presents the morphological characterization of the twelve genotypes of *A. annua* in the germplasm collection at West Virginia University. Plant height, biomass of the aerial system, leaf area as well as adaxial and abaxial trichome densities were measured and compared among the genotypes.

The third chapter shows the GC-MS analysis of non-polar secondary metabolites extracted with hexane. Fifty-two peaks were identified with National Institute of Standards and Technology (NIST) mass spectral library and of those 52 compounds, two of which (camphor and endo-borneol) were confirmed with authentic standards after pre-statistical analysis. The online MetaboAnalyst (4.0) software was used to analyze the relative abundance of compounds.

The fourth chapter deals with the gene expression analysis performed to understand the transcriptional differences of the ART biosynthesis pathway genes among the twelve genotypes in the study. RT-qPCR was used to compare the mRNA levels for each gene between genotypes relative to a housekeeping gene using the SYBR Green fluorescence system.

The fifth chapter discusses how to apply the knowledge gained from this research and already available information in the scientific literature towards the breeding of genotypes with high ART and DHAA yield and how to overcome the inherent barriers associated with *A. annua* genome.

The sixth and the last chapter presents the summary and conclusions reached with the data collected and information produced in this research.

CHAPTER 1

The historical importance and current status of malaria in the world

Malaria is one of the deadliest disease of human history. Nearly half of the world population, 3.2 billion is at the risk in 106 countries. Only in 2016, malaria killed about 445,000 people in the world, 72% of them being children under age 5. It also accounts for US\$ 12 billion dollars direct cost only in Africa (CDC, 2018). Earlier it was believed that humans and chimpanzees each inherited cerebral malaria parasite-like infections from their common ancestor and these parasites co-evolved with their hosts (Escalante and Ayala, 1994). But it is now believed that cerebral malaria parasite infection arose in humans after acquiring it from a gorilla likely within the past 10,000 years (Liu et al., 2010b; Sundararaman et al., 2016). Some researchers believed that half of all human who ever lived died from malaria (Whitfield et al., 2002). A less known fact is that the Center for Disease Control (CDC) was founded in 1946 mainly to control malaria in the U.S. Even though malaria has been eradicated from many countries, it still threatens the lives of about half a million lives annually.

There are five different *Plasmodium* species that cause malaria in humans: *P. falciparum*, *P. malariae*, *P. vivax*, *P. ovale* and *P. knowlesi*. *P. falciparum* is the most common in African region, accounting for 99% of reported cases in 2016 whereas *P. vivax* is the most common on American, South East Asian and Eastern Mediterranean regions (WHO World Malaria Report, 2017). Although the U.S. did not have any reported cases in recent years, there are areas in the country susceptible to the presence of the *Anopheles* mosquito vector, such as: Southern Florida and Hawaii, and not to mention unincorporated territories, such as Puerto Rico and Guam. Therefore, it is of US interest not only to prevent the introduction of the disease in the country, but also to mitigate the disease by providing foreign aid with the aim to eradicate the disease globally.

Malaria is characterized by an acute febrile illness with an incubation period of seven days or longer (WHO International Travel and Health Report, 2018). Between the two most common parasites causing malaria, *P. falciparum* and *P. vivax*, the former causes the most dangerous version of the disease, called the malignant form, whereas *P. vivax* causes less severe symptoms, called the benign form (Vogel, 2013). Initial mild symptoms of falciparum malaria can be fever, chills, headache, muscular aching and weakness, vomiting, cough, diarrhea, and abdominal pain. If not treated, more severe symptoms can occur, such as acute renal failure, pulmonary edema, generalized convulsions, circulatory collapse, following by coma and death (WHO International Travel and Health Report, 2018). Even though the initial mild symptoms of the disease may not be easily recognizable as malaria, it must be treated early because falciparum malaria can be fatal if treatment is delayed beyond 24 hours after the onset of clinical symptoms. Therefore, it is important to assume the possibility of falciparum malaria in cases of

having an unexplained fever, seven days after the first possible exposure to three months or rarely later after the last possible exposure to the malaria (vector) (WHO International Travel and Health Report, 2018). Other forms of malaria are not generally considered as fatal, except for a few cases of severe *P. vivax* recently reported. Additionally, *P. vivax* and *P. ovale* can remain dormant in the liver and cause relapses months after the exposure (WHO, 2018).

Although most temperate countries have eliminated the malaria threat (**Figure 1**), according to the WHO's World Report of Malaria 2017 (WHO, 2018), 216 million cases of malaria have reported in the world last year, with 90% of the cases from Africa whereas 7% and 2% from Southeast Asia and Eastern Mediterranean regions, respectively. In 2016, an estimated 445,000 deaths have been reported in which 91% are from African region followed by 6% from East Asian region (WHO World Malaria Report, 2017) Out of the 91 countries reporting indigenous malaria cases in 2016, 80% of the burden is carried by India and 14 of sub-Saharan African countries (WHO World malaria report, 2017).

Even though anyone can be infected with malaria, the young and the elderly, pregnant women and immunosuppressed individuals are considered to be at greatest risk when living and travelling in areas where the disease is prevalent. It is a well-known fact that, despite the negative effects of the sickle cell disease, this trait protects against malaria infection, even in asymptomatic (heterozygous) individuals (Serjeant, 2010) – it's therefore thought that the trait first evolved as a protection against the disease in endemic regions of Africa.

Quinine (an ingredient of tonic water) had been traditionally used as medicine against malaria and is still listed an essential medicine by WHO (WHO World Malaria Report, 2015). However, over time the protozoan developed resistance against it, making it less effective (Le Bras and Durand, 2003). This quinoline alkaloid was first isolated from the bark of the *Cinchona officinalis* tree (Rubiaceae) and is still used against chloroquine-resistant *P. falciparum* when the artemisinin-derivative (artesunate) is not available (Esu et al., 2014). The current malaria vaccine (called RTS,S) has low efficacy against *P. falciparum* and is not recommended by WHO to prevent the disease in most cases (WHO World Malaria Report, 2016).

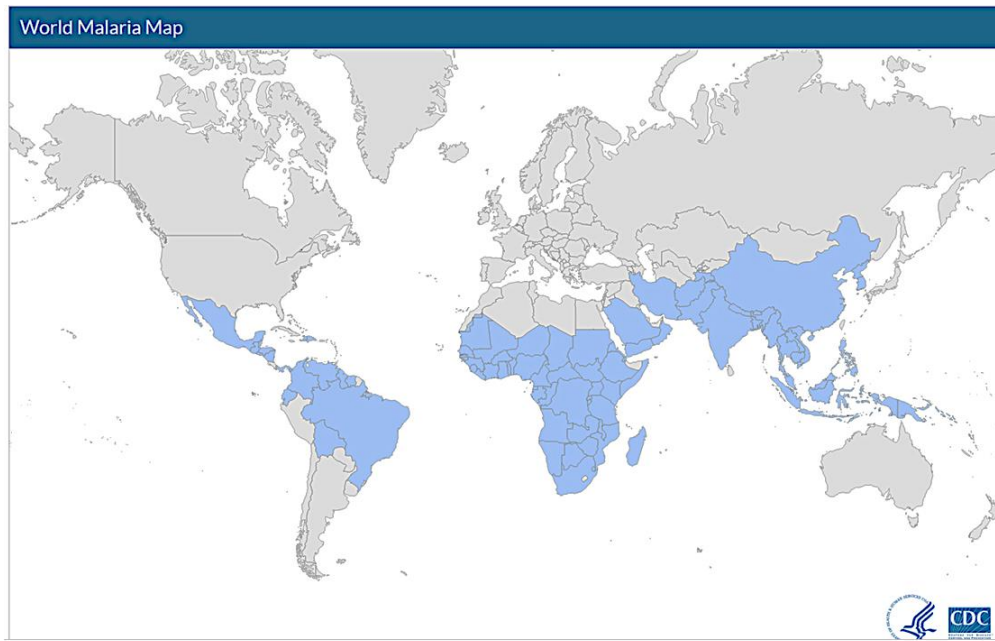


Figure 1. World malaria distribution 2018. Countries under the threat of malaria (in blue). Note that the disease level varies substantially within the 91 countries with reported cases. The Sub-Saharan region in Africa accounts for more than 90% of the global cases, mostly due to *Plasmodium falciparum*. Source: Global Health - Division of Parasitic Diseases and Malaria - Center for Disease Control (CDC) (2018).

***Artemisia annua* and the sesquiterpene lactone, artemisinin**

Artemisia annua is the only commercial source for artemisinin (ART), the precursor for main component of the malaria treatment recommended by WHO against *P. falciparum*: Artemisinin combination therapy (ACT). Due to diverse reasons, this plant remains the most important source of artemisinin in the market. However, the supply has consistently not met the global demand for the medicine (Kindermans et al., 2007; Shretta and Yadav, 2012; Pandey and Pandey-Rai, 2016).

Species of the genus *Artemisia* (family Asteraceae → subfamily Asteroideae → supertribe Asterodae → tribe Anthemideae) are well known for their medicinal, cosmetic, pharmaceutical and culinary values. It consists of a large clade encompassing 200-400 species, including wormwood (*A. absinthium*), an ingredient of absinthe, the culinary herb tarragon (*A. dracunculus*), and the anti-malarial herb, sagewort or sweet wormwood (*A. annua*). The exact number of species is unknown due to a plethora of synonyms. The name of this genus refers to the wilderness goddess Artemis in the Greek pantheon (she is the equivalent to Diana in the Roman mythology).

This genus consists of annual, biennial and perennial herbs and shrubs distributed in Asia, North America, Africa, and Europe (Pandey and Singh, 2017) with a greatest concentration in Asia (Abad et al., 2012). Because of copious secondary metabolite production, many species are

bitter and have strong odor that evolved to overcome herbivory and that have been harnessed to treat human maladies and other uses. Many metabolites synthesized in *Artemisia* spp. is known to have antimalarial, cytotoxic, antibacterial, antifungal, and antioxidant properties (Abad et al., 2012; Pandey and Singh, 2017). And among them, *A. annua* is well known and researched for its principle component Artemisinin (ART), the antimalarial drug precursor for recommended for multi- drug resistance to *P. falciparum*.

Artemisia annua L. is native to temperate Asia (China and Vietnam) used in traditional Chinese medicine for thousands of years as treatment for malaria. It has been naturalized to North America, South America and Europe (Ferreira and Janick, 1995). It is an annual short-day plant with a critical day length of 13.5-h. The species is allogamous and naturally cross-pollinated by wind (Ferreira and Janick, 1995). Currently, *A. annua* is the only natural source of ART, the precursor of semi-synthetic anti-malarial drugs. In 2015, Youyou Tu was awarded the Nobel Prize of Medicine for the discovery of artemisinin (ART) as the anti-malarial active compound in *A. annua* (Su and Miller, 2015).

ART is a highly oxygenated sesquiterpene with a 1,2,4-trioxane ring structure (**Figure 2**) that is believed to be the reaction center that confers its antimalarial activity (Brown, 2010) against the asexual stage of the *P. falciparum* (Bryant et al., 2015). ART-based semi-synthetic drugs are used in the so-called Artemisinin Combination Therapies (ACTs) as the standard treatment for chloroquine- and quinine-resistant *P. falciparum*, which causes cerebral malaria, as well as against other *Plasmodium* species (Weathers et al., 2006; Covello, 2008; Pandey and Pandey-Rai, 2016). ART is not only effective against malaria but also other parasitic diseases, such as schistosomiasis (*Schistosoma* spp.), and leishmaniasis (*Leishmania* spp.) (Krishna et al., 2008). ART was also effective to treat some viral diseases, such as hepatitis B (Nguyen et al., 2011; Rai et al., 2011). Additionally, ART was indicated to be effective against several cancer cell lines and drug-resistant cancers (Weathers et al., 2006).

A. annua genotypes can be classified into two chemotypes based on their geographic origin that produce different levels of artemisinin and its precursors: dihydroartemisinic acid (DHAA) and artemisinic acid (AA). The chemotype clones high in ART and DHAA but low in AA is from Vietnam while the Chinese (which includes the European-American clones) are high in AA but low in ART and DHAA (Wallaart et al., 2000).

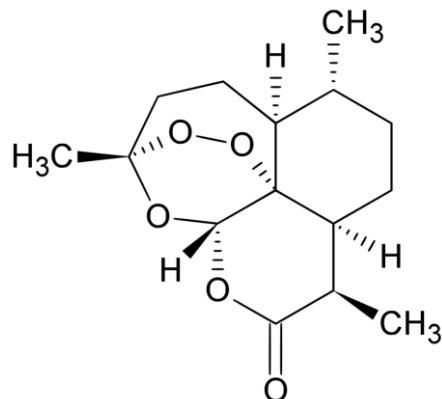


Figure 2. Structure of artemisinin. The endoperoxide group in the 1,2,4-trioxane ring is considered as the active site of this 15-C sesquiterpene molecule.

As the main component of ACT, there is a very high global demand for ART (Shretta and Yadav, 2012). It is estimated that the global demand for ART in 2017 reached 176 metric tons with an expected demand of 216 metric tons for 2021 (Global malaria diagnostic and artemisinin treatment commodities demand forecast, 2018). However, due to the difficulties in cultivating *A. annua* as a field crop, since this species can remain highly undomesticated, ART demands are hardly met (Graham et al., 2010). There have been research efforts to find alternatives to ART production, including the use of biotechnology, such as yeast, bacteria, and algae bioreactors (Ro et al., 2006; Chang et al., 2007; Paddon and Keasling, 2014; Ikram and Simonsen, 2017; Choi et al., 2017; Badshah et al., 2018) as well as tobacco leaves (Farhi et al., 2011; Fuentes et al., 2016). Indeed, the entire artemisinin pathway up to DHAA has been introduced into prokaryotic (*E. coli*) and unicellular eukaryotic systems (budding yeast), as well as moss (*Physcomitrella patens*) and artemisia hairy roots for bioreactor semi-synthesis of DHAA (Paddon and Keasling, 2014; Khairul Ikram et al., 2017; Patra and Srivastava, 2016) in addition to tobacco (Fuentes et al., 2016). However, ART produced in bioreactors becomes too expensive for governments in developing countries and the target populations that need the medicine the most. *In planta* production by *A. annua* field cultivation still remains the cheapest and most economically viable method of producing ART (Delabays et al., 2001; Ikram and Simonsen, 2017).

As a secondary metabolite, ART is synthesized in low quantities in plants (0.01-3% dry weight). Due to the complex nature of its molecular structure, the complete chemical synthesis is not economically viable (Kumar et al., 2004). ART semisynthesis in transgenic yeast has been achieved recently (Paddon et al., 2013). During its semisynthesis in yeast, the natural mevalonic acid (MVA) pathway was improved to increase the biosynthesis of farnesyl pyrophosphate (FPP – the precursor of all sesquiterpenes) and lower the endogenous sterols production (Meadows et al., 2016). Subsequently, by introducing the gene amorpho-4,11-diene synthase (*ADS*) along with a novel cytochrome P450 from *A. annua* (*CYP71AV1*) to first cyclize FPP to amorpho-4,11-diene (AD), which is the first committed metabolite of the ART pathway, and further oxidize it to artemisinic acid (AA), respectively. Finally, the extracted AA can be converted to ART *in vitro*

via catalytic hydrogenation to yield dihydroartemisinic acid (DHAA), and further processed by esterification of the carboxyl bond in DHAA, a reaction with chemically generated singlet oxygen ($^1\text{O}_2$), to produce an allylic hydroperoxide, and finally a series of acid-catalyzed reactions and oxygen to finally yield ART (Paddon et al., 2013). Even though the semisynthesis of ART has been successful industrially, ART extraction from *A. annua* is much more economically viable: the semi-synthetic ART costs US\$350-400/kg vs. US\$250/kg of the naturally extracted from *A. annua* (Peplow, 2016).

Glandular secretory trichomes

Trichomes are epidermal protrusions found on aerial organs of plants to provide physical and chemical protection. As a general feature of the Asteraceae family, *A. annua* bears two types of trichomes: glandular secretory trichomes (GST) (**Figure 3**), as well as filamentous T-shaped non-glandular (tector) trichomes (Ferreira and Janick, 2009). Trichomes are found on the aerial organs of *A. annua* such as stems and leaves (Duke and Paul, 1993) as well as on the corolla and floral receptacles (Ferreira and Janick, 1995). GSTs are evenly distributed on both leaf surfaces, except the leaf margin, but with a higher density on the abaxial surface. *A. annua* GSTs are stalkless (sessile) and consist of ten cells arranged in a biserial manner. Each GST consists of a pair of basal cells, a pair of stalk cells, three pairs of secretory cells, and a bilobed sac made by the cuticle of six secretory cells (Duke and Paul, 1993). Stalk cells contain smaller chloroplasts compared to mesophyll chloroplasts with an average area of $6.1 \mu\text{m}^2$ compared to $6.5 \mu\text{m}^2$ of subapical cells containing large, amorphous plastids lacking starch grains (modified chloroplasts), and the apical cells that contain no chloroplasts (Kjaer et al., 2014). All GST cells have a large amount of smooth endoplasmic reticulum associated with secondary metabolite synthesis and exporting (Duke and Paul, 1993). The main role of GST is considered as a protection against various biotic stresses by producing a large array of secondary metabolites. GSTs protect the plants from herbivores by releasing irritant secondary metabolites by excretion or rupturing the sac upon touch or by ingestion of bitter or poisonous metabolites (Kjaer et al., 2014).

A. annua GSTs are the synthesis site of about 600 secondary metabolites, mostly terpenoids, including ART. Based on evidence of transcription activity of key ART biosynthesis genes, GST are considered as the sole sites of ART biosynthesis. A GST laser dissection study showed that the expression of four genes (ADS, CYP7AV1, DBR2 and ALDH1) occurs in both, apical and subapical cells (Olofsson et al., 2012). Despite of the importance of GST to synthesize and store ART, and plethora of compounds in diverse species, the molecular mechanisms of GST development remain generally poorly understood. Duke and Paul (1993) reported the initiation of GST from foliar cells begins in the youngest leaf primordia and complete the differentiation into the 10-cell structure before the start of the leaf development meaning that the final number of the GST is predetermined regardless of the final leaf size. The GST density peaks in the young, fully expanded leaf and decreases rapidly due to rupturing of the

subcuticular sac (Duke and Paul, 1993) caused by mechanical damage. Singh et al. (2016) showed that the overexpression of β -glucosidase increased the trichome density in *A. annua* (Singh et al., 2016). Similarly, Shi et al. (2018) demonstrated that the overexpression of AaMIXTA1, a MYB transcription factor that is expressed mainly in basal cells of GST, increased the number of GSTs and ART content in *A. annua* (Shi et al., 2018). On the other hand, attempts to increase GST densities by applying abiotic stresses have not been successful (Kjær et al., 2012).

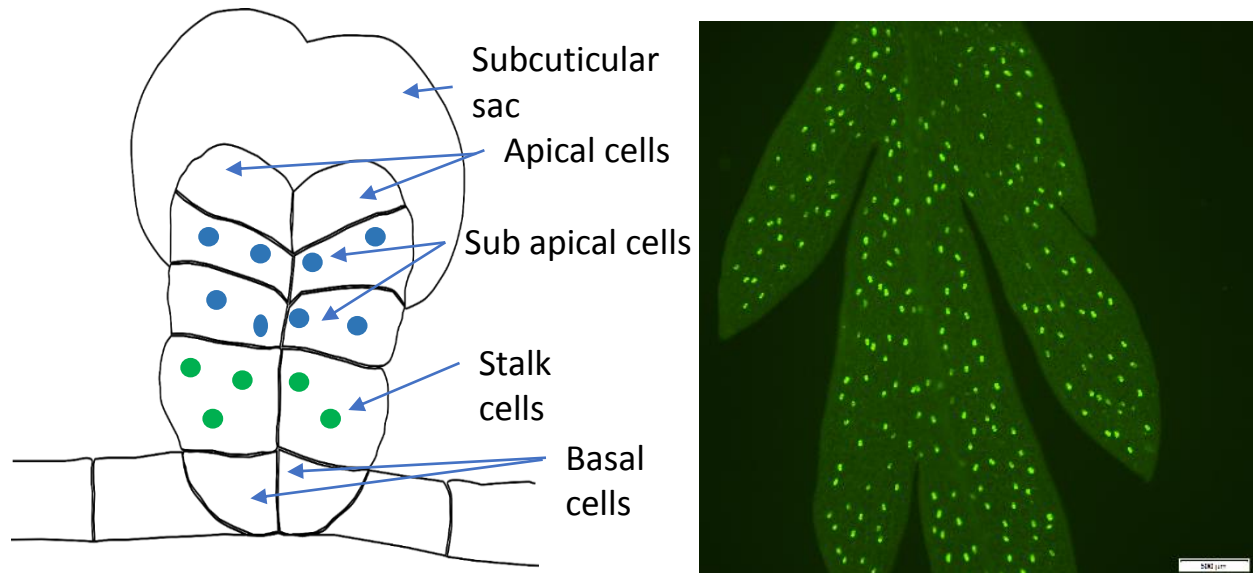


Figure 3. Structure of the glandular trichome of *A. annua*. Left: A schematic diagram of GST indicating cell arrangements. Right: Fluorescence microscopic image of GSTs on the abaxial side of the *A. annua* leaf (Scale bar = 500 μ m).

The artemisinin biosynthesis pathway

ART is a sesquiterpene lactone made up of three isoprene units. It originates in the cytosol from precursors derived from the cytosolic mevalonate (MVA) pathway as well as the plastidial methylerythritol phosphate (MEP) pathway (Weathers et al., 2011) (**Figure 4**). Although there are two independent pathways that synthesize the two basic isoprene units [isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP)], the predominant route in plants is the cytosolic MVA pathway, which originates from acetyl-CoA, whereas the plastidial MEP pathway originates from pyruvate and glyceraldehyde 3-phosphate (G3P) (Weathers et al., 2006; Wen and Yu, 2011).

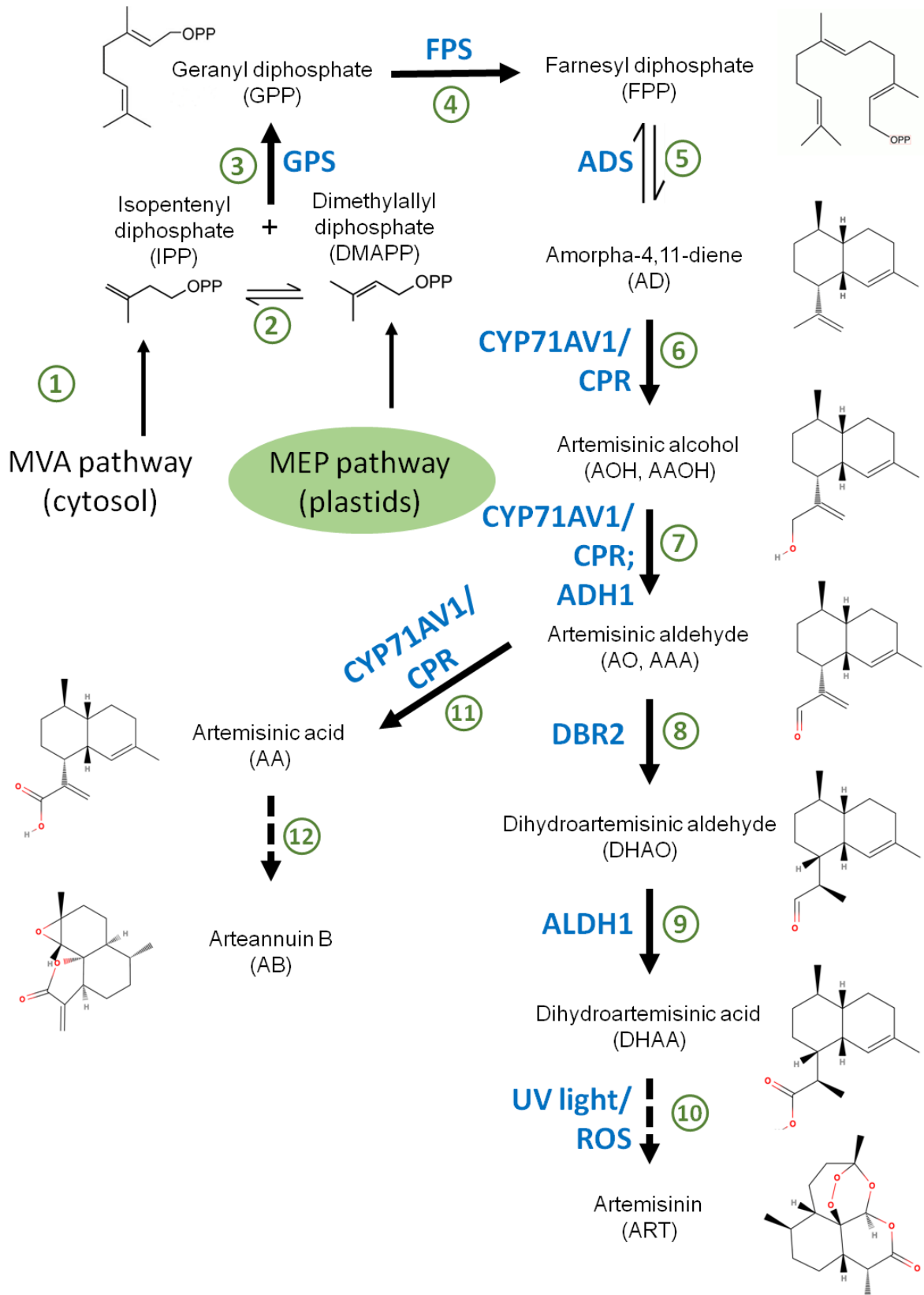


Figure 4. Biosynthesis pathway of the sesquiterpene lactone, artemisinin in glandular trichomes of *Artemisia annua*. **1.** Plant terpenoids are synthesized via two distinct pathways, the cytosolic mevalonate pathway in the cytosol and methyl-erythritol phosphate pathway in plastids, with both contributing with IPP and DMAPP isomeric moieties. **2.** Interconversion of IPP into DMAPP is carried out by IPP isomerase. The steady-state IPP: DMAPP isomeric ratio in *A. annua* cells is estimated in 6:1. Therefore, the availability of the more reactive moiety (DMAPP) is a major bottleneck of the general terpenoid pathway. **3.** In the cytosol, GPP synthase (GPS) fuses IPP and DMAPP to produce GPP, which is the precursor of all terpenoids. **4.** FPP is produced by farnesyl synthase (FPS) by the fusion of GPP and IPP. FPP is the precursor of all ≥ 15 -carbon terpenes. **5.** Amorpho-4,11-diene synthase (ADS) synthesized the first committed product of the artemisinin pathway, AD. ADS is a limiting enzyme and key to increase the flux of the pathway. **6.** Following, the cytochrome P450 monooxygenase CYP71AV1 oxidases AD to AA in three successive reactions (**7** and **11**). CYP71AV1 is the main rate-limiting enzyme of the committed artemisinin pathway. Moreover, the associated cytochrome P450 reductase (CPR) is necessary to restore the active state of CYP71AV1. **7.** Whereas AO can be produced by the CYP71AV1/CRP enzymatic system, the alcohol dehydrogenase 1 (ADH1) can also AO using AOH as a substrate. **8.** AO is disputed by the enzyme artemisinic aldehyde $\Delta 11(13)$ reductase (DBR2), which generates DHAO, and the third reaction of CYP71AV1 (**11**), which is unproductive and generates AA. **9.** DHAO is converted into DHAA by aldehyde dehydrogenase 1 (ALDH1). **10.** The conversion of DHAA into ART is carried non-enzymatically, possibly by ultraviolet light and/or reactive oxygen species generated in the cell by stress. **11.** The third reaction of CYP71AV1 turns AO into AA, deviating the pathway from producing ART. **12.** AA can be converted non-enzymatically into AB. (Chemical structures were captured from www.molview.org).

However, unlike most of other plant terpenes, ART is composed of basic isoprene units originated from both pathways. First, Towler and Weathers (2007) showed that inhibition of MVA and MEP pathways respectively with mevinolin and fosmidomycin, decreased ART biosynthesis (Towler and Weathers, 2007). Later, Schramek et al. (2010) brought up a possible scenario for the contribution of both pathways to synthesize the artemisinin precursor (FPP) by isotopologue profiling of *A. annua* plants with $^{13}\text{CO}_2$. From the ^{13}C NMR spectroscopy results, they proposed that mostly DMAPP of MVA origin is first transferred to the plastid and then used to synthesize geranyl diphosphate (GPP), which is subsequently exported to the cytosol for further elongation to form farnesyl diphosphate (FPP) with a unit of IPP from the MVA pathway (Schramek et al., 2010; Wen and Yu, 2011). Even though the ART pathway is not completely understood yet, as a sesquiterpene, its biosynthesis can be split into two stages, namely the synthesis of the linear precursor (FPP), followed by cyclization and modifications into the final product. While the enzymes for the first stage are evolutionarily conserved across species, enzymes for the second stage are species-specific and result in distinct terpenoid profiles (Wen and Yu, 2011).

Therefore, FPP is synthesized via the activity of several enzymes using precursors from both, MVA and MEP pathways. In the MVA pathway, 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR) catalyzes HMG-CoA to MVA and is considered the first rate limiting enzyme in the terpenoid pathway (Wen and Yu, 2011). Transgenic and chemical studies using plant hormones (e.g., auxin and gibberellin) showed that HMGR is a direct regulator of ART biosynthesis in *A.*

annua (Wen and Yu, 2011). In the MEP pathway, 1-deoxy-D-xylulose-5-phosphate synthase (DXS) and 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXR) catalyze the two rate-limiting steps and were also shown to be regulated in ART biosynthesis (Wen and Yu, 2011).

The first committed step towards ART biosynthesis is the cyclization of FPP to amorpha-4,11-diene by the trichome-specific sesquiterpene cyclase called amorpha-4,11-diene synthase (ADS) (Bouwmeester et al., 1999; Mercke et al., 2000; Wang et al., 2011). This volatile amorpha-4,11-diene is subsequently oxidized to artemisinic alcohol and then to artemisinic aldehyde by the trichome-specific cytochrome P450 monooxygenase, CYP71AV1 along with its redox partner, the cytochrome P450 reductase (CPR) (Teoh et al., 2006; Fuentes et al., 2016). Artemisinic aldehyde can be oxidized further by the same enzyme pair and continue to synthesize arteannuin B (AB) via artemisinic acid (AA). AB and AA are side products of the pathway without an anti-malarial effect. Therefore, they compete with the productive ART pathway.

Alternatively, artemisinic aldehyde can be used as a substrate to artemisinic aldehyde D11 (13) reductase (DBR2) in order to form dihydroartemisinic aldehyde (Zhang et al., 2008). Dihydroartemisinic aldehyde is then oxidized by aldehyde dehydrogenase (ALDH1) to form dihydroartemisinic acid (DHAA) (Zhang et al., 2008), which is considered the immediate precursor for ART *in vivo*. Conversion of DHAA to ART is believed to be a non-enzymatic photooxidative reaction that occurs in the subcuticular space of glandular trichomes (Sy and Brown, 2002; Tang et al., 2014). The presence of large number of oxygenated terpenoid compounds in *A. annua*, such as terpene allylic hydroperoxides and endoperoxides, is evidence used to support this hypothesis, because these compounds are assumed to be intermediates of auto-oxidized terpene precursors (Brown, 2010; Fuentes et al., 2016). On the other hand, some reports suggest the possibility of enzyme-mediated reactions in the final steps of ART biosynthesis (Ferreira et al., 2018).

Genetic regulation of the ART biosynthesis pathway in *Artemisia annua*

Apart from the structural genes that code for enzymes directly involved in the biosynthesis pathway, there are a number of transcription factors found to control artemisinin biosynthesis and accumulation (Shen et al., 2016). The first transcription factor to be isolated and characterized in *A. annua* was AaWRKY1, which is highly expressed in glandular trichomes. AaWRKY1 is strongly induced by methyl jasmonate and has the ability to bind to the W-box of the ADS promoter in order to activate gene expression in both, transgenic tobacco as well as transiently in *A. annua* leaf system (Ma et al., 2009). Also, AaWARKY1 overexpression enhanced the expression of the trichome-specific CYP71AV1. HPLC metabolite profiling of AaWRKY1-overexpressing *A. annua* leaves showed a twofold increase of ART compared to the control line (Han et al., 2014).

In *A. annua*, four AP2/ERF transcription factors have been related to ART biosynthesis: AaORA, AaERF1, AaERF2, and AaTAR1. AaORA is highly expressed in both GSTs as well as tector trichomes. Its overexpression in *A. annua* highly induced the transcription of ADS, CYP71AV1 and

DBR2 genes (Lu et al., 2013). Two other jasmonic acid response AP2/ERF transcription factors, *AaERF1* and *AaERF2*, identified from GST cDNA libraries are highly expressed in the inflorescence and have a similar spatial expression pattern as *ADS* and *CYP71AV1* genes (Yu et al., 2012). Lastly, *TAR1* is expressed at high levels in young leaves and inflorescence and is involved with wax deposition, cuticle permeability, trichome morphology as well as ART biosynthesis. *TAR1* overexpression resulted in increased ART, DHAA and AA in leaves and flower buds, demonstrating its important role as a regulator of ART biosynthesis and trichome development (Tan et al., 2015).

Additional transcription factors are involved in ART biosynthesis. *AaBbZIP1* regulates ART biosynthesis by binding to ABRE motifs in the promoters of *ADS* and *CYP7AV1* to increase their expression (Zhang et al., 2015). *AabHLH1* enhances ART biosynthesis by binding to E-box *cis*-elements in *ADS* and *CYP71AV1* promoters (Ji et al., 2014). *AaMYC2* binds to the G-box-like *cis*-elements in *CYP71AV1* and *DBR2* gene promoters in order to boost the ART biosynthesis in *A. annua* (Tang et al., 2014).

Flavonoids in *A. annua*

A. annua is a species abundant in secondary metabolites. So far about 600 secondary metabolites belonging to eight categories have been described in *A. annua*. Out of these, terpenoids form the largest group whereas flavonoids are the second largest category (Brown, 2010). In fact, *A. annua* is reported to be one of the four medicinal plants with highest ORAC (oxygen radical absorbance capacity) value, with 1,125 and 1,234 μ moles of Trolox equivalents/g in leaf and inflorescence extracts, respectively (Ferreira et al., 2010). The high antioxidant capacity of *A. annua* is attributed to the high phenolic content of *A. annua*, greatly encompassed by flavones and flavonols (Ferreira et al., 2010). Eleven flavones and 29 flavonols have been reported mainly constituting the flavonoid portion of the profile, including diverse polymethoxylated flavonoids (Ferreira et al., 2010).

Different flavonoids are known to play different biological functions, including protection against ultraviolet radiation and as ROS scavengers during biotic and abiotic stress conditions (Agati et al., 2011). Flavonoids offer an antioxidative protection by preventing the formation of ROS through their ability to chelate transition metal ions such as Fe^{2+} and Cu^{2+} , as well as scavenging ROS once they are generated (Di Ferdinando et al., 2012), and by inhibiting ROS generating enzymes (Kumar and Pandey, 2013). Although the flavonoid *in vivo* mode of action regarding ROS scavenging remains unclear, their subcellular localization, on or in close proximity to the location of ROS generation, is consistent with their ROS scavenging role (Di Ferdinando et al., 2012).

Similar to ART, DHAA and AA, the flavonoid content of *A. annua* is related to the genetic composition of the two chemotypes (Ferreira et al., 2010). Therefore, their capacity to scavenge the ROS produced during stress conditions vary depending on the chemotype. Recently,

lyophilized powder of *A. annua* whole plants was shown to be efficient in treating malaria caused by artemisinin-resistant plasmodium, as well as to slow the evolution of drug resistance (Elfawal et al., 2015; Daddy et al., 2017). It is thought that, a set of secondary metabolites in the plant chemical profile, potentially including flavonoids, is acting synergistically to enhance the anti-malarial effect of ART. This is a testable hypothesis that deserves a systematic analysis.

Goal of this Research

This thesis aimed to characterize the differences between the twelve genotypes of the *A. annua* germplasm collection at West Virginia University regarding their physiological, biochemical and genetic properties. This characterization will lead us to better understand the genetic underpinnings leading to the development of favorable characteristics in order to improve breeding efforts of *A. annua* as a resilient crop in the field with maximum ART yield.

CHAPTER 2

Morphological characterization of *Artemisia annua* germplasm collection at West Virginia University

INTRODUCTION

Artemisia annua is an Asteraceae short-day annual herb that naturally grows to about 30-100 cm in height with alternate branches. The aromatic leaves of the plant are pinnately dissected with a varying number of sections per leaf, depending on the maturity of the organ. The length and width of the leaves can vary between 2.5-5 and 1-3 cm, respectively. The flowers are miniscule and yellow, arranged in loose panicles with capitula (World Health Organization, 2006). Plants prefer cross pollination via insects and wind over selfing, making them highly allogamous (Ferreira and Janick, 1995).

Even though *A. annua* is the main commercial source for artemisinin (ART), the main component of *Plasmodium falciparum* malaria treatment, current commercial hybrids produce ART in relatively low quantities, 1.4% d.w. (Ferreira et al., 2018). Therefore, understanding the properties of the plant favorable toward a high ART yield is essential. As dihydroartemisinic acid (DHAA), which is the immediate precursor of ART, can be easily converted to ART *in vitro*, aiming for *A. annua* genotypes with high dihydroartemisinic acid (DHAA) and ART will not only improve the natural ART yield from plants but also the total yield by also taking into account DHAA yield (Ferreira et al., 2018).

West Virginia University (WVU) is home to twelve different genotypes of *A. annua* (**Table 1, Figure 1**) generated from different origin cultivars. Originally, the germplasm collection was compiled and maintained by Dr. Jorge F. S. Ferreira when working at USDA/ARS in Beaver, WV. Seeds of a high ART yielding Brazilian cultivar, 3M and the Swiss cultivar, Artemis were grown and plants that were high in ART and DHAA were selected. Similarly, seeds from cv. Sandeman were grown and a genotype was selected with high artemisinic acid (AA) levels. All of these seed-generated genotypes were maintained vegetatively under the greenhouse conditions. The collection was passed down to Prof. Vagner Benedito at WVU when the research station closed down in 2012. The most striking difference among these genotypes is in the accumulation levels of ART and its precursor DHAA, as well as the non-productive byproduct AA. The twelve genotypes were named A to L for the ease of use and their genotype identification number and the origin of the parental hybrid stocks are provided in **Table 1**.

Table1. Identification codes for *A. annua* genotypes in the germplasm collection at the West Virginia University.

Genotype	Code	Origin
3M13	A	Brazilian
3M29	B	Brazilian
3M39	C	Brazilian
3M49	D	Brazilian
3M51	E	Brazilian
3M83	F	Brazilian
3M85	G	Brazilian
Sandeman	H	United Kingdom
MDP4	I	Swiss
MDP7	J	Swiss
MDP31	K	Swiss
MDP11	L	Swiss

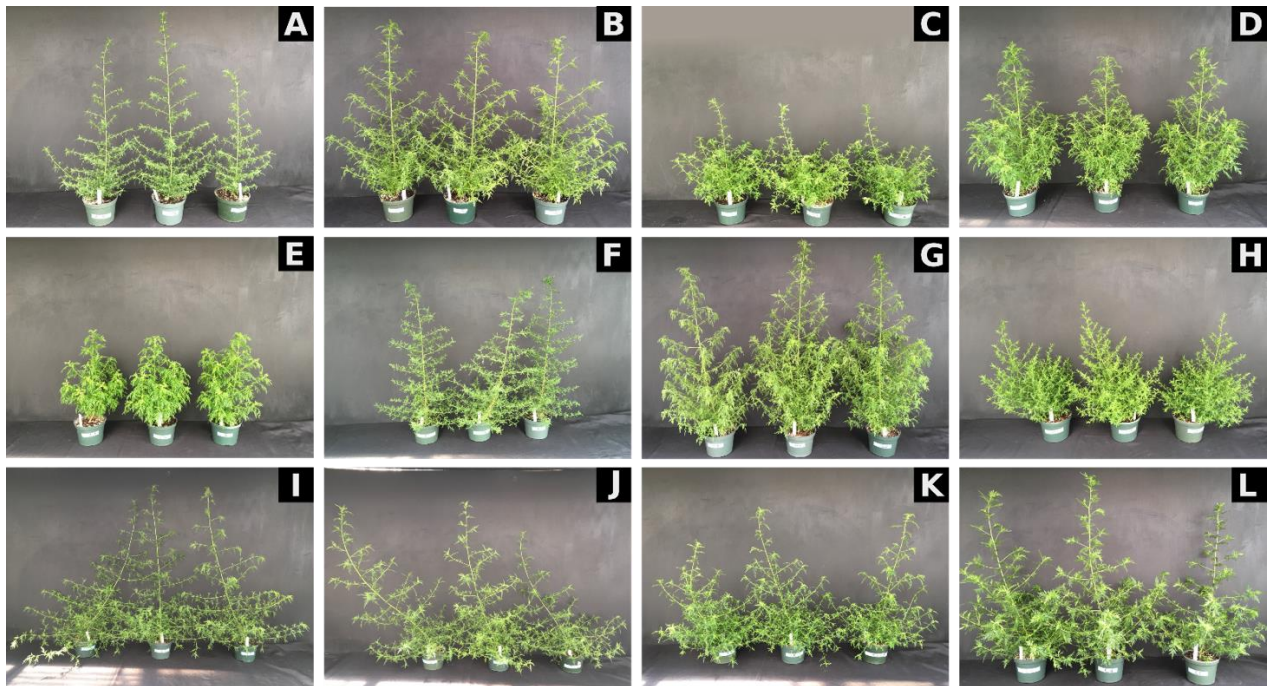


Figure 1. Germplasm collection of *A. annua* (Asteraceae) at the West Virginia University. The images show representative plants from each genotype grown for 45 days after rooting of 10-cm high cuttings.

Morphological analyses of *A. annua* genotypes in the WVU germplasm collection: Four main morphological characteristics that are important in determining ART production were analyzed: height of the plants, biomass of the aerial system of the plants, leaf shape and leaf area, and trichome densities.

MATERIAL AND METHODS

Maintenance of germplasm collection: stock plants were cultivated in 27-cm diameter round plastic pots with soilless medium (Sunshine Mix #1; Sun Gro Horticulture) under greenhouse conditions (25/20°C, average relative humidity 50%) and long-day regimen (16-18h light, depending on the season) to repress flowering and subsequent senescence. Plants were irrigated daily, as needed. Fertilization was carried out every other day with 20-10-20 fertilizer at 200 ppm. The major pests (mealybug, thrips, and spider mites) were treated with insecticides by spraying or drenching of Safari® systemic insecticide, as needed. Alternatively, biological control with the mite predator, *Mesoseiulus longipes*, was used.

Clonal propagation of stem cuttings: 10 cm cuttings were taken from healthy shoots of stock plants, dipped in rooting powder (Hormodin #3), transferred to 6-cell trays filled with soilless medium (Sunshine Mix #1; Sun Gro Horticulture) and placed in a mist room (21.6±2.6 °C/ 19.4±1.9 °C day/night (mean±SD), average daytime relative humidity 57.5% ±10%, 18 h light, on 18±2°C heating pad) for four weeks.

Experimental design: Six well-rooted cuttings (around 13-15 cm height) from each genotype were transferred to pots (15-cm diameter) with soilless medium (Sunshine Mix #1; Sun Gro Horticulture) and arranged into a randomized block design in greenhouse for 45 days. Average greenhouse temperatures were 23.8±1.7/ 19.8±1.6 °C day/night (mean±SD) and the average relative daytime humidity was 38.9±10.5%. The plants were irrigated with 20-10-20 (NPK) fertilizer at 200 ppm, daily to the field capacity. Harvest was carried out after 45 days of growing.

Plant height: Each plant was measured from the soil surface to apical meristem to the nearest millimeter after 45 days after transplanting.

Biomass of aerial system: Plants were harvested at the soil level, put in paper bags, and placed in a 40 °C oven for 15 days, until the dry weights were constant. Biomass of dry material was carried out in an analytical scale.

Leaf measurements: Leaf area was measured using young, fully expanded leaves (n=5) collected from random, lateral shoots of three representative stock plants for each genotype. Leaves were imaged with a HP6000 scanner and analyzed with ImageJ software (version 1.8.0).

Glandular secretory trichome (GST) density: Cuttings were taken and their stems were quickly placed in water to avoid wilting. GSTs were quantified with small leaf sections removed from the

tips of ten fully expanded leaves collected from three stock plants for each genotype. Epifluorescence images of both adaxial and abaxial surfaces of the same leaf section for each leaf were taken with an Olympus MVX10 fluorescence microscope with GFP filter (2X Objective, 2X slider and 1.2X magnification). ImageJ (Version 1.8.0) software was used for counting the fluorescent spots (denoting individual glandular trichomes) and to measure the area of the leaf section in order to calculate the trichome densities.

Statistical data analyses: Data were collected for each parameter with each plant representing an experimental unit, except for the trichome density measurements, for which each leaf represented an experimental unit. ANOVA and post-hoc multiple comparison analyses were carried out in JMP (v.13.2).

RESULTS AND DISCUSSION

Plant height

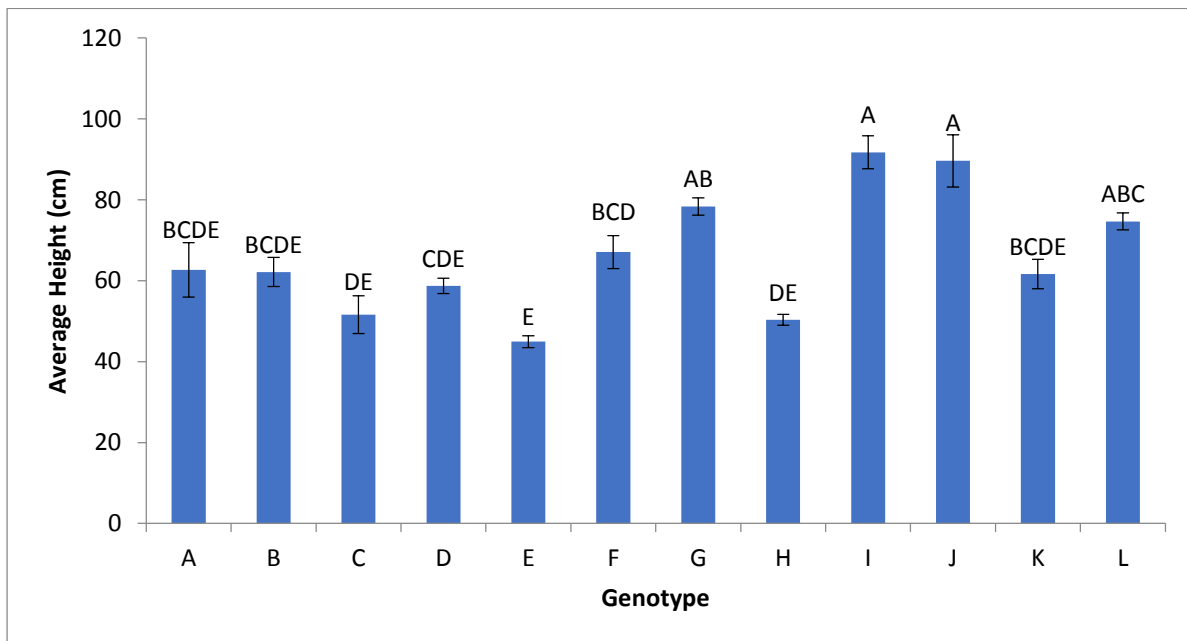


Figure 2. Height of *A. annua* germplasm collection. Height of twelve different genotypes of *A. annua* grown under greenhouse conditions for 45 days were measured from the soil surface to the apical meristem (n=6). Error bars indicate standard error and genotypes not connected by same letter are significantly different (Tukey's test, $\alpha=0.05$).

