
Article:
Czechowski, Tomasz, Larson, TR, Catania, Theresa May orcid.org/0000-0002-9882-3878 et al. (5 more authors) (2018) Detailed phytochemical analysis of high- and low artemisinin-producing chemotypes of Artemisia annua. Frontiers in Plant Science. 641. pp. 1-14. ISSN 1664-462X

https://doi.org/10.3389/fpls.2018.00641

Detailed phytochemical analysis of high- and low artemisinin-producing chemotypes of Artemisia annua

Tomasz Czechowski[1], Tony R. Larson[1], Theresa Catania[1], David Harvey[1], Cenxi Wei[2], Michel Essome[2], Geoffrey D. Brown[2*], Ian A. Graham[1*]

[1]CNAP, Department of Biology, University of York, United Kingdom, [2]Department of Chemistry, University of Reading, United Kingdom

Submitted to Journal:
Frontiers in Plant Science

Specialty Section:
Plant Biotechnology

Article type:
Original Research Article

Manuscript ID:
360089

Received on:
02 Feb 2018

Revised on:
24 Apr 2018

Frontiers website link:
www.frontiersin.org

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

Author contribution statement

IAG planned and supervised the experiments and wrote the manuscript.

Keywords

Artemisia annua, Chemotype, Artemisinin, NMR, Sesquiterpenes, glandular trichomes

Abstract

Word count: 248

Chemical derivatives of aArtemisinin, a sesquiterpene lactone produced by Artemisia annua, is are the active ingredient in the most effective treatment for malaria. Comprehensive phytochemical analysis of two contrasting chemotypes of A. annua resulted in the characterisation of over 80 natural products by NMR, more than 20 of which are novel and described here for the first time. Analysis of high- and low-artemisinin producing (HAP and LAP) chemotypes of A. annua confirmed the latter to have a low level of DBR2 (artemisinic aldehyde Δ11 (13) reductase) gene expression. Here we show that the LAP chemotype accumulates high levels of artemisinic acid, arteannuin B, epi-deoxyarteannuin B and other amorpha-4,11-diene derived sesquiterpenes which are unsaturated at the 11,13-position. By contrast, the HAP chemotype is rich in sesquiterpenes saturated at the 11,13-position (dihydroartemisinic acid, artemisinin and dihydro-epi-deoxyarteannunin B), which is consistent with higher expression levels of DBR2, and also with the presence of a HAP-chemotype version of CYP71AV1 (amorpha-4,11-diene C-12 oxidase). Our results indicate that the conversion steps from artemisinic acid to arteannuin B, epi-deoxyarteannuin B and artemisitene in the LAP chemotype are non-enzymatic and parallel the non-enzymatic conversion of DHAA to artemisinin and dihyro-epi-deoxyarteannuin B in the HAP chemotype. Interestingly, artemisinic acid in the LAP chemotype preferentially converts to arteannuin B rather than the endoperoxide bridge containing artemisitene. In contrast, in the HAP chemotype, DHAA preferentially converts to artemisinin. Broader metabolomic and transcriptomic profiling revealed significantly different terpenoid profiles and related terpenoid gene expression in these two morphologically distinct chemotypes.

Funding statement

We acknowledge financial support for this project from The Bill and Melinda Gates Foundation as well as from The Garfield WestonFoundation for the Centre for Novel Agricultural Products This work was also supported by the Biotechnology and Biological Sciences Research Council Grant BB/G008744/1 (to GDB), “The Biosynthesis of Artemisinin”.

Ethics statements

(Authors are required to state the ethical considerations of their study in the manuscript, including for cases where the study was exempt from ethical approval procedures)

Does the study presented in the manuscript involve human or animal subjects: No

Detailed phytochemical analysis of high- and low artemisinin-producing chemotypes of *Artemisia annua*

Tomasz Czechowski[1], Tony R. Larson[1], Theresa M. Catania[1], David Harvey[1], Cenxi Wei[2], Michel Essome[2], Geoffrey D. Brown[2*] and Ian A. Graham[1*]

[1]Centre for Novel Agricultural Products, Department of Biology, University of York, Heslington, York YO10 5DD, United Kingdom

[2]Department of Chemistry, University of Reading, Reading RG6 6AD, United Kingdom

***Correspondence:** Geoffrey D. Brown: g.d.brown@reading.ac.uk and Ian A. Graham: ian.graham@york.ac.uk

Keywords: Artemisia annua, chemotype, artemisinin, NMR, sesquiterpenes, glandular trichomes

Abstract

Chemical derivatives of a~~A~~rtemisinin, a sesquiterpene lactone produced by *Artemisia annua,* ~~is~~ are the active ingredient in the most effective treatment for malaria. Comprehensive phytochemical analysis of two contrasting chemotypes of *A. annua* resulted in the characterisation of over 80 natural products by NMR, more than 20 of which are novel and described here for the first time. Analysis of high- and low-artemisinin producing (HAP and LAP) chemotypes of *A. annua* confirmed the latter to have a low level of *DBR2* (artemisinic aldehyde $\Delta^{11\,(13)}$ reductase) gene expression. Here we show that the LAP chemotype accumulates high levels of artemisinic acid, arteannuin B, *epi*-deoxyarteannuin B and other amorpha-4,11-diene derived sesquiterpenes which are unsaturated at the 11,13-position. By contrast, the HAP chemotype is rich in sesquiterpenes saturated at the 11,13-position (dihydroartemisinic acid, artemisinin and dihydro-*epi*-deoxyarteannunin B), which is consistent with higher expression levels of *DBR2*, and also with the presence of a HAP-chemotype version of CYP71AV1 (amorpha-4,11-diene C-12 oxidase). Our results indicate that the conversion steps from artemisinic acid to arteannuin B, *epi*-deoxyarteannuin B and artemisitene in the LAP chemotype are non-enzymatic and parallel the non-enzymatic conversion of DHAA to artemisinin and dihyro-*epi*-deoxyarteannuin B in the HAP chemotype. Interestingly, artemisinic acid in the LAP

chemotype preferentially converts to arteannuin B rather than the endoperoxide bridge containing artemisitene. In contrast, in the HAP chemotype, DHAA preferentially converts to artemisinin. Broader metabolomic and transcriptomic profiling revealed significantly different terpenoid profiles and related terpenoid gene expression in these two morphologically distinct chemotypes.

1 Introduction

~~The~~ Chemical derivatives of the sesquiterpene lactone, artemisinin, such as: artesunate, artemether or dihydroartemisinin ~~is~~ are ~~the~~ one of several active ingredients in artemisinin-combination therapies (ACTs) - the most effective treatment for malaria currently available. Biosynthesis of artemisinin occurs in specialized 10-celled biseriate glandular trichomes present on the leaves, stems and inflorescences of *Artemisia annua* (Duke and Paul, 1993; Duke et al., 1994; Ferreira and Janick, 1995). Concentrations of artemisinin can range from 0.01% to 1.4% of leaf dry weight (Larson et al., 2013). The biosynthetic pathway from artemisinin precursors has been fully elucidated over the past decade (Figure 1 C). It starts from the cyclization of farnesyl pyrophosphate (FPP) to amorpha-4,11-diene (A-4,11-D) by amorph-4,11-diene synthase (AMS) (Bouwmeester et al., 1999; Mercke et al., 2000) followed by the three-step oxidation of A-4,11-D by amorpha-4,11-diene C-12 oxidase (CYP71AV1), to artemisinic alcohol (AAOH), artemisinic aldehyde (AAA), and artemisinic acid (AA) (Ro et al., 2006; Teoh et al., 2006). ADH1 – NAD-dependent alcohol dehydrogenase with specificity towards artemisinic alcohol plays a role in the formation of artemisinic aldehyde in the artemisinin pathway of *A. annua* (Paddon et al., 2013). The *ADH1* gene has been used to improve yields of artemisinic acid production in yeast. (Paddon et al., 2013). Artemisinic aldehyde Δ11(13) reductase (DBR2) catalyzes the formation of dihydroartemisinic aldehyde (DHAAA) from AAA (Zhang et al., 2008). DHAAA is subsequently oxidized in the final enzymatic reaction to dihydroartemisinic acid (DHAA) by ~~alcohol~~ aldehyde dehydrogenase ALDH1 (Teoh et al., 2009; ~~Zhang et al., 2009~~). ~~(Paddon et al., 2013)~~Genes encoding all of these biosynthetic enzymes have been shown to be highly expressed in apical and sub-apical cells of *A. annua* glandular trichomes (Olsson et al., 2009; Soetaert et al., 2013). Recent studies have revealed that the conversion of DHAA to artemisinin and dihydro-*epi*-deoxyarteannuin B (DHEDB) proceeds *via* a series of non-enzymatic and spontaneous photochemical reactions, involving the highly reactive tertiary allylic

hydroperoxide of dihydroartemisinic acid, DHAAOOH (Wallaart et al., 1999; Sy and Brown, 2002; Brown and Sy, 2004). Similarly, it has previously been proposed that AA is photochemically converted to arteannuin B (ArtB) *via* the tertiary allylic hydroperoxide of artemisinic acid (Brown and Sy, 2007).