Average height for the collection ranged more than two folds, from 44.93 cm to 91.73 cm, with genotypes I and E, and J and E had the two most height differences among all twelve genotypes, 46.80 cm and 44.72 cm, respectively ($p<0.001$). *A. annua* naturally grows to a height of 30-100 cm and cultivated plants can grow up to 2 m (WHO monograph on GACP for *A. annua* L., 2006).

An interesting work by Thu et al. (2011) indicates the importance of environmental factors over the genetics, on vegetative growth and ART content (Thu et al., 2011). This study showed that two *A. annua* clones (TC1 and TC2) with characteristic late flowering and high ART yield (1-2%) responded differently to high (26.2-32.8 °C) and low (10.5-18.0 °C) temperatures by plants growing about four to five fold taller (TC1: 48.9 cm vs. 255.5 cm, TC2: 56.6 cm vs. 258.8 cm) for both clones and 4.3 and 3.4 fold higher ART content in TC1 and TC2, respectively, thus correlating to the robustness of the plants. They reasoned that the results attributed to high transpirational rates in warmer conditions caused the plants to grow shorter and spend more resources on root growth to increase water absorption (Thu et al., 2011). In an earlier experiment conducted by Elhag et al. (1992) to identify morphological characteristics associated with ART producing *A. annua* genotypes, they concluded that high ART chemotypes are associated with tall, robust plants with long internodes, open-branching, dense leaves and thick stems (Elhag et al., 1992), which is compatible with the data from our research. Interestingly, however, is that a recent research conducted by Czechowski et al. (2018) comparing the high-ART producing (HAP) hybrid Artemis and an open pollinated low-ART producing (LAP) *A. annua* variety mentioned that LAP plants are distinctly different from HAP for having taller and long internode plants with smaller leaves (Czechowski et al., 2018). It is likely that plant height, biochemical profiles, and ART accumulation are genetically unlinked traits, although there is a possibility that distinct loci involved in these processes are physically linked in the chromosomes of *A. annua*. Previous trait genetic mapping efforts (Graham et al., 2010) and the recent genome sequence and annotation (Shen et al., 2017) will likely help resolve this question and improve trait mapping in this species.

Dry weight of aerial plant system

The average dry weights (**Figure 3**) ranged more than three folds: from genotype A weighing an average of 7.51 g to genotype L weighing 23.81 g per plant. Half of the collection (L, I, G, D, J and H) were not significantly different amongst themselves regarding biomass but were significantly distinct from genotypes A, F, C and E ($p < 0.001$). Genotypes B and K were the most similar in terms of biomass with only 0.173 g difference on average. To obtain a more accurate idea about the useful biomass of the plants, it will be ideal to rather measure specifically the weight of leaves without stems in future experiments since the weight of stems does not significantly contribute to ART yield.

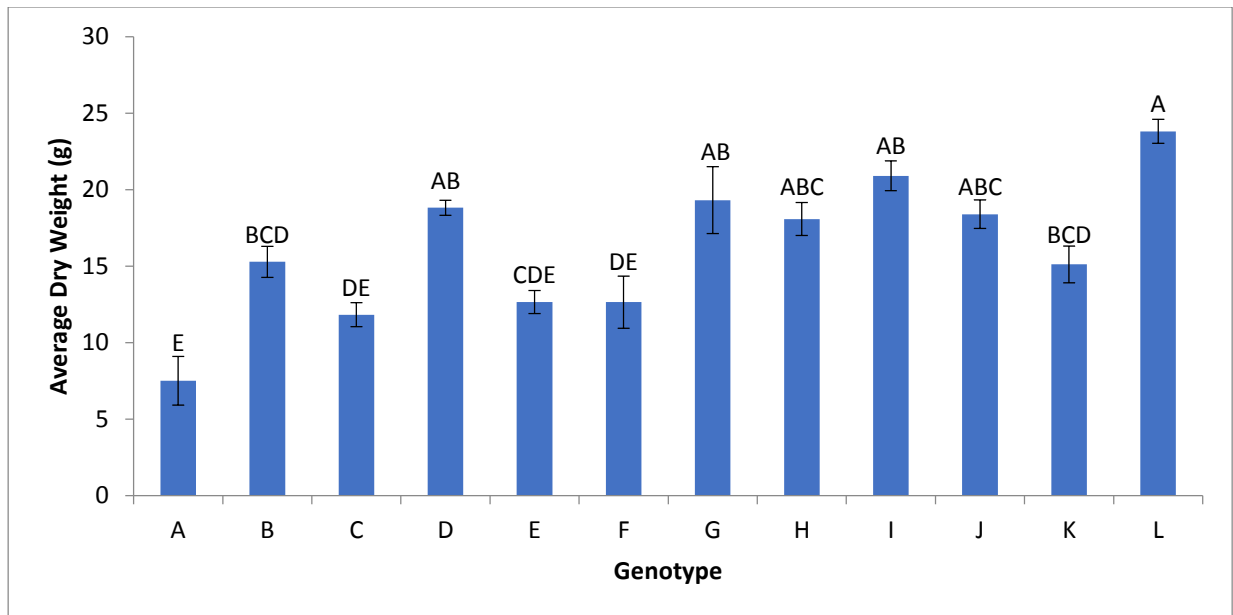


Figure 3. Dry weight of *A. annua* germplasm collection. Aerial parts of twelve genotypes of *A. annua* grown under greenhouse conditions for 45 days were oven-dried and average weight (n=6) was determined. Error bars indicate standard error and levels not connected by the same letter are significantly different (Tukey's test, $\alpha=0.05$).

Various factors affect biomass in both positive and negative ways. In a study conducted by Ferreira using the Artemis hybrid under greenhouse conditions (Ferreira, 2007), plants were provided 2 mg/kg soluble N, 28 mg/kg extractable P and 82, 439 and 192 mg/kg exchangeable K, Ca and Mg, respectively, with 5.08 soil pH level and grown for 90-102 days. These plants produced the highest biomass (70.3 g/plant) compared to untreated plants (6.18 g/plant). It was further shown that no-K treatment resulted in about 70% of leaf biomass of the complete nutrition treatment with higher ART levels (but not significant compared to complete treatment) whereas no-N supplementation resulted in about 20% of the leaf biomass of the complete treatment, which indicates the importance of N for biomass. Another study indicated that gradual increase of N application from 6 mg/L to 106 mg/L increased leaf and stem biomass as well as ART content. However, increasing N further up to about 206 mg/L and 306 mg/L did not increase biomass but impacted ART content negatively (Davies et al., 2009). In the same study, the authors showed that increased application of K had a positive impact on biomass without any effect on ART yield. Therefore, soil N level is likely to be a critical factor determining the biomass and the ART yield. Application of gibberellic acid has increased the biomass by stem elongation and leaf expansion with increased ART content (Banyai et al., 2011). Short-term drought did not affect negatively the biomass of container-grown *A. annua* plants but increased ART content (Marchese et al., 2010). Yang et al. confirmed this result by showing increased levels of ADS in shoots induced by drought (Yang et al., 2010). Certain minerals such as B have little to no effect on biomass yet improved ART content (Aftab et al., 2010). Soil pH level in the range of 5.5-6 was found to be best for maximum biomass yield and ART content (Laughlin, 1994; Ferreira, 2007). Since ART synthesis

depends directly on leaf area, a high biomass production is an essential factor contributing to increase ART yield regardless of other factors. Therefore, in breeding programs, it is important to select genotypes with high biomass along with other factors that positively contribute to the high ART content.

Leaf phenotypes

Artemisinin accumulates in glandular trichomes (GTs) mostly on the leaf epidermis. Theoretically, a large leaf area can create more space for the development of GTs, and hence increase the accumulation of ART and precursors (Graham et al., 2010). Since the leaf is the most important tissue of the *A. annua* plant as a crop, identifying genotypes with larger and less serrated leaves can be helpful for breeding programs. Moreover, finding plants with high trichome densities is another key factor in achieving superior *A. annua* plants. Therefore, we analyzed the collection for differences in leaf shape, size and trichome density among the genotypes in the WVU collection.

Leaf Morphology

A. annua develops small leaves, usually presenting five to seven deep serrated sections (leaflets). This compound shape is true for all genotypes of the collection. The smallest leaves in the collection are developed in genotypes A and H (**Figure 4**). A less serrated leaf with more lamina area is ideal for *A. annua*, since these leaves could potentially bear more GTs per plant. Therefore, more research on external and internal factors that determines the leaf shape in the Asteraceae clade will be potentially helpful in breeding for a better crop.

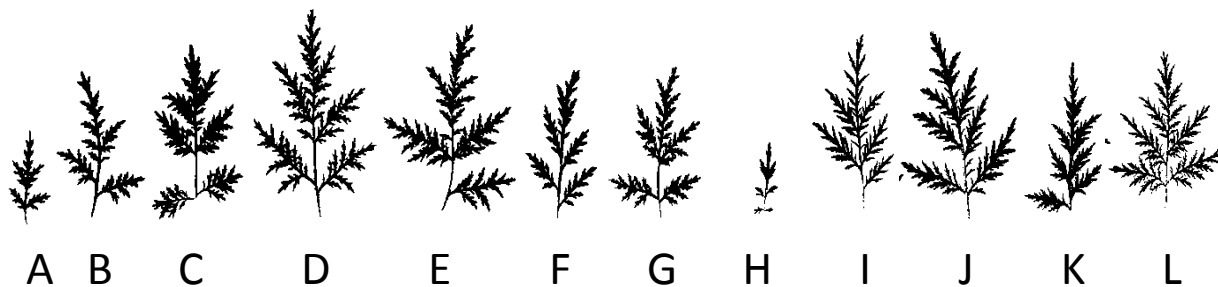


Figure 4. Leaf morphologies of *A. annua* germplasm collection. Scanned images of typical young, fully expanded leaves show the differences among the leaf size and shape of the genotypes.

Leaf Area

The trait leaf area ranged more than 20-fold in the germplasm collection (**Figure 5**). Out of the twelve genotypes, genotype H and genotype A have significantly the lowest leaf areas, 31.72 mm² and 141.27 mm², respectively. On the other hand, genotype D, E, J and K had the highest leaf area. Regarding the leaf area, genotypes from both origins (Brazil and Switzerland) showed great phenotypical variation. This points out to the great phenotypical variability in leaf area in this collection, which should be followed up by the determination of genetic and environmental factors controlling the development of this trait in *A. annua*.

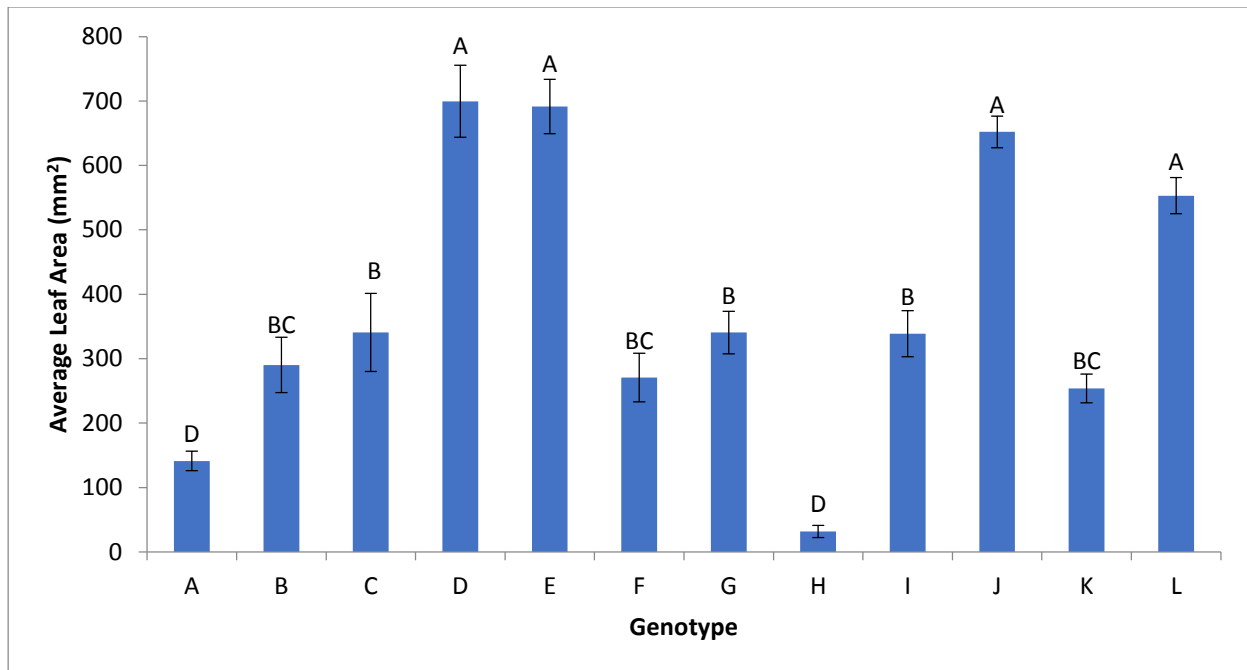


Figure 5. Leaf area of *A. annua* genotypes. Measurements were determined using fully developed leaves (n=5). Leaves were scanned and area was determined using ImageJ software. Error bars indicate standard errors and levels not connected by same letter are significantly different (Tukey's test, $\alpha=0.05$).

Glandular Trichome Densities

Glandular secretory trichomes (GSTs) are the exclusive site of ART biosynthesis and accumulation in *A. annua*. Therefore, GSTs are considered the most important structure of the plant from the economic point of view. This way, GST density is an essential factor when breeding for high ART yields – the total number of trichomes per area, along with how full these trichomes are and the chemical profile stored in them, will determine the ultimate plant yield. We determined both the adaxial and abaxial trichome densities for each of the twelve genotypes in the collection (**Figures 6 and 7** - notice the different scale of the y-axes).

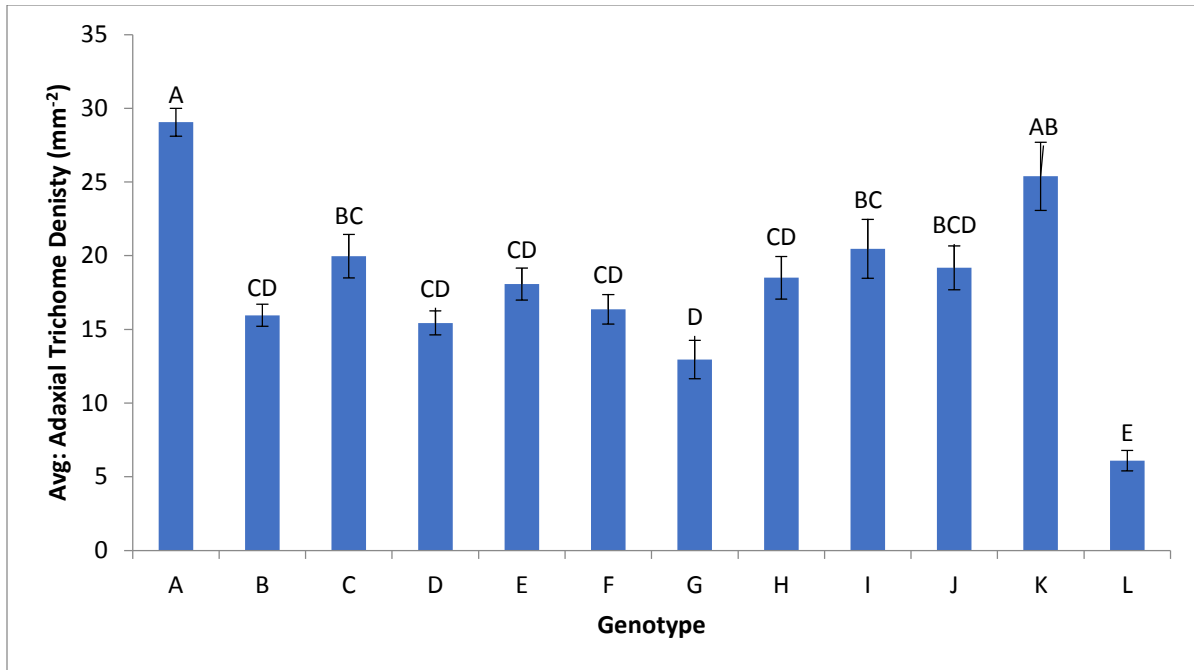


Figure 6. Adaxial trichome density of *A. annua* genotypes. Adaxial trichome density of young, fully expanded leaves (n=10) from three plants representing the twelve *A. annua* genotypes. Trichomes were determined using epifluorescence images taken using Olympus MVX10 fluorescence microscope with the GFP filter. Error bars indicate standard error and levels not connected by same letter are significantly different (Tukey's test, $\alpha=0.05$).

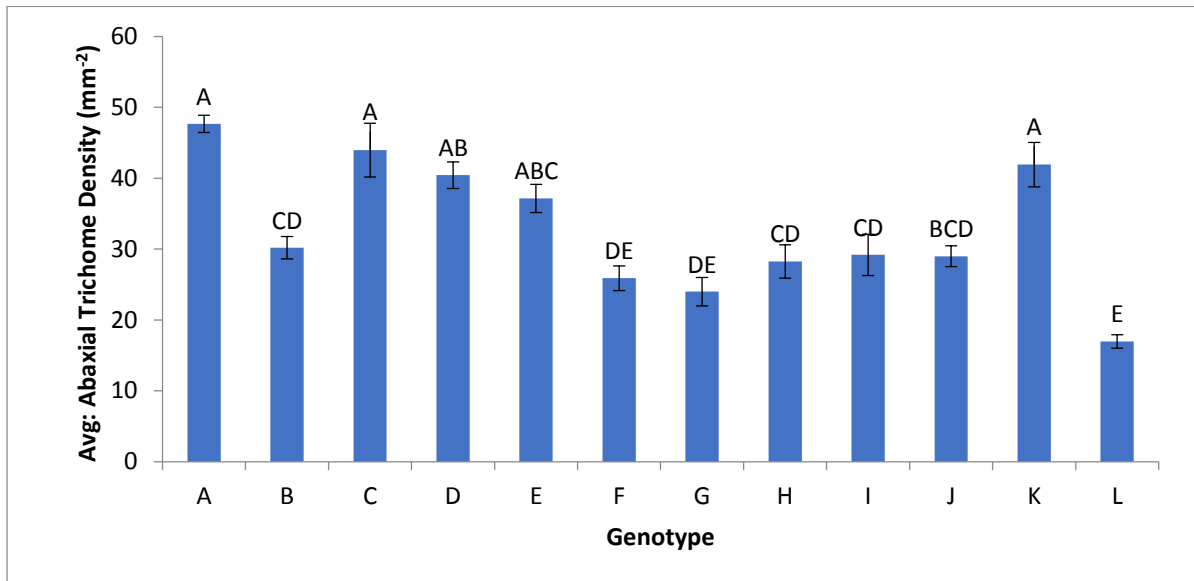


Figure 7. Abaxial trichome density of *A. annua* genotypes. Abaxial trichome density of young fully expanded leaves (n=10) from three plants representing the twelve *A. annua* genotypes. Trichomes were determined using epifluorescence images taken using Olympus MVX10 fluorescence microscope with the GFP filter. Error bars indicate standard error and levels not connected by same letter are significantly different (Tukey's test, $\alpha=0.05$).

In general, for *A. annua* genotypes in the collection, abaxial trichome densities were higher than the adaxial values (**Figures 6-8**). The adaxial trichome density ranged almost 5-fold, from 6.09 to 29.05 mm⁻² whereas abaxial trichome density ranged almost three-fold, from 16.95 to 47.66 mm⁻². On both surfaces, genotype L consistently ranked the lowest.

Fluorescence microscopy is becoming popular to easily visualize GSTs as they auto-fluoresce due to the essential oils stored in them (e.g., Graham et al., 2010). We used the same approach to quantify GSTs in the *A. annua* collection. The epifluorescence micrographs indicated an even distribution of GSTs on the adaxial leaf surface whereas, GSTs are distributed leaving a groove like space around the leaf veins on the abaxial leaf surface (**Figure 8**).

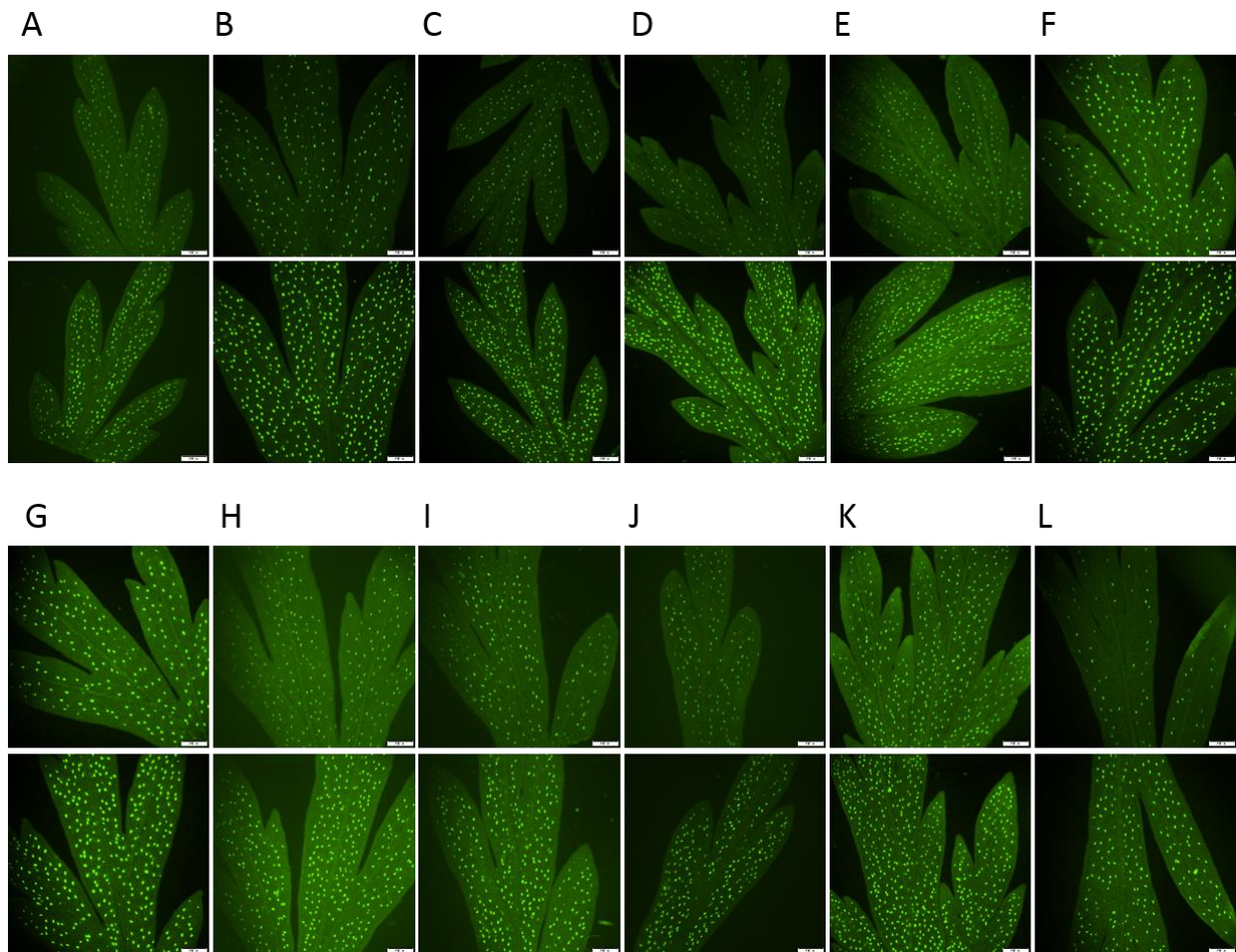


Figure 8. Leaf adaxial and abaxial trichome densities of *A. annua* germplasm collection. For each genotype (A-L), top rows indicate adaxial trichome and bottom rows indicate abaxial trichome densities. Scale Images were taken with Olympus MVX 10 fluorescence microscope with the GFP filter (2X objective, 2X slider and 1.2X magnification) and bars at the right bottom corners are 500 μ m.

The average trichome densities reported in more recent literature are between 10-75 mm⁻² (Kjaer et al., 2014) with significantly higher GST density on abaxial than the adaxial surface (Kjær et al., 2012), which is in agreement with our data. However, Hu et al. (1993) reported the

opposite pattern and the results of Arsenault et al. (2010) showed no difference between these densities (Hu et al., 1993; Arsenault et al., 2010).

CONCLUSIONS

In our morphological analysis of genotypes in the collection, the genotypes I, J, L, and G had the highest height and dry weight values. While genotype H was among the shortest plants, it showed high dry weight, which reveals a very dense plant architecture. Leaf area presented high variation among genotypes, with D, E, J and L showing highest values whereas A and H displayed the smallest leaf areas. Glandular trichomes were present on both leaf surfaces, but they were clearly more abundant in the abaxial surface. Genotypes A and K showed the highest adaxial densities whereas the genotypes A, K along with C, D and E are grouped with the highest abaxial trichome densities. Overall, genotype J had the best anatomical properties to support high ART yield.

Graham et al. (2010) described several anatomical and physiological factors that determine the ART yield, which is a product of ART concentration and plant fresh weight (Graham et al., 2010). They documented high phenotypic variation in the Artemis pedigree, such as ART and other metabolite concentration, leaf area, glandular trichome density, and plant fresh weight. These are the main traits targeted for increasing ART yield. They also confirmed the phenotypic variation with high degree of genetic variation with a mean SNP frequency of 1 in 104 base pairs.

Agronomic factors can also impact the secondary metabolite production in plants, since it is associated with the plant's interaction with its environment (Davies et al., 2009). Therefore, there is a high likelihood for physiological characteristics to be dependent on the interaction of plant with its environment than the genetics of the plant. Even for high yielding varieties, the yield potential is closely related to field management standards in *A. annua* (RBM/UNITAID/WHO Artemisinin Conference Final Report, 2011). Literature related to *A. annua* physiology and environment often presents contrasting results regarding the same trait. The main reason for such discrepancies could be due to how this genetically and chemically complex plant interacts with its macro- and micro-environment. Also, as secondary metabolites are not directly involved with the normal growth and the development of the plant, but they are rather synthesized as a defense mechanism against biotic and abiotic stresses, manipulation of growth conditions in the field is important to force plants to use their genetic potential to synthesize DHAA and ART in maximum quantities.

Currently, China and Vietnam produce about 80% of the global supply, with East Africa countries producing the remaining 20% (Shretta and Yadav, 2012). The average yield of *A. annua* varies between 1.5-2 tons of dry matter per hectare with varying 0.5-1.2% ART content, depending on seed quality, climate, altitude, soil, planting density (stand), and the farmers' knowledge and technology used. Considering a 50-80% extraction efficiency and purification yield, 6-14 kg of ART, the starting material for active pharmaceuticals is produced on average per hectare of *A.*

annua (Kindermans et al., 2007). Improving efficiencies in each step is a *sine qua non* condition to alleviate global supply shortages, especially including the use of high-yield varieties.

CHAPTER 3

Biochemical profiling of *Artemisia annua* genotypes in the germplasm collection at WVU

INTRODUCTION

Artemisia annua (Sweet wormwood, Qinghao) is a metabolite rich aromatic ethnomedicinal herb (Bhakuni et al., 2001). It has been used in Chinese culture since ancient times due to its strong aroma and in Chinese traditional medicine for more than 2,000 years to treat various ailments, such as cough, cold, and agues, including malaria (Wang et al., 2011). *A. annua* has been subject to intensive phytochemical evaluations since the discovery in 1970's of artemisinin (ART), the precursor to current WHO recommended treatment against *Plasmodium falciparum* malaria (Wang et al., 2011). Apart from ART, *A. annua* is a rich source of essential oils that gives the characteristic smell to the plant, and is used commercially in perfumery and cosmetics (Jain et al., 1996). *A. annua* essential oil is known to have insecticidal properties (Jain et al., 1996). It is rich in secondary metabolites with almost 600 known compounds dominated by terpenoids, particularly sesquiterpene lactone, in addition to flavonoids, coumarins and other shikimate metabolites (Brown, 2010).

Metabolomics can be defined as identification and quantification of metabolic responses of living systems to stimuli or genetic modifications (Worley and Powers, 2013). Therefore, metabolomics is the result of genomic, transcriptomic and proteomic responses in an organism to a given stimulus or change. It encompasses the final process in the cell resulted directly from enzymatic activities. Thus, metabolites are more proximal to the phenotype than genes, transcripts and proteins. Indeed, minimal changes in gene expression or protein accumulation does not necessarily correlate with variation in protein activity, but alterations in metabolite accumulation only occurs through such changes (Worley and Powers, 2013).

To study *A. annua* metabolites, various extraction, separation and identification methods have been tested (Christen and Veuthey, 2001; Van Nieuwerburgh et al., 2006; Peng et al., 2006; Czechowski et al., 2018). Among them, GC-MS approach is widely used (Ma et al., 2008). Various statistical methods are also used to simplify and extract more information from complex metabolomics data to understand key factors driving the metabolome. In our study, we used GC-MS to separate and identify compounds and PLS-DA analysis to select and extract information to best describe the twelve genotypes of *A. annua*.

Multivariate analysis methods are used in metabolomics studies to identify biologically relevant spectral features and two of the most common such methods are, principal component analysis (PCA) and partial least squares projection to latent structures (PLS) (Worley and Powers, 2013).

In this research, we used partial least squares discriminant analysis (PLS-DA) to differentiate the genotypes and to identify key features in characterizing each genotype.

MATERIAL AND METHODS

Plant material: The twelve genotypes that compose the *Artemisia annua* germplasm collection were used in this experiment. The experimental set up is described in Chapter 2. Juvenile and young, fully expanded leaves, up to about 12th leaf from apical meristem of the main stem and from one branch from the upper third of the plant was carefully stripped from the stem and packed into a 50 mL plastic vial. Meristem from the main stem was not included as it was harvested for gene expression analysis. The material was flash frozen with liquid nitrogen and stored at -80 °C and freeze-dried within two weeks of storage.

Metabolite extraction and analysis: Non-polar metabolites of 0.100 g freeze-dried leaf material were extracted with 5 mL hexane containing 50 µg/L tetracosane (C₂₄H₅₀) as an internal standard (IS). Extractions were analyzed using a GC-MS (Trace 1310 GC, Thermo Fisher Scientific) coupled to an MS detector system (ISQ QD, Thermo Fisher Scientific) with autosampler (Triplus RSH, Thermo Fisher Scientific). A capillary column (Rxi-5Sil MS, Restek, Bellefonte, PA, USA; 30 m × 0.25 mm × 0.25 µm capillary column w/10 m Integra-Guard Column) was used to detect polar metabolites. After an initial temperature hold at 60 °C, the oven temperature was increased to 200 °C at 7 °C/min, then increased to 320 °C at 15 °C/min and held for 5 min. Injector, MS transfer line, and ion source temperatures were set at 275 °C, 280 °C, and 300 °C, respectively. An aliquot of 1 µL was injected with the split ratio of 1:4. The helium carrier gas was kept at a constant flow rate of 1.2 mL/min. The mass spectrometer was operated in positive electron impact mode (EI) at 70.0 eV ionization energy at *m/z* 40–500 scan range. Metabolite identification was based on mass spectra of standard compounds and retention time or comparison with the in the National Institute of Standards and Technology (NIST). The metabolites ART, DHAA, AA, camphor and endo-borneol were confirmed using standards. The areas of the identified peaks were normalized to the IS area and the relative abundance of each predicted compound were calculated and normalized to the exact dry weight measurements. The data were statistically analyzed with MetaboAnalyst software (v.4.0) (Chong et al., 2018).

RESULTS AND DISCUSSION

A total of 52 peaks were identified using the NIST library database. During initial assessment, camphor and endo-borneol appeared as the two main metabolites that could be used to characterize the collection due to their significantly high *F* values and significantly low *p* values (**Appendix 2: Supplemental Table 1**) and due to high VIP (Variable Importance in Performance)

values (**Appendix 2: Supplemental Table 3**). Therefore, those two metabolites were further confirmed with authentic standards.

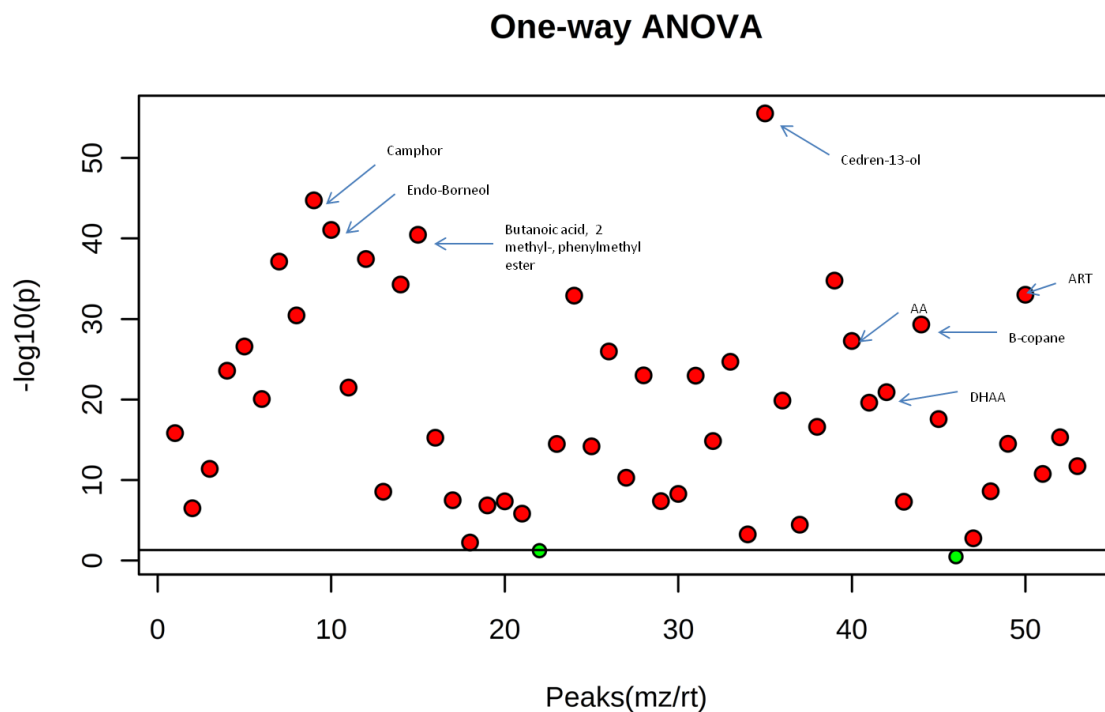


Figure 1. Relative abundance of 52 compounds detected by GC-MS analysis from 12 different *Artemisia annua* genotypes. Compounds on or below the solid line on the graph are not significantly different by Tukey's HSD, $\alpha=0.05$.

From the 52 compounds identified, one-way analysis of ANOVA and Tukey's post-hoc test at $\alpha=0.05$ (**Figure 1**) revealed that 50 compounds were significantly different among the genotypes. Based on the PLS-DA analysis, the first, second and third components described 49.3%, 19.8% and 9.7% of the variability in the dataset respectively (**Figure 2**). Camphor, endo-borneol, and (potentially) eucalyptol had the three highest impacts on the first three components with 0.8893, -0.92878 and 0.62501 loading values, respectively (**Appendix 2: Supplemental Table 2**). Loading values are coefficients of the linear combination of original variables that makes up the components. Therefore, they indicate the weights that each original variable has on the principal components in PLS-DA analysis; i.e., the higher the loading value, the higher the significance of the corresponding variable.

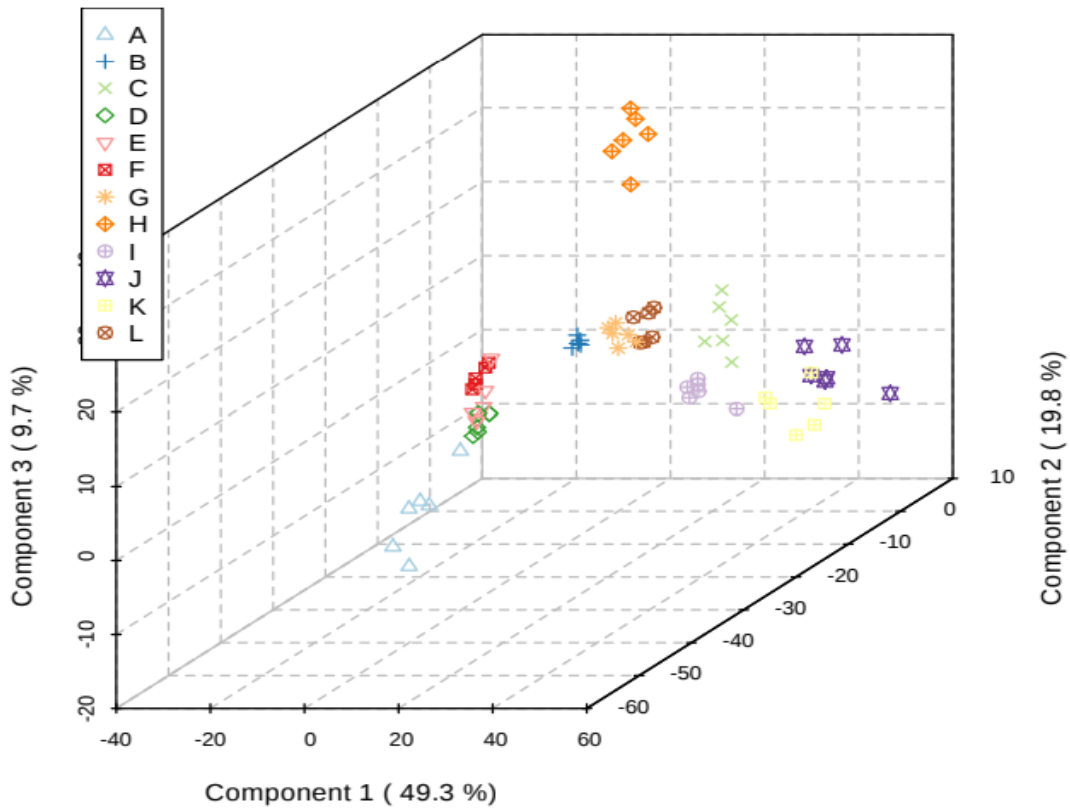


Figure 2. 3D score plot from PLS-DA analysis for the 12 different genotypes of the *Artemisia annua* germplasm collection. The twelve genotypes are well clustered indicating their unique metabolomes.

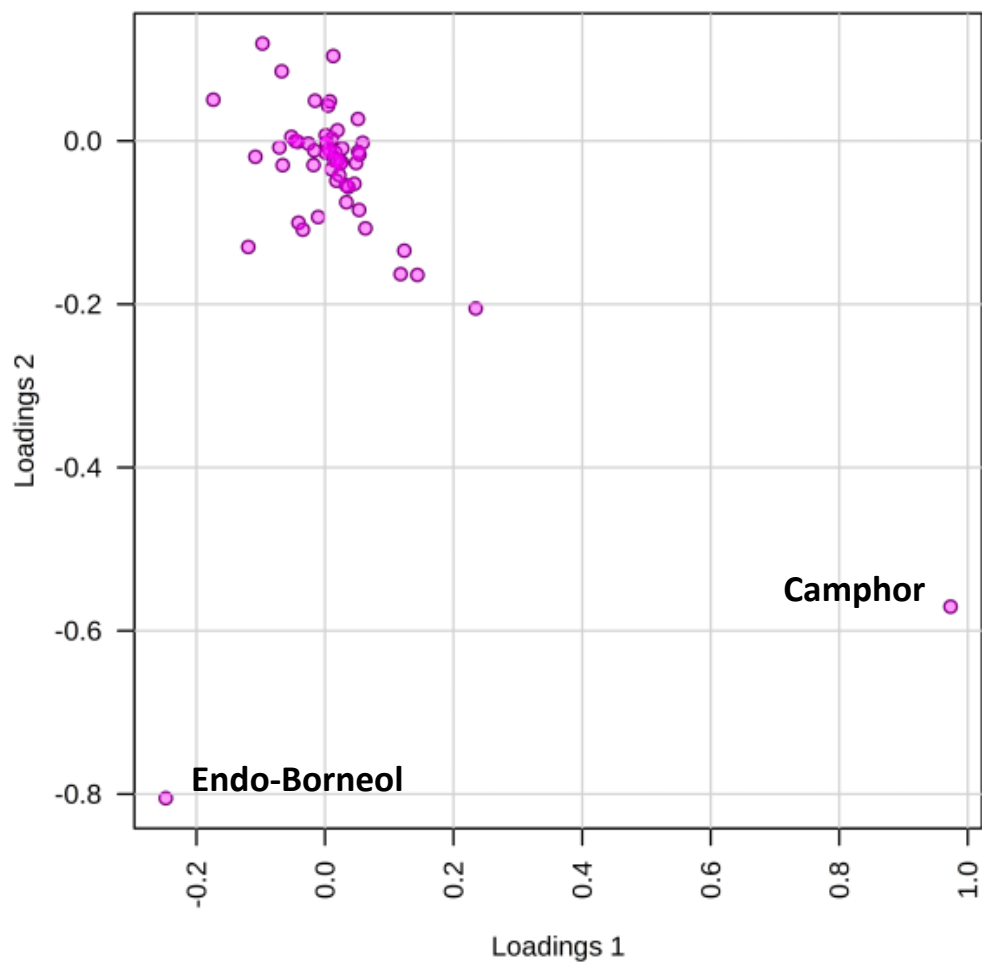


Figure 3. Loading plots from PLS-DA analysis for the 12 different genotypes of the *Artemisia annua* germplasm collection. Camphor and endo-borneol had a significant contribution in describing the germplasm collection

The importance of camphor and endo-borneol on genotypic identification of the collection was further shown with 5.3055 and 3.6694 VIP (Variable Importance in Projection) scores, respectively (**Appendix 2: Supplemental Table 3**). For small datasets (<100), VIP values above 1 are considered as marker compounds (Wang et al., 2009).

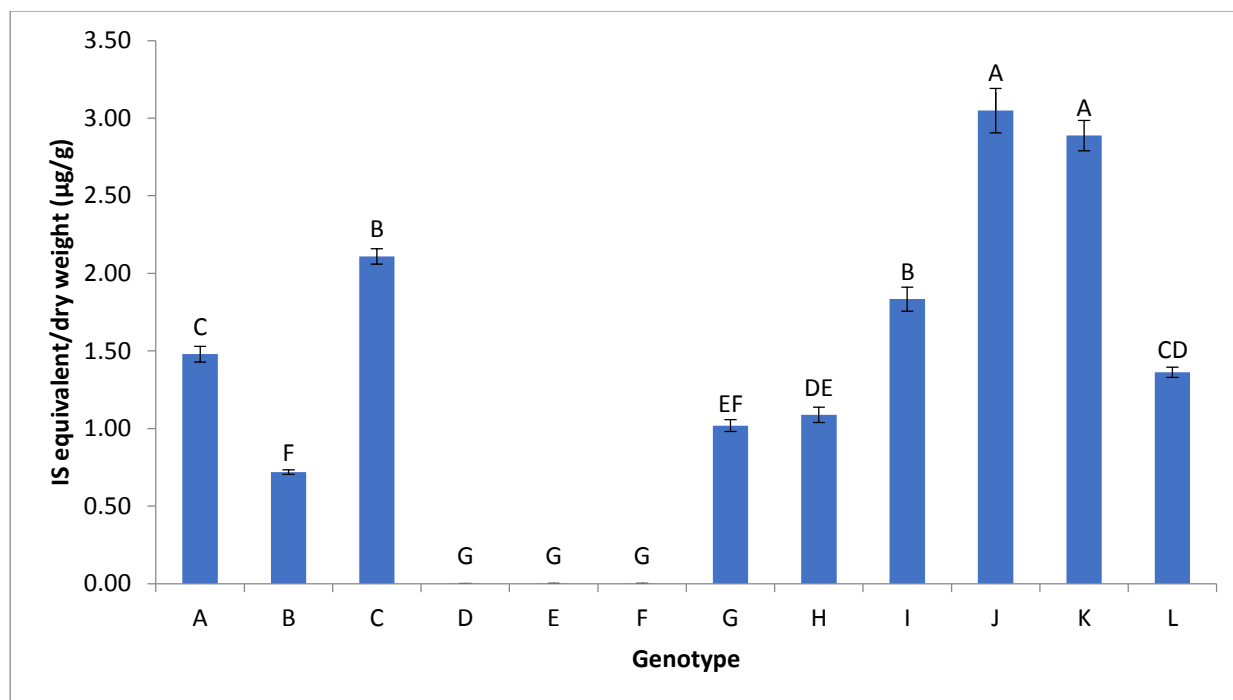


Figure 4. Relative abundance of the monoterpenoid camphor ($C_{10}H_{16}O$) in *A. annua* genotypes. Relative abundance of camphor was measured by comparing the peak intensity with the internal standard and normalized to dry weight. Average camphor content ($n=6$) within genotypes was statistically analyzed and error bars indicate standard error. Genotypes not connected by same letter are significantly different (Tukey's test, $\alpha=0.05$).

(-)-Camphor ($C_{10}H_{16}O$) is a bicyclic monoterpene ketone with strong odor naturally found in essential oils from diverse taxonomic plant groups. Camphor has been present in the human pharmacopoeia since ancient times due to its numerous pharmacological uses as an anti-inflammatory, anti-oxidant, and decongestant agent (Tian et al., 2015). Camphor is derived from geranyl diphosphate (GPP) undergoing three enzymatic reactions. First, the linear GPP is cyclized to bornyl diphosphate with bornyl diphosphate synthase and then hydrolyzed to borneol with borneol synthase. Finally, borneol is oxidized to camphor via borneol dehydrogenase (BDH).

Camphor relative abundance values varied from a maximum of 3.048 $\mu\text{g/g}$ to undetectable levels for the genotypes in the germplasm collection (**Figure 4; Appendix 2: Supplemental Table 6**). The genotypes D, E and F did not show any detectable amount of camphor in the analysis.

In a recent work by Czechowski et al., camphor was the most abundant monoterpene in Artemis, the commercial hybrid recently introduced by the Swiss company Mediplant, which is considered as a high ART chemotype (Czechowski et al., 2018). Indeed, we also observed high camphor levels in the high-ART chemotypes of our collection. All but one of the Swiss genotypes, I, J, K showed the highest levels of camphor in the collection. The only exception was the genotype L. The

relationship between camphor and ART has been well documented in *A. annua* metabolite analysis studies. Detailed studies on essential oils from *A. annua* from Vietnamese and Chinese origins showed that camphor and germacrene D as the main components in the Vietnamese oil whereas artemisia ketone as the main component in the Chinese oil (Bilia et al., 2014). This difference in essential oil composition is thought to reflect ART content in Vietnamese plants and Chinese plants with 1.0% and 0.17% dry weight, respectively (Brown, 2010).