Based on the content of artemisinin and its precursors, two contrasting chemotypes of *A. annua* have been described: a low-artemisinin production (LAP) chemotype and a high-artemisinin production (HAP) chemotype (Wallaart et al., 2000). Both chemotypes contain artemisinin, but the HAP chemotype has a relatively high content of DHAA and artemisinin, whereas the LAP chemotype has a high content of AA and ArtB (Lomenn *et al.,* 2006; Aresenault *et al.,* 2006, (Larson et al., 2013). Recent studies have concluded that a major factor in determining the biochemical phenotype of HAPs and LAPs is the differential expression of *DBR2* - with low expression in LAP chemotypes correlating with a number of insertions/deletions in the *DBR2* promoter sequence (Yang et al., 2015). We have recently shown that the overall pathway to artemisinin biosynthesis is under strict developmental control with early steps in the pathway occurring in young leaves and later steps in mature leaves (Czechowski et al., 2016). In the present study, we have used both metabolomics and transcriptomics to investigate the developmental regulation of sesquiterpene biosynthesis in HAP and LAP chemotypes. Using a combination of NMR and UPLC-/GC-MS techniques we have characterised a number of amorphane and cadinane sesquiterpenes in addition to other terpenes isolated from leaf glandular trichomes. We have also extended the transcript analysis in HAPs and LAPs beyond the genes encoding artemisinin-pathway enzymes. Our findings suggest profound differences in general terpenoid metabolism between HAP and LAP chemotypes that extend well beyond altered *DBR2* expression and artemisinin content.

2. Materials and methods

2.1 Plant material

Artemis is an F1 hybrid variety of *A. annua* developed by Mediplant (Conthey, Switzerland), produced by crossing C4 and C1 parental material of East Asian origin (Delabays et al., 2001). Artemisinin content has been reported to reach 1.4% of the leaf dry weight when grown in the field, and its metabolite profile is typical for the HAP chemotype (Larson *et al.*, 2013). NCV

88 ("non-commercial variety"), an "open-pollinated" variety of European origin was also provided
89 by Mediplant, and has the lowest reported artemisinin content from any *A. annua* germplasm in
90 addition to a metabolite profile characteristic of the LAP chemotype (Larson *et al*., 2013). Plants
91 were grown from seeds in glasshouse conditions as previously described (Graham et al., 2010).

92 **2.2 Leaf area measurements**

93 The leaf area of glasshouse-grown plants was measured by scanning for leaves 14-16 (counting
94 from the apical meristem), followed by calculation of the leaf area using LAMINA software
95 (Bylesjo et al., 2008)

96 **2.3 Trichome density measurements**

97 Trichome density was quantified on the abaxial surface of the terminal leaflets of leaves 14-16
98 (counting from the apical meristem). Trichomes were visualised using a Zeiss fluorescent
99 dissecting microscope (fitted with a 470/40nm excitation filter/ 525/50nm emission filter).
100 Images were recorded using AxioVision 4.7 software (Carl Zeiss Ltd. Herts., UK). Trichome
101 number was counted manually across a 3 x 0.5 mm^2 leaflet sample area and the average (mean)
102 trichome density was then calculated for the whole leaf.

103 **2.4 NMR structural data for natural compounds from** Artemis **and NCV.**

104 Leaf and stem material from Artemis (5 Kg) was extracted in $CHCl_3$ (20 L). The organic solvent
105 was removed by rotary evaporation and a portion of the residual dark green aromatic plant
106 extract (*ca* 2.5% w/w) was "dry-loaded" on to a silica column for gradient column
107 chromatography (see Table 2.4.1).

108 **2.4.1 Gradient Column Chromatography of the Artemis variety of *A. annua***

Solvent	Fraction
10% EtOAc/hexane	A, B*, C and D*
20% EtOAc/hexane	E, F, G, H, I* and J
30% EtOAc/hexane	K, L, M, N and O*
50% EtOAc/hexane	P and Q
EtOAc	R, S, T*, U and V
Methanol	W, X and Y

109

Each of the fractions A-Y from gradient column chromatography of Artemis were then further purified by isocratic preparative normal-phase HPLC (*fractions B, D, I, O and T were also subjected to a second round of isocratic column chromatography prior to prep. HPLC); and individual metabolites were then characterized by NMR, as listed in Fig. 1A and the Supplemental Table 1 (1D- and 2D-NMR data for all metabolites which are novel as natural products is also given in the Supplementary Section). Selected fractions were analysed by UPLC-APCI-high resolution MS to verify molecular weights and chemical formulae. Confirmed annotations were used to update m/z and retention time reference data, to enable reporting of these compounds from plant extracts by UPLC-MS.

Leaf and stem material from the NCV variety of *A. annua* (780 g) was extracted in $CHCl_3$ (4 L). The organic solvent was then removed by rotary evaporation and the residual dark green aromatic plant extract (16.6 g; *ca* 2% w/w) was dry-loaded onto a silica column for gradient column chromatography (see Table 2.4.2).

2.4.2 Gradient Column Chromatography of the NCV variety of *A. annua*

Solvent	Fraction
2% EtOAc/hexane	A, B and C
10% EtOAc/hexane	D, E, F and G
20% EtOAc/hexane	G, H and I
40% EtOAc/hexane	J, K and L
EtOAc	M and N
Methanol	N

Each of the fractions A-N from gradient column chromatography of NCV were then further purified by isocratic preparative normal-phase HPLC; individual metabolites were then characterized by NMR, as listed in Fig. 1B and the Supplemental Table 1 (1D- and 2D-NMR data for all metabolites which are novel as natural products are also given in the Supplementary Section). Selected fractions were analysed by UPLC-APCI-high resolution MS to verify molecular weights and chemical formulae. Confirmed annotations were used to update m/z and retention time reference data, to enable reporting of these compounds from plant extracts by UPLC-MS.

2.5 Metabolite analysis by UPLC-MS and GC-MS

Metabolite analysis by UPLC- and GC-MS were performed as described previously (Czechowski *et al.,* 2016). Fifteen plants from each of five genotype classes were grown from seeds in 4-inch pots under 16 h days for 12 weeks. Metabolite profiles were generated from 50 mg fresh weight (FW) pooled samples of leaves collected at two different developmental stages: 1-5 (counted from the apical meristem), representing the juvenile stage; and leaves 11-13, representing the mature, expanded stage (Figure 1A). Fresh leaf samples were stored at -80°C, pending analysis. In addition, dry leaf material was also obtained from 14-week old plants, cut just above the zone of senescing leaves, and dried for 14 days at 40°C. Leaves were stripped from the plants, and leaf material sieved through 5 mm mesh to remove small stems. Trichome-specific metabolites were extracted as described previously (Czechowski *et al.*, 2016) with minor modifications. Briefly, 50 mg of fresh material was extracted by gentle shaking in 500 µl chloroform for 1 h. Supernatant was taken out and remaining plant material was fully dried in a centrifugal evaporator (GeneVac® Ez-2 plus, Genevac Ltd, Ipswich, UK). Weight of the extracted and dried material was taken and used to quantify abundance of the specific compounds per unit of extracted dry weight. Dry leaf material (0.5 g) was ground to a fine powder using a TissueLyser II ball mill fitted with stainless steel grinding jars (Qiagen, Crawley, UK) operated at 25 Hz for 1 min. Ten mg sub-samples of dry leaf material were extracted in 9:1 (v/v) chloroform:ethanol with gentle shaking for 1 h and then analysed as per fresh material.

For UPLC-MS analysis of sesquiterpenes, a diluted (1:5 (v/v) extract:ethanol) 2 µL aliquot was injected on an Acquity UPLC system (Waters, Elstree, UK) fitted with a Luna 50 × 2 mm 2.5 µm HST column (Phenomenex, Macclesfield, UK). Metabolites were eluted at 0.6 mL/min and 60°C using a linear gradient from 60% to 100% A:B over 2.5 min, where A = 5% (v/v) aqueous MeOH and B = MeOH, with both A and B containing 0.1% (v/v) formic acid. Pseudomolecular $[M+H]^+$ ions were detected using a Thermo Fisher LTQ-Orbitrap (ThermoFisher, Hemel Hempstead, UK) mass spectrometer fitted with an atmospheric pressure chemical ionization source operating in positive ionisation mode under the control of Xcalibur 2.1 software. Data was acquired over the m/z range 100 - 1000 in FTMS centroid mode with resolution set to 7500 FWHM at m/z 400. Data extraction and analysis was performed using packages and custom scripts in R 3.2.2 (https://www.R-project.org/). XCMS (Smith et al., 2006) incorporating the centWave algorithm (Tautenhahn et al., 2008) was used for untargeted peak extraction. Deisotoping, fragment and adduct removal was performed using CAMERA (Kuhl et al., 2012).