In our analysis, correlation coefficient between camphor and ART was 0.63271 (**Appendix 2: Supplemental Tables 4 and 5**). On the other hand, camphor and DHAA had a correlation coefficient of 0.3698. Considering that DHAA is the immediate precursor to ART, and the current hypothesis that DHAA to ART conversion is non-enzymatic (Czechowski et al., 2016), the association between ART and camphor was not expected to be independent of DHAA levels, since ART production depends only on DHAA and ROS levels. Therefore, a strong positive correlation between camphor and DHAA was also expected. The lack of strong correlation between DHAA and camphor can be interpreted as a possible enzymatic intervention of DHAA to ART reaction, in which the genes responsible for this enzyme could be genetically linked to key camphor biosynthesis genes. A recent work by Wang et al. indicated that the terpene synthases responsible for borneol and camphor biosynthesis are located in the plastidial genome of *Amomum villosum* (Zingiberaceae) (Wang et al., 2018). If this is true for *A. annua* as well, then camphor and ART pathway precursor competition may not be intense, as ART uses 5C precursors 2:1 from MVA and MEP pathways, respectively (Schramek et al., 2010) whereas camphor uses both 5C precursors from MEP pathway (Wang et al., 2018). Further, the relationship between camphor and ART is shown in genotypes E, F and G with lack of camphor and low accumulation of ART. Therefore, it could be worth investigating the positive correlation between camphor and ART and use this information in breeding for high *A. annua* genotypes.

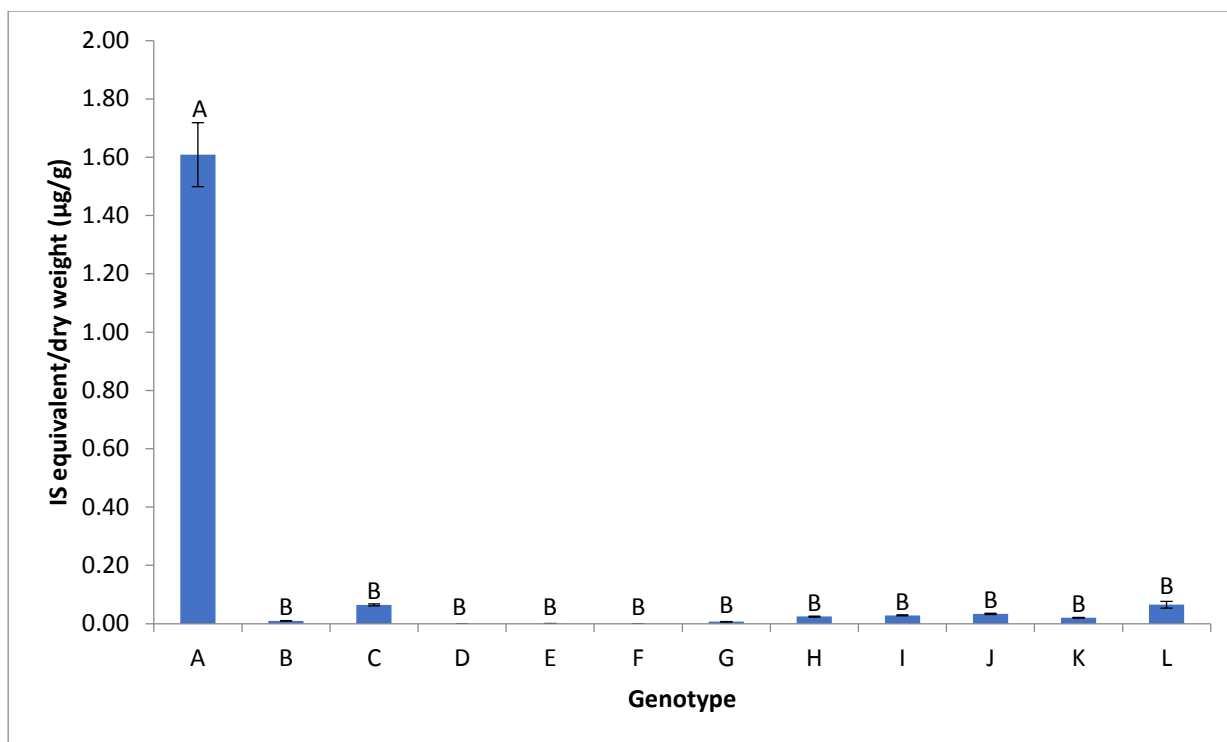


Figure 5. Relative abundance of the monoterpene endo-borneol ($C_{10}H_{16}O$) in *A. annua* genotypes. Relative abundance of endo-borneol was measured by comparing the peak intensity with the internal standard and normalized to dry weight. Average endo-borneol content ($n=6$) within genotypes was statistically analyzed and error bars indicate standard error and genotypes not connected by same letter are significantly different (Tukey's test, $\alpha=0.05$).

endo-Borneol is a bicyclic monoterpene alcohol and an immediate biosynthesis precursor of camphor. It was also identified as a distinctive compound in our biochemical profiling study (**Figure 5**). The maximum mean relative abundance found amongst the genotypes in our collection was $1.61 \mu\text{g/g}$ (genotype A). Similar to camphor, the genotypes D, E and F had no detectable levels of endo-borneol (**Appendix 2: Supplemental Table 6**), indicating that the impairment of the camphor pathway lies either in the very first or this second enzymatic step.

Tian et al. (2015) characterized the gene encoding for borneol dehydrogenase in *A. annua* (AaBDH). They found that the enzyme is specific for the substrate borneol and acts in the last step of camphor biosynthesis. It has a higher activity compared to alcohol dehydrogenase (AaADH2), which is an enzyme specific to glandular trichomes that catalyzes the production of various secondary alcohols *in vitro*, including borneol (Polichuk et al., 2010). Therefore, we suggest that genotype A may lack a functional AaBDH allele but has the activity of AaADH2 retained to carry out borneol hydrolysis to produce camphor. On the other hand, other

genotypes (B-C, G-L) with high levels of camphor may have higher AaBDH activity so that they have significantly lower borneol levels.

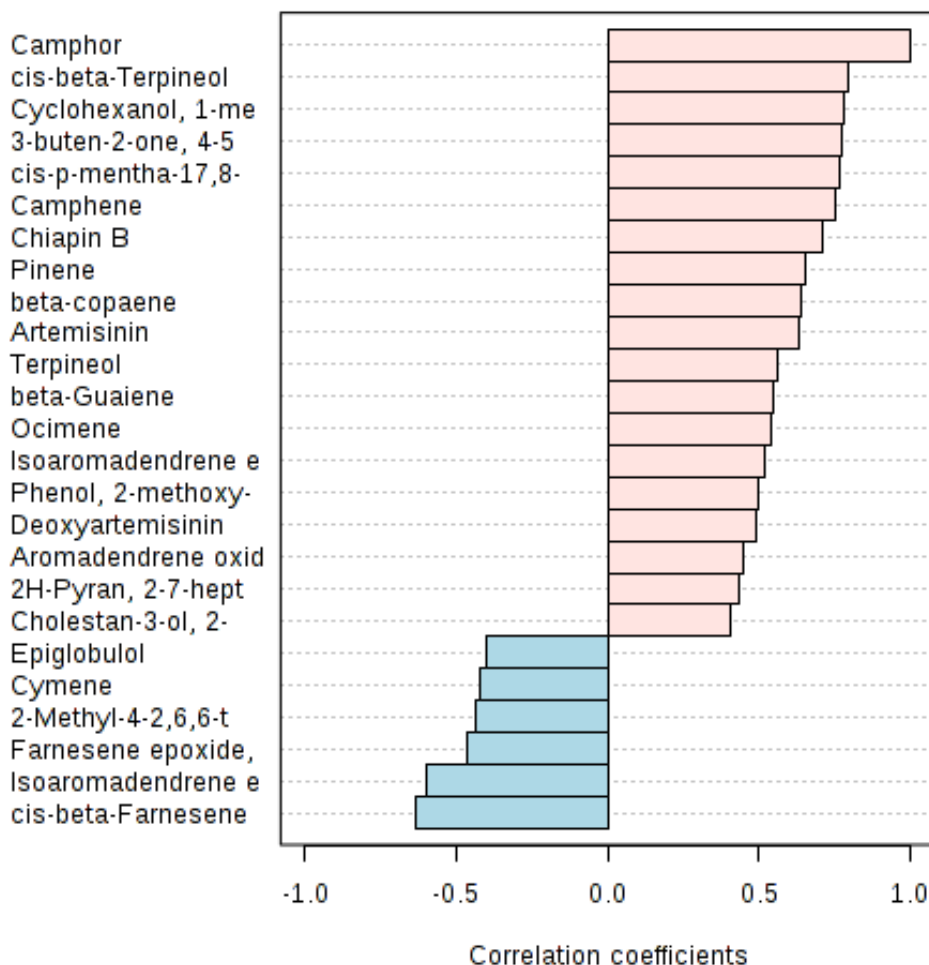


Figure 6. Correlation coefficients between camphor and other metabolites in *Artemisia annua* germplasm collection. The graph indicates the top 25 tentatively identified compounds correlated with camphor.

Among other tentatively identified compounds, two potential monoterpenes (camphene, cymene) and two sesquiterpenes (germacrene D, cedren-13-ol) also showed VIP scores above the threshold value of 1 (**Appendix 2, Supplemental Table 3**). Two additional compounds that were highlighted as markers still need further identification. The compounds highlighted in the VIP score plot are mono- and sesquiterpenes. These results underscore the terpenoid-rich nature of the *A. annua* metabolome. Surprisingly, none of the compounds of particular focus in our study, i.e. ART and its chemical relatives, had a significant impact on the genotype separation of the germplasm collection. The VIP values for ART, DHAA and AA were 0.24454, 0.35902, and 0.042541, respectively (**Appendix 2, Supplemental Table 3**). The compounds describing the

genotypes are components of the *A. annua* essential oil, mainly monoterpenoids synthesize potentially in larger quantities compared to ART, DHAA and AA (Brown 2010).

Generally, there was a positive correlation between ART and camphor levels (**Figure 6**). The positive correlation coefficient (PCC=0.632) and the correlation probability of $4.20E^{-09}$ indicate a clear positive relationship between these compounds (**Appendix 2, Supplemental Tables 4 and 5**).

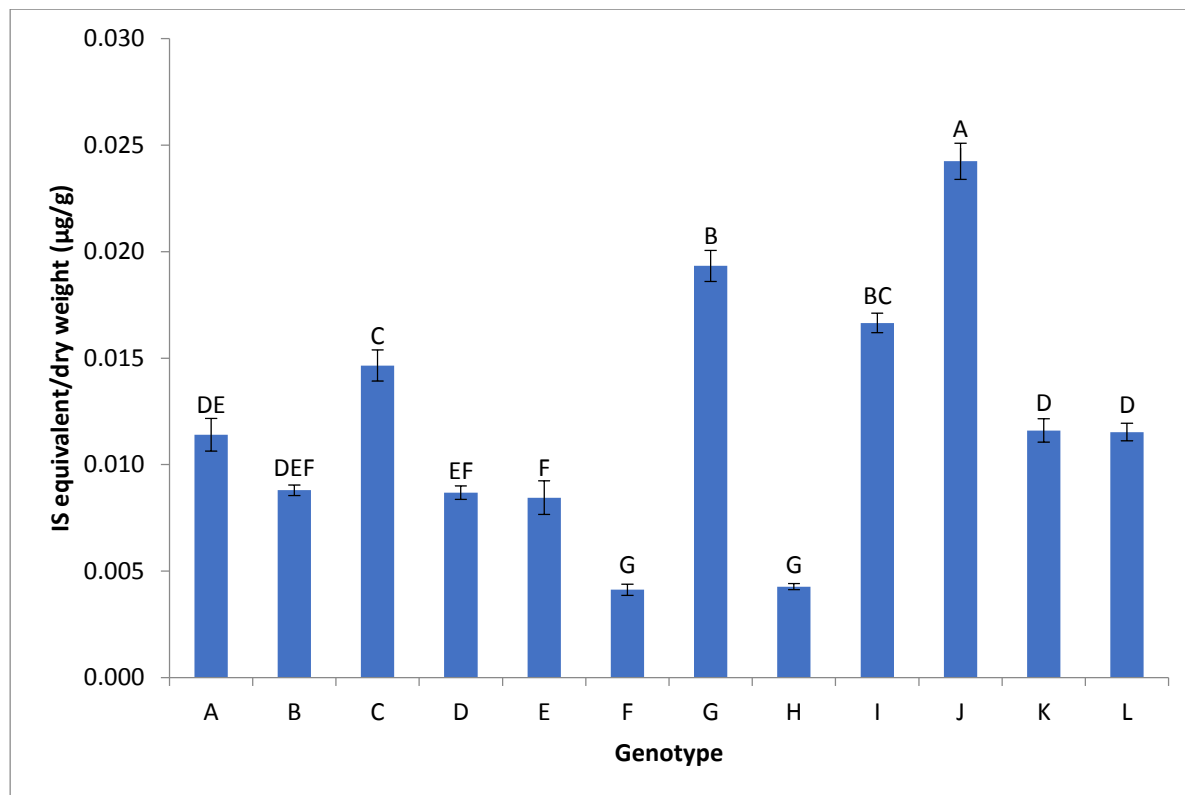


Figure 7. Relative abundance of the anti-malarial sesquiterpene, artemisinin (ART), in the twelve genotypes of the *Artemisia annua* germplasm collection at the West Virginia University. Relative abundance of ART content was measured by comparing the peak intensity with the internal standard and normalized to dry weight. Average ART content (n=6) within genotypes was statistically analyzed and error bars indicate standard error and genotypes not connected by the same letter are significantly different (Tukey's test, $\alpha=0.05$).

ART levels within the collection ranged from 0.024 $\mu\text{g/g}$ (genotype J) to 0.004 $\mu\text{g/g}$ (genotype F) (**Figure 7**). The genotype with highest average level of ART was J. Genotypes F and H had the least levels of ART. ART has a strong positive Pearson correlation with DHAA (PCC=0.75296). Additionally, chiapin B, β -copaene, deoxyartemisinin, and camphor showed high positive correlations (**Figure 8**). On the other hand, ART and AA had an overall low, significant negative relationship (PCC= -0.27403, $p > 0.021702$) (Figure 9).

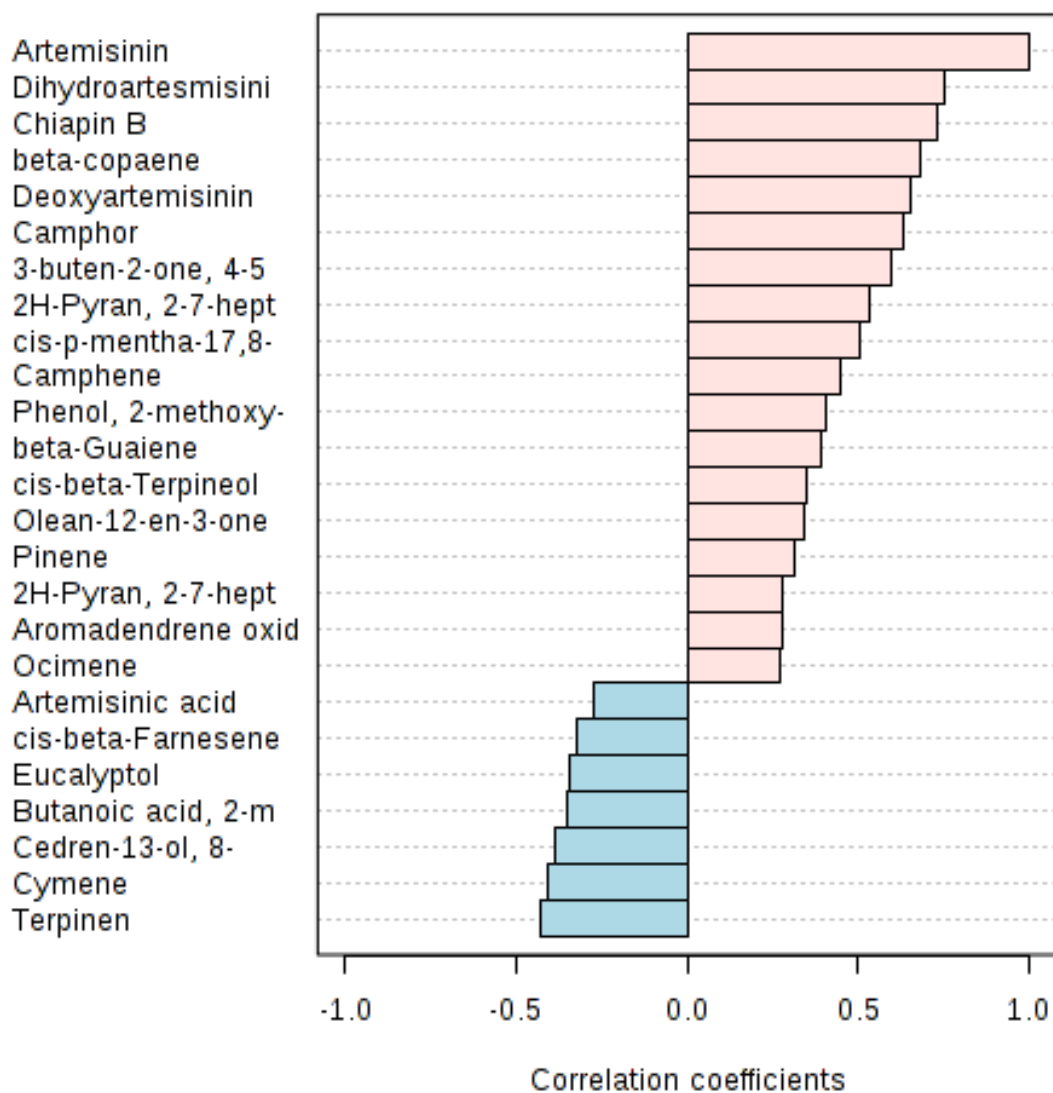


Figure 8. Correlation coefficients between artemisinin and other metabolites in *Artemisia annua* germplasm collection. The graph shows the top 25 tentatively identified compounds correlated with artemisinin.

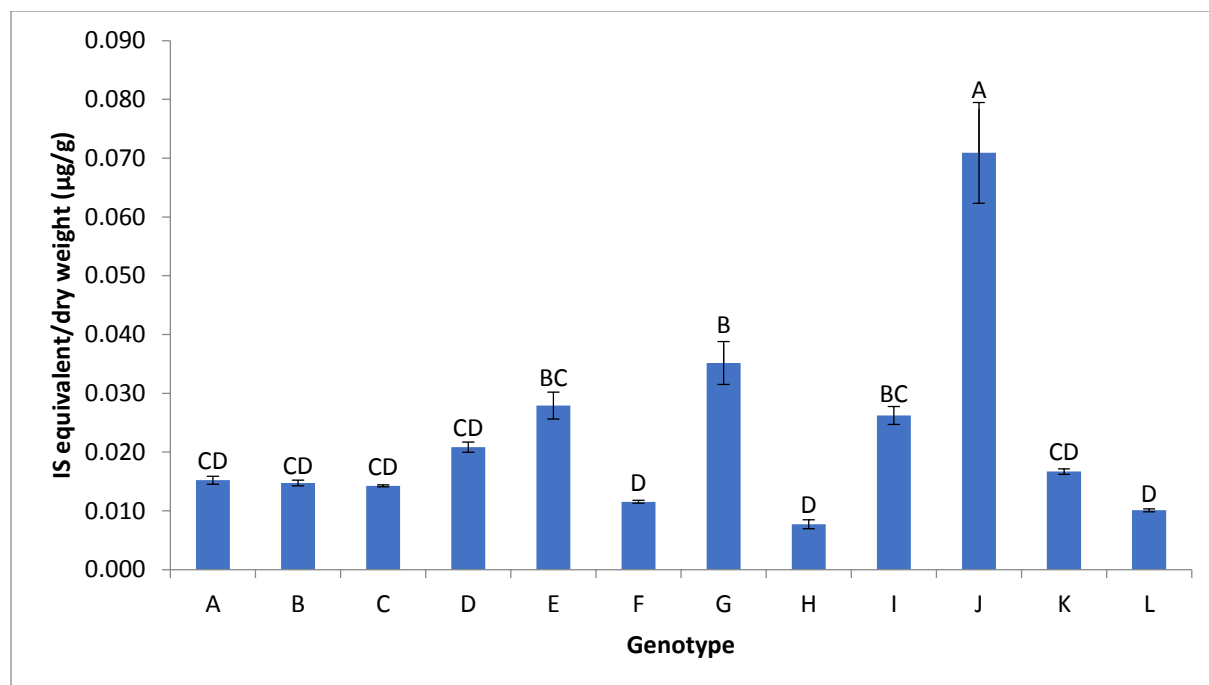


Figure 9. Relative abundance of the artemisinin precursor, dihydroartemisinic acid (DHAA) in *A. annua* genotypes. Relative abundance of DHAA content was measured by comparing the peak intensity with the internal standard and normalized to dry weight. Average DHAA content (n=6) within genotypes was statistically analyzed and error bars indicate standard error and genotypes not connected by the same letter are significantly different (Tukey's test, $\alpha=0.05$).

The levels of the ART immediate precursor, dihydroartemisinic acid (DHAA) within the genotypes in the collection varied more than nine-fold, from 0.0077 to 0.071 $\mu\text{g/g}$. Genotype J showed distinctively the highest level (**Figure 9**). Together with biomass (Chapter 2), Genotype J is the best genotype in terms of ART and DHAA yield because of its high biomass and ART high ART and DHAA accumulations. Next to genotype J, genotypes G and I had next highest levels of ART yield.

DHAA and ART showed a similar, highly correlated accumulation pattern in the metabolic profiling (**Figure 10**). DHAA is converted to ART via a potentially non-enzymatic oxidation step that forms the peroxide bridge directly involved in the mechanism of action against the *Plasmodium* parasite (Wen and Yu, 2011)

In a recent publication by Ferreira et al., the potential of high DHAA accessions for semi-synthesis of ART, as well as plant breeding for high ART producing varieties, was discussed (Ferreira et al., 2018). High DHAA accumulation indicates a strong pull of C skeletons from competing terpenoid pathways in general, including specifically other concurrent sesquiterpenoid compounds, toward the relevant compound, ART.

The very final step of ART conversion still needs further elucidation. If this is indeed a non-enzymatic process triggered by reactive oxygen species (ROS) generated by O₂ in the presence of light, and one would expect most of the accumulation to occur as ART in physiological conditions, not as DHAA. But it has shown *ex-vivo* and suggested the hydrophobic environment in the subcuticular cavity favours the conversion of DHAA to ART (Czechowski et al., 2016).

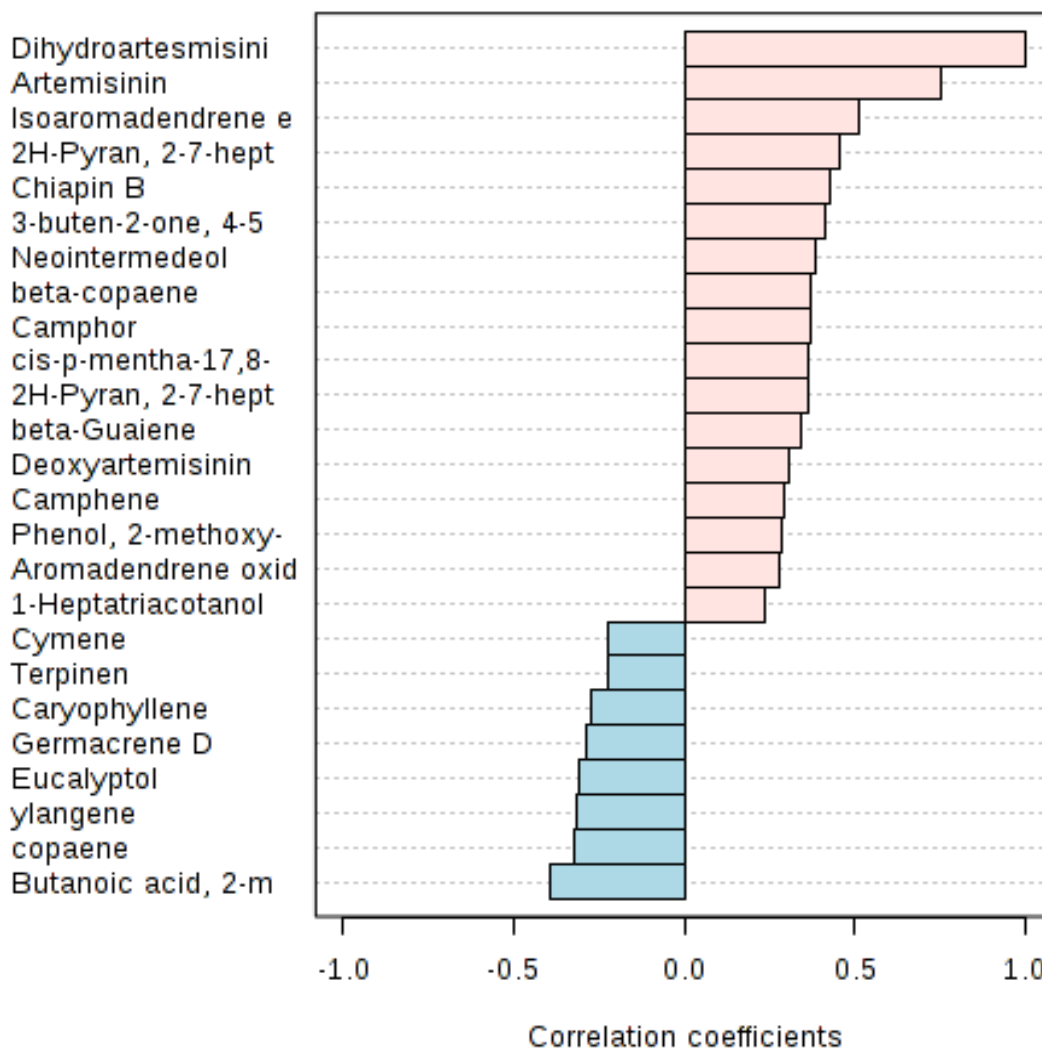


Figure 10. Correlation coefficients between dihydroartemisinic acid (DHAA) and other metabolites in 12 different *Artemisia annua* genotypes. The graph shows the top 25 tentatively identified compounds correlated with DHAA.

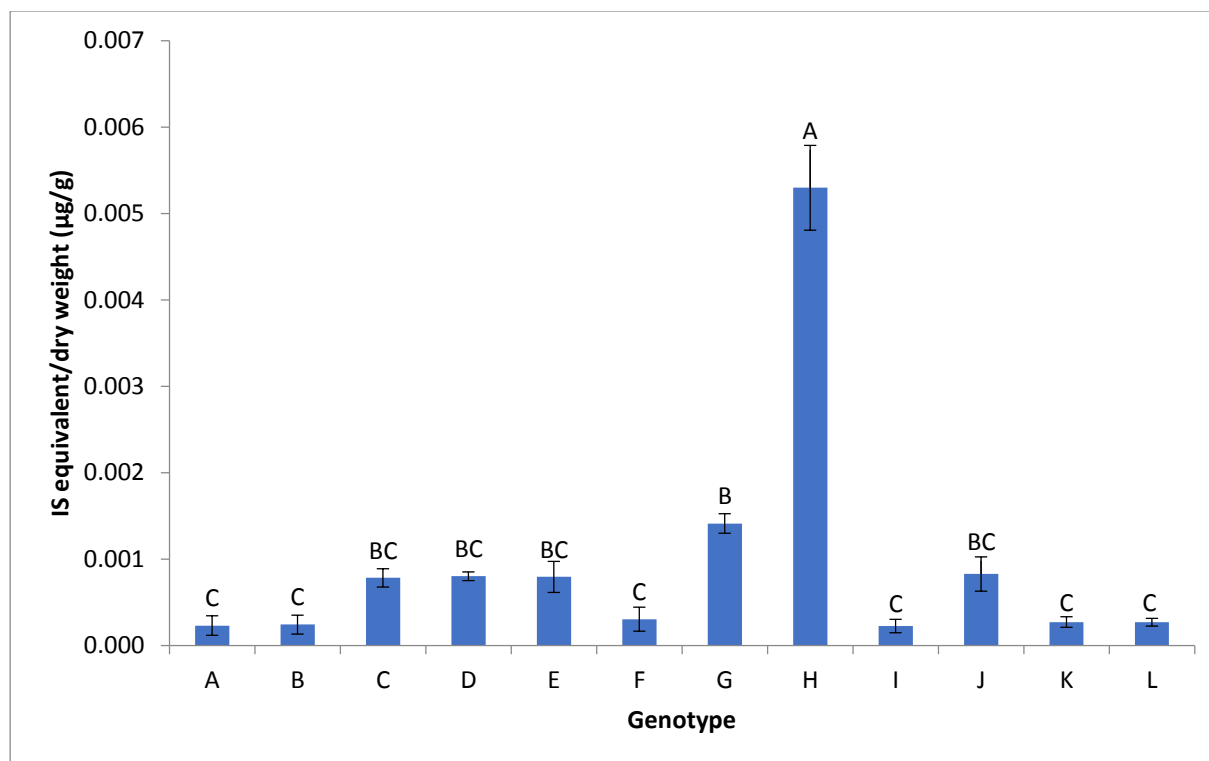


Figure 11. Relative abundance of the non-productive artemisinic acid (AA) in *A. annua* genotypes. Relative abundance of AA content was measured by comparing the peak intensity with the internal standard and normalized to dry weight. Average AA content (n=6) within genotypes was statistically analyzed and error bars indicate standard error and genotypes not connected by the same letter are significantly different (Tukey's test, $\alpha=0.05$).

Artemisinin acid (AA) biosynthesis competes for substrate with the ART pathway at the level of artemisinic aldehyde. To maximize ART accumulation levels, AA is expected to be low. Genotype H accumulated significantly higher levels of AA (0.005 $\mu\text{g/g}$) compared to all other genotypes analyzed, which have relatively low AA levels (**Figure 11**). Like the Chinese genotypes, Sandman (genotype H) can be categorized as a low ART-producing chemotype because of its high AA and low ART levels. The rest of the collection can be categorized as high ART-producing chemotypes because of their low AA and high ART profiles. Even though genotypes such as D, E, and F have low ART levels, they are still categorized under the high ART-producing chemotypes because of their low AA content. As previously mentioned, the levels of ART and other metabolites are known to vary widely due to high degree of genetic variation in this species. For example, a variation between 0.93 and 20.65 $\mu\text{g/mg}$ was observed in the famous high ART-producing Artemis hybrid cultivar (Graham et al., 2010).

Eucalyptol and a yet unidentified compound showed strong positive correlation coefficients with AA - 0.81179 and 0.77667, respectively (**Figure 12**). The strong positive correlation between AA and eucalyptol may be explained by the availability of GPP (the precursor of all terpenoids), or

the coincidence that genotypes that utilize most of artemisinic aldehyde towards AA conversion also have strong alleles (e.g., coding for functional enzymes with enhanced kinetic properties) for monoterpene biosynthesis.

The deviation of the ART precursors toward the non-productive formation of AA is a complex problem to tackle. The cytochrome P450 monooxygenase CYP71AV1 plays a role in two consecutive conversions from amorpha-4, 11-diene (AD) to artemisinic aldehyde (Shen et al., 2012). However, it can convert artemisinic aldehyde to AA, thus redirecting substrates to the non-productive pathway. Ideally, breeding programs will select for CYP71AV1 alleles with high-affinity for substrates in the first two reactions, but with low activity in converting artemisinic aldehyde to AA. It is possible that genotype J (highest ART and highest DHAA, and low AA) carries an interesting CYP71AV1 allele for breeding purposes. An enzymological analysis of natural variation of the CYP71AV1 gene in *A. annua* germplasm could be valuable in finding ideal alleles for breeding. Moreover, for biotechnological purposes (e.g., ART production in bioreactors) the use of directed evolution to engineer an optimal cytochrome P450 monooxygenase with high affinity to productive substrates but with very little affinity to artemisinic aldehyde could start with a gene from a genotype with a profile similar to that of genotype J.

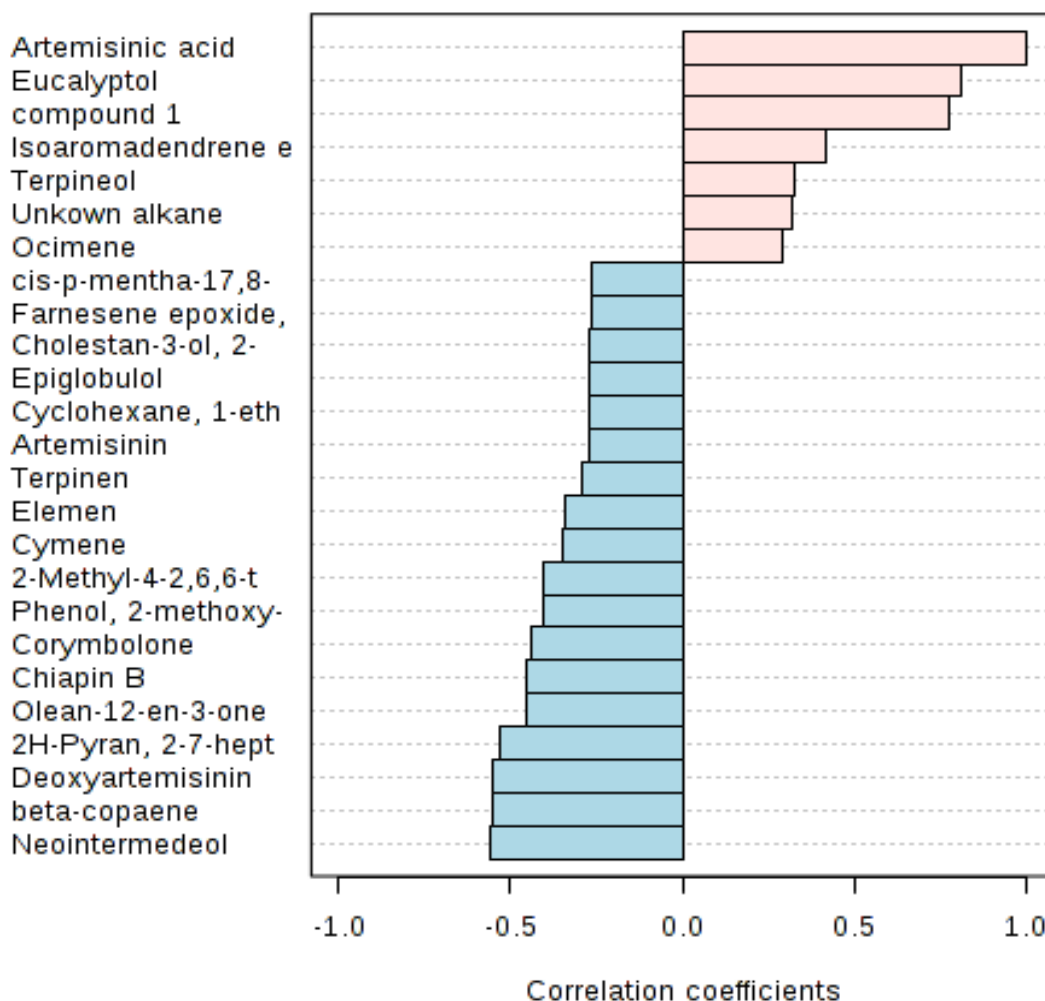


Figure 12. Correlation coefficients between artemisinin acid (AA) and other metabolites in 12 *A. annua* germplasm collection. The graph shows the top 25 tentatively identified compounds correlated with AA.

CONCLUSION

Genotype J had the highest accumulation of ART and DHAA compare to the rest of the collection. Furthermore, genotypes G, I and C showed high ART content whereas G, I and E presented high DHAA content. Genotype H had the highest accumulation of AA compared to the rest of the collection, which showed low AA accumulation levels. Because of high AA accumulation and low DBR2 gene expression (Chapter 3), we concluded that H belongs to the low ART-producing chemotype. Even though there are genotypes with low ART in the collection, they are still considered as high ART-producing genotypes because of the low AA levels. Along with other published research, we also observed metabolite variation within descendants of *A. annua* in the

data collected in this research as well. For example, the Brazilian genotypes are descendants of the cultivar 3M while the Swiss genotypes are descendants of the cultivar Artemis (Chapter 1). The relationship between camphor and ART was also highlighted from the data collected. The two heterozygous parent lines (C1 and C4) that is used in producing Artemis are of Vietnamese origin (Graham et al., 2010, Czechowski et al., 2016) and the Vietnamese genotypes are known be high in camphor and ART (Brown, 2010). Futher High levels of monoterpenoids indicates the retaining potential of the *A. annua* metabolome that can be used achieve high DHAA and ART genotypes content by improving the C flux towards ART biosynthesis via breeding.

Altogether, this chapter presented and discussed the biochemical profiles of the genotypes in the *A. annua* germplasm collection at WVU. This profiling is the key to characterize the germplasm collection as well as provide information on how to use the phenotypical variability within to propose breeding opportunities for enhanced ART accumulation.

CHAPTER 4

Transcriptional analysis of genes involved in the artemisinin pathway in the *Artemisia annua* germplasm collection at WVU

INTRODUCTION

The enzymes, genes, and many transcriptional regulators involved in the biosynthesis pathway for artemisinin (ART) are quite well known (Lu et al., 2013), except for the last reaction, which is believed to be non-enzymatic (Czekowski et al., 2016), i.e. the conversion of dihydroartemisinic acid (DHAA) to ART. Even though there is a large body of research showing the possibility of auto-oxidation of DHAA via dihydroartemisinic acid tertiary peroxide (DHAAOOH) into ART in the presence of light (Sy and Brown, 2002; Wallaart et al., 1999; Brown and Sy, 2004; Czechowski et al., 2016), strong evidence demonstrating that this reaction happens *in vivo* remains missing. ART is a sesquiterpenoid that accumulates in the subcuticular space of glandular secretory trichomes (GSTs). The immediate ART precursor (DHAA) is synthesized in the GST sub-apical and apical cells (Olofsson et al., 2011).

Briefly, the ART biosynthesis pathway starts in the general terpenoid pathway. Dimethylallyl diphosphate (DMAPP) from the cytosolic mevalonate pathway (MVA) is transferred to plastid and condenses with another 5-C isoprenoid (IPP) from the plastidial methylerythritol pathway (MEP) to generate geranyl pyrophosphate (GPP), which is subsequently imported by the cytosol for further elongation with an isopentyl pyrophosphate (IPP) from the MVA pathway to form the 15-C farnesyl pyrophosphate (FPP), the precursor of all sesquiterpenes in the plant (Wen and Yu, 2011). In the ART committed pathway, FPP is cyclized to amorpha-4,11-diene (AD) by amorpha diene synthase (ADS) and oxidized with the CYP7AV1/CPR pair to form artemisinic alcohol (AAOH). At this point, the pathway divides into two branches leading to the artemisinic acid (a non-productive compound) via artemisinic aldehyde with additional activity of ADH1. Meanwhile in the productive branch, the double bonds at 11 and 13 positions of AAOH are reduced by DBR2 and then oxidized to DHAA by ALDH1 (Weathers et al., 2017).

Herein, we analyzed the transcriptional activity of 16 genes in the 12 genotypes of the *Artemisia annua* germplasm collection at the West Virginia University - 10 structural genes from the biosynthetic pathway, 5 transcription factors previously described as regulators of this pathway, as well as GAPDH as the constitutive housekeeping gene for data normalization. By comparing gene expression of key genes in the ART pathway with the metabolite profile of these genotypes under the same conditions, we are able to better understand how this pathway is regulated at the genetic and biochemical levels.

MATERIAL AND METHODS

Sampling and cDNA synthesis

The same samples collected for the biochemical profiling (Chapter 2) were used for this experiment. In brief, young leaves with the apical meristem were flash frozen in liquid N₂ and stored at -80 °C until processing. Total RNA was extracted with TRIzol reagent and purified with Direct-zol™ RNA Miniprep plus kit (Zymo Research) following the manufacture's instruction. Isolated total RNA was treated with DNase I (DNA-free TURBO kit, Ambion), quantified with a NanoDrop™ spectrophotometer (Nanodrop), and used for cDNA synthesis. cDNA synthesis was carried out with 1 µg of DNase treated RNA and SuperScript®IV first-strand synthesis system (Invitrogen) with oligod(T) primer, according to the manufacture's instruction in 10-µL reactions. Synthesized cDNA was diluted 10 times with RNase free water to a final volume of 100 µL and stored at -80 °C.

Quantitative RT-PCR

Primers were designed as described by (Mock et al., 2017) (**Table 1**). cDNA was used in qPCR amplification using SybrGreen fluorescence. Fifteen genes were amplified and GAPDH was used as an internal control for normalization. qPCR was performed using 1 µL diluted cDNA, 0.5 µL RNA free water, 2.5 µL 2X Power SYBR® Green master mix (ThermoFisher Scientific) and 1 µL mix of 10 mM each forward and reverse primers. Each sample was run in technical triplicates, and each treatment was run in three biological replicates. The reactions were performed with ABI 7500 instrument (Applied Biosystems) and CFX96 Touch™ (Bio-Rad) with appropriate clear 96 well plates.

Table 1. Primers used in gene expression analysis in this study.

Gene ID	Functional Annotation	Forward primer sequence 5'-3'	Reverse primer sequence 5'-3'	Amplicon size (bp)
AF182286.2	1-deoxy-D-xylulose-5-phosphate synthase (AaDXS1)	GCAGATCAATTAGCCGAAGC	CATGACCTCCAAAGCCTCTC	96
AF182287.2	1-deoxy-D-xylulose-5-phosphate reductoisomerase	ACCGGACAACGTGAAATACC	GGCTTTCTCGTTAGCTGCAC	100
AF142473.1	HMG-CoA reductase (AaHMGR1)	ATTGCGTCAAAGGTGTGAAAG	ACCCTTCAAACCGCACAC	101
GQ420346.1	Farnesyl pyrophosphate synthase (AaFPS)	TGATTTGAGTCTAGCGAGTGG	TACAGACAACATCGGCTTGC	70
KR445687.1	Amorpha-4,11-diene synthase (AaADS)	TGATCTCATGACCCACAAGG	TCAAGGTTTGGGCATACTCC	107

KJ609178.1	Amorpha-4,11-diene monooxygenase (AaCYP71AV1)	CAACTGGAAACTCCCAATG	CTCGGTCTTTCTTGCATCG	85
JN594507.1	Cytochrome P450 reductase (AaCPR)	TTGTATGTTTGGCGGTGATGC	TAGAGCTCCGCCTTTGAGG	107
JF910157.1	Alcohol dehydrogenase 1 (ADH1), from	GGCTAGTGTGGCCATTAACAG	GGTTCTTACGCCACCAAAG	79
EU704257.1	Artemisinic aldehyde delta-11(13) reductase (AaDBR2)	CTGATTTGGTCGCTTTTGG	CCTATTTAAAGGTGCGTTGAGC	83
FJ809784.1	Aldehyde dehydrogenase 1 (AaALDH1)	CGAGATCTATTCGAGCAGGTG	TCCTTGTTCTCGTCCAAACC	107
JQ797708.1	Octadecanoid-derivative responsive AP2 domain protein	TGAGGAGCTTCGGATGGTAG	GTCATGATCTGTTCCACCTC	79
FJ390842.1	WRKY transcription factor 1 (AaWRKY1)	GATTGAACCTGCTTGATAGTTGG	TGCAGAAAGATACGTCTCAAG	67
JN162091.1	AP2/ERF transcription factor (AaERF1)	GCCGAGTGGTAGATAGGATTTG	AAATCCAGACCCGATACCATC	101
KP119607.1	MYC2 transcription factor	GGTAAGCTCGTCTTCTCATGG	CACCCAAATCCTGCTTTTTG	61
JQ617313.1	ABA receptor (AaPYL9)	CGAGTCATTTGTGGTGGATG	CTCTGAGACATCAGCCAACG	109
GQ870632.1	glyceraldehyde-3-phosphate dehydrogenase	TGGTGACAGTAGGTCAAGCATC	GGGTGCTGTATCCCCATTC	107

Data analysis

The geometric means of Ct values were used to calculate the relative mRNA abundance according to a formula provided in Jacometo et al. with a slight modification by assuming 100% amplification of mRNA instead of using the primer efficiency in the equation (Jacometo et al., 2014). The dataset was subjected to ANOVA and, when appropriate, to post-hoc testing using Tukey's range test ($\alpha=0.05$).

RESULTS AND DISCUSSION

A description of the enzymes involved in the ART pathway is reported in the Introduction section of this thesis. To understand the relationship between biochemical phenotypes of the genotypes, ten enzymatic genes and five genes encoding transcription factors were analyzed.

Genes which expressed were found not significantly different among the genotypes in the collection were HMGR from the general MVA pathway, DXR1 and DXS1 from the MEP pathway, and the ART biosynthesis genes, CYP7AV1 and ADH1 along with all five transcription factors,

WRKY1, MYC2, PYL9, TAR1 and ERF1 (cf. Appendix 3). Among the ten structural genes of biosynthesis pathway, only three (FPS, CPR, and DBR2) showed significant differences in mRNA abundance among the genotypes tested.

According to (Olofsson et al., 2012), most ART biosynthesis genes are expressed in sub- and apical cells of glandular trichomes (GST). FPS is the key enzyme to enter the sesquiterpenoid pathway and acts on the synthesis of 15-C farnesyl pyrophosphate (FPP) by condensation of geranyl pyrophosphate (GPP) and IPP in the cytosol. Even though FPP is the starting point for wide variety of other sesquiterpenes, such as squalene and β -farnesene, the overexpression of FPS showed a moderate to high influence on the ART yield (Wang et al., 2011; Banyai et al., 2011). The relative abundance of FPS expression among the genotypes in our materials ranged from 0.480468 (genotype C) to 2.439918 (genotype K) units, compared to GAPDH expression level, which is set as 1 unit by definition (**Figure 1**). According to our data, FPS mRNA abundance of genotypes K and L were significantly higher compared to genotypes B, J, A, H and C. Considering the alleles present in all genotypes have comparable kinetic levels (e.g., substrate affinity, turnover rates, inhibitor levels), we would expect a stronger C flow towards the general sesquiterpene synthesis and a lower monoterpene accumulation in genotypes K and L compared to A, B, C, H, and J but we could not observe this assumption. Therefore we can assume unfavorable allelic versions of genes encoding ART pathway enzymes in K and L that are inefficient in converting the C flow created by FPS towards DHAA and ART.

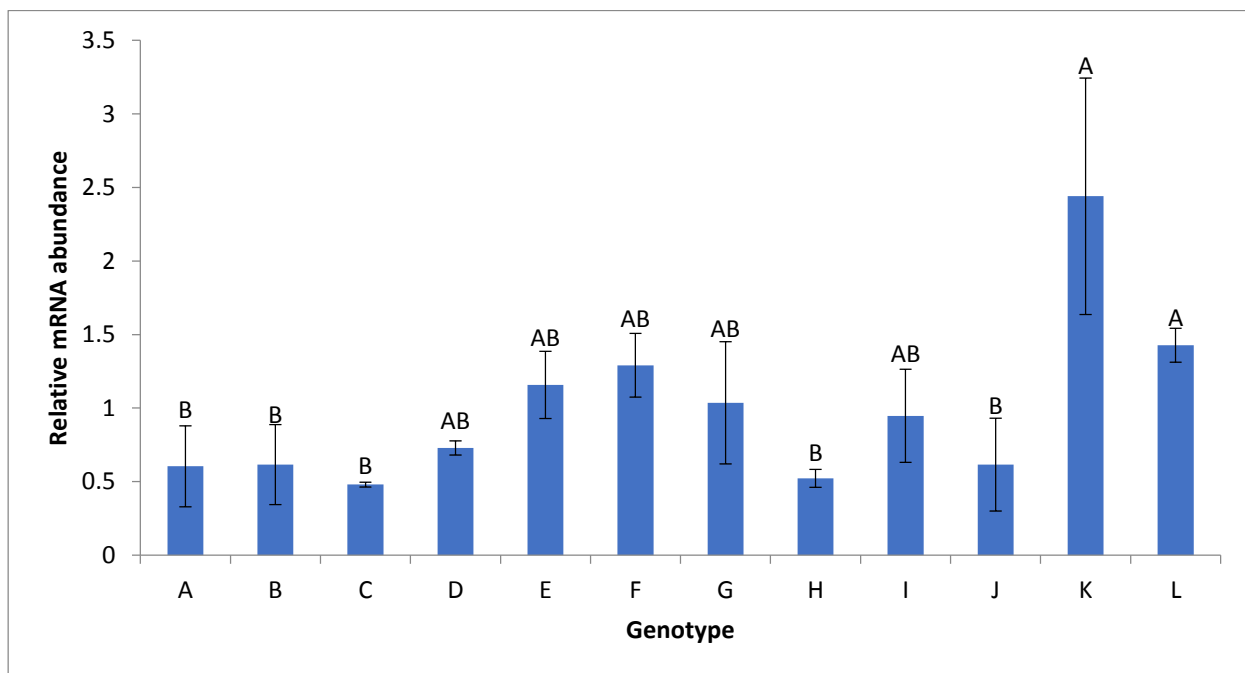


Figure 1. Relative mRNA abundance of FPS gene within the *A. annua* genotypes of the WVU germplasm collection. mRNA abundance was measured and averaged ($n=3$) with young, unexpanded leaf materials collected from plants grown for 45 days. Error bars indicate standard error and genotypes not connected by the same letter are significantly different (Tukey's test, $\alpha=0.05$).

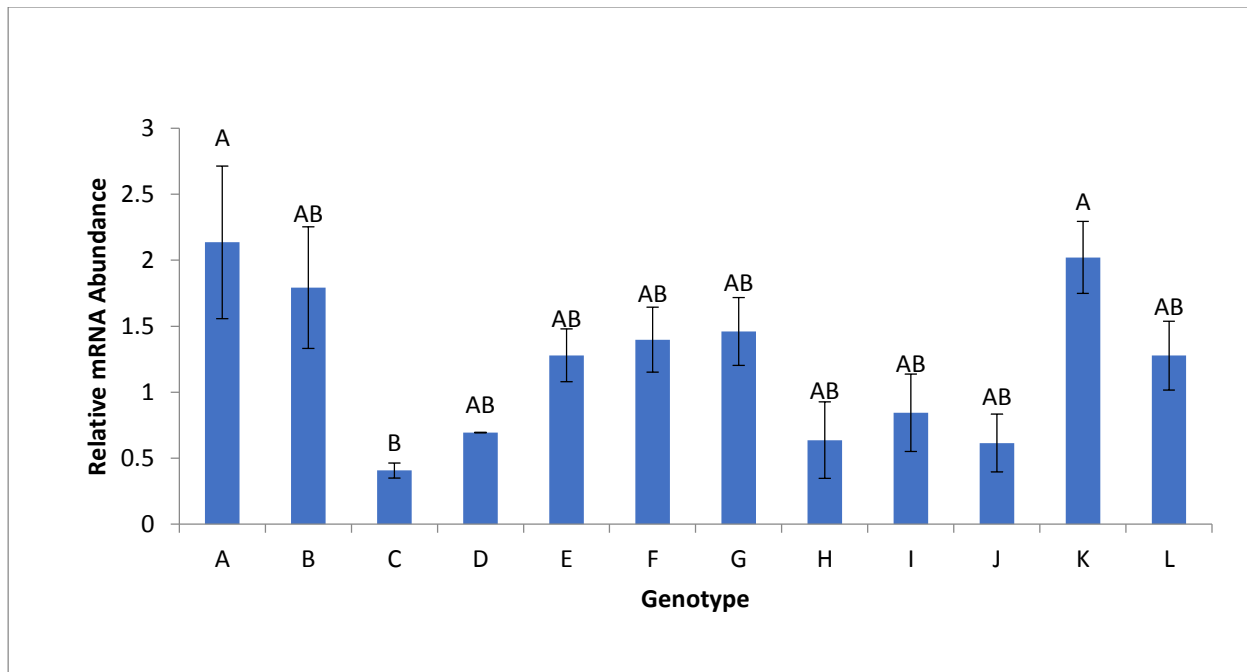


Figure 2. Relative mRNA abundance of CPR gene within the *A. annua* genotypes of the WVU germplasm collection. mRNA abundance was measured and averaged (n=3) with young, unexpanded leaf materials collected from plants grown for 45 days. Error bars indicate standard error and genotypes not connected by the same letter are significantly different (Tukey's test, $\alpha=0.05$).

The cytochrome P450 reductase (CPR) (**Figure 2**) is the redox partner of the cytochrome P450 monooxygenase (CYP71AV1) that carries out the oxidation steps from amorpha-4,11-diene into artemisinic aldehyde (Shen et al., 2012) as well as a third, non-productive step to artemisinic acid (AA) (Shen et al., 2012). The CPR reduction activity restores the catalytic properties of CYP7AV1 and can be seen a single enzymatic complex. Whereas we have not observed transcriptional differences for CYP71AV1 amongst the genotypes in the collection, we did for CPR. The largest CPR mRNA abundance differences were between the genotypes expressing the highest levels (A and K) compared to the group with lowest transcriptional activity (C).

Since genotype H had significantly high content of AA compared to the rest of the collection (cf. Chapter 3), one would expect significantly higher expression levels of enzymes CYP7AV1, CPR and ALDH1 genes. However, none of these genes were significantly high for genotype H. Therefore, we reason that the high AA accumulation in genotype H is more due to allelic variation with enhanced properties towards conversion of artemisinic aldehyde to AA, which is highly undesirable for high-ART biosynthesis.

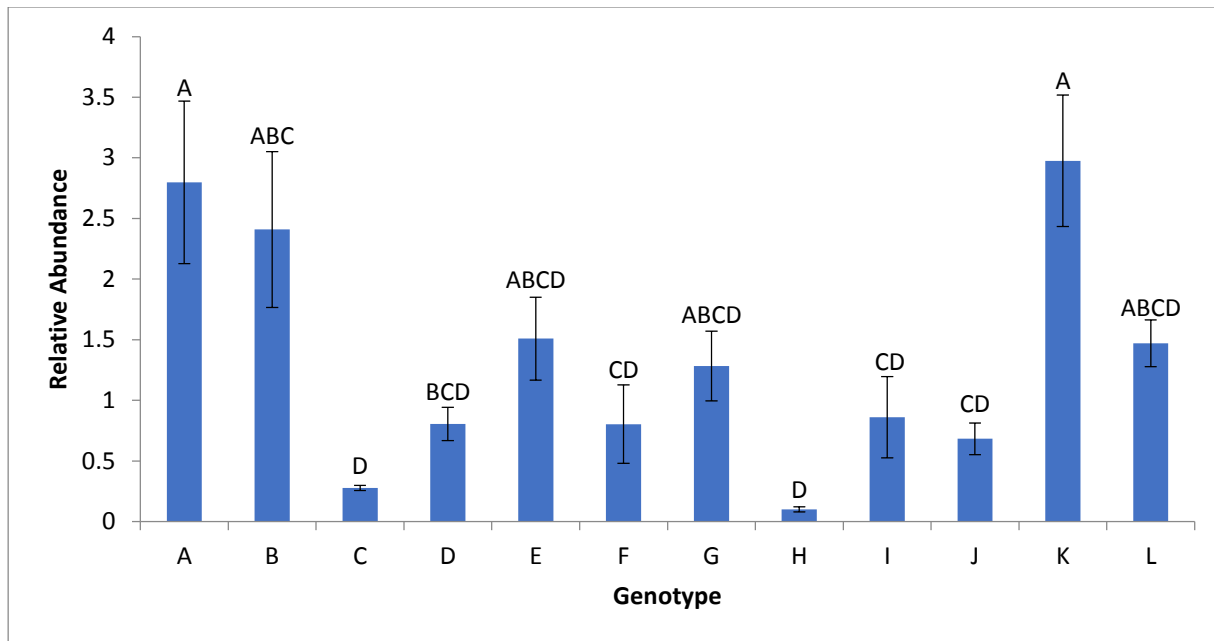


Figure 3. Relative mRNA abundance of DBR2 gene within the *A. annua* genotypes of the WVU germplasm collection. mRNA abundance was measured and averaged (n=3) with young, unexpanded leaf materials collect from plants grown for 45 days. Error bars indicate standard error and genotypes not connected by same letter are significantly different (Tukey’s test, $\alpha=0.05$).

Among all the genes tested, the artemisinic aldehyde- $\Delta 11$ (13) reductase (DBR2) was the most drastically differently expressed amongst the genotypes. DBR2 mRNA abundance values ranged 30-fold, from 2.97599 (genotype K) to 0.101044 relative units (genotype H). *A. annua* is divided into two chemotypes based on the content of ART and its precursors (Wallaart et al., 2000). ART, DHAA, and AA content along with DBR2 expression levels are parameters used to categorize *A. annua* chemotypes, the high-ART (HAP) and low ART-producing (LAP) groups. In LAP, the AA content is high (Lommen et al., 2006; Arsenault et al., 2010; Larson et al., 2013; Czechowski et al., 2018). A study by Yang et al. deduced that the difference between HAP and LAP chemotypes are due to DBR2 differential expression and specifically due to the insertions and deletions of certain *cis*-elements in the DBR2 promoter in LAP, leading to a lower transcriptional activity of the gene coding for this enzyme (Yang et al., 2015). In our germplasm collection, our biochemical profiling data categorized the genotype H as a LAP, which is congruent with the lowest expression level of the DBR2 gene seen in this study for H (**Figure 3**). The remaining genotypes of the collection can be categorized under HAP chemotype because of their low AA levels and high DBR2 expression compared to genotype H. Notwithstanding, the DBR2 expression level of H was not statistically different from genotypes C-J and L, possibly due to the high data variation amongst the biological replicates and warrant the inclusion of additional replicates and a re-analysis of the data in future.

Notwithstanding, given that the ART committed pathway is localized in glandular trichomes, the significant differences in trichome densities amongst the genotypes (cf. Chapter 2) may be contributing to muddle the gene expression results. Currently, our group is considering the addition of another level of normalization to the gene expression dataset according to glandular trichome density values. For that, we will need to establish the number of glandular trichomes occurs per dry weight mass of sampled leaves.

ART biosynthesis pathway is considered to correlate with leaf development (Liu et al., 2016). The early steps of the pathway occur in glandular trichomes of young juvenile leaves and the later steps happening in mature leaves (Czechowski et al., 2016). Another study by the same group tested for the gene expression of young leaves of HAP and LAP chemotypes and found significantly higher expression of DBR2 in HAP compared to LAP and, conversely, significantly higher expression of ALDH1 in LAP compared to HAP. Additionally, they also showed that there was no significant differences between ADS and CYP7AV1 gene expression in young leaves of both chemotypes (Czechowski et al., 2018), which corroborate our findings. Expression of ALDH1 specific promoter with GUS expression in different tissues showed strong expression in GSTs of young leaves and reduced to no expression in mature and old GSTs, respectively. In our study, we used young, juvenile leaf samples to study gene expression, and our results are compatible with recent reports: the genotype H can be classified as chemotype LAP (chapter 3) and shows low DBR2 expression levels whereas a high transcription of ALDH1. On the other hand, the transcriptional levels for genotypes belonging to the HAP chemotype group are also compatible with the results reported by Czechowski et al. (2018), however with some deviations that could be attributed to environmental variations (Ferreira et al., 2018).

CONCLUSION

Considering the results from biochemical analysis of the genotypes, genotype J was expected to have significantly higher mRNA levels for biosynthesis pathway genes compared to the rest of the collection, specifically for DBR2, since DHAA levels of the genotype J was significantly high. However, transcriptional levels were not necessarily high for genotype J. This could be because of the different (multi)alleles may be encoding for enzyme isoforms with different levels of affinities to the substrate and other catalytic properties. An experiment conducted by Graham et al. using the Artemis hybrid showed that the mean SNP frequency between plants within this hybrid was extremely high, 1 in 104 bp (Graham et al., 2010). Therefore, depending on the SNP locations, the process from gene expression to protein structure and catalytic activity can be affected. The other factor can be the highly convoluted nature of the terpene biosynthesis pathway. The high biochemical complexity of *A. annua* is attributed to promiscuous enzymes encoded by multiple members of gene families expressed in glandular trichomes (Wang et al., 2009; Graham et al., 2010; Soetaert et al., 2013). On the other hand, terpenoid pathway intermediates act as substrates for different types of terpene synthases. For example, DHAA is the precursor to 16 other compounds including ART (Brown, 2010). This means that there can

be other unidentified enzymes contributing to ART biosynthesis pathway as well as unidentified enzymes that use intermediates of ART biosynthesis pathway reducing the flow towards the end products such as ART, DHAA and AA. Therefore, the chemical phenotypes and the genotypes may not be compatible with each other in these plants.

Further, mRNA abundance of none of the transcription factors which we tested were significantly different among the genotypes analyzed. Importantly, some of the primers designed for transcription factors did not amplify consistently within genotypes. This could be due to the genotypical constitution of the genotypes that deviates from the sequences available for primer design. In conclusion, apart from DBR2, the gene expression data does not completely explain the biochemical phenotype of the collection. Therefore, an analysis of the allelic variation of the genes of interest in each genotype through whole-genome sequencing could be a valid approach to explain the biochemical phenome of this germplasm collection.

CHAPTER 5

Genetic improvement of *Artemisia annua*: honing domestication and boosting yields of artemisinin and dihydroartemisinic acid

INTRODUCTION

Artemisia annua is an annual herb with hermaphrodite composite flowers (Wetzstein et al., 2014) and a self-incompatible outcrossing reproduction strategy (allogamy) (Ferreira et al., 2005). Despite of some breeding efforts to improve it as a viable crop, one can consider it as an undomesticated species that has been selected from crossings of wild accessions growing in South Asia. The main features that typifies this species as undomesticated are especially its flowering induction promptly triggered by short days, the high degree of genetic polymorphism, an incapability of self-pollinating (which would facilitate cultivation and development of true varieties by stabilizing the genotype), and the high variability of economically relevant traits, such as leaf morphology and size (especially total leaf area).

In this diploid species ($2n=2X=18$; 1.74Gb per haploid set; (Shen et al., 2018), flowering is promptly induced under a photoperiod threshold of 16 hours or less, depending on the genotype (Ferreira et al., 1995), highly restricting the latitude of field cultivation. The genetic identity of this plant cannot be preserved by self-pollination, and the genotype (i.e., the exact allelic combination of an organism) is permanently lost via sexual propagation. Due to the high degree of genetic heterozygosity conferred by the obligatory allogamy, artemisia breeding is challenging. Notwithstanding, artemisia breeding is feasible, as already illustrated by many successful examples in allogamous crops in the Asteraceae family (e.g., sunflower, artichoke) as well as a plethora of cases in other plant families (e.g., corn, rye, alfalfa, tobacco, cotton), and especially in species which value, like artemisia, resides in the accumulation of terpenoids inside of glandular trichomes (e.g., hop, cannabis).

Conventional plant breeding implicates basically the creation of genome-wide allelic combinations for all genes contributing to the relevant phenotypes, followed by selection of the best assortment. In conventional crop breeding of allogamous species, the strategies used employ either highly inbred lines to generate commercial F_1 hybrids (like corn) or use an open-pollination strategy with two populations with stable alleles for the most important traits of the crop (Souza Jr., 2011). For artemisia, given that self-crossing is not a viable option for most genotypes, the second case applies. The parental lines of commercial hybrids remain highly heterozygous and must be maintained by vegetative propagation, or as two distinct populations. Although many efforts have been made on the molecular genetics front to understand the ART biosynthesis and genetic regulation of the pathway, conventional breeding programs of artemisia have traditionally been performed by only a few companies and institutions around the world, such as Médiplant in Switzerland (Simonnet et al., 2008), Bionexx in Madagascar (Townsend et

al., 2013; Suberu et al., 2016), the University of York in the UK (Townsend et al., 2013; Czechowski et al., 2016), North Carolina State University in the U.S. (Alejos-Gonzalez et al., 2011), and some places in China, such as the Institute of Medicinal Plants in Kunming (Yang et al., 2010; Liu et al., 2010a) and the Academy of Chinese Materia in Chongqing (Li et al., 2010, 2017). In order to generate commercial lines able to achieve the genetic potential of the species for ART production in different cropping environments, more concerted efforts and investment are required.

ART yield is a function of leaf ART content as well as leaf biomass production. Moreover, besides the plant genetic constitution, crop management and environmental conditions also influence the final yield (Ferreira, 2007). Breeding goals are to boost ART content from 1-1.4% (g ART per 100 g total leaf dry weight) to ~2% (Delabays et al., 2001; van der Kooy and Sullivan, 2013; Suberu et al., 2016) or even more. Two Chinese varieties, 'Anamed' and 'Chongqing', produce ~1% ART (Yang et al., 2015; Han et al., 2014). The F₁ hybrid 'Artemis' achieves 1.3% ART (Simonnet et al., 2008) and a reported yield of 37.8 kg/ha. 'Artemis' is produced by crossing of two Vietnamese populations (Graham et al., 2010). It was developed by the Centre for Novel Agricultural Products (CNAP) at the University of York, UK in collaboration with East-West Seed International. Most recently, 'Hyb8001r' was developed and introduced by CNAP as a new hybrid, which is now commercialized by East-West Seed International. In field trials worldwide, 'Hyb8001r' produced high ART concentration (up to 1.44%), high biomass (up to 4.4 kg/ha dry leaf), reaching high ART yields (54 kg/ha) (Suberu et al., 2016 – in this work, 'Hyb8001r' is referred to as CNAP8001).

Being the raw material of conventional breeding, the genetic basis of artemisia is fortunately quite broad (Li et al., 2017). Artemisia shows great phenotypic diversity for important traits, including leaf morphology, trichome density, plant size, as well as terpenoid profile, and artemisinin (ART) yield. As an example of this advantage for genetic breeding, a basic genetic map was produced with a segregating population developed from a crossing between two commercial parental populations that are used to produce the 'Artemis' F₁ hybrid (Graham et al., 2010). In the segregating population, ART content ranged extensively between 0.1-2.1%. Furthermore, doubling the artemisia ploidy (4X=36), may have a positive effect on yield, although the results are not conclusive (Lin et al., 2011; Wallaart et al., 1999), and the impact on susceptibilities to biotic and abiotic stresses or flower induction remain untested.

Boosting the terpenoid as well as the committed ART biosynthesis in glandular trichomes can be achieved by selecting alleles with strong promoters or optimized protein sequences that lead to a stronger flux of photosynthates towards these pathways as well as most substrates flowing towards sesquiterpene biosynthesis, and ensuring that the pathway does not branch off to synthesis of other products than DHAA and ART.

ART is a sesquiterpene lactone that holds an uncharacteristic peroxide bridge. It is synthesized in multicellular glandular trichomes that develop on the epidermis of aerial organs of the plant (Olofsson et al., 2012). The co-accumulation of its precursor dihydroartemisinic acid (DHAA) is often an overlooked aspect in artemisia breeding. DHAA is converted non-enzymatically into ART by UV light exposure, even during post-harvest sun drying (Ferreira and Luthria, 2010). Thus, it is reasonable to consider both, ART and DHAA, in genetic gain models for ART yields (Ferreira et al., 2018). In considering breeding for high ART yields, it is important to appreciate the genetic aspects involved in the ART pathway and consider not only the allelic variation and regulation of structural genes coding for enzymes or enzymatic units, but also the regulatory components, such as transcription factors.

Below, we briefly introduce the most relevant aspects of the general terpenoid biosynthesis pathway and leading to the committed ART pathway.

The general isoprenoid/terpenoid biosynthesis pathway

Quite exceptionally, plant terpenoids (a.k.a. isoprenoids) are synthesized via two distinct pathways in different compartments of the plant cell from precursors derived from sugar catabolism. The mevalonate pathway (MVA) uses acetyl-CoA as a precursor and occurs in the cytosol, possibly with some contribution by the ER. There have been suggestions that the MVA pathway also receives contributions from peroxisomes (Sapir-Mir et al., 2008; Clastre et al., 2011). In turn, the methylerythritol phosphate (MEP or DOXP) pathway occurs in plastids and utilizes glyceraldehyde-3-P (G3P) and pyruvate as precursors. Importantly for breeding purposes, the structural genes of the plastidial MEP pathway are encoded in the nucleus and the enzymes are imported into the plastid. Therefore, they display Mendelian segregation and are not cytoplasmically (unilaterally) inherited. In both cases, the general terpenoid pathways produce the 5-carbon compounds isopentenyl diphosphate (IPP or IDP) and its isomer dimethylallyl diphosphate (DMAPP) as building blocks of all isoprenoids. The steady state of IPP to DMAPP in plastids is estimated to be 6:1 (Eisenreich et al., 2004; Ma et al., 2016), which makes DMAPP a limiting substrate for terpenoid biosynthesis. Following the synthesis of IPP and DMAPP, the remaining steps of the terpene biosynthesis occurs in the cytosol for sesqui- and triterpenes, whereas mono-, di- and tetraterpenes are synthesized in the plastid (Chen et al., 2011). IPP and DMAPP are used to synthesize the 10-carbon compound geranyl diphosphate (GPP), the 15-carbon compound farnesyl diphosphate (FPP), and the 20-carbon compound geranylgeranyl diphosphate (GGPP) as terpenoid building blocks. In the first steps committed to sesquiterpene synthesis in the cytosol, GPP is produced by the fusion of IPP and DMAPP, to which another IPP unit is added to generate FPP.

The committed ART biosynthesis pathway

Since FPP is the precursor of all sesquiterpenes, for maximum accumulation, it is important that in glandular trichomes most FPP rather flows towards FPP synthesis and then ART synthesis. The entry point of the committed ART pathway occurs with FPP conversion to amorpha-4,11-diene (AD) by its synthase (ADS) in the cytosol. Consequently, it is the ADS sink strength - as defined by its gene expression level and catalytic properties of the enzyme (AD affinity and turnover rate) – that will determine how much of the cytosolic FPP pool will flow towards ART biosynthesis. Terpene synthases (TPS) encompass a large class of enzymes that utilize GPP, FPP, or GGPP to synthesize mono-, sesqui- and tri-, or di- and tetraterpenes, respectively. Four phylogenetically TPS clades contain enzymes with sesquiterpene synthase activities: TPS-a-1, some in TPS-c, TPS-e/f, and TPS-g (Chen et al., 2011). Although an ADS gene from artemisia responsible for AD synthesis has been identified (Lommen et al., 2007; Zhang et al., 2015), attention must be paid for alternate ADS isozymes to could cooperate with FPP flux to ART, as well as other TPS competing for substrate in glandular trichomes. In the next step in the pathway, the cytochrome P450 monooxygenase CYP71AV1, in association with the cytochrome P450 reductase (CPR), oxidizes AD to AA in three successive reactions via the intermediates artemisinic alcohol (AOH) and artemisinic aldehyde (AO). CYP71AV1 is the main rate-limiting enzyme of the committed ART biosynthesis (Lv et al., 2016; Chen et al., 2017). Although the CYP71AV1/CPR enzymatic system does not produce dihydroartemisinic aldehyde (DHAO), it can convert DHAO into dihydroartemisinic acid (DHAA). Therefore, ART biosynthesis competes with the AA pathway for AO. Manipulating this step is key to invert the high-AA, low-ART profile of certain genotypes. In the ART pathway, AO is reduced by the AO Δ 11(13) reductase encoded by the DOUBLE BOND REDUCTASE 2 (DBR2) gene into DHAO, which can be further reduced by either the CYP71AV1/CPR system or AOH dehydrogenase (ALDH1) into DHAA. In these convoluted, promiscuous enzymatic systems involved in the late pathway, it is critical to define allelic variants with higher affinity to substrates largely favoring ART biosynthesis, rather than AA. Finally, the last step of ART biosynthesis is the conversion of DHAA into ART, which is thought to occur non-enzymatically in the trichome subcuticular space (oil sac) through photo-oxidation by UV light (Czechowski et al., 2016). Alternatively, reactive oxygen species (ROS) that generated by abiotic stresses may be also contributing to this conversion (Nguyen et al., 2011).

Furthermore, several transcription factors activating the expression of structural genes of the ART biosynthesis, especially between FPP and AO, have been identified, such as TAR1, AabZIP1, AaERAF1/2, AaORA, AaWORKY, AabHLH1, and AaMYC2 (Shen et al., 2016; Lv et al., 2017). These are obvious candidates for having a large impact on ART synthesis and accumulation and mapping of their genes will potentially help selection of optimal combinations of allelic variants.

Despite being a crucial element in this pathway, the plasma membrane transporter that secretes DHAA from the apical secretory cells to the subcuticular sac has not been characterized yet.

Transmembrane transport - plastid to cytosol delivery of IPP and artemisinin efflux to the secretory reservoir

Plastidial contribution with IPP to the cytosolic terpenoid biosynthesis is facilitated by unidentified proton symporters present in the envelope membrane (Bick and Lange, 2003; Dudareva et al., 2005; Pick and Weber, 2014; Henry et al., 2015). At a fundamental level, the molecular identification of such plastidial transporters could help streaming the IPP pool towards sesquiterpene biosynthesis. Moreover, the fact that ART, or its immediate precursor DHAA, is synthesized in the secretory cells of glandular trichomes and then stored in the subcuticular reservoir, a transmembrane efflux system for ART or DHAA is required but remains unidentified. In plants, two members of the ABC transporter family were characterized as terpene transporters at the plasma membrane: ABC1 *Nicotiana plumbaginifolia* (Jasiński et al., 2001) and ABCG/PDR from *N. benthamiana* (Shibata et al., 2016).

Indeed, the coordination of the ten cells that compose the peltate glandular trichome towards ART biosynthesis in artemisia is not fully understood (Xiao et al., 2016). Only the four sub-apical cells contain chloroplasts, and it appears that only the two apical secretory cells express the late enzymes of the ART pathway (Nguyen et al., 2011; Xiao et al., 2016). It would be valuable to have a better resolution of gene expression of each enzyme in specific cell types of the glandular structure, such as by using RNA *in situ* hybridization or confocal microscopy with a fluorescent protein fusion under native promoter. How substrates are trafficked from one cell to another is also largely unknown, including how ART is secreted to the trichome subcuticular oil storage cavity (i.e., the cuticular sac or gland tip), as aforementioned. In the Chinese herb *Tripterygium wilfordii*, an MDR transporter (TwMDR1) was recently identified with affinity to the sesquiterpene pyridine alkaloids, wilforine and wilforgine. This opens up the possibility that an MDR transporter, in addition to ABC transporters, could be involved in the export of DHAA to the subcuticular sac (Miao et al., 2017). An analysis of membrane transporters exclusively expressed in secretory cells of glandular trichomes could lead to potential candidates that play these essential yet neglected functions in ART accumulation.

Breeding for high ART accumulation via targeted allelic selection of natural variants

By and large, for plant selection programs aiming at boosting ART+DHAA in artemisia should primarily focus on the following main traits:

1. **Enhanced general terpenoid and ART accumulation in glandular trichomes:** Most of the rate-limiting enzymes and many transcriptional regulators of the ART pathway have already been

characterized. The first key enzyme in the MVA pathway is 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR), which gene expression is positively correlated with ART content in artemisia (Alam and Abdin, 2011). The transcription factor AabHLH1 was shown to induce HMGR as well as some enzymes of the committed ART pathway, such as ADS and CYP71AV1 (Ji et al., 2014). To tackle DMAPP limitations for sesquiterpene biosynthesis, over-expression of the type-I IPP isomerase (AalPPI1) in the cytosol in artemisia leaf cells led to a 1:7 IPP-to-DMAPP ratio, which considerably increased the levels of arteannuin B (4%) along with yields of ~0.20% ART and 1% AA (Ma et al., 2016).

The sesquiterpene biosynthesis occurs in the cytosol with a terpene synthase fusing two IPP units to form farnesyl diphosphate (FPP or FDP) the primordial sesquiterpene (15C) molecule. Then, the first committed reaction of ART pathway is a cyclization catalyzed by ADS to produce amorpha-4,11-diene by a synthase (ADS). Virtually all structural genes of this pathway have been identified, along with various regulators. However, additional unknown genes may be contributing or competing with the ART pathway. For example, in the tomato genome, there are 44 terpene synthases (TPS) with 14 expressing in trichomes (Matsuba et al., 2013). The functional characterization of potential sesquiterpene synthases expressed in tomato glandular trichomes as a model organism could lead to the identification of TPS genes coding for enzymes with sesquiterpene synthase activity as well as their transcriptional regulators. Undoubtedly, the genome-wide identification of all TPS members in the artemisia genome would be highly beneficial and is now within reach with the availability of high-quality genome sequence and annotation (Shen et al., 2018).

Mapping of these genes on the artemisia genome will enable selection of allelic versions that may function as molecular markers for high ART accumulation. Combining an optimal set of alleles for enzymes and transcriptional regulators in the committed biosynthesis pathway may lead to major gains in yield. Moreover, the availability of a gene expression atlas for all different organs and main structures, the generation of regulatory gene networks and genome-wide association studies (GWAS) will be highly beneficial to further explore its genome for breeding purposes.

2. Improved final ART+DHAA profile and low accumulation of side products: For best results, the pathway must flow towards producing ART+DHAA, and not stray precursors to make compounds with no or little anti-malarial effects. The main products that drain artemisinic aldehyde (AO) from the final steps of ART biosynthesis, lowering the yield, are artemisinic acid (AA) and arteannuin B (AB). Expectedly, ART+DHAA yields are negatively correlated with AA+AB content. The accessions for which data are available show phenotypical variation leading to massive yield differences for each of these compounds. For example, a recent study showed very

distinct chemical profiles of Chinese accessions (Zhang et al., 2017). Therefore, selection of allelic variants with high ART+DHAA and low in AA+AB is feasible. This goal can be accomplished by optimizing gene expression and/or catalytic activity of DBR2 to compete for the AO substrate in order to maximally produce DHAA. Moreover, the use of high AA variants to compare at the gene variation (at expression and/or protein sequence levels) potentially responsible for the variation in chemical profiles could reveal the most important genetic variations that define ART/AA branching of the pathway.

3. **High glandular trichome density on leaves and floral calyx:** Glandular trichomes are multicellular epidermal appendices on aerial organs of a plant. They are generally regarded as chemical factories of the plant (Glas et al., 2012; Huchelmann et al., 2017). The development of higher densities of glandular trichomes on the surface of aerial organs directly causes higher accumulation of ART. The density of glandular trichomes in *Artemisia annua* can vary from none (glandless mutants) up to over 25 glands per mm² for the adaxial side of fully expanded leaves. However, it will be interesting in future to compute trichome densities on both, adaxial and abaxial sides of the leaf, in mathematical models predicting ART yields. Recently, the closely-related species, *A. deserti* was reported as having much higher glandular trichome densities amongst five species analyzed in the genus (Salehi et al., 2018), which could make it an ideal material for introgression of the trait in breeding programs.

Much has been unveiled in recent years about the developmental pathway of non-glandular (tector) trichomes in Arabidopsis at the molecular genetics level (Pattanaik et al., 2014). Although relatively little has been achieved regarding glandular trichome developmental pathways, recent advances are quickly changing this landscape. For example, the MYB transcription factor AmMYBML3 from *Antirrhinum majus* and the bHLH transcription factor PIGMENT GLAND FORMATION (PGF) from cotton were identified as key regulators of glandular trichome formation (Jaffé et al., 2007; Ma et al., 2016). Very recently, the R2R3 MYB transcription factor AaMIXTA1 was identified in artemisia as an inducer of glandular trichome development as well as cuticle biosynthesis (Shi et al., 2018). Additionally, besides inducing enzymatic genes, the AP2 transcription factor AaTAR1 is also involved in trichome development (Tan et al., 2015). Which other players are involved with glandular trichome development and whether these pathways are conserved in *Artemisia annua* remains to be discovered. Studies using a related Asteraceae model species with genome sequence available, such as sunflower (Aschenbrenner et al., 2015; Badouin et al., 2017), which belongs to the same Asteroideae subfamily, could help elucidate these questions.

4. **Total leaf area (size and shape):** Leaf size is a highly variable trait in *Artemisia annua*, varying at least ten-fold, from about 500 to 5,000 mm² (Graham et al., 2010). A key QTL for leaf

area was mapped to the linkage group 8, although the exact gene remains to be identified (Graham et al., 2010). Further studies to identify the genetic determination of leaf size in *Artemisia annua* will be very informative to increase ART yield.

As a common feature found in members of the Asteraceae family, the shape of the leaf lamina in *Artemisia annua* is highly subdivided into 5-7 leaflets connected by the rachis. This conformation dramatically reduces the total area available for the development of glandular trichomes, not to mention limiting photosynthesis for biomass accumulation. Leaf complexity is a function of continued meristematic activity at the leaf margins in the shoot apical meristem and in plant models, such as *Arabidopsis*, involves a complex genetic network coordinated by several transcription factors, including NO APICAL MERISTEM (NAM), ARABIDOPSIS TRANSCRIPTION ACTIVATION FACTOR 1 and 2 (ATAF1/2), CUP-SHAPED COTYLEDON 2 (CUC2), not to mention small RNA regulators, such as miR164 (Wang et al., 2016). A better understanding of the molecular basis of compound leaf development in *Artemisia annua* could lead to the development of tools to generate a genotype with a simpler leaf pattern with increased lamina area. Therefore, recognizing the functional homologs of these transcription factors in the Asteraceae, and especially in *Artemisia annua* could lead to an important breakthrough in leaf reshaping for leaf yield increases. Moreover, genome sequencing of *A. annua* and further narrowing of loci associated with relevant traits to ART yield via genome-wide association studies (GWAS) will greatly speed up gene discovery and breeding.

5. **High biomass production and high leaf-to-stem ratio:** Along with leaf parameters (area, perimeter and specific area), the total plant size (height and stature) and fresh weight were associated with artemisinin yields (Graham et al., 2010). A QTL for stature fell in linkage group 2 where a gene homolog of the shoot branching regulator and strigolactone biosynthesis, MORE AXILLARY BRANCHING 3 (MAX3/CCD7), is located (Bennett et al., 2006). Meanwhile, a QTL for fresh weight was mapped in linkage group 4, albeit no obvious gene was found in that genomic interval (Graham et al., 2010). In this species, flowering is induced under short-day conditions. The day length threshold for flower induction depends on the genotype and show genetic variability in the range of 12 to 16 hours (Ferreira et al., 1995). Thus, a better characterization of the genetic pathways associated with vegetative and reproductive development in *A. annua*, starting with the identification of flowering gene homologs will lead to the design of molecular tools to advance breeding to expand the latitude range of this crop.

6. **Resistance to biotic and abiotic stresses:** in addition to high ART yields, crop production requires plants to be tolerant to edaphoclimatic stresses (extreme temperatures, poor soils, drought) and pests. So far, there has been little evaluation of tolerance of accessions to diverse

stresses. The main diseases of artemisia in the field have not been reported. In greenhouse conditions, the main pests we observed are mealybugs, thrips, and spider mites. Selecting for accessions tolerant to stresses in field conditions is essential to generating successful field varieties.

The genome size of *A. annua* is 1.74 Gb and its sequence has been published recently (Shen et al., 2018). The artemisia genome is only slightly larger than the human genome, twice that of tomato and almost 13-fold of Arabidopsis. Very recently, the sequence of its plastid genome was published (Shen et al., 2017), which can be useful to explore the MEP pathway that contributes to ART biosynthesis. An initial genetic mapping of loci contributing to important traits was produced with a segregating population (Graham et al., 2010), but the resolution was low and in most cases, it did not allow for the identification of candidate genes within each locus. Revisiting this dataset in the light of the newly published genome sequence might be revealing. The use of genetic transformation can be used to boost ART and DHAA yields (Elfahmi et al., 2014; Tang et al., 2014). Genetic transformation and *in vitro* regeneration protocols for *A. annua* are available (Alok et al., 2016; Alam et al., 2014). For example, substrate channeling was obtained by expressing ADS-FDS fused enzymatic system under constitutive expression or the trichome-specific promoter of the CYP71AV1 gene in order to converge the general terpene intermediates (IDP and DMADP) into directly producing amorpho-4,11-diene. This strategy more than doubled ART yields (Han et al., 2016). Further, making a gene expression atlas available would be invaluable to move forward with gene discovery and potentially lead to leaps and bounds of breeding advancement in the crop.

Genome editing as a potential approach

The CRISPR/Cas9 genome editing system has revolutionized the way we think of crop breeding. Now it is feasible to quickly engineer a plant genome to result in the desirable phenotypes. And this can be accomplished without the final product being transgenic, especially when only gene knockout is desired. Once the transgenic cassette carrying out the T-DNA coding for the Cas9 and gRNAs (and, when the case, other additional elements), once the genome is edited (usually already in the T₀ generation, which the T-DNA is in hemizygosity), the transgenic regenerant line can be allowed to sexually reproduce and segregate the T-DNA out, thus resulting in a non-transgenic plant (for all practical and scientific purposes, although policies and politics may dispute this notion) with its genome edited. Albeit still challenging, but within reach in near future, manipulation of expression of inherent genes under certain conditions (inducing, repressing transcription) will become feasible. In this case, CRISPR/Cas9-driven systems with modified nucleases (fused of other proteins and with the nuclease activity dead) still able to recognize and bind to promoter regions of the genes of interest but instead of cutting the DNA,

they call (or hinder) the transcriptional machinery to the promoter, or base-editing enzymes to the site to be mutated (Pandiarajan and Grover, 2018; Murovec et al., 2017; Demirci et al., 2018).

The *de novo* domestication of crop species was suggested as a potential novel approach to create varieties that are highly resilient to environmental stresses (Zsögön et al., 2017). More recently, the approach was experimentally demonstrated for tomato (Zsögön et al., 2018) and gooseberry (Lemmon et al., 2018). In the first case, knockout of just six genes of the closest relative of the cultivated tomato (*Solanum pimpinellifolium*) resulted in a more condensed plant architecture (half of the wild-type stature, and with half of the sympodial index) that produced more flowers (six-fold increase), and larger fruits (three times heavier) with a preferred shape (more oblong) and higher nutritional content (five-fold increase in lycopene).

For genome editing of plant species, sequence of target genes (preferably, information on the whole genome is useful to prevent off-targeting) as well as efficient transformation and regeneration protocols are required. The nuclear and plastidial genome sequences of artemisia has been made available recently (Shen et al., 2017, 2018). An efficient method of genetically transforming artemisia via *Agrobacterium tumefaciens* has also been described (Alam et al., 2014).

In artemisia, key genes to be targeted are the early structural genes in the terpenoid pathway that compete for C skeletons with the committed ART pathway – the main structures drained from the main ART pathway are IPP/DMAPP to make hemiterpenes, GPP to make monoterpenes, and FPP to make other sesquiterpenes and higher-order terpenoid compounds. The main enzymes identified to deviate FPP from ART biosynthesis are CPS (towards β -carophyllene), SQS (towards squalene and sterols), BFS (towards β -farnesene), and GAS (towards germacrene A) (Chen et al., 2011). In addition, blocking the RED1 enzyme, which competes for dihydroartemisinic aldehyde (Shi et al., 2017) could also be beneficial to increase the ART biosynthesis flux.

In addition, enabling the plant to self-pollinate would be highly beneficial to create true varieties with stable genetics (i.e., in homozygosity to virtually the whole gene set) that can be propagated by seeds. A self-pollinating genotype was reported for *A. annua* (Alejos-Gonzalez et al., 2011), although the molecular genetics mechanism behind this trait remains unknown. In tomato (Solanaceae), pollen rejection and self-incompatibility involves an S-RNase (gene Solyc01g055200) and HT proteins (Tovar-Méndez et al., 2017) – such mechanisms should be identified in the Asteraceae ((Allen et al., 2011; Koseva et al., 2017) in order to design self-compatible *A. annua* plants.

Boosting ART and DHAA yields in *Artemisia annua* through crop management and post-harvest practices

Moreover, understanding the phenology of the crop is important to optimize biochemical yields. For example, glandular trichomes reach their plumpness maximum after the leaf reaches maturity and tend to thin out as the organ gets older. The best harvesting time for maximum ART yield has been identified as at the anthesis (Ferreira et al., 1995; Arsenault et al., 2010). The nutritional and physiological conditions of the crop can affect ART accumulation. A mild K deficiency increased 75% ART content (Ferreira, 2007). Cold shock treatment (4°C for 6h) improved ART yields threefold to about 5% (Liu et al., 2017). Sun drying the leaves for one to three weeks was effective in converting almost all DHAA into ART and increased ART yield in almost 50% compared to freeze-drying. On the other hand, drying leaves in a forced-air oven at 45°C for 12-16h led to 40% conversion of DHAA into ART (Ferreira and Luthria, 2010).

Conclusions

Compared to crops that have been domesticated over millennia, *Artemisia annua* can be considered still a wild species that requires concerted efforts to be tamed as a viable, easy to grow crop. Metabolic engineering and plant design via genome editing can be of great help to speed up the process of making it self-compatible, less-inducible to flowering, and maximally channel C skeletons towards ART biosynthesis. ART yields of 5% have been reported and may be an achievable goal.

In toto, combining ideal varieties, environmental conditions, crop management and post-harvest practices will make artemisia a more viable cash crop to supply the demand for ART to the anti-malarial needs of the world.

SUMMARY AND CONCLUSION

In our study, the genotype J showed best performance regarding the anatomical characters important for high ART yield. J was among the tallest genotypes, with heaviest aerial biomass, and also among the genotypes with largest leaf areas. However, it was not among the genotypes with highest trichome densities.

Next to J, the genotypes I and G also showed high performance in height, biomass and leaf area categories, but they were not among the genotypes with high trichome densities neither. The genotypes with highest trichome density were A, C, D, E and K. Genotype D was among the one with highest biomass and leaf area whereas the genotype E was amongst those with highest leaf area.

We identified camphor and endo-borneol as key compounds to distinguish the genotypes in the collection by biochemical profiling. We could also classify genotype H as the only one belonging to chemotype LAP, whereas all other genotypes are considered of chemotype HAP.

The genotype J had the highest accumulation of ART+DHAA accumulation compared to the rest of the collection. Considering these active compounds separately, the genotypes G, I and C had the highest ART content whereas G, I and E had the highest DHAA content.

From the 15 genes tested for gene expression variation among the collection, only three (FPS, CPR, and DBR2) were differentially expressed. However, the biological variation observed within genotypes was too high to identify genotypes with the highest transcriptional activities.

Based on anatomical, biochemical and genotypic data, the genotype J were the best performer for high ART and DHAA yield as well as morphological features. Next, I and G were best genotypes in terms of ART and DHAA yield.

We discussed parameters to be considered in breeding programs of *Artemisia annua*, make a case that this species remains undomesticated and propose biotechnological tools to speed up the process to make it a viable crop.

REFERENCES

- Abad, M.J., Bedoya, L.M., Apaza, L., and Bermejo, P.** (2012). The artemisia L. Genus: a review of bioactive essential oils. *Molecules* **17**: 2542–2566.
- Aftab, T., Khan, M.M.A., Idrees, M., Naeem, M., Singh, M., and Ram, M.** (2010). Stimulation of crop productivity, photosynthesis and artemisinin production in *Artemisia annua* L. by triacontanol and gibberellic acid application. *Journal of Plant Interactions* **5**: 273–281.
- Agati, G., Biricolti, S., Guidi, L., Ferrini, F., Fini, A., and Tattini, M.** (2011). The biosynthesis of flavonoids is enhanced similarly by UV radiation and root zone salinity in *L. vulgare* leaves. *J. Plant Physiol.* **168**: 204–212.
- Alam, P. and Abdin, M.Z.** (2011). Over-expression of HMG-CoA reductase and amorpho-4,11-diene synthase genes in *Artemisia annua* L. and its influence on artemisinin content. *Plant Cell Rep* **30**: 1919–1928.
- Alam, P., Mohammad, A., Ahmad, M.M., Khan, M.A., Nadeem, M., Khan, R., Akmal, M., Ahlawat, S., and Abdin, M.Z.** (2014). Efficient method for *Agrobacterium* mediated transformation of *Artemisia annua* L. *Recent Pat Biotechnol* **8**: 102–107.
- Alejos-Gonzalez, F., Qu, G., Zhou, L.-L., Saravitz, C.H., Shurtleff, J.L., and Xie, D.-Y.** (2011). Characterization of development and artemisinin biosynthesis in self-pollinated *Artemisia annua* plants. *Planta* **234**: 685–697.
- Allen, A.M., Thorogood, C.J., Hegarty, M.J., Lexer, C., and Hiscock, S.J.** (2011). Pollen-pistil interactions and self-incompatibility in the Asteraceae: new insights from studies of *Senecio squalidus* (Oxford ragwort). *Ann. Bot.* **108**: 687–698.
- Alok, A., Shukla, V., Pala, Z., Kumar, J., Kudale, S., and Desai, N.** (2016). In vitro regeneration and optimization of factors affecting *Agrobacterium* mediated transformation in *Artemisia Pallens*, an important medicinal plant. *Physiol Mol Biol Plants* **22**: 261–269.
- Arsenault, P.R., Vail, D., Wobbe, K.K., Erickson, K., and Weathers, P.J.** (2010). Reproductive Development Modulates Gene Expression and Metabolite Levels with Possible Feedback Inhibition of Artemisinin in *Artemisia annua*. *Plant Physiology* **154**: 958–968.
- Aschenbrenner, A.-K., Amrehn, E., Bechtel, L., and Spring, O.** (2015). Trichome differentiation on leaf primordia of *Helianthus annuus* (Asteraceae): morphology, gene expression and metabolite profile. *Planta* **241**: 837–846.
- Badouin, H. et al.** (2017). The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. *Nature* **546**: 148–152.
- Badshah, S.L., Ullah, A., Ahmad, N., Almarhoon, Z.M., and Mabkhot, Y.** (2018). Increasing the Strength and Production of Artemisinin and Its Derivatives. *Molecules* **23**: 100.