Artemisinin was quantified using the standard curve of the response ratio of artemisinin (Sigma, Poole, UK) to internal standard (β-artemether; Hallochem Pharmaceutical, Hong Kong) that was previously added to extracts and standards. Metabolites were identified with reference to authentic standards or NMR-resolved structures and empirical mass formulae calculated using the R package rcdk (Guha, 2007) within 10 ppm error and elemental constraints of: C = 1–100, H = 1-200, O = 0–20, N = 0 – 1. Peak concentrations were calculated using bracketed response curves, where standard curves were run every ~30 samples. Metabolite concentrations were expressed as a proportion of the residual dry leaf material following extraction.

For analysis of monoterpenes and volatile sesquiterpenes from fresh leaf samples, an aliquot of chloroform extract (prior to dilution with ethanol for UPLC analysis) was taken for GC-MS analysis using an Agilent 6890 GC interfaced to a Leco Pegasus IV TOF MS (Leco, Stockport, UK). A 1 µL aliquot was injected into a CIS4 injector (Gerstel, Mülheim an der Ruhr, Germany) fitted with a 2 mm ID glass liner containing deactivated glass wool at 10°C. The injector was ramped from 10°C to 300°C at 12°C/s then held at 300°C for 5 min. The carrier gas was He at a constant flow of 1 mL/min and the injection split ratio was 1:10. Peaks were eluted using a Restek Rxi-5Sil MS column, 30 m x 0.25 mm ID x 0.25 µm film thickness (Thames Restek, Saunderton, UK). The following temperature gradient was used: isothermal 40°C 2 min; ramp at 20°C/min to 320°C then hold for 1 min; total run time ~20 min. The transfer line was maintained at 250°C and the MS used to collect -70 eV EI scans over the m/z range 20-450 at a scan rate of 20 spectra/second. Acquisition was controlled by ChromaTof 4.5 software (Leco). ChromaTof was used to identify peaks and deconvolute spectra from each run, assuming a peak width of 3 s and a minimum *s/n* of 10. Peak areas were reported as deconvoluted total ion traces (DTIC). Further analyses including annotation against authentic standards, between-sample peak alignment, grouping, consensus DTIC reporting, and missing value imputation were performed using custom scripts in R.

R was used for all statistical data analysis using the stats base package, nlme (http://CRAN.R-project.org/package=nlme) and pcaMethods (Stacklies et al., 2007)

2.6 RNA isolation, cDNA synthesis and quantitative RT-PCR

Leaf tissue from juvenile and mature-stage leaves sampled as described above was ground to a fine powder using Qiagen Retsch MM300 TissueLyser (Qiagen, Hilden, Germany) and total RNA extracted using the RNAeasy kit (Qiagen, Hilden, Germany). RNA was quantified using NanoDrop-1000 (NanoDrop products, Wilmington, USA) and integrity was checked on 2200 Tape Station Instrument (Agilent, Santa Clara, CA, USA). Only samples scoring RIN number ≥7.0 were taken for further analysis. Removal of genomic DNA was performed by treating with TURBO DNA-free™ (Life Technologies Ltd, Paisley, UK) following manufacturer's instructions. 5 ug of total RNA, pooled from 4 individual plants, representing 3 biological replicates, was reversely transcribed using SuperScript II kit (Life Technologies Ltd, Paisley, UK) and Oligo(dT)12-18 Primer (Life Technologies Ltd, Paisley , UK) according to manufacturer's instructions. PCR using primers (AMS_Ex4 for 5'-GGCTGTCTCTGCACCTCCTC-3', AMS_Ex5 for 5'- CAGCCATCAATAACGGCCTTG -3') designed spanning intron 4 of the *AMS* gene (GenBank: AF327527). Only samples that resulted in amplification of the 251bp fragment from cDNA and not the 363 bp fragment from genomic DNA were taken for further qPCR analysis.

Expression levels of amorpha-4,11-diene synthase (AMS), amorpha-4,11-diene C-12 oxidase (CYP71AV1), cytochrome P450 reductase (CPR), artemisinic aldehyde Δ 11 (13) reductase (DBR2) and aldehyde dehydrogenase (ALDH1), relative to ubiquitin (UBI) were determined by qPCR. Reactions were run in 3 technical replicates. Gene-specific primers used were: AMS for 5'- GGGAGATCAGTTTCTCATCTATGAA- 3'; AMS_Rev 5'-CTTTTAGTAGTTGCCGCACTTCTT-3'; 5'ALDH1 for 5'- GATGTGTGTGGCAGGGTCTC-3'; ALDH1_Rev 5- ACGAGTGGCGAGATCAAAAG-3'; CYP71AV1 for 5'-TCAACTGGAAACTCCCCAATG-3'; CYP71AV1_Rev 5'- CGGTCATGTCGATCTGGTCA-3'; CPR_For 5'- GCTCGGAACAGCCATCTTATTCTT-3', CPR_Rev 5'-GAAGCCTTCTGAGTCATCTTGTGT-3', DBR2 for 5'- GAACGGACGAATATGGTGGG-3'; DBR2_Rev 5'- GCAGTATGAATTTGCAGCGGT-3', UBI for 5'-TGATTGGCGTCGTCTTCGA-3' and UBI_Rev 5'-CCCATCCTCCATTTCTAGCTCAT-3'. Reactions condtions and qPCR analysis were peformed as above, 1 ul of 1/20 first strand cDNA dilutiuon was used instead of genomic DNA. Background subtraction, average PCR efficiency for each amplicon and N0 values were calculated using LinRegPCR ver. 2012 software (Ruijter

et al 2009). Expression levels for each sample and gene of interest (GOI) were represented as N0 GOI / N0 UBI.

3 Results

3.1 NMR spectroscopic analysis uncovers novel metabolites in both HAP and LAP chemotypes

The natural products found in *A. annua* have previously been grouped into eight broad categories, including: i) monoterpenes; ii) sesquiterpenes; iii) diterpenes, iv) sterols and triterpenes; v) aliphatic hydrocarbons, alcohols, aldehydes and acids; vi) aromatic alcohols, ketones and acids; vi) phenylpropanoids; and viii) flavonoids (Brown, 2010). In the present work we have used the Artemis variety of *A. annua* as a representative of the HAP chemotype and NCV as a representative of the LAP chemotype (Larson et al., 2013). Our initial investigations using NMR analysis of leaf extracts of Artemis resulted in the isolation of 41 metabolites (6 of which were novel) representing all eight classes of natural products (Figure 1A, Supplemental List 1). The structures of all compoudns were determined by 1D- and 2D- NMR spectroscopy (detailed NMR data in Supplementary Section). Novel compounds which have not been isolated before as natural products include four new 11,13-dihydroamorphanes: 5β-hydroperoxy-eudesma-4(15),11-diene **(4),** 11-Hydroxy-arteannuin I **(18),** 6a-Hydroxy-arteannuin J **(19)** , Arteannuin P **(20);** the ketal form of arteannuin Q, **(26)** and abeo-Amorphane sesquiterpene **(27)**. Artemisinin **(22)** was the most abundant metabolite in this analysis (Figure 2, Supplemental List 1, and Supplemental Table 1); but the Artemis extract also contained two other sesquiterpenes: dihydroartemisinic acid (DHAA, **8**), and dihydro-*epi*-deoxyarteannuin B (DHEDB, **12**) in substantial amounts (Figure 2, Supplemental List 1 and Supplemental Table 1). In addition, a further nine known 11,13-dihydroamorphanes (α-epoxy-dihydroartemisinic acid **(10)**; 4α,5α-epoxy-6α-hydroxyamorphan-12-oic acid **(11)**; dihydroarteannuin B **(14)**; arteannuin M **(15)**; arteannuins H, I and J (**21, 16** and **17**); deoxyartemisinin **(23)**; and a 4,5-*seco*-4,5-diketo-amorphan-12-oic acid **(24)** (see Figure 1A, Supplemental Figure 1 and Supplemental Table 1) were also isolated as minor components from the Artemis leaf extracts (Figure 2, Supplemental List 1 and Supplemental Table 1)

Phytochemical investigation of the NCV variety by NMR yielded 57 metabolites, 20 of which were novel (Figure 1B Supplemental List 2), representing 7 of the 8 categories above. Novel metabolites from NCV variety are depicted on Figure 1B and include: *(E)*-7-Hydroxy-2,7-dimethylocta-2,5-dien-4-one (**43**), (*E*)-7-Hydroperoxy-2,7-dimethylocta-2,5-dien-4-on (**44**), 6,7-Epoxy-6,7-dihydro-β-farnesene (**45**), 6-Hydroxy-γ-humulene (**48**), 7α-Hydroxy-artemisinic acid (**52**), Arteannuin R (**54**), Arteannuin S (**55**), 4α, 5α-Epoxy-6α-hydroxyartemisinic acid methyl ester (**57**), Dehydroarteannuin L (**59**), *epi*-11-Hydroxy-arteannuin I (**64**), Artemisinic acid, 6α-peroxy ester (**65**), Deoxyartemistene (**67**) (novel as a natural product[1]), Arteannuin T (**69**), Arteannuin U (**70**), Arteannuin V (**72**), Arteannuin W (**73**), Arteannuin Y (**74**), Isoarteannuin A (**77**), Arteannuin Z (**78**) and 3-(2-(2,5-Dihydrofuran-3-yl)ethyl)-2,2-dimethyl-4-methylenecyclohexan-1-one (**79**).