- Banyai, W., Mii, M., and Supaibulwatana, K.** (2011). Enhancement of artemisinin content and biomass in *Artemisia annua* by exogenous GA3 treatment. *Plant Growth Regul* **63**: 45–54.
- Bennett, T., Sieberer, T., Willett, B., Booker, J., Luschnig, C., and Leyser, O.** (2006). The Arabidopsis MAX pathway controls shoot branching by regulating auxin transport. *Curr. Biol.* **16**: 553–563.
- Bhakuni, R.S., Jain, D.C., Sharma, R.P., and Kumar, S.** (2001). Secondary metabolites of *Artemisia annua* and their biological activity. *Current Science* **80**: 35–48.
- Bick, J.A. and Lange, B.M.** (2003). Metabolic cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis: unidirectional transport of intermediates across the chloroplast envelope membrane. *Arch. Biochem. Biophys.* **415**: 146–154.
- Bilia, A.R., Santomauro, F., Sacco, C., Bergonzi, M.C., and Donato, R.** (2014). Essential Oil of *Artemisia annua* L.: An Extraordinary Component with Numerous Antimicrobial Properties. *Evid Based Complement Alternat Med* **2014**.
- Bouwmeester, H.J., Wallaart, T.E., Janssen, M.H.A., van Loo, B., Jansen, B.J.M., Posthumus, M.A., Schmidt, C.O., De Kraker, J.-W., König, W.A., and Franssen, M.C.R.** (1999). Amorpha-4,11-diene synthase catalyses the first probable step in artemisinin biosynthesis. *Phytochemistry* **52**: 843–854.
- Brown, G.D.** (2010). The biosynthesis of artemisinin (Qinghaosu) and the phytochemistry of *Artemisia annua* L. (Qinghao). *Molecules* **15**: 7603–7698.
- Brown, G.D. and Sy, L.-K.** (2004). In vivo transformations of dihydroartemisinic acid in *Artemisia annua* plants. *Tetrahedron* **60**: 1139–1159.
- Bryant, L., Flatley, B., Patole, C., Brown, G.D., and Cramer, R.** (2015). Proteomic analysis of *Artemisia annua* – towards elucidating the biosynthetic pathways of the antimalarial pro-drug artemisinin. *BMC Plant Biology* **15**: 175.
- Chang, Y.-C., Chang, T.-J., Jiang, Y.-D., Kuo, S.-S., Lee, K.-C., Chiu, K.C., and Chuang, L.-M.** (2007). Association Study of the Genetic Polymorphisms of the Transcription Factor 7-Like 2 (TCF7L2) Gene and Type 2 Diabetes in the Chinese Population. *Diabetes* **56**: 2631–2637.
- Chen, J.-L., Fang, H.-M., Ji, Y.-P., Pu, G.-B., Guo, Y.-W., Huang, L.-L., Du, Z.-G., Liu, B.-Y., Ye, H.-C., Li, G.-F., and Wang, H.** (2011). Artemisinin biosynthesis enhancement in transgenic *Artemisia annua* plants by downregulation of the β -caryophyllene synthase gene. *Planta Med.* **77**: 1759–1765.
- Chen, X., Zhang, C., and Too, H.-P.** (2017). Multienzyme Biosynthesis of Dihydroartemisinic Acid. *Molecules* **22**: 1422.
- Choi, B.R., Kim, H.K., and Park, J.K.** (2017). Penile Erection Induced by Scoparone from *Artemisia capillaris* through the Nitric Oxide-Cyclic Guanosine Monophosphate Signaling Pathway. *World J Mens Health* **35**: 196–204.
- Christen, P. and Veuthey, J.L.** (2001). New trends in extraction, identification and quantification of artemisinin and its derivatives. *Curr. Med. Chem.* **8**: 1827–1839.

- Clastre, M., Papon, N., Courdavault, V., Giglioli-Guivarc'h, N., St-Pierre, B., and Simkin, A.J.** (2011). Subcellular evidence for the involvement of peroxisomes in plant isoprenoid biosynthesis. *Plant Signal Behav* **6**: 2044–2046.
- Covello, P.S.** (2008). Making artemisinin. *Phytochemistry* **69**: 2881–2885.
- Czechowski, T., Larson, T.R., Catania, T.M., Harvey, D., Brown, G.D., and Graham, I.A.** (2016). *Artemisia annua* mutant impaired in artemisinin synthesis demonstrates importance of nonenzymatic conversion in terpenoid metabolism. *Proc Natl Acad Sci U S A* **113**: 15150–15155.
- Czechowski, T., Larson, T.R., Catania, T.M., Harvey, D., Wei, C., Essome, M., Brown, G.D., and Graham, I.A.** (2018). Detailed Phytochemical Analysis of High- and Low Artemisinin-Producing Chemotypes of *Artemisia annua*. *Front. Plant Sci.* **9**.
- Daddy, N.B., Kalisya, L.M., Bagire, P.G., Watt, R.L., Towler, M.J., and Weathers, P.J.** (2017). *Artemisia annua* dried leaf tablets treated malaria resistant to ACT and i.v. artesunate: Case reports. *Phytomedicine* **32**: 37–40.
- Davies, M.J. et al.** (2009). Enhancement of artemisinin concentration and yield in response to optimization of nitrogen and potassium supply to *Artemisia annua*. *Ann Bot* **104**: 315–323.
- Delabays, N., Simonnet, X., and Gaudin, M.** (2001). The genetics of artemisinin content in *Artemisia annua* L. and the breeding of high yielding cultivars. *Curr. Med. Chem.* **8**: 1795–1801.
- Demirci, Y., Zhang, B., and Unver, T.** (2018). CRISPR/Cas9: An RNA-guided highly precise synthetic tool for plant genome editing. *J. Cell. Physiol.* **233**: 1844–1859.
- Di Ferdinando, M., Brunetti, C., Fini, A., and Tattini, M.** (2012). Flavonoids as Antioxidants in Plants Under Abiotic Stresses. In *Abiotic Stress Responses in Plants: Metabolism, Productivity and Sustainability*, P. Ahmad and M.N.V. Prasad, eds (Springer New York: New York, NY), pp. 159–179.
- Dudareva, N., Andersson, S., Orlova, I., Gatto, N., Reichelt, M., Rhodes, D., Boland, W., and Gershenzon, J.** (2005). The nonmevalonate pathway supports both monoterpene and sesquiterpene formation in snapdragon flowers. *Proc. Natl. Acad. Sci. U.S.A.* **102**: 933–938.
- Duke, S.O. and Paul, R.N.** (1993). Development and Fine Structure of the Glandular Trichomes of *Artemisia annua* L. *International Journal of Plant Sciences* **154**: 107–118.
- Eisenreich, W., Bacher, A., Arigoni, D., and Rohdich, F.** (2004). Biosynthesis of isoprenoids via the non-mevalonate pathway. *Cell. Mol. Life Sci.* **61**: 1401–1426.
- Elfahmi, Suhandono, S., and Chahyadi, A.** (2014). Optimization of genetic transformation of *Artemisia annua* L. Using *Agrobacterium* for Artemisinin production. *Pharmacognosy Magazine* **10**: 176.
- Elfawal, M.A., Towler, M.J., Reich, N.G., Weathers, P.J., and Rich, S.M.** (2015). Dried whole-plant *Artemisia annua* slows evolution of malaria drug resistance and overcomes resistance to artemisinin. *PNAS* **112**: 821–826.

- Elhag, H.M., El-Domiaty, M.M., El-Feraly, F.S., Mossa, J.S., and El-Olemy, M.M.** (1992). Selection and micropropagation of high artemisinin producing clones of *Artemisia annua* L. *Phytotherapy Research* **6**: 20–24.
- Escalante, A.A. and Ayala, F.J.** (1994). Phylogeny of the malarial genus *Plasmodium*, derived from rRNA gene sequences. *Proc Natl Acad Sci U S A* **91**: 11373–11377.
- Esu, E., Effa, E.E., Opie, O.N., Uwaoma, A., and Meremikwu, M.M.** (2014). Artemether for severe malaria. *Cochrane Database Syst Rev*: CD010678.
- Farhi, M., Marhevka, E., Ben-Ari, J., Algamas-Dimantov, A., Liang, Z., Zeevi, V., Edelbaum, O., Spitzer-Rimon, B., Abeliovich, H., Schwartz, B., Tzfira, T., and Vainstein, A.** (2011). Generation of the potent anti-malarial drug artemisinin in tobacco. *Nat. Biotechnol.* **29**: 1072–1074.
- Ferreira, J. and Janick, J.** (2009). Annual Wormwood (*Artemisia annua* L.): 9.
- Ferreira, J.F., Simon, J.E., and Janick, J.** (1995). Developmental studies of *Artemisia annua*: flowering and artemisinin production under greenhouse and field conditions. *Planta Med.* **61**: 167–170.
- Ferreira, J.F.S.** (2007). Nutrient Deficiency in the Production of Artemisinin, Dihydroartemisinic Acid, and Artemisinic Acid in *Artemisia annua* L. *J. Agric. Food Chem.* **55**: 1686–1694.
- Ferreira, J.F.S., Benedito, V.A., Sandhu, D., Marchese, J.A., and Liu, S.** (2018). Seasonal and Differential Sesquiterpene Accumulation in *Artemisia annua* Suggest Selection Based on Both Artemisinin and Dihydroartemisinic Acid may Increase Artemisinin in planta. *Front Plant Sci* **9**.
- Ferreira, J.F.S. and Janick, J.** (1995). Floral Morphology of *Artemisia annua* with Special Reference to Trichomes. *International Journal of Plant Sciences* **156**: 807–815.
- Ferreira, J.F.S., Laughlin, J.C., Delabays, N., and de Magalhães, P.M.** (2005). Cultivation and genetics of *Artemisia annua* L. for increased production of the antimalarial artemisinin. *Plant Genetic Resources* **3**: 206–229.
- Ferreira, J.F.S. and Luthria, D.L.** (2010). Drying Affects Artemisinin, Dihydroartemisinic Acid, Artemisinic Acid, and the Antioxidant Capacity of *Artemisia annua* L. Leaves. *J. Agric. Food Chem.* **58**: 1691–1698.
- Ferreira, J.F.S., Luthria, D.L., Sasaki, T., and Heyerick, A.** (2010). Flavonoids from *Artemisia annua* L. as antioxidants and their potential synergism with artemisinin against malaria and cancer. *Molecules* **15**: 3135–3170.
- Fuentes, P., Zhou, F., Erban, A., Karcher, D., Kopka, J., and Bock, R.** (2016). A new synthetic biology approach allows transfer of an entire metabolic pathway from a medicinal plant to a biomass crop. *eLife* **5**: e13664.
- Glas, J.J., Schimmel, B.C.J., Alba, J.M., Escobar-Bravo, R., Schuurink, R.C., and Kant, M.R.** (2012). Plant Glandular Trichomes as Targets for Breeding or Engineering of Resistance to Herbivores. *International Journal of Molecular Sciences* **13**: 17077–17103.

- Global malaria diagnostic and artemisinin treatment commodities demand forecast (2018).**
- Graham, I.A. et al. (2010).** The genetic map of *Artemisia annua* L. identifies loci affecting yield of the antimalarial drug artemisinin. *Science* **327**: 328–331.
- Han, J., Wang, H., Kanagarajan, S., Hao, M., Lundgren, A., and Brodelius, P.E. (2016).** Promoting Artemisinin Biosynthesis in *Artemisia annua* Plants by Substrate Channeling. *Mol Plant* **9**: 946–948.
- Han, J., Wang, H., Lundgren, A., and Brodelius, P.E. (2014).** Effects of overexpression of AaWRKY1 on artemisinin biosynthesis in transgenic *Artemisia annua* plants. *Phytochemistry* **102**: 89–96.
- Henry, L.K., Gutensohn, M., Thomas, S.T., Noel, J.P., and Dudareva, N. (2015).** Orthologs of the archaeal isopentenyl phosphate kinase regulate terpenoid production in plants. *PNAS* **112**: 10050–10055.
- Hu, S., Xu, Z., Pan, J., and Hou, Y. (1993).** The glandular trichomes of *Artemisia annua* L. and their secretions. *Journal of Research and Education in Indian Medicine* **July–September**: 9–15.
- Huchelmann, A., Boutry, M., and Hachez, C. (2017).** Plant Glandular Trichomes: Natural Cell Factories of High Biotechnological Interest. *Plant Physiology* **175**: 6–22.
- Ikram, N.K.B.K. and Simonsen, H.T. (2017).** A Review of Biotechnological Artemisinin Production in Plants. *Front Plant Sci* **8**.
- Jacometo, C.B., Schmitt, E., Pfeifer, L.F.M., Schneider, A., Bado, F., da Rosa, F.T., Halfen, S., Del Pino, F.A.B., Loor, J.J., Corrêa, M.N., and Dionello, N.J.L. (2014).** Linoleic and α -linolenic fatty acid consumption over three generations exert cumulative regulation of hepatic expression of genes related to lipid metabolism. *Genes Nutr* **9**.
- Jaffé, F.W., Tattersall, A., and Glover, B.J. (2007).** A truncated MYB transcription factor from *Antirrhinum majus* regulates epidermal cell outgrowth. *J Exp Bot* **58**: 1515–1524.
- Jain, D.C., Mathur, A.K., Gupta, M.M., Singh, A.K., Verma, R.K., Gupta, A.P., and Kumar, S. (1996).** Isolation of high artemisinin-yielding clones of *Artemisia annua*. *Phytochemistry* **43**: 993–1001.
- Jasiński, M., Stukkens, Y., Degand, H., Purnelle, B., Marchand-Brynaert, J., and Boutry, M. (2001).** A plant plasma membrane ATP binding cassette-type transporter is involved in antifungal terpenoid secretion. *Plant Cell* **13**: 1095–1107.
- Ji, Y., Xiao, J., Shen, Y., Ma, D., Li, Z., Pu, G., Li, X., Huang, L., Liu, B., Ye, H., and Wang, H. (2014).** Cloning and characterization of AabHLH1, a bHLH transcription factor that positively regulates artemisinin biosynthesis in *Artemisia annua*. *Plant Cell Physiol.* **55**: 1592–1604.
- Kindermans, J.-M., Pilloy, J., Olliaro, P., and Gomes, M. (2007).** Ensuring sustained ACT production and reliable artemisinin supply. *Malar J* **6**: 125.

- Kjær, A., Grevsen, K., and Jensen, M.** (2012). Effect of external stress on density and size of glandular trichomes in full-grown *Artemisia annua*, the source of anti-malarial artemisinin. *AoB Plants* **2012**.
- Kjaer, A., Grevsen, K., and Jensen, M.** (2014). Trichomes in *Artemisia annua*: Initiation, Development, Maturation and the Possibilities to Influence these Factors. In *Artemisia annua - Pharmacology and Biotechnology*, T. Aftab, J.F.S. Ferreira, M.M.A. Khan, and M. Naeem, eds (Springer Berlin Heidelberg: Berlin, Heidelberg), pp. 113–123.
- van der Kooy, F. and Sullivan, S.E.** (2013). The complexity of medicinal plants: the traditional *Artemisia annua* formulation, current status and future perspectives. *J Ethnopharmacol* **150**: 1–13.
- Koseva, B., Crawford, D.J., Brown, K.E., Mort, M.E., and Kelly, J.K.** (2017). The genetic breakdown of sporophytic self-incompatibility in *Tolpis coronopifolia* (Asteraceae). *New Phytol.* **216**: 1256–1267.
- Krishna, S., Bustamante, L., Haynes, R.K., and Staines, H.M.** (2008). Artemisinins: their growing importance in medicine. *Trends Pharmacol. Sci.* **29**: 520–527.
- Kumar, S., Gupta, S.K., Singh, P., Bajpai, P., Gupta, M.M., Singh, D., Gupta, A.K., Ram, G., Shasany, A.K., and Sharma, S.** (2004). High yields of artemisinin by multi-harvest of *Artemisia annua* crops. *Industrial Crops and Products* **19**: 77–90.
- Kumar, S. and Pandey, A.K.** (2013). Chemistry and Biological Activities of Flavonoids: An Overview. *The Scientific World Journal*.
- Larson, T.R., Branigan, C., Harvey, D., Penfield, T., Bowles, D., and Graham, I.A.** (2013). A survey of artemisinic and dihydroartemisinic acid contents in glasshouse and global field-grown populations of the artemisinin-producing plant *Artemisia annua* L. *Industrial Crops and Products* **45**: 1–6.
- Laughlin, J.C.** (1994). Agricultural production of artemisinin—a review. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **88**: 21–22.
- Le Bras, J. and Durand, R.** (2003). The mechanisms of resistance to antimalarial drugs in *Plasmodium falciparum*. *Fundam Clin Pharmacol* **17**: 147–153.
- Lemmon, Z.H., Reem, N.T., Dalrymple, J., Soyk, S., Swartwood, K.E., Rodriguez-Leal, D., Eck, J.V., and Lippman, Z.B.** (2018). Rapid improvement of domestication traits in an orphan crop by genome editing. *Nature Plants* **4**: 766.
- Li, L., Josef, B.A., Liu, B., Zheng, S., Huang, L., and Chen, S.** (2017). Three-Dimensional Evaluation on Ecotypic Diversity of Traditional Chinese Medicine: A Case Study of *Artemisia annua* L. *Front. Plant Sci.* **8**.
- Li, L., Wu, Y., Ma, P., Cui, G., Zhong, G., Wang, M., and Li, F.** (2010). [Breeding and spreading of new variety “Yu-Qing No. 1” of *Artemisia annua*]. *Zhongguo Zhong Yao Za Zhi* **35**: 2516–2522.

- Liu, D., Yang, M., Shao, A., Wang, X., Huang, L., Tang, S., Jin, H., and Fan, Z.** (2010a). [Correlation and path analysis on artemisinin content and yield with different agronomic traits of *Artemisia annua*]. *Zhongguo Zhong Yao Za Zhi* **35**: 2801–2807.
- Liu, M., Shi, P., Fu, X., Brodelius, P.E., Shen, Q., Jiang, W., He, Q., and Tang, K.** (2016). Characterization of a trichome-specific promoter of the aldehyde dehydrogenase 1 (ALDH1) gene in *Artemisia annua*. *Plant Cell Tiss Organ Cult* **126**: 469–480.
- Liu, W. et al.** (2010b). Origin of the human malaria parasite *Plasmodium falciparum* in gorillas. *Nature* **467**: 420–425.
- Liu, W., Wang, H., Chen, Y., Zhu, S., Chen, M., Lan, X., Chen, G., and Liao, Z.** (2017). Cold stress improves the production of artemisinin depending on the increase in endogenous jasmonate. *Biotechnology and Applied Biochemistry* **64**: 305–314.
- Lommen, W.J.M., Elzinga, S., Verstappen, F.W.A., and Bouwmeester, H.J.** (2007). Artemisinin and sesquiterpene precursors in dead and green leaves of *Artemisia annua* L. crops. *Planta Med.* **73**: 1133–1139.
- Lommen, W.J.M., Schenk, E., Bouwmeester, H.J., and Verstappen, F.W.A.** (2006). Trichome dynamics and artemisinin accumulation during development and senescence of *Artemisia annua* leaves. *Planta Med.* **72**: 336–345.
- Lu, X., Shen, Q., Zhang, L., Zhang, F., Jiang, W., Lv, Z., Yan, T., Fu, X., Wang, G., and Tang, K.** (2013). Promotion of artemisinin biosynthesis in transgenic *Artemisia annua* by overexpressing ADS, CYP71AV1 and CPR genes. *Industrial Crops and Products* **49**: 380–385.
- Lv, Z., Zhang, F., Pan, Q., Fu, X., Jiang, W., Shen, Q., Yan, T., Shi, P., Lu, X., Sun, X., and Tang, K.** (2016). Branch Pathway Blocking in *Artemisia annua* is a Useful Method for Obtaining High Yield Artemisinin. *Plant Cell Physiol* **57**: 588–602.
- Lv, Z., Zhang, L., and Tang, K.** (2017). New insights into artemisinin regulation. *Plant Signaling & Behavior* **12**: e1366398.
- Ma, C., Wang, H., Lu, X., Xu, G., and Liu, B.** (2008). Metabolic fingerprinting investigation of *Artemisia annua* L. in different stages of development by gas chromatography and gas chromatography-mass spectrometry. *J Chromatogr A* **1186**: 412–419.
- Ma, D. et al.** (2016). Genetic basis for glandular trichome formation in cotton. *Nature Communications* **7**: 10456.
- Ma, D., Pu, G., Lei, C., Ma, L., Wang, H., Guo, Y., Chen, J., Du, Z., Wang, H., Li, G., Ye, H., and Liu, B.** (2009). Isolation and characterization of AaWRKY1, an *Artemisia annua* transcription factor that regulates the amorpha-4,11-diene synthase gene, a key gene of artemisinin biosynthesis. *Plant Cell Physiol.* **50**: 2146–2161.
- Marchese, J.A., Ferreira, J.F.S., Rehder, V.L.G., and Rodrigues, O.** (2010). Water deficit effect on the accumulation of biomass and artemisinin in annual wormwood (*Artemisia annua* L., Asteraceae). *Brazilian Journal of Plant Physiology* **22**: 1–9.

- Matsuba, Y. et al.** (2013). Evolution of a complex locus for terpene biosynthesis in solanum. *Plant Cell* **25**: 2022–2036.
- Meadows, A.L. et al.** (2016). Rewriting yeast central carbon metabolism for industrial isoprenoid production. *Nature* **537**: 694–697.
- Mercke, P., Bengtsson, M., Bouwmeester, H.J., Posthumus, M.A., and Brodelius, P.E.** (2000). Molecular cloning, expression, and characterization of amorpho-4,11-diene synthase, a key enzyme of artemisinin biosynthesis in *Artemisia annua* L. *Arch. Biochem. Biophys.* **381**: 173–180.
- Miao, G.-P., Han, J., Zhang, J.-F., Zhu, C.-S., and Zhang, X.** (2017). A MDR transporter contributes to the different extracellular production of sesquiterpene pyridine alkaloids between adventitious root and hairy root liquid cultures of *Tripterygium wilfordii* Hook.f. *Plant Mol. Biol.* **95**: 51–62.
- Mock, T. et al.** (2017). Evolutionary genomics of the cold-adapted diatom *Fragilariopsis cylindrus*. *Nature* **541**: 536–540.
- Murovec, J., Pirc, Ž., and Yang, B.** (2017). New variants of CRISPR RNA-guided genome editing enzymes. *Plant Biotechnol. J.* **15**: 917–926.
- Nguyen, K.T., Arsenault, P.R., and Weathers, P.J.** (2011). Trichomes + roots + ROS = artemisinin: regulating artemisinin biosynthesis in *Artemisia annua* L. *In Vitro Cell. Dev. Biol., Plant* **47**: 329–338.
- Olofsson, L., Engström, A., Lundgren, A., and Brodelius, P.E.** (2011). Relative expression of genes of terpene metabolism in different tissues of *Artemisia annua*L. *BMC Plant Biology* **11**: 45.
- Olofsson, L., Lundgren, A., and Brodelius, P.E.** (2012). Trichome isolation with and without fixation using laser microdissection and pressure catapulting followed by RNA amplification: expression of genes of terpene metabolism in apical and sub-apical trichome cells of *Artemisia annua* L. *Plant Sci.* **183**: 9–13.
- Paddon, C.J. et al.** (2013). High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **496**: 528–532.
- Paddon, C.J. and Keasling, J.D.** (2014). Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nature Reviews Microbiology* **12**: 355–367.
- Pandey, A.K. and Singh, P.** (2017). The Genus *Artemisia*: A 2012–2017 Literature Review on Chemical Composition, Antimicrobial, Insecticidal and Antioxidant Activities of Essential Oils. *Medicines (Basel)* **4**.
- Pandey, N. and Pandey-Rai, S.** (2016). Updates on artemisinin: an insight to mode of actions and strategies for enhanced global production. *Protoplasma* **253**: 15–30.
- Pandiarajan, R. and Grover, A.** (2018). In vivo promoter engineering in plants: Are we ready? *Plant Sci.* **277**: 132–138.

- Pattanaik, S., Patra, B., Singh, S.K., and Yuan, L.** (2014). An overview of the gene regulatory network controlling trichome development in the model plant, *Arabidopsis*. *Front. Plant Sci.* **5**.
- Peng, C.A., Ferreira, J.F.S., and Wood, A.J.** (2006). Direct analysis of artemisinin from *Artemisia annua* L. using high-performance liquid chromatography with evaporative light scattering detector, and gas chromatography with flame ionization detector. *Journal of Chromatography A* **1133**: 254–258.
- Peplow, M.** (2016). Synthetic biology's first malaria drug meets market resistance. *Nature News* **530**: 389.
- Pick, T.R. and Weber, A.P.M.** (2014). Unknown components of the plastidial permeome. *Front. Plant Sci.* **5**.
- Polichuk, D.R., Zhang, Y., Reed, D.W., Schmidt, J.F., and Covello, P.S.** (2010). A glandular trichome-specific monoterpene alcohol dehydrogenase from *Artemisia annua*. *Phytochemistry* **71**: 1264–1269.
- Rai, R., Meena, R.P., Smita, S.S., Shukla, A., Rai, S.K., and Pandey-Rai, S.** (2011). UV-B and UV-C pre-treatments induce physiological changes and artemisinin biosynthesis in *Artemisia annua* L. – An antimalarial plant. *Journal of Photochemistry and Photobiology B: Biology* **105**: 216–225.
- RBM/UNITAID/WHO Artemisinin Conference Final Report** (2011). (Hanoi, Vietnam).
- Ro, D.-K. et al.** (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**: 940–943.
- Salehi, M., Karimzadeh, G., Naghavi, M.R., Naghdi Badi, H., and Rashidi Monfared, S.** (2018). Expression of artemisinin biosynthesis and trichome formation genes in five *Artemisia* species. *Industrial Crops and Products* **112**: 130–140.
- Sapir-Mir, M., Mett, A., Belausov, E., Tal-Meshulam, S., Frydman, A., Gidoni, D., and Eyal, Y.** (2008). Peroxisomal Localization of *Arabidopsis* Isopentenyl Diphosphate Isomerases Suggests That Part of the Plant Isoprenoid Mevalonic Acid Pathway Is Compartmentalized to Peroxisomes. *Plant Physiology* **148**: 1219–1228.
- Schramek, N., Wang, H., Römisch-Margl, W., Keil, B., Radykewicz, T., Winzenhörlein, B., Beerhues, L., Bacher, A., Rohdich, F., Gershenzon, J., Liu, B., and Eisenreich, W.** (2010). Artemisinin biosynthesis in growing plants of *Artemisia annua*. A 13CO₂ study. *Phytochemistry* **71**: 179–187.
- Serjeant, G.R.** (2010). One hundred years of sickle cell disease. *Br. J. Haematol.* **151**: 425–429.
- Shen, Q. et al.** (2018). The Genome of *Artemisia annua* Provides Insight into the Evolution of Asteraceae Family and Artemisinin Biosynthesis. *Mol Plant* **11**: 776–788.
- Shen, Q., Chen, Y.F., Wang, T., Wu, S.Y., Lu, X., Zhang, L., Zhang, F., Jiang, W.M., Wang, G.F., and Tang, K.X.** (2012). Overexpression of the cytochrome P450 monooxygenase (*cyp71av1*) and cytochrome P450 reductase (*cpr*) genes increased artemisinin content in *Artemisia annua* (Asteraceae). *Genetics and molecular research : GMR* **11**: 3298–3309.

- Shen, Q., Yan, T., Fu, X., and Tang, K.** (2016). Transcriptional regulation of artemisinin biosynthesis in *Artemisia annua* L. *Science Bulletin* **61**: 18–25.
- Shen, X., Wu, M., Liao, B., Liu, Z., Bai, R., Xiao, S., Li, X., Zhang, B., Xu, J., and Chen, S.** (2017). Complete Chloroplast Genome Sequence and Phylogenetic Analysis of the Medicinal Plant *Artemisia annua*. *Molecules* **22**: 1330.
- Shi, P. et al.** (2018). The roles of AaMIXTA1 in regulating the initiation of glandular trichomes and cuticle biosynthesis in *Artemisia annua*. *New Phytol.* **217**: 261–276.
- Shi, P., Fu, X., Liu, M., Shen, Q., Jiang, W., Li, L., Sun, X., and Tang, K.** (2017). Promotion of artemisinin content in *Artemisia annua* by overexpression of multiple artemisinin biosynthetic pathway genes. *Plant Cell Tiss Organ Cult* **129**: 251–259.
- Shibata, Y., Ojika, M., Sugiyama, A., Yazaki, K., Jones, D.A., Kawakita, K., and Takemoto, D.** (2016). The Full-Size ABCG Transporters Nb-ABCG1 and Nb-ABCG2 Function in Pre- and Postinvasion Defense against *Phytophthora infestans* in *Nicotiana benthamiana*. *The Plant Cell* **28**: 1163–1181.
- Shretta, R. and Yadav, P.** (2012). Stabilizing supply of artemisinin and artemisinin-based combination therapy in an era of wide-spread scale-up. *Malaria Journal* **11**: 399.
- Simonnet, X., Quennoz, M., and Carlen, C.** (2008). NEW ARTEMISIA ANNUA HYBRIDS WITH HIGH ARTEMISININ CONTENT. *Acta Horticulturae*: 371–373.
- Singh, N.D., Kumar, S., and Daniell, H.** (2016). Expression of β -glucosidase increases trichome density and artemisinin content in transgenic *Artemisia annua* plants. *Plant Biotechnol. J.* **14**: 1034–1045.
- Soetaert, S.S.A., Van Neste, C.M.F., Vandewoestyne, M.L., Head, S.R., Goossens, A., Van Nieuwerburgh, F.C.W., and Deforce, D.L.D.** (2013). Differential transcriptome analysis of glandular and filamentous trichomes in *Artemisia annua*. *BMC Plant Biol.* **13**: 220.
- Souza Jr., C.L.D.** (2011). Cultivar development of allogamous crops. *Crop Breeding and Applied Biotechnology* **11**: 8–15.
- Su, X.-Z. and Miller, L.H.** (2015). The discovery of artemisinin and the Nobel Prize in Physiology or Medicine. *Sci China Life Sci* **58**: 1175–1179.
- Suberu, J., Gromski, P.S., Nordon, A., and Lapkin, A.** (2016). Multivariate data analysis and metabolic profiling of artemisinin and related compounds in high yielding varieties of *Artemisia annua* field-grown in Madagascar. *J Pharm Biomed Anal* **117**: 522–531.
- Sundararaman, S.A. et al.** (2016). Genomes of cryptic chimpanzee *Plasmodium* species reveal key evolutionary events leading to human malaria. *Nature Communications* **7**: 11078.
- Sy, L.-K. and Brown, G.D.** (2002). The mechanism of the spontaneous autoxidation of dihydroartemisinic acid. *Tetrahedron* **58**: 897–908.

- Tan, H., Xiao, L., Gao, S., Li, Q., Chen, J., Xiao, Y., Ji, Q., Chen, R., Chen, W., and Zhang, L.** (2015). TRICHOME AND ARTEMISININ REGULATOR 1 Is Required for Trichome Development and Artemisinin Biosynthesis in *Artemisia annua*. *Mol Plant* **8**: 1396–1411.
- Tang, K., Shen, Q., Yan, T., and Fu, X.** (2014). Transgenic approach to increase artemisinin content in *Artemisia annua* L. *Plant Cell Rep.* **33**: 605–615.
- Teoh, K.H., Polichuk, D.R., Reed, D.W., Nowak, G., and Covello, P.S.** (2006). *Artemisia annua* L. (Asteraceae) trichome-specific cDNAs reveal CYP71AV1, a cytochrome P450 with a key role in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin. *FEBS Lett.* **580**: 1411–1416.
- Thu, B.T.T., Van Minh, T., Lim, B.P., and Keng, C.L.** (2011). Effects of Environmental Factors on Growth and Artemisinin Content of *Artemisia annua* L. *Trop Life Sci Res* **22**: 37–43.
- Tovar-Méndez, A., Lu, L., and McClure, B.** (2017). HT proteins contribute to S-RNase-independent pollen rejection in *Solanum*. *The Plant Journal* **89**: 718–729.
- Towler, M.J. and Weathers, P.J.** (2007). Evidence of artemisinin production from IPP stemming from both the mevalonate and the nonmevalonate pathways. *Plant Cell Rep.* **26**: 2129–2136.
- Townsend, T., Segura, V., Chigeza, G., Penfield, T., Rae, A., Harvey, D., Bowles, D., and Graham, I.A.** (2013). The Use of Combining Ability Analysis to Identify Elite Parents for *Artemisia annua* F1 Hybrid Production. *PLOS ONE* **8**: e61989.
- Van Nieuwerburgh, F.C.W., Vande Castele, S.R.F., Maes, L., Goossens, A., Inzé, D., Van Bocxlaer, J., and Deforce, D.L.D.** (2006). Quantitation of artemisinin and its biosynthetic precursors in *Artemisia annua* L. by high performance liquid chromatography-electrospray quadrupole time-of-flight tandem mass spectrometry. *J Chromatogr A* **1118**: 180–187.
- Vogel, G.** (2013). The forgotten malaria. *Science* **342**: 684–687.
- Wallaart, null, van Uden W, null, Lubberink, null, Woerdenbag, null, Pras, null, and Quax, null** (1999). Isolation and identification of dihydroartemisinic acid from *artemisia annua* and its possible role in the biosynthesis of artemisinin. *J. Nat. Prod.* **62**: 430–433.
- Wallaart, T.E., Pras, N., Beekman, A.C., and Quax, W.J.** (2000). Seasonal variation of artemisinin and its biosynthetic precursors in plants of *Artemisia annua* of different geographical origin: proof for the existence of chemotypes. *Planta Med.* **66**: 57–62.
- Wang, H., Ma, D., Yang, J., Deng, K., Li, M., Ji, X., Zhong, L., and Zhao, H.** (2018). An Integrative Volatile Terpenoid Profiling and Transcriptomics Analysis for Gene Mining and Functional Characterization of AvBPPS and AvPS Involved in the Monoterpenoid Biosynthesis in *Amomum villosum*. *Front Plant Sci* **9**.
- Wang, H., Olofsson, L., Lundgren, A., and Brodelius, P.E.** (2011). Trichome-Specific Expression of Amorpha-4,11-Diene Synthase, a Key Enzyme of Artemisinin Biosynthesis in *Artemisia annua* L., as Reported by a Promoter-GUS Fusion. *American Journal of Plant Sciences* **02**: 619.

- Wang, Q., Hasson, A., Rossmann, S., and Theres, K.** (2016). Divide et impera: boundaries shape the plant body and initiate new meristems. *New Phytologist* **209**: 485–498.
- Wang, W., Wang, Y., Zhang, Q., Qi, Y., and Guo, D.** (2009). Global characterization of *Artemisia annua* glandular trichome transcriptome using 454 pyrosequencing. *BMC Genomics* **10**: 465.
- Weathers, P.J., Arsenault, P.R., Covello, P.S., McMickle, A., Teoh, K.H., and Reed, D.W.** (2011). Artemisinin production in *Artemisia annua*: studies in planta and results of a novel delivery method for treating malaria and other neglected diseases. *Phytochem Rev* **10**: 173–183.
- Weathers, P.J., Elkholy, S., and Wobbe, K.K.** (2006). Artemisinin: The Biosynthetic Pathway and Its Regulation in *Artemisia annua*, a Terpenoid-Rich Species. *In Vitro Cellular & Developmental Biology. Plant* **42**: 309–317.
- Wen, W. and Yu, R.** (2011). Artemisinin biosynthesis and its regulatory enzymes: Progress and perspective. *Pharmacogn Rev* **5**: 189–194.
- Wetzstein, H.Y., Porter, J.A., Janick, J., and Ferreira, J.F.S.** (2014). Flower morphology and floral sequence in *Artemisia annua* (Asteraceae)1. *Am. J. Bot.* **101**: 875–885.
- Whitfield, M.L., Sherlock, G., Saldanha, A.J., Murray, J.I., Ball, C.A., Alexander, K.E., Matese, J.C., Perou, C.M., Hurt, M.M., Brown, P.O., and Botstein, D.** (2002). Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol. Biol. Cell* **13**: 1977–2000.
- WHO International Travel and Health Report** (2018). (WHO).
- WHO World Malaria Report** (2017). (WHO).
- WHO World Malaria Report** (2015). (WHO).
- WHO World Malaria Report** (2016). (WHO).
- World Health Organization** (2006). WHO monograph on good agricultural and collection practices (GACP) for *Artemisia annua* L. (archived).
- Worley, B. and Powers, R.** (2013). Multivariate Analysis in Metabolomics. *Curr Metabolomics* **1**: 92–107.
- Xiao, L., Tan, H., and Zhang, L.** (2016). *Artemisia annua* glandular secretory trichomes: the biofactory of antimalarial agent artemisinin. *Science Bulletin* **61**: 26–36.
- Yang, K., Monfared, S.R., Monfared, R.S., Wang, H., Lundgren, A., and Brodelius, P.E.** (2015). The activity of the artemisinic aldehyde $\Delta^{11}(13)$ reductase promoter is important for artemisinin yield in different chemotypes of *Artemisia annua* L. *Plant Mol. Biol.* **88**: 325–340.
- Yang, R.-Y., Zeng, X.-M., Lu, Y.-Y., Lu, W.-J., Feng, L.-L., Yang, X.-Q., and Zeng, Q.-P.** (2010). Senescent leaves of *Artemisia annua* are one of the most active organs for overexpression of artemisinin biosynthesis responsible genes upon burst of singlet oxygen. *Planta Med.* **76**: 734–742.

- Yu, Z.-X., Li, J.-X., Yang, C.-Q., Hu, W.-L., Wang, L.-J., and Chen, X.-Y.** (2012). The jasmonate-responsive AP2/ERF transcription factors AaERF1 and AaERF2 positively regulate artemisinin biosynthesis in *Artemisia annua* L. *Mol Plant* **5**: 353–365.
- Zhang, F., Fu, X., Lv, Z., Lu, X., Shen, Q., Zhang, L., Zhu, M., Wang, G., Sun, X., Liao, Z., and Tang, K.** (2015). A Basic Leucine Zipper Transcription Factor, AabZIP1, Connects Abscisic Acid Signaling with Artemisinin Biosynthesis in *Artemisia annua*. *Molecular Plant* **8**: 163–175.
- Zhang, X., Zhao, Y., Guo, L., Qiu, Z., Huang, L., and Qu, X.** (2017). Differences in chemical constituents of *Artemisia annua* L from different geographical regions in China. *PLOS ONE* **12**: e0183047.
- Zhang, Y., Teoh, K.H., Reed, D.W., Maes, L., Goossens, A., Olson, D.J.H., Ross, A.R.S., and Covello, P.S.** (2008). The molecular cloning of artemisinic aldehyde Delta11(13) reductase and its role in glandular trichome-dependent biosynthesis of artemisinin in *Artemisia annua*. *J. Biol. Chem.* **283**: 21501–21508.
- Zsögön, A., Čermák, T., Naves, E.R., Notini, M.M., Edel, K.H., Weinl, S., Freschi, L., Voytas, D.F., Kudla, J., and Peres, L.E.P.** (2018). *De novo* domestication of wild tomato using genome editing. *Nature Biotechnology*.
- Zsögön, A., Cermak, T., Voytas, D., and Peres, L.E.P.** (2017). Genome editing as a tool to achieve the crop ideotype and de novo domestication of wild relatives: Case study in tomato. *Plant Science* **256**: 120–130.

APPENDIX 1

Biomass

Genotype	Avg Dry weight (g)	SD	SE
A	7.511667	3.885303	1.586168
B	15.288	2.274021	1.016973
C	11.83	1.911565	0.780393
D	18.81833	1.215523	0.496235
E	12.65167	1.852559	0.756304
F	12.65167	4.173055	1.703643
G	19.31667	5.372097	2.193149
H	18.08	2.635762	1.076045
I	20.905	2.368702	0.967019
J	18.38833	2.28124	0.931312
K	15.115	2.966174	1.210936
L	23.81	1.930824	0.788255

Height

Genotype	Height (cm)	SD	SE
A	67.42	16.43706	6.710402
B	62.16667	8.822396	3.601728
C	51.61667	11.40832	4.657425
D	58.75	4.688177	1.91394
E	44.93333	3.509226	1.432635
F	67.1	9.962931	4.06735
G	78.36667	5.273203	2.152776
H	50.35	3.254996	1.328847
I	91.73333	9.978109	4.073546
J	89.65	15.78135	6.442709
K	61.65	8.871246	3.621671
L	74.66667	5.16901	2.11024

Leaf Area

Leaf Area (mm ²)			
Genotype	Average	SD	SE

A	141.266	34.15129	15.27292
B	290.238	96.48973	43.15152
C	340.84	135.2585	60.48942
D	699.6	125.2167	55.99862
E	691.466	94.76308	42.37934
F	270.784	84.45247	37.76829
G	340.65	73.78729	32.99868
H	20.77042	20.77042	9.288816
I	338.646	80.14543	35.84213
J	652.01	55.04141	24.61527
K	253.886	49.97252	22.34839
L	552.962	62.89215	28.12622

Adaxial and abaxial trichome density

Genotype	Trichome Density/mm ²)					
	Abaxial	SD	SE	Adaxial	SD	SE
A	47.66733	3.841648	1.214836	29.05378	2.965062	0.937635
B	30.18174	5.0327	1.591479	15.95805	2.398	0.758314
C	43.94468	13.14686	3.795172	20.12748	5.327932	1.606432
D	40.41052	5.907275	1.868044	15.43737	2.578836	0.815499
E	37.13038	6.295646	1.990858	18.0763	3.463788	1.095346
F	25.88249	5.468571	1.729314	16.3587	3.17513	1.004064
G	23.97742	6.301147	1.992598	12.94511	4.117952	1.302211
H	28.24997	7.457773	2.358355	18.50155	4.571628	1.445676
I	29.12951	8.267891	2.755964	20.46048	6.009572	2.003191
J	29.64868	4.004197	1.513444	19.17322	3.933026	1.486544
K	41.90864	8.288223	3.132654	25.55991	5.669946	2.004629
L	16.95694	3.094288	0.893244	6.094229	2.322155	0.670349

APPENDIX 2

Anova table

	f.value	p.value	#NAME?	FDR	Tukey's LSD
Cedren-13-ol, 8-	640.61	2.93E-56	55.533	1.55E-54	F - A; F - B; F - C; F - D; F - E; F - G; F - H; F - I; F - J; F - K; F - L
Camphor	268.32	1.84E-45	44.736	4.87E-44	A - B; C - A; A - D; A - E; A - F; A - G; A - H; I - A; J - A; K - A; C - B; B - D; B - E; B - F; G - B; H - B; I - B; J - B; K - B; L - B; C - D; C - E; C - F; C - G; C - H; C - I; J - C; K - C; C - L; G - D; H - D; I - D; J - D; K - D; L - D; G - E; H - E; I - E; J - E; K - E; L - E; G - F; H - F; I - F; J - F; K - F; L - F; I - G; J - G; K - G; L - G; I - H; J - H; K - H; L - H; J - I; K - I; I - L; J - L; K - L
endo-Borneol	198.89	8.67E-42	41.062	1.53E-40	A - B; A - C; A - D; A - E; A - F; A - G; A - H; A - I; A - J; A - K; A - L
Butanoic acid, 2-methyl-, phenylmethyl ester	189.33	3.46E-41	40.461	4.58E-40	A - B; A - C; A - D; A - E; A - F; A - G; A - H; A - I; A - J; A - K; A - L; B - C; B - D; B - E; B - F; B - G; B - H; B - I; B - J; B - K; B - L; C - F; C - G; C - H; C - I; C - J; C - K; C - L; D - F; D - G; D - H; D - I; D - J; D - K; D - L; E - F; E - G; E - H; E - I; E - J; E - K; E - L; F - G; F - H; F - I; F - J; F - K; G - I; G - J; L - G; H - I; H - J; L - H; K - I; L - I; K - J; L - J; L - K
cis-p-mentha-17,8-dien-2-ol	147.8	3.53E-38	37.452	3.75E-37	A - B; A - D; A - E; A - F; A - G; A - H; A - I; J - A; K - A; A - L; C - B; B - D; B - E; G - B; B - H; B - I; J - B; K - B; C - D; C - E; C - F; C - G; C - H; C - I; J - C; K - C; C - L; G - D; J - D; K - D; F - E; G - E; J - E; K - E; L - E; G - F; F - H; F - I; J - F; K - F; G - H; G - I; J - G; K - G; G - L;

					J - H; K - H; L - H; J - I; K - I; L - I; J - L; K - L
cis-beta-Terpineol	143.94	7.39E-38	37.132	6.52E-37	A - B; A - C; A - D; A - E; A - F; A - G; A - H; A - I; A - J; A - L; C - B; B - D; B - E; B - F; G - B; H - B; J - B; K - B; L - B; C - D; C - E; C - F; C - G; C - H; C - I; C - L; G - D; H - D; I - D; J - D; K - D; L - D; G - E; H - E; I - E; J - E; K - E; L - E; G - F; H - F; I - F; J - F; K - F; L - F; H - G; G - I; J - G; K - G; H - I; J - H; K - H; H - L; J - I; K - I; L - I; K - J; J - L; K - L
2-Methyl-4-2,6,6-trimethylcyclohex-2-enylbut-3-en-2-ol	118.29	1.70E-35	34.771	1.28E-34	B - A; A - C; D - A; E - A; A - F; G - A; A - H; A - I; A - J; A - K; A - L; B - C; D - B; B - E; B - F; B - G; B - H; B - I; B - J; B - K; B - L; D - C; E - C; G - C; C - H; J - C; C - L; D - E; D - F; D - G; D - H; D - I; D - J; D - K; D - L; E - F; E - H; E - I; E - J; E - K; E - L; G - F; F - H; F - L; G - H; G - I; G - J; G - K; G - L; I - H; J - H; K - H; L - H; I - L; J - K; J - L; K - L
Phenol, 2-methoxy-3-2-propenyl-	113.58	5.20E-35	34.284	3.45E-34	C - A; D - A; E - A; G - A; A - H; I - A; J - A; K - A; A - L; C - B; D - B; E - B; G - B; B - H; I - B; J - B; K - B; B - L; C - F; C - H; I - C; J - C; K - C; C - L; D - F; D - H; I - D; J - D; K - D; D - L; E - F; E - H; I - E; J - E; K - E; E - L; G - F; F - H; I - F; J - F; K - F; F - L; G - H; I - G; J - G; K - G; G - L; I - H; J - H; K - H; I - J; K - I; I - L; K - J; J - L; K - L
Artemisinin	102.07	9.74E-34	33.011	5.74E-33	A - B; C - A; A - D; A - E; A - F; G - A; A - H; I - A; J - A; C - B; B - F; G - B; B - H; I - B; J - B; K - B; L - B; C - D; C - E; C - F; G - C; C - H; I - C; J - C; C - K; C - L; D - F; G - D; D - H; I - D; J - D; K - D; L - D; E - F; G - E; E - H; I - E; J - E; K - E; L - E; G - F; I - F; J - F; K - F; L - F; G -

					H; G - I; J - G; G - K; G - L; I - H; J - H; K - H; L - H; J - I; I - K; I - L; J - K; J - L
Caryophyllene oxide	101.14	1.25E-33	32.903	6.63E-33	A - B; A - C; A - D; A - E; A - F; A - G; A - H; A - I; A - J; A - K; A - L; B - D; B - E; F - B; G - B; H - B; I - B; J - B; K - B; L - B; C - D; C - E; F - C; G - C; H - C; I - C; J - C; K - C; L - C; F - D; G - D; H - D; I - D; J - D; K - D; L - D; F - E; G - E; H - E; I - E; J - E; K - E; L - E; G - F; H - F; F - J; F - L; G - H; G - I; G - J; G - K; G - L; H - I; H - J; H - K; H - L; K - J; K - L
Cyclohexanol, 1-methyl-4-1-methylethenyl-, cis-	82.15	3.56E-31	30.448	1.72E-30	A - B; A - C; A - D; A - E; A - F; A - G; A - H; A - I; A - J; A - L; C - B; G - B; H - B; J - B; K - B; L - B; C - D; C - E; C - F; C - G; C - H; C - I; K - C; C - L; G - D; H - D; J - D; K - D; L - D; G - E; H - E; J - E; K - E; L - E; G - F; H - F; J - F; K - F; L - F; H - G; G - I; J - G; K - G; L - G; H - I; J - H; K - H; J - I; K - I; L - I; K - J; J - L; K - L
beta-copaene	74.567	4.82E-30	29.317	2.13E-29	A - B; A - D; A - E; A - F; A - H; K - A; A - L; C - B; B - E; B - F; G - B; B - H; I - B; J - B; K - B; B - L; C - D; C - E; C - F; C - H; K - C; C - L; D - E; D - F; G - D; D - H; I - D; J - D; K - D; D - L; E - F; G - E; E - H; I - E; J - E; K - E; G - F; F - H; I - F; J - F; K - F; L - F; G - H; K - G; G - L; I - H; J - H; K - H; L - H; K - I; I - L; K - J; J - L; K - L
Artemisinic acid	62.475	5.34E-28	27.272	2.18E-27	C - A; D - A; E - A; G - A; H - A; J - A; D - B; E - B; G - B; H - B; J - B; G - C; H - C; C - I; C - L; G - D; H - D; D - I; D - K; D - L; G - E; H - E; E - I; E - K; E - L; G - F; H - F; H - G; G - I; G - J; G - K; G - L; H - I; H - J; H - K; H - L; J - I; J - K; J - L

Eucalyptol	58.844	2.59E-27	26.588	9.79E-27	A - B; A - D; A - E; A - F; A - G; H - A; A - I; A - J; A - K; A - L; B - D; B - E; H - B; B - L; C - D; C - E; C - F; H - C; C - I; C - J; C - L; G - D; H - D; G - E; H - E; K - E; H - F; H - G; G - L; H - I; H - J; H - K; H - L; K - L
Isoaromadendrene epoxide.1	55.728	1.07E-26	25.969	3.80E-26	B - A; D - A; E - A; A - H; A - I; A - K; A - L; B - C; D - B; B - E; B - F; B - G; B - H; B - I; B - J; B - K; B - L; D - C; E - C; F - C; G - C; C - H; C - I; C - K; C - L; D - E; D - F; D - G; D - H; D - I; D - J; D - K; D - L; E - F; E - G; E - H; E - I; E - J; E - K; E - L; F - H; F - I; F - J; F - K; F - L; G - H; G - I; G - J; G - K; G - L; I - H; J - H; J - I; I - K; I - L; J - K; J - L
Corymbolone	49.756	2.04E-25	24.691	6.76E-25	B - A; A - C; D - A; A - F; A - H; A - I; A - J; A - K; A - L; B - C; D - B; B - E; B - F; B - G; B - H; B - I; B - J; B - K; B - L; D - C; E - C; G - C; C - H; C - L; D - E; D - F; D - G; D - H; D - I; D - J; D - K; D - L; E - F; E - H; E - I; E - J; E - K; E - L; G - F; F - H; F - L; G - H; G - I; G - J; G - K; G - L; I - H; J - H; K - H; L - H; I - L; J - L; K - L
Cymene	45.034	2.63E-24	23.579	8.21E-24	A - B; A - C; D - A; E - A; F - A; A - G; A - H; A - J; A - L; D - B; E - B; F - B; I - B; K - B; L - B; D - C; E - C; F - C; I - C; K - C; L - C; D - G; D - H; D - I; D - J; D - K; D - L; E - G; E - H; E - I; E - J; E - K; E - L; F - G; F - H; F - I; F - J; F - K; F - L; I - G; K - G; L - G; I - H; K - H; L - H; I - J; K - J; L - J; K - L
Cubenol	42.74	9.95E-24	23.002	2.93E-23	B - A; A - C; A - E; F - A; I - A; K - A; A - L; B - C; B - D; B - E; B - G; B - H; I - B; B - J; K - B; B - L; D - C; F - C; G - C; H - C; I - C; J - C; K - C; C - L; F - D; H - D; I - D; K - D;

					D - L; F - E; G - E; H - E; I - E; K - E; E - L; F - G; I - F; F - J; K - F; F - L; I - G; K - G; G - L; I - H; H - J; K - H; H - L; I - J; K - I; I - L; K - J; J - L; K - L
Neointermedeol	42.597	1.08E-23	22.965	3.02E-23	A - C; D - A; A - F; A - G; A - H; A - K; A - L; B - C; B - F; B - G; B - H; B - I; B - K; B - L; D - C; E - C; G - C; C - H; I - C; J - C; K - C; C - L; D - E; D - F; D - G; D - H; D - I; D - J; D - K; D - L; E - F; E - G; E - H; E - L; F - H; I - F; J - F; K - F; F - L; G - H; J - G; G - L; I - H; J - H; K - H; L - H; I - L; J - L; K - L
Terpineol	37.195	3.26E-22	21.487	8.63E-22	A - B; A - C; A - D; A - E; A - F; A - G; A - I; A - J; K - A; A - L; B - D; B - E; B - F; H - B; K - B; C - D; C - E; C - F; H - C; K - C; C - L; G - D; H - D; I - D; J - D; K - D; G - E; H - E; I - E; J - E; K - E; G - F; H - F; I - F; J - F; K - F; L - F; H - G; K - G; G - L; H - I; H - J; K - H; H - L; K - I; I - L; K - J; J - L; K - L
Dihydroartemesisinic acid	35.272	1.21E-21	20.916	3.06E-21	E - A; G - A; I - A; J - A; E - B; G - B; I - B; J - B; E - C; G - C; I - C; J - C; D - F; G - D; D - H; J - D; D - L; E - F; E - H; J - E; E - K; E - L; G - F; I - F; J - F; G - H; G - I; J - G; G - K; G - L; I - H; J - H; K - H; J - I; I - K; I - L; J - K; J - L
Terpinen	32.557	8.64E-21	20.064	2.08E-20	A - B; A - C; A - G; A - H; A - I; A - J; A - L; D - B; E - B; F - B; I - B; K - B; D - C; E - C; F - C; I - C; K - C; E - D; F - D; D - G; D - H; D - J; D - L; E - G; E - H; E - I; E - J; E - L; F - G; F - H; F - I; F - J; F - L; I - G; K - G; I - H; K - H; I - J; K - I; I - L; K - J; K - L
Isoaromadendrene epoxide.3	31.972	1.34E-20	19.872	3.09E-20	A - B; A - D; A - E; A - G; H - A; K - A; A - L; C - B; B - E; F - B; G - B; H - B; I - B; J - B; K - B; B - L; C -

					D; C - E; H - C; K - C; C - L; D - E; F - D; G - D; H - D; I - D; J - D; K - D; F - E; G - E; H - E; I - E; J - E; K - E; H - F; K - F; F - L; H - G; K - G; G - L; H - I; H - J; H - L; K - I; I - L; K - J; J - L; K - L
3-buten-2-one, 4-5,5-dimethyl-1-oxaspiro2.5oct-4-yl	31.185	2.46E-20	19.61	5.43E-20	A - B; A - C; A - D; A - E; A - F; J - A; A - L; C - B; G - B; H - B; I - B; J - B; K - B; C - D; C - E; C - F; G - C; J - C; K - C; C - L; G - D; H - D; I - D; J - D; K - D; G - E; H - E; I - E; J - E; K - E; G - F; H - F; I - F; J - F; K - F; J - G; G - L; J - H; K - H; H - L; J - I; I - L; J - L; K - L
Chiapin B	25.6	2.67E-18	17.574	5.66E-18	A - B; A - D; A - E; A - F; A - H; J - A; A - L; C - B; G - B; B - H; I - B; J - B; K - B; C - D; C - E; C - F; C - H; J - C; C - L; D - E; D - F; G - D; D - H; I - D; J - D; K - D; G - E; E - H; I - E; J - E; K - E; L - E; G - F; F - H; I - F; J - F; K - F; L - F; G - H; J - G; G - L; I - H; J - H; K - H; L - H; J - I; K - I; I - L; J - L; K - L
Cholestan-3-ol, 2-methylene-, 3beta,5a-	23.247	2.47E-17	16.607	5.03E-17	A - B; A - D; A - E; A - F; A - G; A - H; A - I; A - J; K - A; A - L; B - H; K - B; B - L; C - E; C - F; C - G; C - H; K - C; C - L; D - H; K - D; D - L; E - H; K - E; E - L; F - H; K - F; F - L; G - H; K - G; G - L; I - H; J - H; K - H; H - L; K - I; I - L; K - J; J - L; K - L
Pinene	21.478	1.48E-16	15.829	2.91E-16	A - B; C - A; A - D; A - E; A - F; A - I; C - B; B - D; B - E; B - F; H - B; J - B; K - B; L - B; C - D; C - E; C - F; C - G; C - H; C - I; C - K; C - L; G - D; H - D; J - D; K - D; L - D; G - E; H - E; J - E; K - E; L - E; G - F; H - F; J - F; K - F; L - F; H - G; G - I; J - G; H - I; J - I; K - I; L - I; J - L
Unkown alkane	20.366	4.85E-16	15.314	9.18E-16	A - B; A - C; A - D; A - E; A - F; A - G; A - H; A - I; A - J; A - K; A - L; C

					- B; D - B; B - F; G - B; H - B; B - I; C - F; H - C; C - I; C - J; C - L; D - E; D - F; H - D; D - I; D - J; D - L; E - F; G - E; H - E; E - I; G - F; H - F; I - F; J - F; K - F; L - F; H - G; G - I; G - J; G - L; H - I; H - J; H - K; H - L; K - I; L - I; K - J
Elemen	20.243	5.55E-16	15.256	1.01E-15	A - C; D - A; E - A; F - A; A - H; A - I; A - L; B - C; E - B; B - G; B - H; B - I; B - K; B - L; D - C; E - C; F - C; J - C; K - C; C - L; D - G; D - H; D - I; D - J; D - K; D - L; E - G; E - H; E - I; E - J; E - K; E - L; F - G; F - H; F - I; F - J; F - K; F - L; G - H; G - L; I - H; J - H; K - H; J - I; I - L; J - L; K - L
Epiglobulol	19.389	1.43E-15	14.845	2.53E-15	D - A; E - A; A - H; A - I; A - J; A - L; B - C; D - B; B - F; B - H; B - I; B - J; B - K; B - L; D - C; E - C; C - H; C - I; C - J; C - L; D - E; D - F; D - G; D - H; D - I; D - J; D - K; D - L; E - F; E - G; E - H; E - I; E - J; E - K; E - L; F - H; F - L; G - H; G - I; G - J; G - L; K - H; K - I; K - J; J - L; K - L
Deoxyartemisinin	18.699	3.15E-15	14.502	5.38E-15	A - B; A - D; A - E; A - F; A - H; A - L; C - B; B - F; G - B; B - H; I - B; J - B; K - B; C - D; C - E; C - F; C - H; C - L; D - F; G - D; D - H; I - D; J - D; K - D; G - E; E - H; I - E; J - E; K - E; G - F; F - H; I - F; J - F; K - F; G - H; G - L; I - H; J - H; K - H; L - H; I - L; J - L; K - L
compound 1	18.66	3.30E-15	14.482	5.46E-15	A - F; H - A; I - A; B - F; H - B; C - F; H - C; D - F; H - D; I - D; E - F; H - E; G - F; H - F; I - F; J - F; K - F; H - G; H - I; H - J; H - K; H - L; I - L
beta-Guaiene	18.083	6.49E-15	14.188	1.04E-14	F - A; G - A; I - A; J - A; K - A; A - L; F - B; G - B; I - B; J - B; K - B; B - L; G - C; C - H; I - C; J - C; K - C;

					C - L; F - D; G - D; I - D; J - D; K - D; D - L; F - E; G - E; I - E; J - E; K - E; F - H; K - F; F - L; G - H; K - G; G - L; I - H; J - H; K - H; K - I; I - L; K - J; J - L; K - L
Olean-12-en-3-one	13.739	1.88E-12	11.725	2.94E-12	B - A; C - A; D - A; A - H; I - A; L - A; B - E; B - F; B - G; B - H; I - B; L - B; C - E; C - F; C - G; C - H; C - J; C - K; D - E; D - F; D - G; D - H; D - J; D - K; E - H; I - E; J - E; L - E; I - F; J - F; K - F; L - F; G - H; I - G; L - G; I - H; J - H; K - H; L - H; I - J; I - K; L - J; L - K
Camphene	13.213	4.06E-12	11.392	6.14E-12	A - B; A - D; A - E; A - F; A - G; J - A; A - L; H - B; I - B; J - B; K - B; C - D; C - E; C - F; I - C; J - C; K - C; G - D; H - D; I - D; J - D; K - D; L - D; G - E; H - E; I - E; J - E; K - E; L - E; G - F; H - F; I - F; J - F; K - F; L - F; H - G; I - G; J - G; K - G; J - H; H - L; I - L; J - L; K - L
Heptacosane	12.255	1.73E-11	10.762	2.55E-11	A - B; A - C; A - D; A - E; A - F; A - G; A - H; A - I; A - J; A - K; A - L; B - F; G - B; B - J; C - D; C - F; C - I; C - J; C - K; C - L; D - F; G - D; D - J; E - F; G - E; E - I; E - J; G - F; H - F; G - H; G - I; G - J; G - K; G - L; H - I; H - J; H - L
Farnesene epoxide, E-	11.563	5.15E-11	10.289	7.37E-11	B - A; D - A; A - H; A - I; A - L; B - C; B - F; B - G; B - H; B - I; B - J; B - K; B - L; D - C; E - C; C - H; C - L; D - F; D - G; D - H; D - I; D - J; D - K; D - L; E - H; E - I; E - J; E - K; E - L; F - H; F - I; F - K; F - L; G - H; G - I; G - K; G - L; J - H; K - H; J - L; K - L
2H-Pyran, 2-7-heptadecyloxytetrahydro-.1	9.3023	2.47E-09	8.6081	3.44E-09	A - B; A - D; A - F; A - H; A - L; G - B; B - H; I - B; J - B; K - B; C - D; C - F; C - H; C - L; G - D; D - H; I - D; J - D; K - D; E - F; E - H; K - E; G - F; F - H; I - F; J - F; K - F; G - H; G

					- L; I - H; J - H; K - H; L - H; I - L; J - L; K - L
Cyclohexane, 1-ethenyl-1-methyl-2-1-methylethenyl-4-1-methyle	9.2301	2.81E-09	8.5507	3.82E-09	A - B; I - A; L - A; C - B; D - B; G - B; I - B; K - B; L - B; C - E; C - F; C - H; I - C; C - J; L - C; D - E; D - F; D - H; I - D; D - J; L - D; I - E; L - E; I - F; L - F; G - H; I - G; L - G; I - H; L - H; I - J; I - K; L - J; L - K
Isoaromadendrene epoxide.2	8.8949	5.23E-09	8.2814	6.93E-09	A - C; A - F; A - H; K - A; A - L; B - C; B - F; B - H; K - B; B - L; D - C; I - C; K - C; C - L; D - F; D - H; K - D; D - L; I - E; K - E; E - L; I - F; K - F; F - L; I - G; K - G; G - L; I - H; K - H; H - L; I - J; I - L; K - J; J - L; K - L
Ylangene	7.9522	3.21E-08	7.4938	4.15E-08	D - A; G - A; H - A; I - A; L - A; B - J; L - B; I - C; L - C; D - E; D - F; D - J; L - D; G - E; H - E; I - E; L - E; G - F; H - F; I - F; L - F; G - J; L - G; H - J; L - H; I - J; I - K; L - I; L - J; L - K
Aromadendrene oxide-2	7.8245	4.14E-08	7.3835	5.22E-08	A - E; A - F; A - J; A - L; B - E; B - F; B - J; B - L; C - E; C - F; C - J; C - L; D - E; D - F; D - J; D - L; H - E; K - E; E - L; H - F; K - F; F - L; K - G; G - L; H - J; H - L; K - I; I - L; K - J; J - L; K - L
cis-beta-Farnesene	7.7941	4.39E-08	7.3571	5.42E-08	B - A; D - A; E - A; F - A; G - A; I - A; L - A; B - C; B - H; B - J; B - K; D - C; F - C; D - H; D - I; D - J; D - K; F - E; E - H; E - J; E - K; F - G; F - H; F - I; F - J; F - K; F - L; G - H; G - J; G - K; I - H; L - H; I - J; I - K; L - J; L - K
Isoaromadendrene epoxide.4	7.7366	4.93E-08	7.307	5.94E-08	A - B; A - C; A - H; A - I; J - A; A - K; A - L; D - B; E - B; F - B; G - B; J - B; E - C; F - C; G - C; J - C; E - D; D - H; J - D; D - K; E - H; E - I; E - K; E - L; F - H; F - I; F - K; F - L; G

					- H; G - I; J - G; G - K; G - L; J - H; J - I; J - K; J - L
Copaene	7.2251	1.40E-07	6.8539	1.65E-07	D - A; G - A; H - A; I - A; L - A; B - J; L - B; L - C; D - E; D - F; D - J; L - D; G - E; H - E; I - E; L - E; G - F; H - F; I - F; L - F; G - J; L - G; H - J; L - H; I - J; L - I; L - J; L - K
Ocimene	6.835	3.18E-07	6.4982	3.66E-07	A - D; A - E; A - F; A - I; C - B; H - B; J - B; C - D; C - E; C - F; C - I; G - D; H - D; J - D; K - D; L - D; G - E; H - E; J - E; K - E; L - E; G - F; H - F; J - F; K - F; L - F; G - I; H - I; J - I; K - I
Germacrene D	6.1279	1.48E-06	5.8299	1.67E-06	D - A; G - A; H - A; I - A; L - A; B - J; L - B; I - C; L - C; D - E; D - F; D - J; L - D; G - E; H - E; I - E; L - E; G - F; H - F; I - F; L - F; G - J; L - G; H - J; L - H; I - J; L - I; L - J; L - K
cis-Z-alpha-Bisabolene epoxide	4.7692	3.50E-05	4.4564	3.86E-05	A - C; A - D; A - G; E - B; F - B; I - B; K - B; E - C; F - C; I - C; J - C; K - C; L - C; E - D; F - D; I - D; K - D; E - G; E - H; F - G; F - H; I - G; K - G; I - H; K - H
2H-Pyran, 2-7-heptadecyloxytetrahydro-	3.6658	0.000558	3.2536	0.00060325	A - B; A - F; A - H; A - L; E - B; J - B; K - B; J - C; J - D; E - F; E - H; E - L; J - F; K - F; J - G; J - H; K - H; J - I; K - I; J - L; K - L
Aromadendrene oxide-2.1	3.2349	0.001717	2.7653	0.0018198	A - B; A - D; A - F; A - L; C - B; G - B; I - B; J - B; K - B; C - D; C - F; C - L; G - D; I - D; J - D; K - D; G - F; H - F; I - F; J - F; K - F; G - L; H - L; I - L; J - L; K - L
Caryophyllene	2.7765	0.005775	2.2384	0.0060018	D - A; L - A; D - B; L - B; D - C; L - C; D - E; D - G; D - H; D - J; F - J; L - G; K - H; L - H; I - J; K - J; L - J

PLS-DA loadings

	Comp 1	Comp 2	Comp 3	Comp 4	Comp 5	Comp 6	Co
Pinene	0.15539	-0.04311	0.22212	0.11636	-0.09019	0.1623	-0
Ocimene	0.13127	-0.03066	0.21406	0.095667	-0.1071	0.13923	-
Camphene	0.23821	-0.05493	0.076524	-0.14344	0.041644	-0.22284	
Cymene	-0.13499	-0.12051	-0.35505	-0.35159	0.33068	-0.28554	-0
Eucalyptol	0.029457	-0.04424	0.62501	-0.19478	-0.00361	-0.40318	0.0
Terpinen	-0.04571	-0.07608	-0.13548	-0.24581	0.082954	-0.135	-0
cis-beta-Terpineol	0.080037	-0.05563	0.040733	0.010408	-0.01441	0.013961	-0
Cyclohexanol, 1-methyl-4-1-methylethenyl-, cis-	0.066173	-0.04645	0.025478	-0.01781	0.020649	0.001819	-
Camphor	0.8893	0.097887	-0.1497	-0.09759	0.082124	0.042853	-0
endo-Borneol	0.088222	-0.92878	-0.00019	0.040971	0.10807	0.12239	0.0
Terpineol	0.041736	-0.03187	0.04964	-0.0487	0.010841	-0.13071	-0
cis-p-mentha-17,8-dien-2-ol	0.043943	-0.02312	-0.03239	-0.01099	-0.05461	0.035848	-0
Cyclohexane, 1-ethenyl-1-methyl-2-1-methylethenyl-4-1-methyle	0.011648	0.007621	-0.03348	0.098257	0.21991	0.008364	0.0
Phenol, 2-methoxy-3-2-propenyl-	0.042979	0.035682	-0.14171	-0.07182	-0.00902	-0.16252	0.0
Butanoic acid, 2-methyl-, phenylmethyl ester	-0.04512	-0.18484	-0.01279	0.091187	-0.05794	0.031316	-0
Elemen	-0.03414	-0.02119	-0.09908	-0.05112	-0.11051	-0.06368	-0
ylangene	-0.0083	0.029671	0.044088	0.13366	0.2627	0.00106	0
Caryophyllene	-0.02427	0.008635	-0.06701	0.046353	0.29578	-0.07149	0
copaene	-0.00876	0.025329	0.034593	0.11346	0.23024	-0.00275	0
cis-beta-Farnesene	-0.07569	0.027982	-0.05553	0.026657	0.095168	0.069141	0
Germacrene D	-0.01997	0.060256	0.076734	0.3058	0.5847	-0.04132	0
Isoaromadendrene epoxide	0.015426	0.003677	0.01257	0.024344	-0.02454	-0.02225	-0
compound 1	0.006532	0.009259	0.068179	-0.01393	-0.00927	-0.08289	-0
Caryophyllene oxide	0.023489	-0.09862	0.048198	-0.04423	0.004606	0.018485	0
beta-Guaiene	0.022238	0.009677	-0.0429	-0.05912	-0.03083	-0.03302	0.0
Isoaromadendrene epoxide.1	-0.04123	-0.01379	-0.05353	0.053011	-0.11393	-0.02767	0.0
Farnesene epoxide, E-	-0.03216	-0.01366	-0.05991	0.037653	-0.11017	-0.02764	0.0
Cubenol	0.042447	0.013268	-0.06779	-0.24782	-0.00799	-0.31302	
Aromadendrene oxide-2	0.005672	-0.01058	-0.00034	-0.01226	-0.01633	-0.05775	0.0
Isoaromadendrene epoxide.2	0.007183	-0.00363	-0.02591	-0.01755	-0.02132	-0.06982	0.0
Neointermedeol	-0.00458	-0.0278	-0.12177	0.015185	-0.09162	-0.0855	0.0
Epiglobulol	-0.01864	-0.01136	-0.04568	0.025365	-0.05727	-0.06552	-0
Corybolone	-0.04085	-0.04533	-0.15644	0.090919	-0.17288	-0.13653	0.0
2H-Pyran, 2-7-heptadecynyloxytetrahydro-	0.018011	-0.00934	-0.02058	-0.00464	-0.04535	-0.01275	-0
Cedren-13-ol, 8-	-0.10485	0.016688	-0.05346	-0.55063	0.014982	0.55842	0

Isoaromadendrene epoxide.3	0.027353	-0.00981	0.040381	-0.07512	-0.00981	-0.04407	0.0
cis-Z-alpha-Bisabolene epoxide	0.003393	-0.00885	-0.03131	-0.07652	0.036612	-0.0077	0.0
Cholestan-3-ol, 2-methylene-, 3beta,5a-	0.017834	-0.02133	-0.04032	-0.04212	-0.05001	-0.066	0.0
2-Methyl-4-2,6,6-trimethylcyclohex-2-enylbut-3-en-2-ol	-0.07632	-0.05126	-0.22807	0.16152	-0.25294	-0.18787	0
Artemisinic acid	-0.00159	0.007181	0.061771	-0.01215	-0.01148	-0.03594	-0
3-buten-2-one, 4-5,5-dimethyl-1-oxaspiro2.5oct-4-yl	0.04967	-0.01595	0.004783	-0.00853	-0.03836	-0.0594	0.0
Dihydroartemisinic acid	0.042234	0.035244	-0.08376	0.032567	-0.18588	0.048915	0.0
Isoaromadendrene epoxide.4	-0.00857	-0.01822	-0.03121	-0.01906	-0.07922	0.10074	-0
beta-copaene	0.13552	-0.05259	-0.30549	0.12953	-0.09853	-0.14614	0
Chiapin B	0.02584	-0.00795	-0.04012	0.018366	-0.01254	0.010665	0
1-Heptatriacotanol	0.009999	-0.00793	0.0063	-0.03273	-0.0281	0.019288	0.0
Aromadendrene oxide-2.1	0.021556	-0.00785	0.003714	-0.00018	-0.02752	-0.02783	0.0
2H-Pyran, 2-7-heptadecyloxytetrahydro-.1	0.029007	-0.01853	-0.07876	0.05103	-0.03012	0.01098	0.0
Deoxyartemisinin	0.047284	-0.0376	-0.12777	0.083603	-0.03686	-0.01318	0
Artemisinin	0.042639	0.015261	-0.05153	0.073991	-0.05618	0.040156	0
Heptacosane	0.001487	-0.10667	0.038271	0.079432	-0.08786	-0.02515	-0
Unkown alkane	0.009206	-0.11216	0.10645	0.11068	-0.05708	-0.17672	-0
Olean-12-en-3-one	0.028023	0.058051	-0.15134	0.31603	0.2331	0.046794	0

PLSDA VIP values

	Comp. 1	Comp. 2	Comp. 3	Comp. 4	Comp. 5
Pinene	0.33469	0.81716	0.77008	0.7311	0.73908
Ocimene	0.43504	0.62129	0.54747	0.66974	0.6661
Camphene	1.2714	1.1818	1.1996	1.1618	1.1684
Cymene	0.40051	0.99114	1.6688	1.6065	1.5835
Eucalyptol	0.10973	0.27355	0.27825	0.79406	0.79848
Terpinen	0.2312	0.36534	0.72525	0.69297	0.69056
cis-beta-Terpineol	0.09971 4	0.39461	0.34975	0.33364	0.34058
Cyclohexanol, 1-methyl-4-1-methylethenyl-, cis-	0.17679	0.27069	0.30171	0.28615	0.30709
Camphor	5.3055	4.915	4.3541	4.1359	4.0742
endo-Borneol	3.6694	3.7187	3.8495	3.6489	3.595
Terpineol	0.11197	0.16704	0.19833	0.18923	0.28729
cis-p-mentha-17,8-dien-2-ol	0.02723 9	0.25545	0.23802	0.22569	0.23676
Cyclohexane, 1-ethenyl-1-methyl-2-1-methylethenyl-4-1-methyle	0.32761	0.34408	0.41982	0.49703	0.50955

Phenol, 2-methoxy-3-2-propenyl-	0.37822	0.32687	0.30027	0.32459	0.32482
Butanoic acid, 2-methyl-, phenylmethyl ester	1.5518	1.5139	1.4119	1.4822	1.4709
Elemen	0.45861	0.41324	0.44457	0.42718	0.42298
Ylangene	0.43769	0.55647	0.65801	0.6285	0.62531
Caryophyllene	0.20573	0.38534	0.45244	0.81836	0.80752
Copaene	0.36449	0.47197	0.55478	0.5362	0.53235
cis-beta-Farnesene	0.34819	0.33421	0.44229	0.87296	0.88198
Germacrene D	0.89222	1.1483	1.3497	1.3201	1.3078
Isoaromadendrene epoxide	0.09594 9	0.24365	0.40092	0.47988	0.47476
compound 1	0.08518 2	0.07330 2	0.06460 4	0.06177 3	0.06268 9
Caryophyllene oxide	0.15903	0.16128	0.53287	0.62886	0.72794
beta-Guaiene	0.16831	0.14646	0.1294	0.13238	0.24361
Isoaromadendrene epoxide.1	0.61063	0.57682	0.72313	0.70804	0.69969
Farnesene epoxide, E-	0.51525	0.49202	0.60269	0.57285	0.56731
Cubenol	0.35597	0.30444	0.27372	0.37165	0.851
Aromadendrene oxide-2	0.12319	0.17913	0.23344	0.32696	0.32285
Isoaromadendrene epoxide.2	0.01897	0.06791 3	0.08121 6	0.08568	0.11014
Neointermedeol	0.36123	0.39604	0.43763	0.41683	0.41679
Epiglobulol	0.33775	0.33015	0.4092	0.41003	0.40399
Corymbolone	0.908	0.90145	1.04	0.99254	0.98411
2H-Pyran, 2-7-heptadecynloxytetrahydro-	0.03096 4	0.08631	0.07729 9	0.12562	0.15901
Cedren-13-ol, 8-	0.21344	0.56296	0.51854	1.0661	1.0725
Isoaromadendrene epoxide.3	0.10582	0.12627	0.1127	0.17248	0.20487
cis-Z-alpha-Bisabolene epoxide	0.18676	0.24213	0.42099	0.44033	0.44404
Cholestan-3-ol, 2-methylene-, 3beta,5a-	0.13843	0.25474	0.28934	0.34129	0.34199
2-Methyl-4-2,6,6-trimethylcyclohex-2-enylbut-3-en-2-ol	1.3015	1.2399	1.429	1.3546	1.3534
Artemisinic acid	0.04254 1	0.05193 8	0.04954	0.07245 6	0.07138 2
3-buten-2-one, 4-5,5-dimethyl-1-oxaspiro2.5oct-4-yl	0.17441	0.22574	0.20663	0.23876	0.28025
Dihydroartemisinic acid	0.35902	0.30807	0.27724	0.87405	0.86185
Isoaromadendrene epoxide.4	0.08311 6	0.08244	0.18113	0.42988	0.42928
beta-copaene	0.05839 8	0.86027	0.887	0.8509	0.90279
Chiapin B	0.06563	0.12715	0.11312	0.11094	0.11036
1-Heptatriacotanol	0.02121 2	0.04106	0.04560 3	0.04380 4	0.14279

Aromadendrene oxide-2.1	0.026438	0.12429	0.12347	0.12617	0.15664
2H-Pyran, 2-7-heptadecyloxytetrahydro-.1	0.016728	0.19216	0.18051	0.18153	0.18257
Deoxyartemisinin	0.12033	0.40312	0.40868	0.39058	0.3989
Artemisinin	0.24454	0.23854	0.2116	0.39021	0.3852
Heptacosane	0.75598	0.81042	0.74468	0.70753	0.70246
Unkown alkane	0.61706	0.65747	0.57997	0.61957	0.63301
Olean-12-en-3-one	0.32813	0.28412	0.38382	0.56196	0.60996

Metabolite correlation coefficients

	cis-beta-Farnesene	Cedren-13-ol, 8-	Cymene	Terpinen	Butanoic acid, 2-m	Neointermedeol	Isoaromadendrene	Farnesene epoxide,	Elemen
cis-beta-Farnesene	1.00	0.43	0.34	0.17	0.13	0.16	0.49	0.34	0.34
Cedren-13-ol, 8-	0.43	1.00	0.33	0.40	0.00	-0.06	0.12	0.09	0.27
Cymene	0.34	0.33	1.00	0.87	0.26	0.40	0.26	0.29	0.55
Terpinen	0.17	0.40	0.87	1.00	0.27	0.38	0.16	0.20	0.59
Butanoic acid, 2-m	0.13	0.00	0.26	0.27	1.00	0.30	0.44	0.40	0.34
Neointermedeol	0.16	-0.06	0.40	0.38	0.30	1.00	0.66	0.59	0.77
Isoaromadendrene	0.49	0.12	0.26	0.16	0.44	0.66	1.00	0.81	0.71
Farnesene epoxide,	0.34	0.09	0.29	0.20	0.40	0.59	0.81	1.00	0.62
Elemen	0.34	0.27	0.55	0.59	0.34	0.77	0.71	0.62	1.00
Epiglobulol	0.31	-0.04	0.40	0.32	0.43	0.69	0.75	0.63	0.72
Corymbolone	0.32	-0.07	0.39	0.32	0.48	0.86	0.84	0.75	0.79
2-Methyl-4-2,6,6-t	0.39	-0.05	0.37	0.27	0.42	0.85	0.88	0.78	0.77
cis-p-mentha-17,8-	-0.49	-0.15	-0.22	-0.08	0.07	0.25	-0.22	-0.09	0.00
cis-beta-Terpineol	-0.65	-0.37	-0.40	-0.25	0.13	-0.13	-0.50	-0.38	-0.38
Cyclohexanol, 1-me	-0.65	-0.31	-0.25	-0.11	0.11	-0.15	-0.56	-0.42	-0.33
3-buten-2-one, 4-5	-0.58	-0.37	-0.38	-0.23	-0.24	0.10	-0.35	-0.28	-0.24
Camphene	-0.52	-0.32	-0.34	-0.18	-0.28	-0.03	-0.49	-0.47	-0.32
Camphor	-0.63	-0.37	-0.42	-0.31	-0.36	-0.04	-0.60	-0.47	-0.40
Pinene	-0.56	-0.36	-0.61	-0.51	-0.04	-0.36	-0.49	-0.40	-0.51
Ocimene	-0.48	-0.31	-0.54	-0.43	-0.11	-0.30	-0.48	-0.33	-0.42
Terpineol	-0.57	-0.34	-0.19	0.03	0.03	-0.13	-0.50	-0.35	-0.26
Isoaromadendrene	-0.47	0.06	-0.27	-0.05	-0.20	-0.20	-0.45	-0.36	-0.30
Isoaromadendrene	-0.10	-0.16	0.16	0.32	-0.02	0.58	0.10	0.11	0.43
Cholestan-3-ol, 2-	-0.23	-0.04	0.18	0.38	0.20	0.58	0.09	0.16	0.43

Cubenol	-0.12	0.11	0.15	0.34	-0.20	0.30	-0.18	-0.07	0.19
Aromadendrene oxid	-0.05	-0.14	0.04	0.18	0.23	0.29	0.11	0.13	0.21
Dihydroartemisini	-0.21	-0.17	-0.22	-0.22	-0.39	0.38	0.12	0.07	0.18
Artemisinin	-0.33	-0.39	-0.40	-0.43	-0.36	0.26	-0.08	-0.07	-0.16
Phenol, 2-methoxy-	-0.17	-0.14	0.23	0.30	-0.39	0.47	-0.13	0.03	0.21
beta-Guaiene	-0.23	0.12	0.04	0.21	-0.37	0.40	-0.14	-0.07	0.22
2H-Pyran, 2-7-hept	-0.17	-0.26	-0.08	-0.04	0.01	0.38	-0.04	0.06	0.04
Chiapin B	-0.35	-0.28	-0.14	-0.10	-0.11	0.34	-0.17	0.00	-0.08
beta-copaene	-0.26	-0.35	-0.04	0.00	-0.01	0.59	-0.01	0.09	0.12
Deoxyartemisinin	-0.18	-0.30	-0.04	-0.03	0.12	0.55	0.09	0.16	0.06
Olean-12-en-3-one	0.06	-0.34	-0.04	-0.26	-0.10	0.11	-0.01	0.06	-0.23
Caryophyllene	0.59	0.07	0.39	0.21	-0.02	0.07	-0.01	-0.08	0.05
Cyclohexane, 1-eth	0.29	-0.23	0.08	-0.10	-0.15	-0.12	-0.27	-0.32	-0.36
Germacrene D	0.43	-0.21	-0.06	-0.24	-0.14	-0.30	-0.19	-0.30	-0.37
Ylangene	0.39	-0.20	-0.09	-0.28	-0.15	-0.37	-0.24	-0.34	-0.44
Copaene	0.42	-0.18	-0.06	-0.25	-0.15	-0.36	-0.22	-0.31	-0.42
compound 1	-0.33	-0.31	-0.35	-0.27	-0.26	-0.42	-0.31	-0.37	-0.27
Eucalyptol	-0.29	-0.12	-0.37	-0.23	-0.03	-0.55	-0.32	-0.35	-0.41
Artemisinic acid	-0.23	-0.13	-0.35	-0.29	-0.24	-0.56	-0.22	-0.27	-0.34
cis-Z-alpha-Bisabo	-0.08	0.24	0.38	0.47	-0.10	0.09	-0.21	-0.09	0.10
Isoaromadendrene e	0.05	0.26	0.08	0.11	-0.01	0.22	0.24	0.16	0.31
Heptacosane	-0.19	-0.22	-0.09	0.00	0.59	0.05	0.13	0.13	-0.02
Unkown alkane	-0.34	-0.45	-0.11	-0.07	0.47	-0.09	0.01	0.03	-0.13
endo-Borneol	-0.23	-0.09	0.18	0.28	0.73	0.21	0.02	0.03	0.07
Caryophyllene oxide	-0.33	0.02	-0.08	0.11	0.36	-0.06	-0.25	-0.23	-0.18
Isoaromadendrene e	-0.20	-0.19	-0.22	-0.17	0.02	-0.10	-0.09	-0.09	-0.13
Aromadendrene oxid	-0.31	-0.26	-0.28	-0.14	-0.16	0.00	-0.27	-0.21	-0.17
2H-Pyran, 2-7-hept	-0.22	-0.20	-0.06	0.02	-0.09	0.24	-0.04	-0.05	0.13
1-Heptatriacotanol	-0.15	0.09	-0.12	0.00	-0.15	-0.07	-0.19	-0.14	-0.04

	Epiglobulol	Corymbolone	2-Methyl-4-2,6,6-t	cis-p-mentha-1,7,8-	cis-beta-Terpineol	Cyclohexanol, 1-me	3-buten-2-one, 4-5	Camphene	Camphor
cis-beta-Farnesene	0.31	0.32	0.39	-0.49	-0.65	-0.65	-0.58	-0.52	-0.63
Cedren-13-ol, 8-	-0.04	-0.07	-0.05	-0.15	-0.37	-0.31	-0.37	-0.32	-0.37
Cymene	0.40	0.39	0.37	-0.22	-0.40	-0.25	-0.38	-0.34	-0.42
Terpinen	0.32	0.32	0.27	-0.08	-0.25	-0.11	-0.23	-0.18	-0.31
Butanoic acid, 2-m	0.43	0.48	0.42	0.07	0.13	0.11	-0.24	-0.28	-0.36

Neointermedeol	0.69	0.86	0.85	0.25	-0.13	-0.15	0.10	-0.03	-0.04
Isoaromadendrene e	0.75	0.84	0.88	-0.22	-0.50	-0.56	-0.35	-0.49	-0.60
Farnesene epoxide,	0.63	0.75	0.78	-0.09	-0.38	-0.42	-0.28	-0.47	-0.47
Elemen	0.72	0.79	0.77	0.00	-0.38	-0.33	-0.24	-0.32	-0.40
Epiglobulol	1.00	0.85	0.84	0.00	-0.26	-0.26	-0.17	-0.41	-0.40
Corymbolone	0.85	1.00	0.97	0.06	-0.27	-0.29	-0.14	-0.31	-0.34
2-Methyl-4-2,6,6-t	0.84	0.97	1.00	-0.03	-0.38	-0.41	-0.19	-0.39	-0.43
cis-p-mentha-17,8-	0.00	0.06	-0.03	1.00	0.82	0.80	0.67	0.49	0.77
cis-beta-Terpineol	-0.26	-0.27	-0.38	0.82	1.00	0.97	0.73	0.63	0.80
Cyclohexanol, 1-me	-0.26	-0.29	-0.41	0.80	0.97	1.00	0.67	0.59	0.78
3-buten-2-one, 4-5	-0.17	-0.14	-0.19	0.67	0.73	0.67	1.00	0.73	0.77
Camphene	-0.41	-0.31	-0.39	0.49	0.63	0.59	0.73	1.00	0.75
Camphor	-0.40	-0.34	-0.43	0.77	0.80	0.78	0.77	0.75	1.00
Pinene	-0.39	-0.39	-0.47	0.57	0.80	0.74	0.54	0.55	0.66
Ocimene	-0.38	-0.34	-0.40	0.46	0.64	0.59	0.49	0.48	0.54
Terpineol	-0.17	-0.24	-0.34	0.49	0.73	0.76	0.65	0.55	0.56
Isoaromadendrene e	-0.28	-0.38	-0.43	0.42	0.58	0.56	0.70	0.57	0.52
Isoaromadendrene e	0.40	0.43	0.35	0.29	0.13	0.13	0.39	0.29	0.25
Cholestan-3-ol, 2-	0.39	0.44	0.32	0.66	0.46	0.47	0.52	0.35	0.41
Cubenol	0.05	0.08	0.02	0.20	0.14	0.17	0.39	0.36	0.30
Aromadendrene oxid	0.34	0.29	0.22	0.32	0.34	0.34	0.42	0.21	0.19
Dihydroartemisini	0.02	0.15	0.18	0.36	0.09	0.01	0.41	0.30	0.37
Artemisinin	-0.13	0.04	0.05	0.51	0.35	0.25	0.60	0.45	0.63
Phenol, 2-methoxy-	0.14	0.22	0.17	0.38	0.12	0.15	0.43	0.33	0.50
beta-Guaiene	0.06	0.14	0.06	0.55	0.30	0.31	0.60	0.42	0.55
2H-Pyran, 2-7-hept	0.15	0.21	0.19	0.52	0.34	0.27	0.44	0.25	0.43
Chiapin B	0.03	0.13	0.08	0.75	0.57	0.53	0.66	0.47	0.71
beta-copaene	0.22	0.37	0.31	0.70	0.49	0.44	0.65	0.44	0.64
Deoxyartemisinin	0.18	0.33	0.31	0.58	0.40	0.34	0.52	0.31	0.49
Olean-12-en-3-one	-0.03	0.05	0.08	-0.04	-0.05	-0.07	-0.06	-0.01	0.19
Caryophyllene	0.17	0.09	0.13	-0.18	-0.20	-0.13	-0.17	-0.17	-0.17
Cyclohexane, 1-eth	-0.18	-0.17	-0.14	-0.10	0.03	0.05	0.03	0.10	0.17
Germacrene D	-0.17	-0.22	-0.16	-0.34	-0.12	-0.10	-0.15	-0.06	-0.09
Ylangene	-0.22	-0.29	-0.23	-0.36	-0.11	-0.09	-0.15	-0.08	-0.08
Copaene	-0.20	-0.27	-0.21	-0.36	-0.13	-0.10	-0.16	-0.09	-0.10
compound 1	-0.22	-0.41	-0.41	-0.19	0.16	0.14	0.29	0.30	0.10
Eucalyptol	-0.32	-0.49	-0.50	-0.15	0.27	0.24	0.23	0.23	0.03
Artemisinic acid	-0.27	-0.44	-0.40	-0.26	0.06	0.04	0.13	0.07	-0.10
cis-Z-alpha-Bisabo	-0.09	-0.10	-0.14	0.03	-0.05	0.04	0.08	0.16	0.10
Isoaromadendrene e	0.14	0.12	0.19	0.15	-0.12	-0.15	0.02	-0.10	-0.13
Heptacosane	0.25	0.16	0.14	0.24	0.37	0.30	0.18	-0.02	-0.05

Unkown alkane	0.18	0.06	0.03	0.17	0.44	0.40	0.26	0.09	0.00
endo-Borneol	0.05	0.16	0.09	0.35	0.48	0.46	0.27	0.23	0.08
Caryophyllene oxide	-0.21	-0.14	-0.18	0.35	0.55	0.52	0.51	0.39	0.21
Isoaromadendrene e	0.00	-0.04	-0.09	0.20	0.29	0.26	0.16	0.07	0.22
Aromadendrene oxid	-0.04	-0.16	-0.18	0.41	0.44	0.38	0.60	0.42	0.45
2H-Pyran, 2-7-hept	0.06	0.04	0.04	0.47	0.31	0.28	0.33	0.29	0.35
1-Heptatriacotanol	-0.08	-0.21	-0.19	0.19	0.16	0.13	0.28	0.22	0.18

	Pinene	Ocimene	Terpineol	Isoaromadendrene e	Isoaromadendrene e e	Cholestan-3-ol, 2-	Cubanol	Aromadendrene oxid	Dihydroartemisinin
cis-beta-Farnesene	-0.56	-0.48	-0.57	-0.47	-0.10	-0.23	0.12	-0.05	-0.21
Cedren-13-ol, 8-	-0.36	-0.31	-0.34	0.06	-0.16	-0.04	0.11	-0.14	-0.17
Cymene	-0.61	-0.54	-0.19	-0.27	0.16	0.18	0.15	0.04	-0.22
Terpinen	-0.51	-0.43	0.03	-0.05	0.32	0.38	0.34	0.18	-0.22
Butanoic acid, 2-m	-0.04	-0.11	0.03	-0.20	-0.02	0.20	0.20	0.23	-0.39
Neointermedeol	-0.36	-0.30	-0.13	-0.20	0.58	0.58	0.30	0.29	0.38
Isoaromadendrene e	-0.49	-0.48	-0.50	-0.45	0.10	0.09	0.18	0.11	0.12
Farnesene epoxide,	-0.40	-0.33	-0.35	-0.36	0.11	0.16	0.07	0.13	0.07
Elemen	-0.51	-0.42	-0.26	-0.30	0.43	0.43	0.19	0.21	0.18
Epiglobulol	-0.39	-0.38	-0.17	-0.28	0.40	0.39	0.05	0.34	0.02
Corymbolone	-0.39	-0.34	-0.24	-0.38	0.43	0.44	0.08	0.29	0.15
2-Methyl-4-2,6,6-t	-0.47	-0.40	-0.34	-0.43	0.35	0.32	0.02	0.22	0.18
cis-p-mentha-17,8-	0.57	0.46	0.49	0.42	0.29	0.66	0.20	0.32	0.36
cis-beta-Terpineol	0.80	0.64	0.73	0.58	0.13	0.46	0.14	0.34	0.09
Cyclohexanol, 1-me	0.74	0.59	0.76	0.56	0.13	0.47	0.17	0.34	0.01
3-buten-2-one, 4-5	0.54	0.49	0.65	0.70	0.39	0.52	0.39	0.42	0.41
Camphene	0.55	0.48	0.55	0.57	0.29	0.35	0.36	0.21	0.30
Camphor	0.66	0.54	0.56	0.52	0.25	0.41	0.30	0.19	0.37
Pinene	1.00	0.87	0.52	0.43	-0.10	0.17	0.07	0.15	0.09
Ocimene	0.87	1.00	0.49	0.40	-0.05	0.13	0.02	0.11	0.14
Terpineol	0.52	0.49	1.00	0.75	0.42	0.53	0.56	0.59	-0.17
Isoaromadendrene e	0.43	0.40	0.75	1.00	0.30	0.45	0.57	0.52	-0.03
Isoaromadendrene e	-0.10	-0.05	0.42	0.30	1.00	0.73	0.73	0.52	0.19