As might have been expected, the most striking difference between the NCV and *Artemis* varieties was the almost complete absence of artemisinin, DHAA (**8**) and DHEDB (**12**) in the former (Supplemental Table 1). The NCV variety did, however, have relatively high levels of three 11-13-unsaturated amorphanes, which were found only as minor components in the *Artemis* variety, namely: artemisinic acid (AA, **9**), arteannuin B (ArtB, **60**) and *epi*-deoxyarteannuin B (EDB, **13**) (Figure 2, Supplemental Figure 2 and Supplemental Table 1). All the other amorphane sesquiterpenes isolated and characterized from the NCV variety by NMR shared this same trait: *i.e.* possession of an 11,13-unsaturated methylene group (Figure 1B, Figure 2 and Supplemental Table 1), and there is an almost complete absence of 11,13-dihydro-amorphanes from NCV, that contrasts with the abundance of these compounds in the *Artemis* variety (Supplemental List 2 and Supplemental Table 1). It is interesting to note that there are ten examples where 11,13-dihdyro/ 11,13-dehydro amorphanolides seem to occur as "pairs" between *Artemis* and NCV depicted on Figure 2. These include: DHAA **(8) /** AA **(9)**; artemisinin **(22) /** artemisitene **(65)**; dihydro-*epi*-deoxyarteannuin B **(12) /** *epi*-deoxyarteannuin B **(13)**; α-epoxy-dihydroartemisinic acid **(10)** / α-epoxy-artemisinic acid **(56)**; dihydroarteannuin B **(14) /** arteannuin B **(60);** arteannuin M **(15) /** dehydroarteannuin M **(61)**; arteannuin I **(16) /** annulide **(62)**; arteannuin J **(17) /** isoannulide **(63)**; deoxyartemsinin **(23) /** deoxyartemsitene **(67)**; and 4,5-*seco*-4,5-diketo-amorphan-12-oic acid **(24)** and its 11,13-dehydro-analogue **(68)**. It is also noteworthy that 9 of the 20 novel amorphane and *seco*-amorphane sesquiterpenes isolated and characterized from the NCV variety by NMR, possess an 11, 13-unsaturated methylene group (Figure 1B, Supplemental List 2)

All the above results are consistent with a higher DBR2 activity in the HAP chemotype compared to LAP chemotype (Yang *et al*., 2015). The relative abundances for 8 of these 10 "pairs" are also well matched between the *Artemis* and NCV varieties, suggesting a "shared" further metabolism for DHAA in *Artemis* and AA in NCV. The first exception is arteannuin B (ArtB **60**), which is abundant in NCV, whilst its analogue, dihydroarteannuin B (**14**), is relatively low in *Artemis*. (Supplemental Table 1). The second is artemisitene, the 11,13-dehydro analogue of artemisinin (Acton *et al.,* 1985; (Woerdenbag *et al*., 1994) (Fig 1; Supplemental Table 1) which is a minor compound in NCV, while its 'partner' artemisinin is the most abundant metabolite in *Artemis* (Supplemental Table 1). These observations suggest that while there are many parallels in the pathways that further transform DHAA (**8**) and AA (**9**) in the HAP and LAP chemotypes there are some significant differences.

3.2 Metabolomic and gene expression studies reveal multiple differences between HAP and LAP chemotypes.

Using a leaf maturation time-series, we recently demonstrated that artemisinin levels increase gradually from juvenile to mature leaves and remain stable during the post-harvest drying process in Artemis HAP chemotype plants (Czechowski *et al.,* 2016). Using a similar time-series (which included fresh leaf 1–5 (juvenile), and 11–13 (mature) (counting from the apical meristem); plus oven-dried whole plant-stripped leaves (dry) from 12-week-old glasshouse-grown plants), we have now performed UPLC- and GC-MS based metabolite profiling of extracts from both HAP (Artemis*)* and LAP (NCV) chemotypes. We found that the pathway entry–point metabolite, amorpha-4,11-diene (A-4,11-D), is only detectable in juvenile leaves, and at approximately 2-fold higher concentration in Artemis as compared to NCV (Figure 3Ai; Supplemental Table 3). A much greater difference was seen for the enzymatically-produced artemisinin precursor, dihydroartemisinic acid (DHAA), which was present at a 24-fold higher concentration in juvenile Artemis leaves compared to NCV (Figure 3A ii), Supplemental Table 2). Artemisinic acid (AA) on the other hand accumulated in NCV leaves at a 10-fold higher concentration than in Artemis (Figure 3A iii), Supplemental Table 1). Interestingly the levels of AA in the young leaves of NCV variety are approximately twice the levels of DHAA in young leaves of Artemis (Figure 3A ii) and iii), Supplemental Table 2). The levels of both DHAA and AA dropped sharply beyond the juvenile leaf stage in Artemis and NCV, respectively (Figure 3A

ii) and iii), Supplemental Table 2). These changes in metabolite levels occur during leaf maturation are mirrored by changes in steady state mRNA levels of genes encoding the enzymes involved in their biosynthesis including: amorpha-4,11-diene synthase (AMS), amorpha-4,11-diene C-12 oxidase (CYP71AV1), artemisinic aldehyde $\Delta^{11,(13)}$ reductase (DBR2) and aldehyde dehydrogenase (ALDH1) which are expressed at levels two to three orders of magnitude higher in juvenile than in mature leaves (Figure 3B i), ii), iii), iv)).

Previous work has suggested that *in vivo* conversions beyond DHAA (**8**) (Czechowski *et al.*, 2016) and *in vitro* conversions beyond AA (**9**) (Brown and Sy, 2007) are non-enzymatic. Consistent with this, we have found that mature leaves of NCV contain high levels of *epi*-deoxyarteannuin B (EDB, **13**) and arteannuin B (ArtB, **60**) and (Figure 3A v) and vii), Supplemental Table 2), while Artemis accumulates dihydro-*epi*-deoxyarteannuin B (DHEDB, **12**) and artemisinin (**22**) (Figure 3A iv), vi) Supplemental Table 2) at 20 to 30-fold higher levels than NCV. Both artemisinin (**22**) and arteannuin B (**60**) continue to accumulate in the post-harvest drying process in Artemis and NCV respectively (Figure 3A vi) and vii)). Post-harvest accumulation of artemisinin has been reported before (Ferreira and Luthria, 2010) ~~I~~and it might be related to light-dependent conversion of DHAA. However slightly different batch specific environmental effects during drying might explain difference between artemisinin accumulation pattern shown on Figure 3A vi) and what was previously reported for the Artemis variety (Czechowski et al., 2016). Interestingly, the developmental pattern of DHEDB (**12**) accumulation in Artemis leaves is different to its 11,13-dehydro analog, EDB (**13**) in NCV leaves. DHEDB (**12**) follws the same accumulation pattern as for artemsinin (**22**) in Artemis (Figure 3A iv) and vi)); whereas EDB (**13**) is found predominantly in juvenile leaves of the NCV variety (Figure 3A v)). We have found that production of the artemisinin 11,13-dehydro analog, artemisitene (**66**) in NCV parallels the accumulation of artemisinin (**22**) in Artemis (Supplemental Table 2), albeit at very much reduced levels. The levels of deoxyartemisinin (**23**), another product of non-enzymatic conversion of DHAA through the DHAA allylic hydroperoxide, increase during dry leaf storage, accumulating to 0.1% leaf dry weight (Supplemental Table 2), which is consistent with previous findings (Czechowski *et al.*, 2016). This process is paralleled by accumulation of deoxyartemisitene (**67**) (the 11,13-dehydro analog of deoxyartemisinin) in the NCV variety (Supplemental Table 2).

RT-qPCR analysis confirmed the expression level for *DBR2* to be significantly repressed (8-fold lower) in the juvenile leaves of NCV compared to Artemis, which is consistent with previous findings (Yang et al. 2015). Interestingly, *DBR2* transcript abundance had decreased to the same levels in mature leaves of both chemotypes (Figure 3B iii)), highlighting the importance of developmental timing in regulating flux and partitioning of sesquiterpene metabolites. More surprisingly, *ALDH1* expression is increased in juvenile leaves (2.4-fold) and further increased in mature leaves (40-fold) of NCV (Figure 3B iv)) compared to Artemis. Thus it would appear that in addition to *DBR2* being down-regulated in the NCV (LAP) chemotype, *ALDH1* is up-regulated at the transcriptional level. This could also account for the increase in flux into artemisinic acid and the artemisitene branch of sesquiterpene metabolism. The major differences in metabolite levels and gene expression between Artemis and NCV varieties for the artemisinin biosynthetic pathway are summarized in Figure 3C.