Cholestan-3-ol, 2-	0.17	0.13	0.53	0.45	0.73	1.00	0.67	0.67	0.09
Cubenol	-0.07	-0.02	0.56	0.57	0.73	0.67	1.00	0.52	-0.07
Aromadendrene oxid	0.15	0.11	0.59	0.52	0.52	0.67	0.52	1.00	-0.16
Dihydroartemisini	0.09	0.14	-0.17	-0.03	0.19	0.09	0.07	-0.16	1.00
Artemisinin	0.31	0.27	-0.05	0.04	0.12	0.14	0.07	-0.12	0.75
Phenol, 2-methoxy-	-0.09	-0.07	0.24	0.21	0.65	0.63	0.66	0.32	0.28
beta-Guaiene	0.11	0.07	0.38	0.49	0.66	0.73	0.74	0.36	0.34
2H-Pyran, 2-7-hept	0.11	0.04	0.09	0.02	0.43	0.42	0.13	0.02	0.36
Chiapin B	0.35	0.28	0.25	0.23	0.32	0.48	0.17	0.22	0.43
beta-copaene	0.22	0.16	0.28	0.20	0.54	0.69	0.40	0.35	0.37
Deoxyartemisinin	0.13	0.06	0.14	0.08	0.34	0.52	0.20	0.24	0.31
Olean-12-en-3-one	-0.06	-0.08	-0.26	-0.36	-0.09	-0.20	0.19	-0.24	0.01
Caryophyllene	-0.30	-0.26	-0.06	-0.13	0.12	-0.01	0.10	0.12	-0.27
Cyclohexane, 1-eth	0.00	-0.05	-0.08	-0.17	-0.07	-0.21	0.12	-0.10	-0.15
Germacrene D	-0.03	-0.02	-0.06	-0.16	-0.18	-0.42	0.15	-0.11	-0.29
Ylangene	-0.01	-0.02	-0.08	-0.18	-0.26	-0.49	0.20	-0.15	-0.31
Copaene	-0.03	-0.04	-0.08	-0.17	-0.25	-0.47	0.18	-0.13	-0.32
compound 1	0.30	0.31	0.46	0.44	0.07	-0.09	0.15	0.21	-0.06
Eucalyptol	0.37	0.34	0.54	0.58	-0.11	-0.11	0.10	0.26	-0.31
Artemisinic acid	0.25	0.29	0.33	0.42	-0.20	-0.27	0.06	0.16	-0.12
cis-Z-alpha-Bisabo	-0.24	-0.19	0.07	0.15	0.24	0.17	0.35	-0.10	0.03
Isoaromadendrene	-0.15	-0.08	-0.33	-0.17	0.00	-0.05	0.31	-0.30	0.51
Heptacosane	0.23	0.12	0.25	0.12	0.10	0.21	0.16	0.24	-0.15
Unkown alkane	0.32	0.23	0.44	0.23	0.06	0.14	0.15	0.30	-0.21
endo-Borneol	0.22	0.17	0.34	0.18	0.07	0.31	0.04	0.22	-0.13
Caryophyllene oxide	0.37	0.35	0.54	0.53	0.09	0.28	0.19	0.26	-0.09
Isoaromadendrene	0.21	0.06	0.20	0.14	0.03	0.18	0.00	0.16	-0.07
Aromadendrene oxid	0.31	0.25	0.39	0.42	0.40	0.33	0.18	0.29	0.28
2H-Pyran, 2-7-hept	0.17	0.11	0.15	0.09	0.29	0.29	0.01	0.07	0.46
1-Heptatriacotanol	0.07	0.12	0.13	0.27	0.21	0.10	0.05	0.01	0.23

	Arte misinin	Phenol , 2- methoxy-	beta - Guaiene	2H- Pyran , 2-7- hept	Chiapin B	beta- copaene	Deoxy artemisinin	Olean- 12-en- 3-one	Caryo phyllene
cis-beta-Farnesene	-0.33	-0.17	-0.23	-0.17	-0.35	-0.26	-0.18	0.06	0.59
Cedren-13-ol, 8-	-0.39	-0.14	0.12	-0.26	-0.28	-0.35	-0.30	-0.34	0.07
Cymene	-0.40	0.23	0.04	-0.08	-0.14	-0.04	-0.04	-0.04	0.39
Terpinen	-0.43	0.30	0.21	-0.04	-0.10	0.00	-0.03	-0.26	0.21
Butanoic acid, 2-m	-0.36	-0.39	-0.37	0.01	-0.11	-0.01	0.12	-0.10	-0.02
Neointermedeol	0.26	0.47	0.40	0.38	0.34	0.59	0.55	0.11	0.07
Isoaromadendrene e	-0.08	-0.13	-0.14	-0.04	-0.17	-0.01	0.09	-0.01	-0.01
Farnesene epoxide,	-0.07	0.03	-0.07	0.06	0.00	0.09	0.16	0.06	-0.08
Elemen	-0.16	0.21	0.22	0.04	-0.08	0.12	0.06	-0.23	0.05
Epiglobulol	-0.13	0.14	0.06	0.15	0.03	0.22	0.18	-0.03	0.17
Corymbolone	0.04	0.22	0.14	0.21	0.13	0.37	0.33	0.05	0.09
2-Methyl-4-2,6,6-t	0.05	0.17	0.06	0.19	0.08	0.31	0.31	0.08	0.13
cis-p-mentha-17,8-	0.51	0.38	0.55	0.52	0.75	0.70	0.58	-0.04	-0.18
cis-beta-Terpineol	0.35	0.12	0.30	0.34	0.57	0.49	0.40	-0.05	-0.20
Cyclohexanol, 1-me	0.25	0.15	0.31	0.27	0.53	0.44	0.34	-0.07	-0.13
3-buten-2-one, 4-5	0.60	0.43	0.60	0.44	0.66	0.65	0.52	-0.06	-0.17
Camphene	0.45	0.33	0.42	0.25	0.47	0.44	0.31	-0.01	-0.17
Camphor	0.63	0.50	0.55	0.43	0.71	0.64	0.49	0.19	-0.17
Pinene	0.31	-0.09	0.11	0.11	0.35	0.22	0.13	-0.06	-0.30
Ocimene	0.27	-0.07	0.07	0.04	0.28	0.16	0.06	-0.08	-0.26
Terpineol	-0.05	0.24	0.38	0.09	0.25	0.28	0.14	-0.26	-0.06
Isoaromadendrene e	0.04	0.21	0.49	0.02	0.23	0.20	0.08	-0.36	-0.13
Isoaromadendrene e	0.12	0.65	0.66	0.43	0.32	0.54	0.34	-0.09	0.12
Cholestan-3-ol, 2-	0.14	0.63	0.73	0.42	0.48	0.69	0.52	-0.20	-0.01
Cubenol	-0.07	0.66	0.74	0.13	0.17	0.40	0.20	-0.19	0.10
Aromadendrene oxid	-0.12	0.32	0.36	0.02	0.22	0.35	0.24	-0.24	0.12
Dihydroartemisini	0.75	0.28	0.34	0.36	0.43	0.37	0.31	0.01	-0.27
Artemisinin	1.00	0.41	0.39	0.54	0.73	0.68	0.65	0.34	-0.22
Phenol, 2-methoxy-	0.41	1.00	0.79	0.49	0.57	0.73	0.55	0.19	0.12
beta-Guaiene	0.39	0.79	1.00	0.42	0.55	0.66	0.46	-0.13	-0.03
2H-Pyran, 2-7-hept	0.54	0.49	0.42	1.00	0.68	0.73	0.67	0.28	0.07
Chiapin B	0.73	0.57	0.55	0.68	1.00	0.85	0.77	0.31	-0.04
beta-copaene	0.68	0.73	0.66	0.73	0.85	1.00	0.87	0.34	0.08

Deoxyartemisinin	0.65	0.55	0.46	0.67	0.77	0.87	1.00	0.40	0.03
Olean-12-en-3-one	0.34	0.19	-0.13	0.28	0.31	0.34	0.40	1.00	0.27
Caryophyllene	-0.22	0.12	-0.03	0.07	-0.04	0.08	0.03	0.27	1.00
Cyclohexane, 1-eth	0.20	0.14	-0.06	0.21	0.23	0.24	0.26	0.59	0.73
Germacrene D	-0.09	-0.19	-0.30	-0.10	-0.14	-0.12	-0.11	0.37	0.76
Ylangene	-0.08	-0.23	-0.33	-0.15	-0.15	-0.16	-0.13	0.37	0.69
Copaene	-0.09	-0.22	-0.32	-0.15	-0.15	-0.15	-0.12	0.36	0.72
compound 1	-0.18	-0.17	-0.06	-0.34	-0.32	-0.34	-0.41	-0.30	-0.21
Eucalyptol	-0.35	-0.39	-0.20	-0.47	-0.39	-0.45	-0.45	-0.43	-0.19
Artemisinic acid	-0.27	-0.41	-0.24	-0.53	-0.45	-0.55	-0.55	-0.46	-0.21
cis-Z-alpha-Bisabo	-0.07	0.31	0.31	0.14	0.07	0.07	0.05	-0.13	-0.02
Isoaromadendrene e	0.16	-0.13	0.03	0.41	0.13	-0.03	0.02	-0.22	-0.13
Heptacosane	-0.05	-0.23	-0.18	0.40	0.14	0.14	0.19	-0.16	-0.15
Unkown alkane	-0.12	-0.30	-0.26	0.14	0.02	0.02	0.03	-0.15	-0.12
endo-Borneol	-0.03	-0.19	-0.09	0.23	0.24	0.23	0.32	-0.14	-0.05
Caryophyllene oxide	0.05	-0.14	0.16	0.14	0.23	0.19	0.23	-0.37	-0.13
Isoaromadendrene e	0.16	0.10	0.14	0.14	0.15	0.16	0.16	0.11	-0.12
Aromadendrene oxid	0.28	0.28	0.33	0.65	0.41	0.37	0.31	-0.08	-0.05
2H-Pyran, 2-7-hept	0.28	0.22	0.23	0.57	0.32	0.30	0.21	-0.13	-0.11
1-Heptatriacotanol	0.06	0.06	0.16	0.50	0.16	0.06	0.00	-0.18	-0.05

	Cyclohexane, 1-eth	Germacrene D	ylangene	copaene	compound 1	Eucalyptol	Artemisinic acid	cis-Z-alpha-Bisabo	Isoaromadendrene e
cis-beta-Farnesene	0.29	0.43	0.39	0.42	-0.33	-0.29	-0.23	-0.08	0.05
Cedren-13-ol, 8-	-0.23	-0.21	-0.20	-0.18	-0.31	-0.12	-0.13	0.24	0.26
Cymene	0.08	-0.06	-0.09	-0.06	-0.35	-0.37	-0.35	0.38	0.08
Terpinen	-0.10	-0.24	-0.28	-0.25	-0.27	-0.23	-0.29	0.47	0.11
Butanoic acid, 2-m	-0.15	-0.14	-0.15	-0.15	-0.26	-0.03	-0.24	-0.10	-0.01
Neointermedeol	-0.12	-0.30	-0.37	-0.36	-0.42	-0.55	-0.56	0.09	0.22
Isoaromadendrene e	-0.27	-0.19	-0.24	-0.22	-0.31	-0.32	-0.22	-0.21	0.24
Farnesene epoxide,	-0.32	-0.30	-0.34	-0.31	-0.37	-0.35	-0.27	-0.09	0.16
Elemen	-0.36	-0.37	-0.44	-0.42	-0.27	-0.41	-0.34	0.10	0.31
Epiglobulol	-0.18	-0.17	-0.22	-0.20	-0.22	-0.32	-0.27	-0.09	0.14
Corymbolone	-0.17	-0.22	-0.29	-0.27	-0.41	-0.49	-0.44	-0.10	0.12
2-Methyl-4-2,6,6-t	-0.14	-0.16	-0.23	-0.21	-0.41	-0.50	-0.40	-0.14	0.19
cis-p-mentha-17,8-	-0.10	-0.34	-0.36	-0.36	-0.19	-0.15	-0.26	0.03	0.15
cis-beta-Terpineol	0.03	-0.12	-0.11	-0.13	0.16	0.27	0.06	-0.05	-0.12
Cyclohexanol, 1-me	0.05	-0.10	-0.09	-0.10	0.14	0.24	0.04	0.04	-0.15
3-buten-2-one, 4-5	0.03	-0.15	-0.15	-0.16	0.29	0.23	0.13	0.08	0.02

Camphene	0.10	-0.06	-0.08	-0.09	0.30	0.23	0.07	0.16	-0.10
Camphor	0.17	-0.09	-0.08	-0.10	0.10	0.03	-0.10	0.10	-0.13
Pinene	0.00	-0.03	-0.01	-0.03	0.30	0.37	0.25	-0.24	-0.15
Ocimene	-0.05	-0.02	-0.02	-0.04	0.31	0.34	0.29	-0.19	-0.08
Terpineol	-0.08	-0.06	-0.08	-0.08	0.46	0.54	0.33	0.07	-0.33
Isoaromadendrene e	-0.17	-0.16	-0.18	-0.17	0.44	0.58	0.42	0.15	-0.17
Isoaromadendrene e	-0.07	-0.18	-0.26	-0.25	0.07	-0.11	-0.20	0.24	0.00
Cholestan-3-ol, 2-	-0.21	-0.42	-0.49	-0.47	-0.09	-0.11	-0.27	0.17	-0.05
Cubenol	-0.12	-0.15	-0.20	-0.18	0.15	0.10	-0.06	0.35	-0.31
Aromadendrene oxid	-0.10	-0.11	-0.15	-0.13	0.21	0.26	0.16	-0.10	-0.30
Dihydroartemisini	-0.15	-0.29	-0.31	-0.32	-0.06	-0.31	-0.12	0.03	0.51
Artemisinin	0.20	-0.09	-0.08	-0.09	-0.18	-0.35	-0.27	-0.07	0.16
Phenol, 2-methoxy-	0.14	-0.19	-0.23	-0.22	-0.17	-0.39	-0.41	0.31	-0.13
beta-Guaiene	-0.06	-0.30	-0.33	-0.32	-0.06	-0.20	-0.24	0.31	0.03
2H-Pyran, 2-7-hept	0.21	-0.10	-0.15	-0.15	-0.34	-0.47	-0.53	0.14	0.41
Chiapin B	0.23	-0.14	-0.15	-0.15	-0.32	-0.39	-0.45	0.07	0.13
beta-copaene	0.24	-0.12	-0.16	-0.15	-0.34	-0.45	-0.55	0.07	-0.03
Deoxyartemisinin	0.26	-0.11	-0.13	-0.12	-0.41	-0.45	-0.55	0.05	0.02
Olean-12-en-3-one	0.59	0.37	0.37	0.36	-0.30	-0.43	-0.46	-0.13	-0.22
Caryophyllene	0.73	0.76	0.69	0.72	-0.21	-0.19	-0.21	-0.02	-0.13
Cyclohexane, 1-eth	1.00	0.82	0.81	0.82	-0.19	-0.23	-0.27	-0.02	-0.25
Germacrene D	0.82	1.00	0.98	0.98	0.04	0.05	0.05	-0.14	-0.31
Ylangene	0.81	0.98	1.00	1.00	0.04	0.08	0.08	-0.16	-0.34
Copaene	0.82	0.98	1.00	1.00	0.03	0.06	0.07	-0.16	-0.34
compound 1	-0.19	0.04	0.04	0.03	1.00	0.78	0.78	-0.07	-0.19
Eucalyptol	-0.23	0.05	0.08	0.06	0.78	1.00	0.81	-0.17	-0.29
Artemisinic acid	-0.27	0.05	0.08	0.07	0.78	0.81	1.00	-0.22	-0.14
cis-Z-alpha-Bisabo	-0.02	-0.14	-0.16	-0.16	-0.07	-0.17	-0.22	1.00	0.13
Isoaromadendrene e	-0.25	-0.31	-0.34	-0.34	-0.19	-0.29	-0.14	0.13	1.00
Heptacosane	-0.15	-0.18	-0.19	-0.19	0.07	0.18	0.02	-0.17	0.29
Unkown alkane	-0.12	-0.04	-0.05	-0.05	0.36	0.44	0.32	-0.23	0.07
endo-Borneol	0.02	-0.14	-0.15	-0.15	-0.14	0.09	-0.16	0.13	0.15
Caryophyllene oxide	-0.05	-0.09	-0.08	-0.08	0.12	0.33	0.14	0.15	0.10
Isoaromadendrene e	0.06	-0.09	-0.08	-0.09	0.15	0.14	0.09	-0.23	-0.19
Aromadendrene oxid	0.00	-0.13	-0.18	-0.18	0.28	0.12	0.07	0.09	0.42
2H-Pyran, 2-7-hept	-0.13	-0.25	-0.30	-0.29	0.03	-0.12	-0.11	0.11	0.61
1-Heptatriacotanol	-0.13	-0.19	-0.24	-0.24	0.17	0.06	0.06	0.21	0.66

	Hepta cosane	Unkown alkane	endo- Borneol	Caryo phyllene oxide	Isoaro madendrene e	Aroma dendrene oxid	2H- Pyran , 2-7- hept	1-Hepta triacotanol
cis-beta-Farnesene	-0.19	-0.34	-0.23	-0.33	-0.20	-0.31	-0.22	-0.15
Cedren-13-ol, 8-	-0.22	-0.45	-0.09	0.02	-0.19	-0.26	-0.20	0.09
Cymene	-0.09	-0.11	0.18	-0.08	-0.22	-0.28	-0.06	-0.12
Terpinen	0.00	-0.07	0.28	0.11	-0.17	-0.14	0.02	0.00
Butanoic acid, 2-m	0.59	0.47	0.73	0.36	0.02	-0.16	-0.09	-0.15
Neointermedeol	0.05	-0.09	0.21	-0.06	-0.10	0.00	0.24	-0.07
Isoaromadendrene e	0.13	0.01	0.02	-0.25	-0.09	-0.27	-0.04	-0.19
Farnesene epoxide,	0.13	0.03	0.03	-0.23	-0.09	-0.21	-0.05	-0.14
Elemen	-0.02	-0.13	0.07	-0.18	-0.13	-0.17	0.13	-0.04
Epiglobulol	0.25	0.18	0.05	-0.21	0.00	-0.04	0.06	-0.08
Corymbolone	0.16	0.06	0.16	-0.14	-0.04	-0.16	0.04	-0.21
2-Methyl-4-2,6,6-t	0.14	0.03	0.09	-0.18	-0.09	-0.18	0.04	-0.19
cis-p-mentha-17,8-	0.24	0.17	0.35	0.35	0.20	0.41	0.47	0.19
cis-beta-Terpineol	0.37	0.44	0.48	0.55	0.29	0.44	0.31	0.16
Cyclohexanol, 1-me	0.30	0.40	0.46	0.52	0.26	0.38	0.28	0.13
3-buten-2-one, 4-5	0.18	0.26	0.27	0.51	0.16	0.60	0.33	0.28
Camphene	-0.02	0.09	0.23	0.39	0.07	0.42	0.29	0.22
Camphor	-0.05	0.00	0.08	0.21	0.22	0.45	0.35	0.18
Pinene	0.23	0.32	0.22	0.37	0.21	0.31	0.17	0.07
Ocimene	0.12	0.23	0.17	0.35	0.06	0.25	0.11	0.12
Terpineol	0.25	0.44	0.34	0.54	0.20	0.39	0.15	0.13
Isoaromadendrene e	0.12	0.23	0.18	0.53	0.14	0.42	0.09	0.27
Isoaromadendrene e	0.10	0.06	0.07	0.09	0.03	0.40	0.29	0.21
Cholestan-3-ol, 2-	0.21	0.14	0.31	0.28	0.18	0.33	0.29	0.10
Cubenol	-0.16	-0.15	-0.04	0.19	0.00	0.18	0.01	0.05
Aromadendrene oxid	0.24	0.30	0.22	0.26	0.16	0.29	0.07	0.01
Dihydroartemisini	-0.15	-0.21	-0.13	-0.09	-0.07	0.28	0.46	0.23
Artemisinin	-0.05	-0.12	-0.03	0.05	0.16	0.28	0.28	0.06
Phenol, 2-methoxy-	-0.23	-0.30	-0.19	-0.14	0.10	0.28	0.22	0.06
beta-Guaiene	-0.18	-0.26	-0.09	0.16	0.14	0.33	0.23	0.16
2H-Pyran, 2-7-hept	0.40	0.14	0.23	0.14	0.14	0.65	0.57	0.50
Chiapin B	0.14	0.02	0.24	0.23	0.15	0.41	0.32	0.16
beta-copaene	0.14	0.02	0.23	0.19	0.16	0.37	0.30	0.06
Deoxyartemisinin	0.19	0.03	0.32	0.23	0.16	0.31	0.21	0.00
Olean-12-en-3-one	-0.16	-0.15	-0.14	-0.37	0.11	-0.08	-0.13	-0.18
Caryophyllene	-0.15	-0.12	-0.05	-0.13	-0.12	-0.05	-0.11	-0.05
Cyclohexane, 1-eth	-0.15	-0.12	0.02	-0.05	0.06	0.00	-0.13	-0.13
Germacrene D	-0.18	-0.04	-0.14	-0.09	-0.09	-0.13	-0.25	-0.19
Ylangene	-0.19	-0.05	-0.15	-0.08	-0.08	-0.18	-0.30	-0.24

Copaene	-0.19	-0.05	-0.15	-0.08	-0.09	-0.18	-0.29	-0.24
compound 1	0.07	0.36	-0.14	0.12	0.15	0.28	0.03	0.17
Eucalyptol	0.18	0.44	0.09	0.33	0.14	0.12	-0.12	0.06
Artemisinic acid	0.02	0.32	-0.16	0.14	0.09	0.07	-0.11	0.06
cis-Z-alpha-Bisabo	-0.17	-0.23	0.13	0.15	-0.23	0.09	0.11	0.21
Isoaromadendrene e	0.29	0.07	0.15	0.10	-0.19	0.42	0.61	0.66
Heptacosane	1.00	0.85	0.63	0.53	0.14	0.46	0.38	0.39
Unkown alkane	0.85	1.00	0.54	0.49	0.16	0.37	0.27	0.22
endo-Borneol	0.63	0.54	1.00	0.79	-0.01	0.16	0.14	0.11
Caryophyllene oxide	0.53	0.49	0.79	1.00	0.02	0.28	0.08	0.19
Isoaromadendrene e	0.14	0.16	-0.01	0.02	1.00	0.22	-0.03	0.01
Aromadendrene oxid	0.46	0.37	0.16	0.28	0.22	1.00	0.66	0.78
2H-Pyran, 2-7-hept	0.38	0.27	0.14	0.08	-0.03	0.66	1.00	0.59
1-Heptatriacotanol	0.39	0.22	0.11	0.19	0.01	0.78	0.59	1.00

	cis-beta-Farnesene	Cedren-13-ol, 8-	Cymene	Terpinen	Butanoic acid, 2-m	Neointermedeol	Isoaromadendrene e	Farnesene epoxide,	Elemen	Epiglobulol
cis-beta-Farnesene	NA	0.00	0.00	0.16	0.27	0.18	0.00	0.00	0.00	0.01
Cedren-13-ol, 8-	0.00	NA	0.00	0.00	0.98	0.61	0.34	0.44	0.02	0.75
Cymene	0.00	0.00	NA	0.00	0.03	0.00	0.03	0.01	0.00	0.00
Terpinen	0.16	0.00	0.00	NA	0.02	0.00	0.17	0.09	0.00	0.01
Butanoic acid, 2-m	0.27	0.98	0.03	0.02	NA	0.01	0.00	0.00	0.00	0.00
Neointermedeol	0.18	0.61	0.00	0.00	0.01	NA	0.00	0.00	0.00	0.00
Isoaromadendrene e	0.00	0.34	0.03	0.17	0.00	0.00	NA	0.00	0.00	0.00
Farnesene epoxide,	0.00	0.44	0.01	0.09	0.00	0.00	0.00	NA	0.00	0.00
Elemen	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	NA	0.00
Epiglobulol	0.01	0.75	0.00	0.01	0.00	0.00	0.00	0.00	0.00	NA
Corymbolone	0.01	0.55	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
2-Methyl-4-2,6,6-t	0.00	0.67	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00
cis-p-mentha-17,8-	0.00	0.20	0.06	0.49	0.59	0.03	0.06	0.47	1.00	0.97
cis-beta-Terpineol	0.00	0.00	0.00	0.04	0.30	0.27	0.00	0.00	0.00	0.03
Cyclohexanol, 1-me	0.00	0.01	0.04	0.38	0.35	0.21	0.00	0.00	0.00	0.03
3-buten-2-one, 4-5	0.00	0.00	0.00	0.06	0.05	0.41	0.00	0.02	0.04	0.17
Camphene	0.00	0.01	0.00	0.13	0.02	0.79	0.00	0.00	0.01	0.00

Camphor	0.00	0.00	0.00	0.01	0.00	0.74	0.00	0.00	0.00	0.00
Pinene	0.00	0.00	0.00	0.00	0.76	0.00	0.00	0.00	0.00	0.00
Ocimene	0.00	0.01	0.00	0.00	0.38	0.01	0.00	0.00	0.00	0.00
Terpineol	0.00	0.00	0.12	0.82	0.78	0.29	0.00	0.00	0.03	0.15
Isoaromadendrene	0.00	0.65	0.02	0.70	0.10	0.10	0.00	0.00	0.01	0.02
Isoaromadendrene	0.42	0.19	0.18	0.01	0.84	0.00	0.39	0.38	0.00	0.00
Cholestan-3-ol, 2-	0.05	0.76	0.13	0.00	0.09	0.00	0.47	0.18	0.00	0.00
Cubenol	0.30	0.38	0.21	0.00	0.10	0.01	0.13	0.58	0.13	0.67
Aromadendrene oxid	0.70	0.24	0.74	0.13	0.05	0.01	0.35	0.30	0.07	0.00
Dihydroartemisini	0.08	0.15	0.07	0.06	0.00	0.00	0.32	0.59	0.13	0.87
Artemisinin	0.01	0.00	0.00	0.00	0.00	0.03	0.49	0.57	0.18	0.29
Phenol, 2-methoxy-	0.16	0.25	0.06	0.01	0.00	0.00	0.30	0.81	0.09	0.25
beta-Guaiene	0.06	0.33	0.75	0.09	0.00	0.00	0.23	0.59	0.07	0.64
2H-Pyran, 2-7-hept	0.15	0.03	0.51	0.74	0.96	0.00	0.73	0.61	0.77	0.23
Chiapin B	0.00	0.02	0.23	0.39	0.38	0.00	0.15	0.99	0.52	0.81
beta-copaene	0.03	0.00	0.77	1.00	0.94	0.00	0.96	0.45	0.33	0.07
Deoxyartemisinin	0.13	0.01	0.75	0.81	0.33	0.00	0.48	0.18	0.65	0.13
Olean-12-en-3-one	0.63	0.00	0.73	0.03	0.42	0.36	0.96	0.64	0.06	0.79
Caryophyllene	0.00	0.55	0.00	0.07	0.87	0.59	0.93	0.49	0.68	0.16
Cyclohexane, 1-eth	0.02	0.06	0.49	0.42	0.21	0.33	0.03	0.01	0.00	0.13
Germacrene D	0.00	0.09	0.64	0.04	0.25	0.01	0.11	0.01	0.00	0.16
Ylangene	0.00	0.11	0.48	0.02	0.20	0.00	0.05	0.00	0.00	0.06
Copaene	0.00	0.14	0.63	0.03	0.22	0.00	0.06	0.01	0.00	0.09
compound 1	0.00	0.01	0.00	0.03	0.03	0.00	0.01	0.00	0.02	0.06
Eucalyptol	0.01	0.33	0.00	0.05	0.81	0.00	0.01	0.00	0.00	0.01
Artemisinic acid	0.05	0.30	0.00	0.01	0.05	0.00	0.07	0.03	0.00	0.02
cis-Z-alpha-Bisabo	0.52	0.05	0.00	0.00	0.43	0.47	0.08	0.45	0.42	0.46
Isoaromadendrene	0.68	0.03	0.49	0.35	0.94	0.07	0.05	0.19	0.01	0.23
Heptacosane	0.11	0.06	0.47	0.97	0.00	0.67	0.27	0.29	0.87	0.04
Unkown alkane	0.00	0.00	0.36	0.54	0.00	0.47	0.90	0.83	0.28	0.13
endo-Borneol	0.06	0.46	0.14	0.02	0.00	0.08	0.90	0.82	0.58	0.67
Caryophyllene oxide	0.01	0.84	0.51	0.35	0.00	0.62	0.03	0.06	0.15	0.08
Isoaromadendrene	0.09	0.12	0.07	0.16	0.86	0.43	0.43	0.46	0.28	0.97
Aromadendrene oxid	0.01	0.03	0.02	0.26	0.18	0.99	0.02	0.07	0.15	0.73

2H-Pyran, 2-7-hept	0.07	0.10	0.65	0.88	0.45	0.04	0.77	0.70	0.27	0.61
1-Heptatriacotanol	0.22	0.47	0.31	0.97	0.22	0.55	0.12	0.26	0.76	0.49

Pval correlation table

	Corymbolone	2-Methyl-4-2,6,6-t	cis-p-mentha-1,7,8-	cis-beta-Terpinol	Cyclohexanol, 1-me	3-buten-2-one, 4-5	Camphene	Camphor	Pine	Ocimene
cis-beta-Farnesene	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cedren-13-ol, 8-	0.55	0.67	0.20	0.00	0.01	0.00	0.01	0.00	0.00	0.01
Cymene	0.00	0.00	0.06	0.00	0.04	0.00	0.00	0.00	0.00	0.00
Terpinen	0.01	0.02	0.49	0.04	0.38	0.06	0.13	0.01	0.00	0.00
Butanoic acid, 2-m	0.00	0.00	0.59	0.30	0.35	0.05	0.02	0.00	0.76	0.38
Neointermedol	0.00	0.00	0.03	0.27	0.21	0.41	0.79	0.74	0.00	0.01
Isoaromadendrene	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Farnesene epoxide,	0.00	0.00	0.47	0.00	0.00	0.02	0.00	0.00	0.00	0.00
Elemen	0.00	0.00	1.00	0.00	0.00	0.04	0.01	0.00	0.00	0.00
Epiglobulol	0.00	0.00	0.97	0.03	0.03	0.17	0.00	0.00	0.00	0.00
Corymbolone	NA	0.00	0.62	0.03	0.02	0.26	0.01	0.00	0.00	0.00
2-Methyl-4-2,6,6-t	0.00	NA	0.77	0.00	0.00	0.12	0.00	0.00	0.00	0.00
cis-p-mentha-1,7,8-	0.62	0.77	NA	0.00	0.00	0.00	0.00	0.00	0.00	0.00
cis-beta-	0.03	0.00	0.00	NA	0.00	0.00	0.00	0.00	0.00	0.00

Terpine ol										
Cyclohexanol, 1-me	0.02	0.00	0.00	0.00	NA	0.00	0.00	0.00	0.00	0.00
3- buten- 2-one, 4-5	0.26	0.12	0.00	0.00	0.00	NA	0.00	0.00	0.00	0.00
Camph ene	0.01	0.00	0.00	0.00	0.00	0.00	NA	0.00	0.00	0.00
Camph or	0.00	0.00	0.00	0.00	0.00	0.00	0.00	NA	0.00	0.00
Pinene	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	NA	0.00
Ocimen e	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	NA
Terpine ol	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Isoaro maden drene e	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Isoaro maden drene e	0.00	0.00	0.02	0.29	0.27	0.00	0.01	0.04	0.41	0.71
Cholest an-3-ol, 2-	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.15	0.27
Cubeno l	0.49	0.89	0.09	0.26	0.16	0.00	0.00	0.01	0.56	0.89
Aroma dendre ne oxid	0.02	0.07	0.01	0.00	0.00	0.00	0.09	0.12	0.22	0.35
Dihydr oartes misini	0.21	0.13	0.00	0.48	0.94	0.00	0.01	0.00	0.48	0.24
Artemi sinin	0.73	0.68	0.00	0.00	0.03	0.00	0.00	0.00	0.01	0.02
Phenol, 2- methox y-	0.07	0.17	0.00	0.32	0.21	0.00	0.00	0.00	0.47	0.58
beta- Guaien e	0.26	0.61	0.00	0.01	0.01	0.00	0.00	0.00	0.36	0.55
2H- Pyran,	0.08	0.12	0.00	0.00	0.02	0.00	0.04	0.00	0.39	0.73

2-7-hept										
Chiapin B	0.29	0.51	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02
beta-copaene	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.18
Deoxyartemisinin	0.01	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.29	0.64
Olean-12-en-3-one	0.65	0.49	0.75	0.66	0.54	0.61	0.96	0.12	0.60	0.51
Caryophyllene	0.48	0.30	0.13	0.09	0.28	0.15	0.16	0.17	0.01	0.03
Cyclohexane, 1-eth	0.17	0.24	0.42	0.83	0.69	0.81	0.40	0.17	0.97	0.70
Germaacrene D	0.07	0.17	0.00	0.31	0.40	0.22	0.60	0.46	0.79	0.86
ylangene	0.02	0.05	0.00	0.36	0.48	0.20	0.51	0.49	0.96	0.86
copaene	0.03	0.08	0.00	0.30	0.41	0.18	0.44	0.41	0.81	0.77
compound 1	0.00	0.00	0.11	0.18	0.25	0.02	0.01	0.39	0.01	0.01
Eucalyptol	0.00	0.00	0.23	0.02	0.05	0.06	0.05	0.80	0.00	0.00
Artemisinic acid	0.00	0.00	0.03	0.64	0.75	0.27	0.59	0.43	0.03	0.02
cis-Z-alpha-Bisabol	0.41	0.25	0.83	0.71	0.75	0.53	0.19	0.43	0.05	0.12
Isoaromadrene e	0.32	0.12	0.21	0.33	0.22	0.88	0.42	0.30	0.20	0.49
Heptacosane	0.17	0.25	0.05	0.00	0.01	0.13	0.86	0.66	0.06	0.34
Unknown alkane	0.60	0.78	0.16	0.00	0.00	0.03	0.44	0.98	0.01	0.05
endo-Borneol	0.19	0.44	0.00	0.00	0.00	0.03	0.06	0.51	0.07	0.16

Caryophyllene oxide	0.23	0.13	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00
Isoaromadrene	0.72	0.48	0.09	0.02	0.03	0.18	0.56	0.07	0.07	0.63
Aromadrene oxide	0.18	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.04
2H-Pyran, 2-7-hept	0.75	0.75	0.00	0.01	0.02	0.00	0.01	0.00	0.17	0.35
1-Heptatriacotanol	0.09	0.11	0.11	0.18	0.28	0.02	0.07	0.15	0.57	0.32

	Terpinol	Isoaromadrene	Isoaromadrene	Cholesterol, 2-	Cubene	Aromadrene oxide	Dihydroartemisinin	Artemisinin	Phenol, 2-methoxy-	beta-Guaiane
cis-beta-Farnesene	0.00	0.00	0.42	0.05	0.30	0.70	0.08	0.01	0.16	0.06
Cedrene-13-ol, 8-	0.00	0.65	0.19	0.76	0.38	0.24	0.15	0.00	0.25	0.33
Cymene	0.12	0.02	0.18	0.13	0.21	0.74	0.07	0.00	0.06	0.75
Terpinene	0.82	0.70	0.01	0.00	0.00	0.13	0.06	0.00	0.01	0.09
Butanoic acid, 2-m	0.78	0.10	0.84	0.09	0.10	0.05	0.00	0.00	0.00	0.00
Neointermedol	0.29	0.10	0.00	0.00	0.01	0.01	0.00	0.03	0.00	0.00
Isoaromadrene	0.00	0.00	0.39	0.47	0.13	0.35	0.32	0.49	0.30	0.23
Farnesene epoxide,	0.00	0.00	0.38	0.18	0.58	0.30	0.59	0.57	0.81	0.59
Elemen	0.03	0.01	0.00	0.00	0.13	0.07	0.13	0.18	0.09	0.07

Epiglobulol	0.15	0.02	0.00	0.00	0.67	0.00	0.87	0.29	0.25	0.64
Corymbolone	0.05	0.00	0.00	0.00	0.49	0.02	0.21	0.73	0.07	0.26
2-Methyl-4-2,6,6-t	0.00	0.00	0.00	0.01	0.89	0.07	0.13	0.68	0.17	0.61
cis-p-mentha-1,7,8-	0.00	0.00	0.02	0.00	0.09	0.01	0.00	0.00	0.00	0.00
cis-beta-Terpinol	0.00	0.00	0.29	0.00	0.26	0.00	0.48	0.00	0.32	0.01
Cyclohexanol, 1-me	0.00	0.00	0.27	0.00	0.16	0.00	0.94	0.03	0.21	0.01
3-buten-2-one, 4-5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Camphene	0.00	0.00	0.01	0.00	0.00	0.09	0.01	0.00	0.00	0.00
Camphor	0.00	0.00	0.04	0.00	0.01	0.12	0.00	0.00	0.00	0.00
Pinene	0.00	0.00	0.41	0.15	0.56	0.22	0.48	0.01	0.47	0.36
Ocimene	0.00	0.00	0.71	0.27	0.89	0.35	0.24	0.02	0.58	0.55
Terpinol	NA	0.00	0.00	0.00	0.00	0.00	0.17	0.69	0.05	0.00
Isoaromadrene e	0.00	NA	0.01	0.00	0.00	0.00	0.81	0.75	0.09	0.00
Isoaromadrene e	0.00	0.01	NA	0.00	0.00	0.00	0.12	0.34	0.00	0.00
Cholestan-3-ol, 2-	0.00	0.00	0.00	NA	0.00	0.00	0.44	0.25	0.00	0.00
Cubenol	0.00	0.00	0.00	0.00	NA	0.00	0.57	0.57	0.00	0.00
Aromadrene oxide	0.00	0.00	0.00	0.00	0.00	NA	0.17	0.33	0.01	0.00

Dihydroartemisinin	0.17	0.81	0.12	0.44	0.57	0.17	NA	0.00	0.02	0.00
Artemisinin	0.69	0.75	0.34	0.25	0.57	0.33	0.00	NA	0.00	0.00
Phenol, 2-methoxy-	0.05	0.09	0.00	0.00	0.00	0.01	0.02	0.00	NA	0.00
beta-Guaienone	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	NA
2H-Pyran, 2-7-hept	0.46	0.89	0.00	0.00	0.28	0.89	0.00	0.00	0.00	0.00
Chiapin B	0.03	0.05	0.01	0.00	0.16	0.07	0.00	0.00	0.00	0.00
beta-copaenone	0.02	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Deoxyartemisinin	0.26	0.48	0.00	0.00	0.10	0.04	0.01	0.00	0.00	0.00
Olean-12-en-3-one	0.03	0.00	0.44	0.09	0.12	0.05	0.96	0.00	0.12	0.29
Caryophyllene	0.60	0.29	0.31	0.93	0.40	0.32	0.02	0.07	0.32	0.82
Cyclohexane, 1-eth	0.50	0.17	0.56	0.08	0.32	0.43	0.21	0.10	0.25	0.60
Germaacrene D	0.60	0.18	0.13	0.00	0.21	0.36	0.02	0.48	0.11	0.01
ylangene	0.53	0.14	0.03	0.00	0.10	0.21	0.01	0.54	0.05	0.01
copaenone	0.51	0.16	0.04	0.00	0.13	0.27	0.01	0.45	0.07	0.01
compound 1	0.00	0.00	0.57	0.46	0.21	0.08	0.60	0.13	0.16	0.60
Eucalyptol	0.00	0.00	0.35	0.37	0.40	0.03	0.01	0.00	0.00	0.09
Artemisinic acid	0.01	0.00	0.10	0.02	0.65	0.19	0.34	0.02	0.00	0.04

cis-Z-alpha-Bisabo	0.56	0.20	0.04	0.16	0.00	0.42	0.84	0.56	0.01	0.01
Isoaromadrene e	0.01	0.16	1.00	0.65	0.01	0.01	0.00	0.17	0.30	0.83
Heptacosane	0.04	0.31	0.43	0.07	0.17	0.04	0.20	0.69	0.05	0.14
Unknown alkane	0.00	0.05	0.63	0.25	0.22	0.01	0.08	0.31	0.01	0.03
endo-Borneol	0.00	0.13	0.55	0.01	0.72	0.06	0.28	0.79	0.12	0.46
Caryophyllene oxide	0.00	0.00	0.46	0.02	0.12	0.03	0.44	0.66	0.24	0.19
Isoaromadrene e	0.10	0.26	0.81	0.14	0.98	0.19	0.54	0.20	0.43	0.24
Aromadrene oxide	0.00	0.00	0.00	0.00	0.13	0.02	0.02	0.02	0.02	0.01
2H-Pyran, 2-7-hept	0.20	0.44	0.01	0.02	0.93	0.56	0.00	0.02	0.06	0.05
1-Heptatriacotanol	0.28	0.02	0.08	0.40	0.66	0.92	0.05	0.63	0.60	0.19

	2H-Pyran, 2-7-hept	Chiapin B	beta-copaene	Deoxyartemisinin	Olean-12-en-3-one	Caryophyllene	Cyclohexane, 1-eth	Germacrene D	ylangene	copaene
cis-beta-Farnesene	0.15	0.00	0.03	0.13	0.63	0.00	0.02	0.00	0.00	0.00
Cedrene-13-ol, 8-	0.03	0.02	0.00	0.01	0.00	0.55	0.06	0.09	0.11	0.14
Cymene	0.51	0.23	0.77	0.75	0.73	0.00	0.49	0.64	0.48	0.63
Terpinen	0.74	0.39	1.00	0.81	0.03	0.07	0.42	0.04	0.02	0.03

Butanoic acid, 2-m	0.96	0.38	0.94	0.33	0.42	0.87	0.21	0.25	0.20	0.22
Neointermedol	0.00	0.00	0.00	0.00	0.36	0.59	0.33	0.01	0.00	0.00
Isoaromadrenene	0.73	0.15	0.96	0.48	0.96	0.93	0.03	0.11	0.05	0.06
Farnesene epoxide,	0.61	0.99	0.45	0.18	0.64	0.49	0.01	0.01	0.00	0.01
Elementol	0.77	0.52	0.33	0.65	0.06	0.68	0.00	0.00	0.00	0.00
Epiglobulol	0.23	0.81	0.07	0.13	0.79	0.16	0.13	0.16	0.06	0.09
Corymbolone	0.08	0.29	0.00	0.01	0.65	0.48	0.17	0.07	0.02	0.03
2-Methyl-4-2,6,6-trimethyl-1,8-cis-p-menthane-10-ol	0.12	0.51	0.01	0.01	0.49	0.30	0.24	0.17	0.05	0.08
1,8-cis-menthane-10-ol	0.00	0.00	0.00	0.00	0.75	0.13	0.42	0.00	0.00	0.00
1,8-cis-menthane-10-ol	0.00	0.00	0.00	0.00	0.66	0.09	0.83	0.31	0.36	0.30
Cyclohexanol, 1-methyl	0.02	0.00	0.00	0.00	0.54	0.28	0.69	0.40	0.48	0.41
3-buten-2-one, 4,5-dimethyl	0.00	0.00	0.00	0.00	0.61	0.15	0.81	0.22	0.20	0.18
Camphene	0.04	0.00	0.00	0.01	0.96	0.16	0.40	0.60	0.51	0.44
Camphor	0.00	0.00	0.00	0.00	0.12	0.17	0.17	0.46	0.49	0.41
Pinene	0.39	0.00	0.06	0.29	0.60	0.01	0.97	0.79	0.96	0.81
Ocimene	0.73	0.02	0.18	0.64	0.51	0.03	0.70	0.86	0.86	0.77
Terpineol	0.46	0.03	0.02	0.26	0.03	0.60	0.50	0.60	0.53	0.51

Isoaromadrenone	0.89	0.05	0.10	0.48	0.00	0.29	0.17	0.18	0.14	0.16
Isoaromadrenone	0.00	0.01	0.00	0.00	0.44	0.31	0.56	0.13	0.03	0.04
Cholestan-3-ol, 2-	0.00	0.00	0.00	0.00	0.09	0.93	0.08	0.00	0.00	0.00
Cubenol	0.28	0.16	0.00	0.10	0.12	0.40	0.32	0.21	0.10	0.13
Aromadrenone oxid	0.89	0.07	0.00	0.04	0.05	0.32	0.43	0.36	0.21	0.27
Dihydroartemisinin	0.00	0.00	0.00	0.01	0.96	0.02	0.21	0.02	0.01	0.01
Artemisinin	0.00	0.00	0.00	0.00	0.00	0.07	0.10	0.48	0.54	0.45
Phenol, 2-methoxy-	0.00	0.00	0.00	0.00	0.12	0.32	0.25	0.11	0.05	0.07
beta-Guaiane	0.00	0.00	0.00	0.00	0.29	0.82	0.60	0.01	0.01	0.01
2H-Pyran, 2-7-hept	NA	0.00	0.00	0.00	0.02	0.55	0.08	0.40	0.21	0.23
Chiapin B	0.00	NA	0.00	0.00	0.01	0.73	0.06	0.26	0.23	0.22
beta-copaene	0.00	0.00	NA	0.00	0.00	0.52	0.05	0.33	0.18	0.20
Deoxyartemisinin	0.00	0.00	0.00	NA	0.00	0.83	0.03	0.37	0.30	0.32
Olean-12-en-3-one	0.02	0.01	0.00	0.00	NA	0.03	0.00	0.00	0.00	0.00
Caryophyllene	0.55	0.73	0.52	0.83	0.03	NA	0.00	0.00	0.00	0.00
Cyclohexane, 1-eth	0.08	0.06	0.05	0.03	0.00	0.00	NA	0.00	0.00	0.00

Germa crene D	0.40	0.26	0.33	0.37	0.00	0.00	0.00	NA	0.00	0.00
ylange ne	0.21	0.23	0.18	0.30	0.00	0.00	0.00	0.00	NA	0.00
Copaen e	0.23	0.22	0.20	0.32	0.00	0.00	0.00	0.00	0.00	NA
compo und 1	0.00	0.01	0.00	0.00	0.01	0.08	0.11	0.74	0.72	0.82
Eucalyp tol	0.00	0.00	0.00	0.00	0.00	0.11	0.06	0.69	0.53	0.60
Artemi sinic acid	0.00	0.00	0.00	0.00	0.00	0.08	0.02	0.68	0.53	0.57
cis-Z- alpha- Bisabo	0.25	0.58	0.59	0.68	0.30	0.87	0.85	0.23	0.19	0.19
Isoaro maden drene e	0.00	0.29	0.83	0.90	0.07	0.27	0.04	0.01	0.00	0.00
Heptac osane	0.00	0.24	0.25	0.11	0.19	0.21	0.23	0.14	0.11	0.12
Unkown alkane	0.24	0.88	0.87	0.84	0.23	0.32	0.32	0.73	0.66	0.65
endo- Borneo l	0.05	0.05	0.05	0.01	0.23	0.71	0.90	0.27	0.21	0.21
Caryop hylene oxide	0.25	0.05	0.12	0.06	0.00	0.28	0.70	0.48	0.51	0.49
Isoaro maden drene e	0.26	0.21	0.19	0.20	0.34	0.33	0.65	0.46	0.50	0.48
Aroma dendre ne oxid	0.00	0.00	0.00	0.01	0.50	0.67	0.99	0.27	0.14	0.14
2H- Pyran, 2-7- hept	0.00	0.01	0.01	0.09	0.28	0.36	0.28	0.04	0.01	0.01
1- Heptatr iacotan ol	0.00	0.18	0.64	0.98	0.13	0.68	0.29	0.11	0.04	0.05

	comp ound 1	Eucalyp tol	Artemi sinic acid	cis-Z- alpha- Bisabo	Isoaro maden drene e	Hepta cosan e	Unkow n alkane	end o- Born eol	Cary ophyl ene oxide	Isoaro maden drene e
cis- beta- Farnes ene	0.00	0.01	0.05	0.52	0.68	0.11	0.00	0.06	0.01	0.09
Cedren -13-ol, 8-	0.01	0.33	0.30	0.05	0.03	0.06	0.00	0.46	0.84	0.12
Cymen e	0.00	0.00	0.00	0.00	0.49	0.47	0.36	0.14	0.51	0.07
Terpine n	0.03	0.05	0.01	0.00	0.35	0.97	0.54	0.02	0.35	0.16
Butanoi c acid, 2-m	0.03	0.81	0.05	0.43	0.94	0.00	0.00	0.00	0.00	0.86
Neoint ermede ol	0.00	0.00	0.00	0.47	0.07	0.67	0.47	0.08	0.62	0.43
Isoaro maden drene e	0.01	0.01	0.07	0.08	0.05	0.27	0.90	0.90	0.03	0.43
Farnes ene epoxid e,	0.00	0.00	0.03	0.45	0.19	0.29	0.83	0.82	0.06	0.46
Elemen	0.02	0.00	0.00	0.42	0.01	0.87	0.28	0.58	0.15	0.28
Epiglob ulol	0.06	0.01	0.02	0.46	0.23	0.04	0.13	0.67	0.08	0.97
Corym bolone	0.00	0.00	0.00	0.41	0.32	0.17	0.60	0.19	0.23	0.72
2- Methyl -4- 2,6,6-t	0.00	0.00	0.00	0.25	0.12	0.25	0.78	0.44	0.13	0.48
cis-p- mentha -17,8-	0.11	0.23	0.03	0.83	0.21	0.05	0.16	0.00	0.00	0.09
cis- beta- Terpine ol	0.18	0.02	0.64	0.71	0.33	0.00	0.00	0.00	0.00	0.02
Cycloh exanol, 1-me	0.25	0.05	0.75	0.75	0.22	0.01	0.00	0.00	0.00	0.03