NMR analysis revealed that metabolite differences between Artemis and NCV are not restricted to artemisinin-related sesquiterpenes. Monoterpenes also vary between the two chemotypes, with for example camphor being most abundant in Artemis while artemisia ketone level is much more abundant in NCV (Supplemental Table 1). Unfortunately, NMR-analysis could only provide approximate information about the relative abundance of the metabolites, therefore metabolite content of both chemotypes was also studied by GC- and UPLC-MS (Supplemental Tables 2 and 3). We were able to detect 75 unique compounds in three leaf types by UPLC-MS of which annotations were assigned to 30 compounds based on NMR-verified standards as described in the Materials and Methods. The majority of the known compounds were sesquiterpenes and flavonoids. GC-MS detected 202 unique compounds in juvenile and mature leaves, of which 33 had assigned annotations. The majority of known GC-MS-detected compounds were mono- and sesquiterpenes. Using principal component analysis, it can be seen that the overall metabolite profile of NCV appears strikingly different to that of Artemis; as much as the difference between the profiles between juvenile leaves and mature- and/or dry leaves. In fact, UPLC- and GC-MS PCA plots show four distinct clusters (Figure 4A and B). Developmental differences are most apparent in juvenile leaf tissue, which show the highest abundance of most of the terpenes described below (Figure 4, Supplemental Tables 2 and 3). Our findings that the metabolite profiles in Artemis and NCV young leaf tissues are considerably different to mature and dry leaves in both varieties are consistent with our previous findings (Czechowski et al., 2016).

There are a number of compounds specifically produced by NCV, mostly in low quantities (Supplemental Tables 2 and 3) which have known medicinal use including, for example, isofraxidin **(39),** which is five-fold more abundant in the juvenile leaves of NCV as compared to Artemis (Supplemental Table 2). Isofraxidin is a coumarin with anti-inflammatory (Niu *et al.*, 2012) and anti-tumour activities (Yamazaki and Tokiwa, 2010). Artemisia ketone **(42)**, an irregular monoterpene found in the essential oil from various *A. annua* varieties displaying antifungal activites (Santomauro et al., 2016) is the most abundant volatile in the juvenile and mature leaves of NCV, but virtually absent in Artemis (Supplemental Table 3). The juvenile and mature leaves of Artemis accumulate velleral, a sesquiterpene dialdehyde which has proposed antibacterial activities (Anke and Sterner, 1991), which is virtually absent in NCV variety (Supplemental Table 3). GC-MS analysis further revealed that several major montoerpenes are also more abundant in juvenile and mature leaves of Artemis, including camphor (3.7-fold higher), camphene (3.4-fold higher), borneol, (16-fold higher), α-pinene (4.6-fold higher) and 1,8-cineole (8-fold higher) (Supplemental Table 3). Some minor monoterpenes detected in the Artemis variety, such as: α-myrcene, α –terpinene, chrysanthenone and α-copaene, are virtually absent in young and mature NCV leaves (Supplemental Table 3). A few striking differences were noted for the level of artemisinin-unrelated abundant sesquiterpenes, such as sabinene and *cis*-sabinene hydrate, which are 7.5-fold and 38-fold (respectively) more abundant in Artemis young leaves than in NCV (Supplemental Table 3). Germacrene A is a sesquiterpene common across the Asteraceae family for which it has been demonstrated that its downstream metabolism parallels artemisinic acid biosynthetic pathway (Nguyen et al., 2010) . Germacrene A levels are 32- and 17-fold higher in NCV young and mature leaves (respectively) making it the most abundant volatile in mature and the second most abundant in young leaves of the NCV variety.

Visualisation of the loadings from the multivariate analyses were used to identify the most influential compounds discriminating chemotypes. PC1 loading plots identified 18 compounds from UPLC- and 20 from GC-MS analysis (Supplemental Figure 1), which were used to create the heatmaps presented in Figure 5. The vast majority of the most influential compounds distinguishing between two chemotypes from UPLC-MS analysis were the amorphane sesquiterpenes (Figure 5A). The mono- and sesquiterpenes mentioned above (together with some unknown compounds) were the most influential GC-MS-detectable metabolites distinguishing between two chemotypes (Figure 5B).

3.3 Morphological difference between two chemotypes of *A. annua*

In addition to having very distinct phytochemical compositions the F1 *Artemis* HAP chemotype and the open pollinated NCV LAP chemotype varieties also have very distinct morphological features (Figure 6). Most strikingly, NCV is much taller with longer internodes but produces smaller leaves than *Artemis*. The density of glandular secretory trichomes, the site of artemisinin synthesis, is similar for both varieties (Figure 6 E), which is consistent with the main difference in artemisinin production is due to an alteration in metabolism rather than trichome density. *A. annua* varieties typically require short day length for flowering (Wetzstein et al., 2014), but we observed that NCV, unlike *Artemis,* can also flower under long days. However, the two chemotypes do cross-pollinate and produce viable progeny.

4 Discussion

This manuscript presents the first detailed phytochemical comparison of high- (HAP) and low-artemisinin production (LAP) chemotypes of *A. annua*.

26 of the 85 metabolites that have been characterized by NMR from the HAP and LAP varieties of *A. annua* in this study are novel as natural products (all are mono- and sesquiterpenes). And of these, 19 are amorphane sesquiterpenes, which is the most diverse and the most abundant sub-class (Supplementary Table 1, Supplementary Lists 1 and 2). The majority of these amorphane sesquiterpenes are highly oxygenated with structures that would be consistent with further oxidative metabolism of DHAA (11,13-saturated, **8**) in the HAP variety and AA (11,13-unsaturated, **9**) in the LAP variety (Figure 1, Figure 2, Supplementary Table 1, Supplemental Lists 1 and 2)

UPLC- and GC-MS analysis of leaf developmental series also revealed amorphanes either saturated or unsaturated at the 11,13-position in the HAP and LAP chemotypes, respectively (Figure 3, Supplemental Table 2). This observation is consistent with the expression of the *DBR*2 gene, which encodes the enzyme responsible for reducing the 11,13-double bond of artemisinic aldehyde (the precursor for 11,13-dihydroamorphane/cadinane sesquiterpenes) being strongly down-regulated in juvenile leaves of NCV (Figure 3B iii). These findings are in complete agreement with the recent report on reduced levels of *DBR2* in LAP compared with HAP chemotypes (Yang *et al.,* 2015). In addition to altered expression of *DBR2*, we also found that

expression of *aldehyde dehydrogenase* (*ALDH1*), which converts artemsinic and dihydroartemsinic aldehydes to acids (Teoh et al., 2009), is significantly elevated in juvenile and mature leaves of NCV compared to Artemis. This may lead to an increased flux from A-4,11-D to AA in NCV when compared with flux from A-4,11-D to DHAA in Artemis which is reflected by significantly higher concentration of AA found in juvenile leafs of NCV when compared to concentration of DHAA in Artemis young leaves (Figure 3A iii) and ii). ~~could enable increased flux from artemisinic aldehyde to artemisinic acid which we have observed in young leaves of NCV variety (Figure 3B iv).~~ The elevated flux from A-4,11-D to AA might also explain lower levels of A-4,11-D found in juvenile leaves of NCV when compared with Artemis (Figure 3A i) as the expression of *amorpha-4,11-diene synthase (AMS)* is at very similar level in both varieties (Figure 3B i). We have also observed that the NCV (LAP) variety expresses a sequence variant of amorpha-4,11-diene C-12 oxidase (*CYP71AV1*) with a 7 amino acid N-extension (Supplemental Figure 2). This LAP-chemotype associated ~~is~~ sequence ~~variant was previously shown~~ variant upon transient expression in *Nicotiana benthamiana*, in combination with the other artemisinin pathway genes resulted ~~(Ting et al., 2013a)~~in a qualitatively different product profile ('chemotype'); that is ~~, in~~ a shift in the ratio between the unsaturated and saturated (dihydro) branch of the pathway (Ting et al., 2013~~a~~). That result strongly suggests the two distinct isoforms of CYP71AV1 are associated with HAP- and LAP-branches of the artemisinin pathway ~~also~~ in *Artemsia annua* (Figure 3 C). ~~to be more efficient in the conversion of amorpha-4,11-diene to artemisinic acid (AA) than to artemisinic aldehyde (AAA) (Ting et al., 2013b). Thus in NCV we find decreased expression of *DBR2*, increased expression of *ALDH1* and the presence of a sequence variant of *CYP71AV1* that favours conversion of AAA to AA. We propose that these factors (Ting et al., 2013a)together lead to an increased flux from A-4,11-D to AA in NCV when compared with flux from A-4,11-D to DHAA in Artemis which is reflected by significantly higher concentration of AA found in juvenile leafs of NCV when compared to concentration of DHAA in Artemis young leaves (Figure 3A iii) and ii). The elevated flux from A-4,11-D to AA might also explain lower levels of A-4,11-D found in juvenile leaves of NCV when compared with Artemis (Figure 3A i) as the expression of *amorpha-4,11-diene synthase (AMS)* is at very similar level in both varieties (Figure 3B i).~~ A number of previous reports described the existence of LAP- and HAP-chemotypes of *A. annua* arising from distinct geographical locations (Lommen et al., 2006); (Arsenault et al., 2010), (Larson et al., 2013). It

would be interesting to establish if sequence variant forms of *CYP71AV1* and differential expression of *DBR2* are generally found between these other LAP- and HAP-chemotypes.

Recent attempts to constitutively overexpress *DBR2* in transgenic *A. annua* resulted in doubling of the artemisinin concentration, which was also accompanied by a significant increase in DHAA and AA production (Yuan et al., 2015). Improvements in artemisinin concentration obtained in these experiments by Yuan *et al.* were significantly better than those achieved by constitutive co-expression of *CYP71AV1* and *CPR* (Shen et al., 2012), where the LAP-sequence variant of *CYP71AV1* was overexpressed in transgenic *A. annua*. Our results suggest the glandular trichome-targeted overexpression of *DBR2* specifically in the HAP-type of *CYP71AV1* might be the more efficient route to improving artemisinin production in transgenic *A. annua*

Although arteannuin B (ArtB) was almost entirely absent from young leaf tissue of the NCV variety, as leaves matured it accumulated to become the most abundant natural product (Figure 3A vii). This first observation seemed to parallel both the accumulation of artemisinin in the mature tissues of Artemis that has been noted above (Figure 3vi), as well as the recently described accumulation of arteannuin X in the mature leaves of the *cyp71av1-1* mutant of *A. annua* (Czechowski et al., 2016). The accumulation of both artemisinin and arteannuin X are considered to be the result of non-enzymatic processes, in which the 4,5-double bond of a precursor sesquiterpene undergoes spontaneous autoxidation with molecular oxygen to produce a tertiary allylic hydroperoxide. The metabolic fate of this hydroperoxide is critically dependant on the identity of the precursor – and in particular on the functionality contained elsewhere in the molecule. Thus, in the case of Artemis, the precursor is DHAA which presents a 12-carboxylic acid group (as well as saturation at the 11,13-position); whilst for the *cyp71av1-1* mutant it is amorpha-4,11-diene (A-4,11-D), which presents a 11,13-double bond (Czechowski et al., 2016). Both *in vivo* and *in vitro* experiments indicate that this difference in functionality is the basis why DHAA-OOH (the tertiary allylic hydroperoxide from DHAA) is converted to artemisinin, whereas A-4,11-D-OOH is converted to arteannuin X (Czechowski et al., 2016).

We therefore hypothesised that the conversion of artemisinic acid (AA) to artemisitene (Art B) in NCV may also be a non-enzymatic process, paralleling the conversion of DHAA into artemisinin in Artemis (Supplemental figure 3A and B) and of amorpha-4,11-diene to arteannuin X in the *cyp71av1-1* mutant (Czechowski et al., 2016). The tertiary allylic hydroperoxide from

494 artemisinic acid (AA-OOH) differs from the two foregoing examples in that it incorporates both
495 a 12-carboxylic acid group and unsaturation at the 11,13-position. In support of this hypothesis,
496 when a sample of AA-OOH (produced by photosensitized oxygenation of AA; and purified by
497 HPLC) was left unattended for several weeks, it was indeed found to have been converted
498 predominantly to ArtB (albeit at a rate that was significantly slower than for the conversion of
499 DHAA-OOH to artemisinin). This unexpected transformation is mostly simply explained by
500 attack of the 12-carboxylic acid group at the allylic position of the hydroperoxide, as is shown in
501 Supplemental figure 3A and B. Further studies will be required to explain why it should be that
502 this (apparently) rather subtle modification to the 12-CO_2H group (i.e. the introduction of 11,13-
503 unsaturation in AA-OOH) has resulted in such a radically different pathway, as compared with
504 DHAA-OOH.

505 The second most abundant product of AA-OOH conversion is *Epi*-deoxyarteannuin B (EDB),
506 which accumulates predominantly in young leaves of NCV. The EDB accumulation pattern is
507 therefore different to DHEDB (the 11,13-saturated anaologue), where the latter's concentration
508 rises from top to mature and dry leaves in Artemis, broadly following the accumulation pattern
509 of artemisinin. We have proposed that the spontaneous conversions of AA into EDB and DHAA
510 into DHEDB progress via very similar molecular mechanisms (Supplemental figure 3C and D).
511 Interestingly we have observed very little EDB arising from the spontaneous conversions of AA-
512 OOH described above, which was predominantly converted to ArtB. It is known that a
513 hydrophobic (lipophilic) environment promotes conversions of DHAA-OOH into artemisinin
514 whereas an aqueous, acidic medium promotes DHAA-OOH conversions to DHEDB (Brown and
515 Sy, 2004). This may also explain the very minor conversion of AA-OOH into EDB which was
516 carried out in a hydrophobic environment (deuterated chloroform), and which promoted AA-
517 OOH conversions to ArtB. This highlights the parallels between artemisinin and arteannuin B
518 biogenesis shown in Supplemental figure 3A and B. It also suggests that *in vivo* conversions of
519 AA-OOH to EDB requires an aqueous intra-cellular environment, which might be expected to be
520 present in young leaf trichomes, but less so in mature leaf trichomes where the sub-apical
521 hydrophobic cavities are predominant (Ferreira and Janick, 1995), or upon cell dehydration (in
522 dried leaf material).

Differences between the LAP and HAP chemotypes extended well beyond artemisinin-related sequiterpenes to other classes of terpenes (Figure 4, 5, Supplemental Tables 1, 2 and 3). This divergence at the level of metabolism is not that surprising given that these chemotypes also exhibit significant differences in their morphology (Figure 5). Artemis is an F1 hybrid derived from HAP parents of East Asian origin (Delabays et al., 2001) while NCV is an open-pollinated variety of Europe origin (personal communication with Dr. Michael Schwerdtfeger, curator of Botanical Garden at the University of Göttingen, Germany). This is consistent with the general trend for the *A. annua* varieties of European and North American which mostly represent the LAP chemotype and the majority of East-Asian origin varieties which represent the HAP chemotype (Wallart *et al.,* 2000), Details of the genetic divergence of these varieties remains a topic for further investigation that could reveal further insight into the sesquiterpene flux into different end products.

5 Conclusion

This first comparative phytochemical analysis of high- (HAP) and low-artemisinin production (LAP) chemotypes of *A. annua* resulted in the characterisation of over 84 natural products by NMR, 26 of which have not previously been described in *A. annua*. We have also shown that the vast majority of *amorphane* sesquiterpenes are unsaturated at the 11,13-position in LAP-chemotype as opposed to the majority of them being saturated at the 11,13-position in HAP-chemotype. This is explained by existence of two sequence variants of *CYP71AV1* in the two investigated chemotypes and differential expression of the key branching enzyme in artemisinin pathway, namely *artemisinic aldehyde Δ 11 (13) reductase (DBR2)*. By highlighting the main points of difference between HAP and LAP chemotypes our findings will help inform strategies for the future improvement of artemisinin production in either *A. annua* or heterologous hosts.

6 Author Contributions

TC planned and performed the experiments, analysed the data, and wrote the manuscript. TRL planned the UPLC-MS and GC-MS experiments, analysed data and reviewed the manuscript. TMC planned and performed morphological plant analysis. DH performed UPLC-MS and GC-MS experiments. CW planned and performed extraction, purifications and NMR experiments and analysed data. ME performed extraction, purifications and NMR experiments. GDB planned

and performed NMR experiments, analysed data, wrote and reviewed manuscript. IAG planned and supervised the experiments and wrote the manuscript.

7 Funding

We acknowledge financial support for this project from The Bill and Melinda Gates Foundation as well as from The Garfield WestonFoundation for the Centre for Novel Agricultural Products This work was also supported by the Biotechnology and Biological Sciences Research Council Grant BB/G008744/1 (to GDB), "The Biosynthesis of Artemisinin".

8 Conflict of Interest Statement

All authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

9 Acknowledgements

We would like to thank: Dr Caroline Calvert for project management; C. Abbot and A. Fenwick for horticulture assistance; X. Simonnet and Médiplant for access to the Artemis and NCV varieties. We acknowledge financial support for this project from The Bill and Melinda Gates Foundation. GDB would like to thank the BBSRC for financial support and the Chemical Analysis Facility (CAF) at the University of Reading for the provision of the 700 MHz NMR spectrometer used in these studies.

10 References

Acton, N.; Klayman, D.L. (1985) Artemisitene, a new sesquiterpene lactone endoperoxide from Artemisia annua. 441-442.

Anke, H., and Sterner, O. (1991). Comparison of the Antimicrobial and Cytotoxic Activities of Twenty Unsaturated Sesquiterpene Dialdehydes from Plants and Mushrooms. *Planta Med* 57(04), 344-346. doi: 10.1055/s-2006-960114.

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 Physiol* 154(2), 958-968. doi: 10.1104/pp.110.162552.

Bouwmeester, H.J., Wallaart, T.E., Janssen, M.H.A., van Loo, B., Jansen, B.J.M., Posthumus, M.A., et al. (1999). Amorpha-4,11-diene synthase catalyses the first probable step in artemisinin biosynthesis. *Phytochemistry* 52(5), 843-854. doi: Doi 10.1016/S0031-9422(99)00206-X.