3-buten-2-one, 4-5	0.02	0.06	0.27	0.53	0.88	0.13	0.03	0.03	0.00	0.18
Camphene	0.01	0.05	0.59	0.19	0.42	0.86	0.44	0.06	0.00	0.56
Camphor	0.39	0.80	0.43	0.43	0.30	0.66	0.98	0.51	0.09	0.07
Pinene	0.01	0.00	0.03	0.05	0.20	0.06	0.01	0.07	0.00	0.07
Ocimene	0.01	0.00	0.02	0.12	0.49	0.34	0.05	0.16	0.00	0.63
Terpineol	0.00	0.00	0.01	0.56	0.01	0.04	0.00	0.00	0.00	0.10
Isoaromadendrene	0.00	0.00	0.00	0.20	0.16	0.31	0.05	0.13	0.00	0.26
Isoaromadendrene	0.57	0.35	0.10	0.04	1.00	0.43	0.63	0.55	0.46	0.81
Cholestan-3-ol, 2-	0.46	0.37	0.02	0.16	0.65	0.07	0.25	0.01	0.02	0.14
Cubenol	0.21	0.40	0.65	0.00	0.01	0.17	0.22	0.72	0.12	0.98
Aromadendrene oxide	0.08	0.03	0.19	0.42	0.01	0.04	0.01	0.06	0.03	0.19
Dihydroartemisinin	0.60	0.01	0.34	0.84	0.00	0.20	0.08	0.28	0.44	0.54
Artemisinin	0.13	0.00	0.02	0.56	0.17	0.69	0.31	0.79	0.66	0.20
Phenol, 2-methoxy-	0.16	0.00	0.00	0.01	0.30	0.05	0.01	0.12	0.24	0.43
beta-Guaiane	0.60	0.09	0.04	0.01	0.83	0.14	0.03	0.46	0.19	0.24
2H-Pyran, 2-7-hept	0.00	0.00	0.00	0.25	0.00	0.00	0.24	0.05	0.25	0.26
Chiapin B	0.01	0.00	0.00	0.58	0.29	0.24	0.88	0.05	0.05	0.21

beta-copaene	0.00	0.00	0.00	0.59	0.83	0.25	0.87	0.05	0.12	0.19
Deoxyartemisinin	0.00	0.00	0.00	0.68	0.90	0.11	0.84	0.01	0.06	0.20
Olean-12-en-3-one	0.01	0.00	0.00	0.30	0.07	0.19	0.23	0.23	0.00	0.34
Caryophyllene	0.08	0.11	0.08	0.87	0.27	0.21	0.32	0.71	0.28	0.33
Cyclohexane, 1-eth	0.11	0.06	0.02	0.85	0.04	0.23	0.32	0.90	0.70	0.65
Geracrene D	0.74	0.69	0.68	0.23	0.01	0.14	0.73	0.27	0.48	0.46
ylangene	0.72	0.53	0.53	0.19	0.00	0.11	0.66	0.21	0.51	0.50
copaene	0.82	0.60	0.57	0.19	0.00	0.12	0.65	0.21	0.49	0.48
compound 1	NA	0.00	0.00	0.54	0.12	0.59	0.00	0.26	0.33	0.20
Eucalyptol	0.00	NA	0.00	0.16	0.01	0.14	0.00	0.48	0.00	0.26
Artemisinic acid	0.00	0.00	NA	0.07	0.25	0.86	0.01	0.19	0.26	0.47
cis-Z-alpha-Bisabol	0.54	0.16	0.07	NA	0.28	0.16	0.05	0.30	0.23	0.06
Isoaromadendrene	0.12	0.01	0.25	0.28	NA	0.01	0.58	0.21	0.43	0.11
Heptacosane	0.59	0.14	0.86	0.16	0.01	NA	0.00	0.00	0.00	0.26
Unknown alkane	0.00	0.00	0.01	0.05	0.58	0.00	NA	0.00	0.00	0.19
endo-Borneol	0.26	0.48	0.19	0.30	0.21	0.00	0.00	NA	0.00	0.91
Caryophyllene oxide	0.33	0.00	0.26	0.23	0.43	0.00	0.00	0.00	NA	0.90
Isoaromadendrene	0.20	0.26	0.47	0.06	0.11	0.26	0.19	0.91	0.90	NA

Aroma dendre ne oxid	0.02	0.32	0.56	0.44	0.00	0.00	0.00	0.19	0.02	0.07
2H- Pyran, 2-7- hept	0.83	0.34	0.39	0.36	0.00	0.00	0.02	0.23	0.53	0.82
1- Heptatr iacotan ol	0.17	0.61	0.63	0.08	0.00	0.00	0.07	0.35	0.11	0.92

	Aroma dendr ene oxid	2H- Pyran, 2-7- hept	1- Heptatr iacotan ol
cis- beta- Farnes ene	0.01	0.07	0.22
Cedren -13-ol, 8-	0.03	0.10	0.47
Cymen e	0.02	0.65	0.31
Terpine n	0.26	0.88	0.97
Butanoi c acid, 2-m	0.18	0.45	0.22
Neoint ermede ol	0.99	0.04	0.55
Isoaro maden drene e	0.02	0.77	0.12
Farnes ene epoxid e,	0.07	0.70	0.26
Elemen	0.15	0.27	0.76
Epiglob ulol	0.73	0.61	0.49
Corym bolone	0.18	0.75	0.09
2- Methyl	0.14	0.75	0.11

-4- 2,6,6-t			
cis-p- mentha -1,7,8-	0.00	0.00	0.11
cis- beta- Terpine ol	0.00	0.01	0.18
Cycloh exanol, 1-me	0.00	0.02	0.28
3- buten- 2-one, 4-5	0.00	0.00	0.02
Camph ene	0.00	0.01	0.07
Camph or	0.00	0.00	0.15
Pinene	0.01	0.17	0.57
Ocimen e	0.04	0.35	0.32
Terpine ol	0.00	0.20	0.28
Isoaro maden drene e	0.00	0.44	0.02
Isoaro maden drene e	0.00	0.01	0.08
Cholest an-3-ol, 2-	0.00	0.02	0.40
Cubeno l	0.13	0.93	0.66
Aroma dendre ne oxid	0.02	0.56	0.92
Dihydr oartes misini	0.02	0.00	0.05
Artemi sinin	0.02	0.02	0.63
Phenol, 2-	0.02	0.06	0.60

methoxy-			
beta-Guaiene	0.01	0.05	0.19
2H-Pyran, 2-7-hept	0.00	0.00	0.00
Chiapin B	0.00	0.01	0.18
beta-copaene	0.00	0.01	0.64
Deoxyartemisinin	0.01	0.09	0.98
Olean-12-en-3-one	0.50	0.28	0.13
Caryophyllene	0.67	0.36	0.68
Cyclohexane, 1-eth	0.99	0.28	0.29
Germa crene D	0.27	0.04	0.11
ylangene	0.14	0.01	0.04
copaene	0.14	0.01	0.05
compound 1	0.02	0.83	0.17
Eucalyptol	0.32	0.34	0.61
Artemisinic acid	0.56	0.39	0.63
cis-Z-alpha-Bisabolo	0.44	0.36	0.08
Isoaromadendrene	0.00	0.00	0.00
Heptacosane	0.00	0.00	0.00

Unkown alkane	0.00	0.02	0.07
endo-Borneol	0.19	0.23	0.35
Caryophyllene oxide	0.02	0.53	0.11
Isoaromadendrene	0.07	0.82	0.92
Aromadendrene oxide	NA	0.00	0.00
2H-Pyran, 2-7-hept	0.00	NA	0.00
1-Heptatriacotanol	0.00	0.00	NA

Metabolite yield

Genotype	ART	SD	SE	AA	SD	SE	DHAA	SD	SE
A	0.011408	0.001871	0.000764	0.000232	0.000277	0.000113	0.015210	0.001708	0.000697
B	0.008801	0.000550	0.000246	0.000244	0.000241	0.000108	0.014743	0.001116	0.000499
C	0.014655	0.001793	0.000732	0.000784	0.000259	0.000106	0.014268	0.000408	0.000167
D	0.008683	0.000786	0.000321	0.000804	0.000122	4.97E-05	0.020835	0.002108	0.000861
E	0.008449	0.001922	0.000784	0.000796	0.000438	0.000179	0.027917	0.005543	0.002263
F	0.004125	0.000595	0.000266	0.000306	0.000311	0.000139	0.011540	0.000528	0.000236
G	0.019329	0.001774	0.000724	0.001413	0.000274	0.000112	0.035164	0.008954	0.003655
H	0.004274	0.000337	0.000138	0.005298	0.001204	0.000491	0.007711	0.001886	0.000770
I	0.016653	0.001138	0.000465	0.000228	0.000190	7.75E-05	0.026225	0.003721	0.001519
J	0.024241	0.002094	0.000855	0.000829	0.000488	0.000199	0.070899	0.020968	0.008560

K	0.011603	0.001346	0.000550	0.000272	0.000147	6.00E-05	0.016682	0.001176	0.000480
L	0.011526	0.001012	0.000413	0.000270	0.000109	4.46E-05	0.010091	0.000650	0.000265

Genotype	Camphor	SD	SE	endo borneol	SD	SE
A	1.478932	0.125588	0.051271	1.608464	0.268925	0.109788
B	0.718140	0.032026	0.014323	0.009653	0.000928	0.000415
C	2.109206	0.119811	0.048913	0.064204	0.007464	0.003047
D	0.000107	0.000261	0.000107	0	0.000000	0.000000
E	0.000807	0.001189	0.000485	8.13E-05	0.000199	0.000081
F	0.000753	0.001226	0.000548	0	0.000000	0.000000
G	1.018489	0.094753	0.038683	0.006622	0.000815	0.000333
H	1.087971	0.121629	0.049655	0.024581	0.003990	0.001629
I	1.834355	0.189417	0.077329	0.028382	0.004165	0.001700
J	3.048455	0.350242	0.142986	0.033854	0.005285	0.002158
K	2.887858	0.239312	0.097699	0.02048	0.003225	0.001317
L	1.362339	0.079286	0.032368	0.064931	0.028079	0.011463

APPENDIX 3

mRNA Abundance

Geno type	ADH1			ADS			ALDH1			CPR		
	Mea n	Std Dev	Std Err	Mea n	Std Dev	Std Err	Mea n	Std Dev	Std Err	Mea n	Std Dev	Std Err
A	1.270 193	0.874 005	0.504 607	4.07 1988	5.45 2242	3.14 7853	1.18 4322	0.77 8017	0.44 9188	2.13 5441	1.00 1263	0.57 8080
B	1.072 725	0.290 407	0.167 666	2.39 4294	2.58 9325	1.49 4948	1.11 1210	0.75 0731	0.43 3435	1.79 2337	0.79 6184	0.45 9677
C	0.119 913	0.041 127	0.023 745	1.37 9114	0.53 0995	0.30 6570	0.14 5262	0.02 3645	0.01 3651	0.40 6942	0.09 8359	0.05 6788
D	0.392 802	0.113 033	0.079 926	2.88 4544	1.03 6341	0.73 2804	0.34 4682	0.06 9879	0.04 9412	0.69 4399	0.00 3065	0.00 2167
E	0.870 636	0.257 766	0.148 821	2.22 3465	0.94 5123	0.54 5667	0.65 1601	0.18 6515	0.10 7684	1.27 9041	0.34 6190	0.19 9873

F	0.664 496	1.026 651	0.592 737	1.77 0003	0.55 4850	0.32 0343	0.99 6560	0.86 8514	0.50 1437	1.39 7736	0.42 6498	0.24 6238
G	0.631 783	0.131 630	0.075 996	3.35 7824	0.72 6694	0.41 9557	0.58 2813	0.14 7484	0.08 5150	1.46 0386	0.44 6328	0.25 7687
H	0.691 213	0.243 268	0.140 451	1.73 7309	1.62 4557	0.93 7939	0.86 5466	0.54 3720	0.31 3917	0.63 6860	0.50 2680	0.29 0223
I	0.441 027	0.165 173	0.095 363	1.68 8239	0.57 0707	0.32 9498	0.36 0669	0.15 4876	0.08 9418	0.84 5266	0.50 8505	0.29 3585
J	0.473 180	0.268 743	0.155 159	2.28 9325	2.31 3696	1.33 5813	0.24 1337	0.12 1755	0.07 0296	0.61 4690	0.37 9554	0.21 9135
K	1.763 454	1.252 188	0.722 951	7.73 7536	3.43 7569	1.98 4681	1.47 6306	0.30 2513	0.17 4656	2.02 0741	0.47 2857	0.27 3004
L	0.582 225	0.184 705	0.106 640	2.52 5460	1.61 1174	0.93 0212	0.41 4653	0.13 4119	0.07 7433	1.27 7369	0.45 0779	0.26 0257

Genotype	CYP7AV1			DBR2			DXR1		
	Mean	Std Dev	Std Err	Mean	Std Dev	Std Err	Mean	Std Dev	Std Err
A	2.1031 83	1.4548 93	0.8399 83	2.7974 36	1.1611 64	0.6703 98	0.2455 50	0.2064 74	0.1192 08
B	1.4193 19	0.7516 39	0.4339 59	2.4090 80	1.1141 17	0.6432 36	0.1447 78	0.1261 08	0.0728 08
C	0.5675 62	0.1412 06	0.0815 26	0.2779 20	0.0371 55	0.0214 52	0.0663 32	0.0114 11	0.0065 88
D	1.3631 77	0.3317 29	0.2345 68	0.8053 06	0.1954 39	0.1381 96	0.1414 66	0.0068 98	0.0048 78
E	1.2162 60	0.3847 77	0.2221 51	1.5094 71	0.5913 27	0.3414 03	0.1547 43	0.1072 30	0.0619 09
F	0.8635 72	0.1973 83	0.1139 59	0.8040 85	0.5581 55	0.3222 51	0.1699 22	0.0125 99	0.0072 74
G	1.2468 98	0.2220 93	0.1282 25	1.2833 35	0.4983 02	0.2876 95	0.1837 24	0.0109 46	0.0063 19
H	0.7096 01	0.2679 70	0.1547 12	0.1010 44	0.0358 63	0.0207 05	0.1610 73	0.0328 13	0.0189 45
I	1.4406 69	0.9313 05	0.5376 89	0.8604 84	0.5814 44	0.3356 97	0.1392 45	0.0452 27	0.0261 12
J	1.0801 31	0.6964 85	0.4021 16	0.6830 17	0.2276 52	0.1314 35	0.0864 63	0.0687 86	0.0397 14
K	1.3797 94	0.2641 37	0.1525 00	2.9759 90	0.9376 28	0.5413 40	0.2324 94	0.1081 90	0.0624 64
L	1.4233 49	0.2752 98	0.1589 44	1.4704 69	0.3323 78	0.1918 98	0.1331 17	0.0286 97	0.0165 68

Genotype	DXS1			ERF1			FPS		
	Mean	Std Dev	Std Err	Mean	Std Dev	Std Err	Mean	Std Dev	Std Err

A	0.8464 56	1.0797 16	0.6233 74	0.4107 64	0.7050 37	0.4070 53	0.6050 01	0.4766 48	0.2751 93
B	0.5935 37	0.4128 68	0.2383 69	0.2996 24	0.4957 96	0.2862 48	0.6159 31	0.4694 93	0.2710 62
C	0.2972 09	0.1101 88	0.0636 17	0.0017 43	0.0016 45	0.0009 49	0.4804 68	0.0274 07	0.0158 23
D	0.2112 82	0.0061 87	0.0043 75	0.0119 69	0.0152 12	0.0107 57	0.7297 50	0.0673 67	0.0476 36
E	0.3708 08	0.1289 91	0.0744 73	0.1613 85	0.2680 43	0.1547 55	1.1576 60	0.3946 46	0.2278 49
F	0.8066 51	0.2584 69	0.1492 27	0.0585 66	0.0708 69	0.0409 16	1.2904 15	0.3746 02	0.2162 76
G	0.6329 42	0.1875 67	0.1082 92	0.0077 66	0.0018 39	0.0010 61	1.0359 00	0.7197 69	0.4155 59
H	0.4377 15	0.1216 41	0.0702 30	0.0475 43	0.0599 07	0.0345 88	0.5215 93	0.1055 73	0.0609 53
I	0.3803 95	0.1679 82	0.0969 84	0.0045 02	0.0044 78	0.0025 85	0.9474 88	0.5473 21	0.3159 96
J	0.1596 73	0.1722 34	0.0994 39	0.0023 24	0.0013 99	0.0008 08	0.6154 22	0.5464 62	0.3155 00
K	0.9647 35	0.4926 95	0.2844 58	0.0004 36	0.0003 33	0.0001 92	2.4399 18	1.3917 91	0.8035 51
L	0.2629 00	0.1038 15	0.0599 37	0.0542 75	0.0590 92	0.0341 17	1.4272 61	0.1984 28	0.1145 63

Genotype	HMGR			MYC2			ORA		
	Mean	Std Dev	Std Err	Mean	Std Dev	Std Err	Mean	Std Dev	Std Err
A	0.4914 07	0.4188 44	0.2418 20	0.5518 09	0.6319 55	0.3648 59	0.4728 50	0.6637 92	0.3832 40
B	0.3366 56	0.4717 46	0.3335 75	0.3284 95	0.4243 59	0.2450 04	0.4561 95	0.6880 99	0.3972 74
C	0.0896 02	0.0266 21	0.0153 69	0.0157 13	0.0046 29	0.0026 73	0.0165 61	0.0044 24	0.0025 54
D	0.9020 78	0.0636 00	0.0449 72	0.0121 10	0.0041 34	0.0029 23	0.1919 57	0.0355 54	0.0251 41
E	2.4054 40	3.4195 37	1.9742 71	0.1066 02	0.0099 67	0.0057 54	0.4493 32	0.5138 79	0.2966 88
F	1.0226 60	0.8539 52	0.4930 30	0.1043 13	0.0892 45	0.0515 25	0.1479 35	0.1211 22	0.0699 30
G	2.7774 41	2.2803 33	1.3165 51	0.0942 59	0.0306 52	0.0176 97	0.0638 92	0.0189 55	0.0109 44
H	0.2324 69	0.0448 44	0.0258 91	0.0208 30	0.0037 51	0.0021 66	0.0325 63	0.0121 74	0.0070 29
I	1.1813 44	0.8949 73	0.5167 13	0.0257 32	0.0060 67	0.0035 03	0.1978 77	0.1345 95	0.0777 09

J	0.5079 27	0.4586 64	0.2648 10	0.0147 17	0.0036 08	0.0020 83	0.0955 04	0.0959 56	0.0554 00
K	0.3312 09	0.1107 03	0.0639 14	0.0586 38	0.0352 42	0.0203 47	0.0803 29	0.0564 05	0.0325 65
L	0.8001 12	0.6575 44	0.3796 33	0.1153 30	0.0120 19	0.0069 39	0.1645 19	0.1195 97	0.0690 49

Genotype	WARKY1			PYL9		
	Mean	Std Dev	Std Err	Mean	Std Dev	Std Err
A	1.5695 54	2.5618 03	1.4790 58	0.1621 86	0.2593 07	0.1497 11
B	0.4409 38	0.7447 34	0.4299 72	0.1580 93	0.2202 98	0.1271 89
C	0.0028 99	0.0016 24	0.0009 38	0.0071 17	0.0048 59	0.0028 05
D	0.0010 83	0.0006 10	0.0004 32	0.0029 98	0.0001 70	0.0001 20
E	0.0185 11	0.0203 28	0.0117 36	0.0606 11	0.0407 87	0.0235 48
F	0.0112 75	0.0145 51	0.0084 01	0.0166 01	0.0155 74	0.0089 92
G	0.0064 41	0.0044 70	0.0025 81	0.0183 13	0.0142 51	0.0082 28
H	0.0015 47	0.0010 46	0.0006 04	0.0035 18	0.0033 53	0.0019 36
I	0.0016 06	0.0006 62	0.0003 82	0.0081 70	0.0055 00	0.0031 76
J	0.0032 21	0.0041 67	0.0024 06	0.0099 31	0.0053 17	0.0030 70
K	0.0124 61	0.0184 93	0.0106 77	0.0259 69	0.0142 26	0.0082 13
L	0.0107 89	0.0048 77	0.0028 16	0.0325 62	0.0037 96	0.0021 91

mRNA abundance graphs

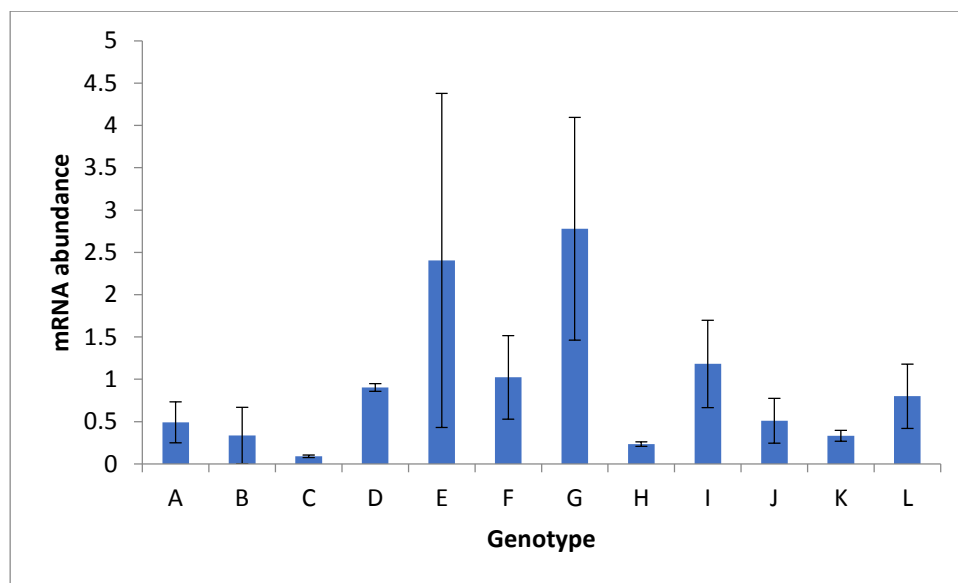


Figure 1: Relative mRNA abundance of HMGR gene within the 12 *A. annua* genotypes of the WVU germplasm collection. mRNA abundance was measured and averaged (n=3) with young, unexpanded leaf materials collect from plants grown for 45 days. Error bars indicate standard error and statistical analysis indicated none of the genotypes are significantly different (Tukey's test, $\alpha=0.05$).

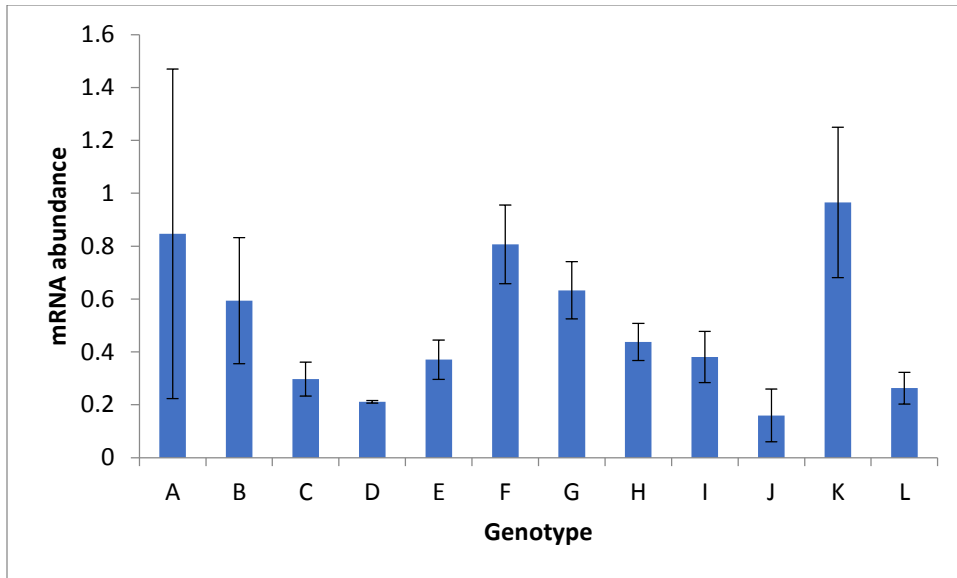


Figure 2: Relative mRNA abundance of DXS1 gene within the 12 *A. annua* genotypes of the WVU germplasm collection. mRNA abundance was measured and averaged (n=3) with young, unexpanded leaf materials collect from plants grown for 45 days. Error bars indicate standard error and statistical analysis indicated none of the genotypes are significantly different (Tukey's test, $\alpha=0.05$).

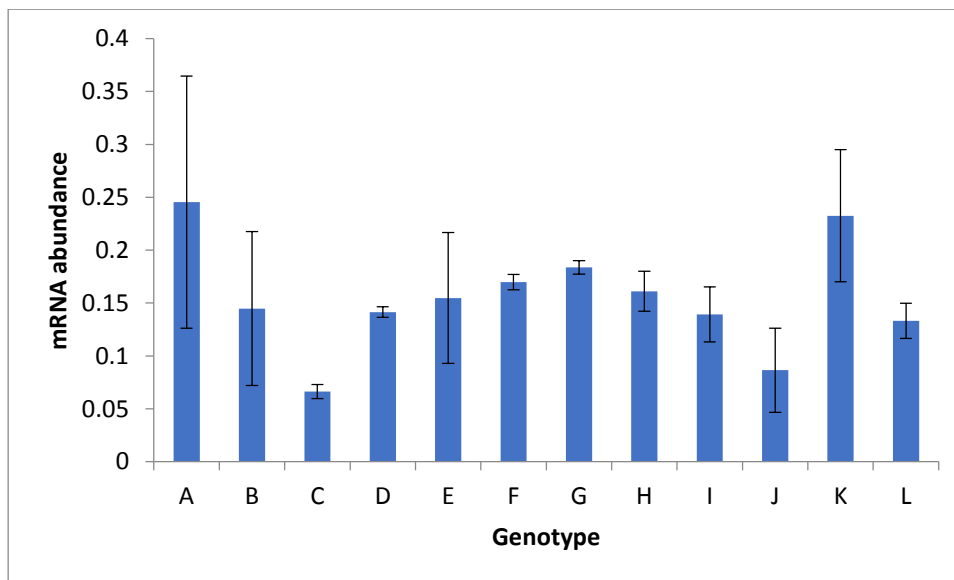


Figure 3: Relative mRNA abundance of DXR1 gene within the 12 *A. annua* genotypes of the WVU germplasm collection. mRNA abundance was measured and averaged (n=3) with young, unexpanded leaf materials collect from plants grown for 45 days. Error bars indicate standard error and statistical analysis indicated none of the genotypes are significantly different (Tukey's test, $\alpha=0.05$).

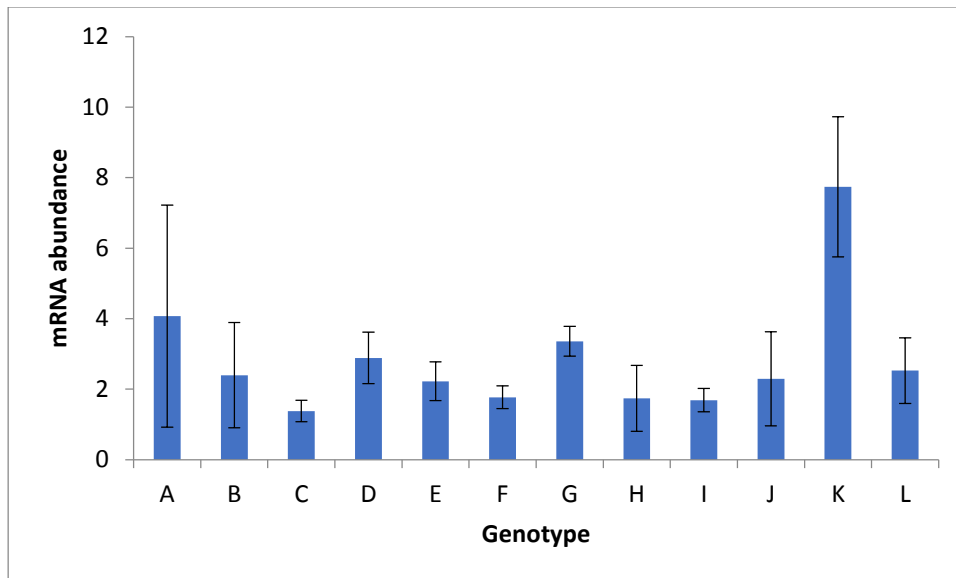


Figure 4: Relative mRNA abundance of ADS gene within the 12 *A. annua* genotypes of the WVU germplasm collection. mRNA abundance was measured and averaged (n=3) with young, unexpanded leaf materials collect from plants grown for 45 days. Error bars indicate standard error and statistical analysis indicated none of the genotypes are significantly different (Tukey's test, $\alpha=0.05$).

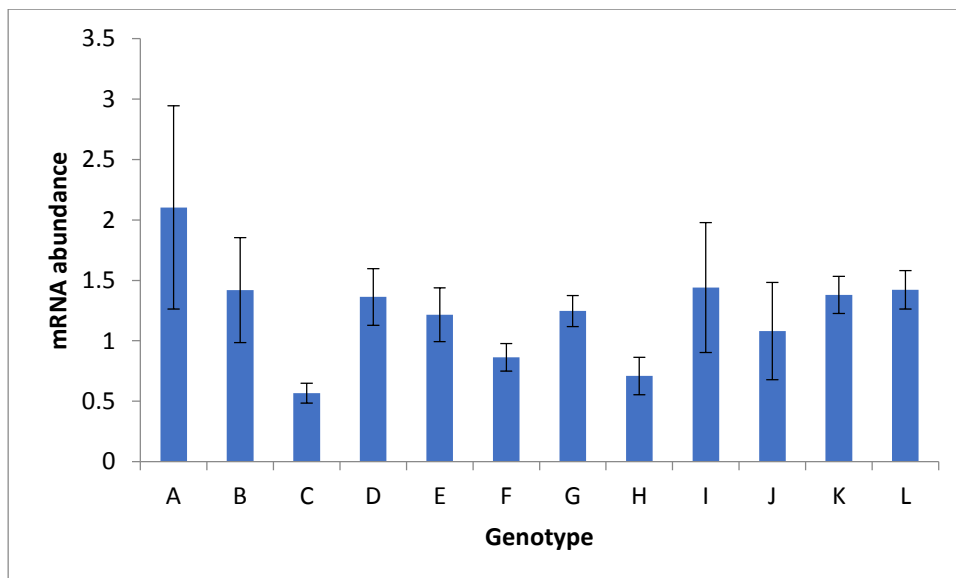


Figure 5: Relative mRNA abundance of CYP7AV1 gene within the 12 *A. annua* genotypes of the WVU germplasm collection. mRNA abundance was measured and averaged (n=3) with young, unexpanded leaf materials collect from plants grown for 45 days. Error bars indicate standard error and statistical analysis indicated none of the genotypes are significantly different (Tukey's test, $\alpha=0.05$).

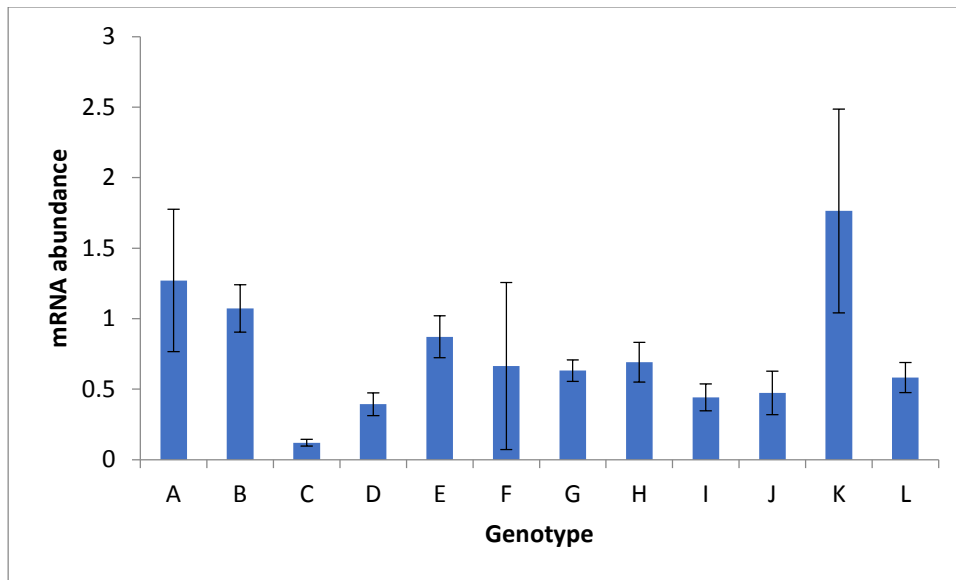


Figure 6: Relative mRNA abundance of ADH1 gene within the 12 *A. annua* genotypes of the WVU germplasm collection. mRNA abundance was measured and averaged (n=3) with young, unexpanded leaf materials collect from plants grown for 45 days. Error bars indicate standard error and statistical analysis indicated none of the genotypes are significantly different (Tukey’s test, $\alpha=0.05$).

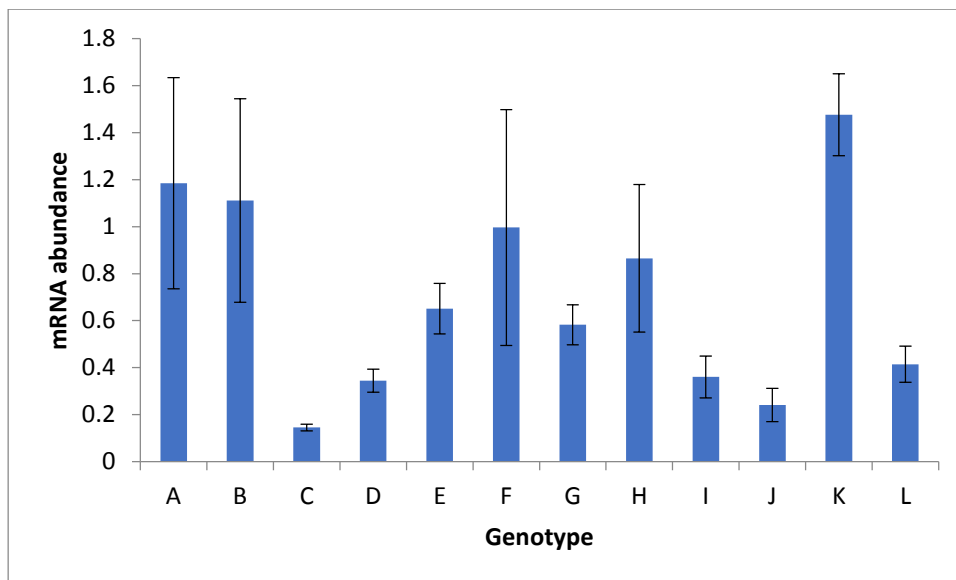


Figure 7: Relative mRNA abundance of ALDH1 gene within the 12 *A. annua* genotypes of the WVU germplasm collection. mRNA abundance was measured and averaged (n=3) with young, unexpanded leaf materials collect from plants grown for 45 days. Error bars indicate standard error and statistical analysis indicated none of the genotypes are significantly different (Tukey’s test, $\alpha=0.05$).

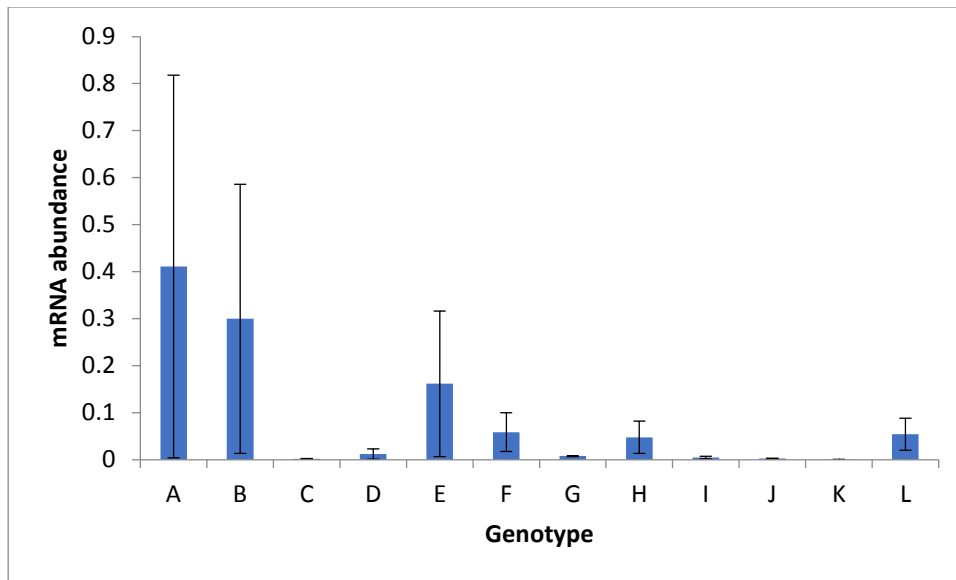


Figure 8: Relative mRNA abundance of ERF1 gene within the 12 *A. annua* genotypes of the WVU germplasm collection. mRNA abundance was measured and averaged (n=3) with young, unexpanded leaf materials collect from plants grown for 45 days. Error bars indicate standard error and statistical analysis indicated none of the genotypes are significantly different (Tukey's test, $\alpha=0.05$).

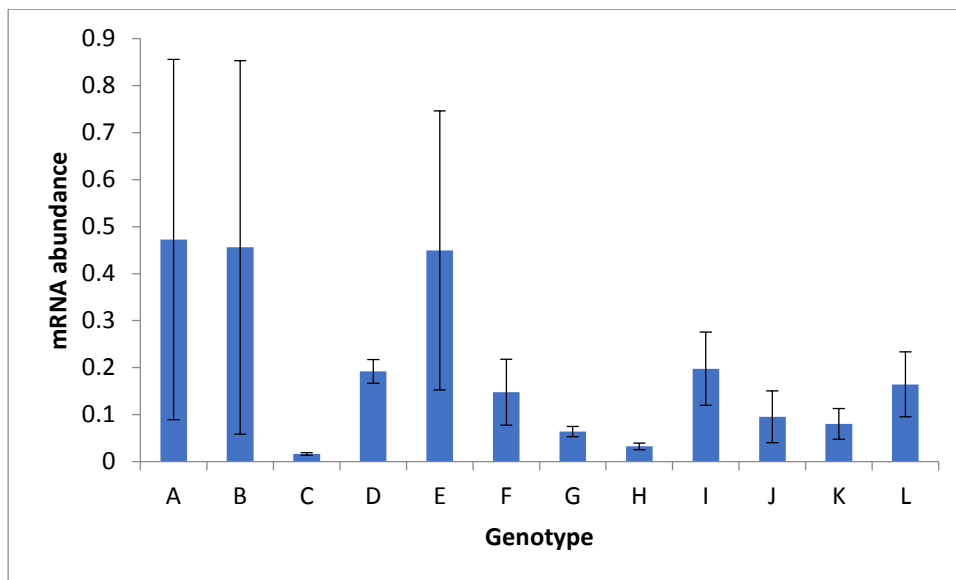


Figure 9: Relative mRNA abundance of ORA gene within the 12 *A. annua* genotypes of the WVU germplasm collection. mRNA abundance was measured and averaged (n=3) with young, unexpanded leaf materials collect from plants grown for 45 days. Error bars indicate standard error and statistical analysis indicated none of the genotypes are significantly different (Tukey's test, $\alpha=0.05$).

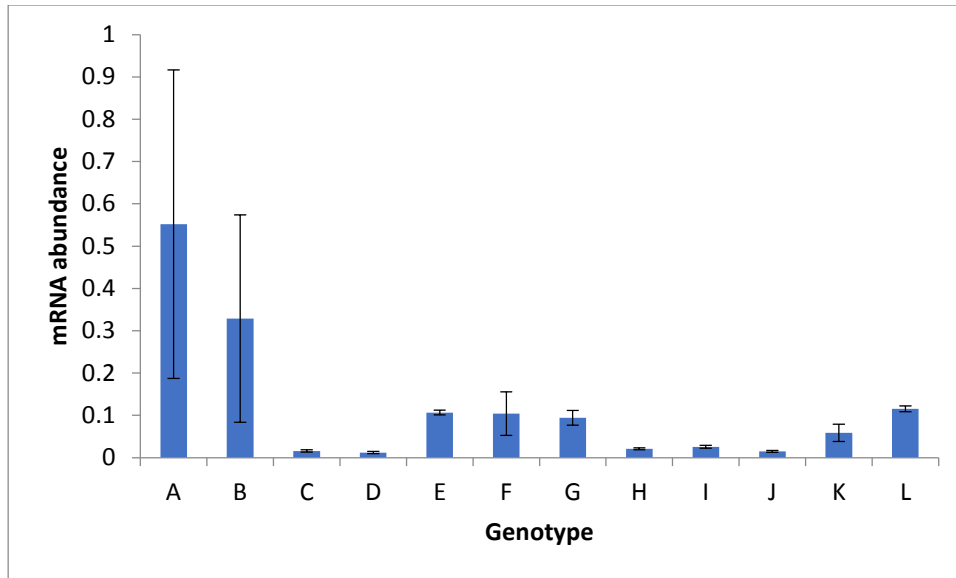


Figure 10: Relative mRNA abundance of MYC2 gene within the 12 *A. annua* genotypes of the WVU germplasm collection. mRNA abundance was measured and averaged (n=3) with young, unexpanded leaf materials collect from plants grown for 45 days. Error bars indicate standard error and statistical analysis indicated none of the genotypes are significantly different (Tukey's test, $\alpha=0.05$).

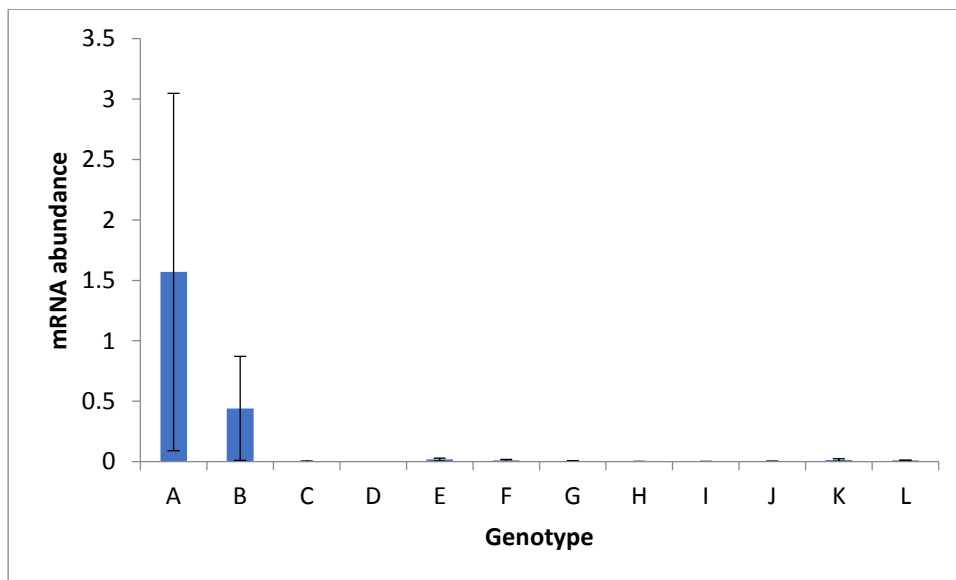


Figure 11: Relative mRNA abundance of WARKY1 gene within the 12 *A. annua* genotypes of the WVU germplasm collection. mRNA abundance was measured and averaged (n=3) with young, unexpanded leaf materials collect from plants grown for 45 days. Error bars indicate standard error and statistical analysis indicated none of the genotypes are significantly different (Tukey's test, $\alpha=0.05$).

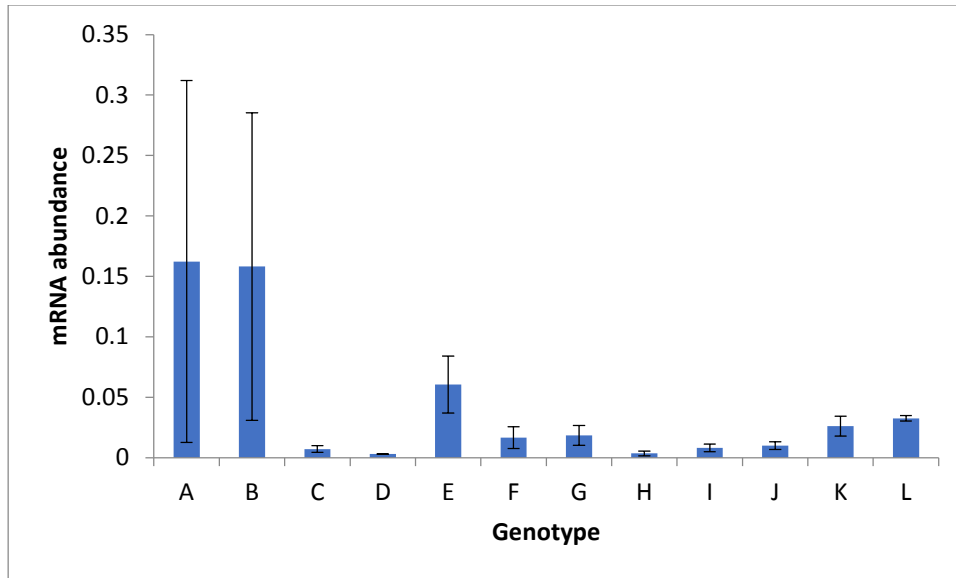


Figure 12: Relative mRNA abundance of PYL9 gene within the 12 *A. annua* genotypes of the WVU germplasm collection. mRNA abundance was measured and averaged (n=3) with young, unexpanded leaf materials collect from plants grown for 45 days. Error bars indicate standard error and statistical analysis indicated none of the genotypes are significantly different (Tukey's test, $\alpha=0.05$).