Brown, G.D. (2010). The biosynthesis of artemisinin (Qinghaosu) and the phytochemistry of Artemisia annua L. (Qinghao). *Molecules* 15(11), 7603-7698. doi: 10.3390/molecules15117603.

583 Brown, G.D., and Sy, L.-K. (2004). In vivo transformations of dihydroartemisinic acid in Artemisia annua
584 plants. *Tetrahedron* 60(5)**,** 1139-1159. doi: 10.1016/j.tet.2003.11.070.
585 Brown, G.D., and Sy, L.-K. (2007). In vivo transformations of artemisinic acid in Artemisia annua plants.
586 *Tetrahedron* 63(38)**,** 9548-9566. doi: 10.1016/j.tet.2007.06.062.
587 Bylesjo, M., Segura, V., Soolanayakanahally, R.Y., Rae, A.M., Trygg, J., Gustafsson, P., et al. (2008).
588 LAMINA: a tool for rapid quantification of leaf size and shape parameters. *BMC Plant Biol* 8**,** 82.
589 doi: 10.1186/1471-2229-8-82.
590 Czechowski, T., Larson, T.R., Catania, T.M., Harvey, D., Brown, G.D., and Graham, I.A. (2016). Artemisia
591 annua mutant impaired in artemisinin synthesis demonstrates importance of nonenzymatic
592 conversion in terpenoid metabolism. *Proc Natl Acad Sci U S A* 113(52)**,** 15150-15155. doi:
593 10.1073/pnas.1611567113.
594 Delabays, N., Simonnet, X., and Gaudin, M. (2001). The genetics of artemisinin content in Artemisia
595 annua L. and the breeding of high yielding cultivars. *Curr Med Chem* 8(15)**,** 1795-1801.
596 Duke, M.V., Paul, R.N., Elsohly, H.N., Sturtz, G., and Duke, S.O. (1994). Localization of Artemisinin and
597 Artemisitene in Foliar Tissues of Glanded and Glandless Biotypes of Artemisia-Annua L.
598 *International Journal of Plant Sciences* 155(3)**,** 365-372. doi: Doi 10.1086/297173.
599 Duke, S.O., and Paul, R.N. (1993). Development and Fine Structure of the Glandular Trichomes of
600 Artemisia annua L. *International Journal of Plant Sciences* 154(1)**,** 107-118. doi:
601 10.2307/2995610.
602 Ferreira, J.F., and Luthria, D.L. (2010). Drying affects artemisinin, dihydroartemisinic acid, artemisinic
603 acid, and the antioxidant capacity of Artemisia annua L. leaves. *J Agric Food Chem* 58(3)**,** 1691-
604 1698. doi: 10.1021/jf903222j.
605 Ferreira, J.F.S., and Janick, J. (1995). Floral Morphology of Artemisia-Annua with Special Reference to
606 Trichomes. *International Journal of Plant Sciences* 156(6)**,** 807-815. doi: Doi 10.1086/297304.
607 Graham, I.A., Besser, K., Blumer, S., Branigan, C.A., Czechowski, T., Elias, L., et al. (2010). The genetic
608 map of Artemisia annua L. identifies loci affecting yield of the antimalarial drug artemisinin.
609 *Science* 327(5963)**,** 328-331. doi: 10.1126/science.1182612.
610 Guha, R. (2007). Chemical Informatics Functionality in R. *2007* 18(5)**,** 16. doi: 10.18637/jss.v018.i05.
611 Kuhl, C., Tautenhahn, R., Bottcher, C., Larson, T.R., and Neumann, S. (2012). CAMERA: an integrated
612 strategy for compound spectra extraction and annotation of liquid chromatography/mass
613 spectrometry data sets. *Anal Chem* 84(1)**,** 283-289. doi: 10.1021/ac202450g.
614 Larson, T.R., Branigan, C., Harvey, D., Penfield, T., Bowles, D., and Graham, I.A. (2013). A survey of
615 artemisinic and dihydroartemisinic acid contents in glasshouse and global field-grown
616 populations of the artemisinin-producing plant Artemisia annua L. *Industrial Crops and Products*
617 45**,** 1-6. doi: 10.1016/j.indcrop.2012.12.004.
618 Lommen, W.J.M., Schenk, E., Bouwmeester, H.J., and Verstappen, F.W.A. (2006). Trichome dynamics
619 and artemisinin accumulation during development and senescence of Artemisia annua leaves.
620 *Planta Medica* 72(4)**,** 336-345. doi: 10.1055/s-2005-916202.
621 Mercke, P., Bengtsson, M., Bouwmeester, H.J., Posthumus, M.A., and Brodelius, P.E. (2000). Molecular
622 cloning, expression, and characterization of amorpha-4,11-diene synthase, a key enzyme of
623 artemisinin biosynthesis in Artemisia annua L. *Arch Biochem Biophys* 381(2)**,** 173-180. doi:
624 10.1006/abbi.2000.1962.
625 Nguyen, D.T., Gopfert, J.C., Ikezawa, N., Macnevin, G., Kathiresan, M., Conrad, J., et al. (2010).
626 Biochemical conservation and evolution of germacrene A oxidase in asteraceae. *J Biol Chem*
627 285(22)**,** 16588-16598. doi: 10.1074/jbc.M110.111757.
628 Niu, X., Xing, W., Li, W., Fan, T., Hu, H., and Li, Y. (2012). Isofraxidin exhibited anti-inflammatory effects
629 in vivo and inhibited TNF-alpha production in LPS-induced mouse peritoneal macrophages in

vitro via the MAPK pathway. *Int Immunopharmacol* 14(2)**,** 164-171. doi: 10.1016/j.intimp.2012.06.022.
Olsson, M.E., Olofsson, L.M., Lindahl, A.L., Lundgren, A., Brodelius, M., and Brodelius, P.E. (2009). Localization of enzymes of artemisinin biosynthesis to the apical cells of glandular secretory trichomes of Artemisia annua L. *Phytochemistry* 70(9)**,** 1123-1128. doi: 10.1016/j.phytochem.2009.07.009.
Paddon, C.J., Westfall, P.J., Pitera, D.J., Benjamin, K., Fisher, K., McPhee, D., et al. (2013). High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* 496(7446)**,** 528-532. doi: 10.1038/nature12051.
Ro, D.K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., et al. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440(7086)**,** 940-943. doi: 10.1038/nature04640.
Santomauro, F., Donato, R., Sacco, C., Pini, G., Flamini, G., and Bilia, A.R. (2016). Vapour and Liquid-Phase Artemisia annua Essential Oil Activities against Several Clinical Strains of Candida. *Planta Med* 82(11/12)**,** 1016-1020. doi: 10.1055/s-0042-108740.
Shen, Q., Chen, Y.F., Wang, T., Wu, S.Y., Lu, X., Zhang, L., et al. (2012). Overexpression of the cytochrome P450 monooxygenase (cyp71av1) and cytochrome P450 reductase (cpr) genes increased artemisinin content in Artemisia annua (Asteraceae). *Genet Mol Res* 11(3)**,** 3298-3309. doi: 10.4238/2012.September.12.13.
Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R., and Siuzdak, G. (2006). XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem* 78(3)**,** 779-787. doi: 10.1021/ac051437y.
Soetaert, S.S., Van Neste, C.M., Vandewoestyne, M.L., Head, S.R., Goossens, A., Van Nieuwerburgh, F.C., et al. (2013). Differential transcriptome analysis of glandular and filamentous trichomes in Artemisia annua. *BMC Plant Biol* 13**,** 220. doi: 10.1186/1471-2229-13-220.
Stacklies, W., Redestig, H., Scholz, M., Walther, D., and Selbig, J. (2007). pcaMethods--a bioconductor package providing PCA methods for incomplete data. *Bioinformatics* 23(9)**,** 1164-1167. doi: 10.1093/bioinformatics/btm069.
Sy, L.K., and Brown, G.D. (2002). The role of the 12-carboxyllic acid group in the spontaneous autoxidation of dihydroartemisinic acid. *Tetrahedron* 58(5)**,** 909-923. doi: Doi 10.1016/S0040-4020(01)01192-9.
Tautenhahn, R., Bottcher, C., and Neumann, S. (2008). Highly sensitive feature detection for high resolution LC/MS. *BMC Bioinformatics* 9**,** 504. doi: 10.1186/1471-2105-9-504.
Teoh, K.H., Polichuk, D.R., Reed, D.W., and Covello, P.S. (2009). Molecular cloning of an aldehyde dehydrogenase implicated in artemisinin biosynthesis inArtemisia annua

Botany 87(6)**,** 635-642. doi: 10.1139/b09-032.
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(5)**,** 1411-1416. doi: 10.1016/j.febslet.2006.01.065.
Ting, H.M., Wang, B., Ryden, A.M., Woittiez, L., van Herpen, T., Verstappen, F.W., et al. (2013a). The metabolite chemotype of Nicotiana benthamiana transiently expressing artemisinin biosynthetic pathway genes is a function of CYP71AV1 type and relative gene dosage. *New Phytol* 199(2)**,** 352-366. doi: 10.1111/nph.12274.
Ting, H.M., Wang, B., Ryden, A.M., Woittiez, L., van Herpen, T., Verstappen, F.W.A., et al. (2013b). The metabolite chemotype of Nicotiana benthamiana transiently expressing artemisinin biosynthetic

pathway genes is a function of CYP71AV1 type and relative gene dosage. *New Phytologist* 199(2), 352-366. doi: 10.1111/nph.12274.

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 Medica* 66(1), 57-62. doi: Doi 10.1055/S-2000-11115.

Wallaart, T.E., Pras, N., and Quax, W.J. (1999). Isolation and identification of dihydroartemisinic acid hydroperoxide from Artemisia annua: A novel biosynthetic precursor of artemisinin. *Journal of Natural Products* 62(8), 1160-1162. doi: Doi 10.1021/Np9900122.

Wetzstein, H.Y., Porter, J.A., Janick, J., and Ferreira, J.F. (2014). Flower morphology and floral sequence in Artemisia annua (Asteraceae)1. *Am J Bot* 101(5), 875-885. doi: 10.3732/ajb.1300329.

Woerdenbag, H.J., Pras, N., Chan, N.G., Bang, B.T., Bos, R., Vanuden, W., et al. (1994). Artemisinin, Related Sesquiterpenes, and Essential Oil in Artemisia-Annua during a Vegetation Period in Vietnam. *Planta Medica* 60(3), 272-275. doi: DOI 10.1055/s-2006-959474.

Yamazaki, T., and Tokiwa, T. (2010). Isofraxidin, a coumarin component from Acanthopanax senticosus, inhibits matrix metalloproteinase-7 expression and cell invasion of human hepatoma cells. *Biol Pharm Bull* 33(10), 1716-1722.

Yang, K., Monafared, R.S., Wang, H., Lundgren, A., and Brodelius, P.E. (2015). The activity of the artemisinic aldehyde Delta11(13) reductase promoter is important for artemisinin yield in different chemotypes of Artemisia annua L. *Plant Mol Biol*. doi: 10.1007/s11103-015-0284-3.

Yuan, Y., Liu, W.H., Zhang, Q.Z., Xiang, L.E., Liu, X.Q., Chen, M., et al. (2015). Overexpression of artemisinic aldehyde Delta 11 (13) reductase gene-enhanced artemisinin and its relative metabolite biosynthesis in transgenic Artemisia annua L. *Biotechnology and Applied Biochemistry* 62(1), 17-23. doi: 10.1002/bab.1234.

~~Zhang, Y., Teoh, K.H., Reed, D.W., and Covello, P.S. (2009). Molecular cloning and characterization of Dbr1, a 2-alkenal reductase fromArtemisia annuaThe nucleotide sequence reported in this article has been deposited in the GenBank database under accession No. FJ750460.This paper is one of a selection of papers published in a Special Issue from the National Research Council of Canada – Plant Biotechnology Institute. *Botany* 87(6), 643-649. doi: 10.1139/b09-033.~~

Zhang, Y., Teoh, K.H., Reed, D.W., Maes, L., Goossens, A., Olson, D.J., et al. (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(31), 21501-21508. doi: 10.1074/jbc.M803090200.

Figure legends:

Figure 1. Novel natural compounds characterised from the Artemis (A) and NCV (B) varieties of *A. annua* by NMR approach.

Numbering of compounds is consistent with Supplemental Lists 1 and 2. Numbering of carbon atoms showed.

Figure 2. Ten pairs of 11,13-dihdyro/ 11,13-dehydro amorphanolides between Artemis (left-hand side) and NCV (right-hand side) varieties of *A. annua* characterised by the NMR approach.

Numbering of compounds is consistent with Supplemental Lists 1 and 2. Numbering of carbon atoms showed. Novel compound indicated by asterisk

Figure 3. Metabolic and transcriptomic comparison of the artemisinin pathway in the low- versus high-artemisinin chemotypes of *A. annua*

A) Level of selected sesquiterpenes were quantified by GC-MS **(i)** and UPLC-MS **(ii)-(vii)** in fresh juvenile leaf 1-5 (Top), fresh mature leaf 11-13 (Mid.) and oven-dried whole plant-stripped leaves (Dry) from 12-weeks old glasshouse-grown Artemis (green bars) and NCV (grey bars) varieties as described in Materials and methods. ~~SI;~~ error bars – SEM (n=15 for Top and Mid. leaf; n=6 for Dry leaf). Letters represent Tukey's range test results after one way ANOVA or REML (see Materials and Methods for details). Groups not sharing letters indicate statistically significant differences. **(B)** – Transcript profiling of enzymes involved in the artemisinin biosynthetic pathway, in two types of leaf material as on panel (A) was done as described in materials and methods, error bars – SE (n=9). Asterisk indicates t-test statically significant difference between Artemis (green bars) and NCV (grey bars) at $p<0.05$. **(C)** Summary of the metabolite and transcriptional differences between Artemis and NCV for the artemisinin biosynthetic pathway: full arrows – known enzymatic steps, dashed arrows – non-enzymatic conversions, red arrows – transcript changes in juvenile leaves of NCV vs. Artemis, green arrows – metabolite changes of NCV vs. Artemis (all types of leaves). DBR2 position in the pathway highlighted in a square.

Metabolite abbreviations: G-3-P – glyceraldehyde-3-phosphate; MEP - 2-C-methylerythritol 4-phosphate; MEcPP - 2-C-methyl-D-erythritol-2,4-cyclopyrophosphate. Cytosolic precursors: HMG-CoA - 3-hydroxy-3-methylglutaryl-CoA; MVA – mevalonate; IPP – isopentenyl pyrophosphate; DMAPP – dimethylallyl pyrophosphate; FPP – farnesyl pyrophosphate; A-4,11-D – amorpha-4,11-diene; AAOH – artemsinic alcohol; AAA – artemsinic aldehyde; AA – artemsinic acid; ArtB – arteannuin B; DHAAA - dihydroartemsinic aldehyde; DHAA - dihydroartemsinic acid; DHAAOOH- dihydroartemsinic acid tertiary hydroperoxide; DHEDB – dihydro-*epi*-deoxyarteanniun B. AAOOH - artemsinic acid tertiary hydroperoxide, EDB – *epi*-deoxyarteannuin B. Enzyme abbreviations: HMGR- 3-hydroxy-3-methylglutaryl coenzyme A reductase, HDR- 4-hydroxy-3-methylbut-2-enyl diphosphate reductase, DXR - 1-deoxy-D-xylulose-5-phosphate reductoisomerase, DXS- 1-deoxy-D-xylulose-5-phosphate synthase, FPS - farnesyl diphosphate synthase. AMS – amorpha-4,11-diene synthase, CYP71AV1 - amorpha-4,11-diene C-12 oxidase, CPR – cytochrome P450 reductase, DBR2 - artemisinic aldehyde Δ 11 (13) reductase, ALDH1 - aldehyde dehydrogenase.

Figure 4. Principal component analysis of UPLC-MS (A) and GC-MS (B) data from different leaf types from Artemis and NCV varieties.

Principal component analysis of 75 UPLC-MS identified peaks (A) and 202 GC-MS identified metabolites (B). Leaf types, corresponding with Figure 2 are represented by symbols: circles – leaf 1-5, triangles – leaf 11-13, crosshairs – oven-dried leaf. Two chemotypes represented by colours green – Artemis and grey – NCV. PCA was performed on log-scaled data and mean-centred data; dotted ellipse = Hotelling's 95% confidence interval.

Figure 5. Heatmaps of influential metabolites from UPLC- and GC-MS PCA analyses.

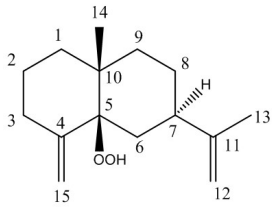
Top-n (UPLC-MS = 18 (A); GC-MS = 20 (B)) metabolites were chosen for visualization based on loadings plots (Supplementary Fig 1~~Y~~) from the PC1 dimensions in the PCA analyses (Fig X). Mean data were log-scaled and then row-scaled for colour intensity plotting (lighter = more abundant). Hierarchical clustering was performed with average linkage, with Euclidean distances for genotypes and 1-absolute values of correlations as distances for metabolites. Metabolite names are abbreviated where necessary for clarity and are given in full in Supplementary tables 2 and 3.

Figure 6. Morphological characterisation of low- and high-artemisinin natural chemotypes of *A. annua*

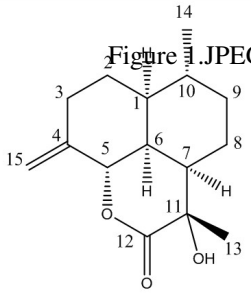
(A) Photographs show four representative 12-week old plants from each two chemotypes of *A. annua*, ruler scaled in cm showed on both sides; Plant height (B), internode length (C), leaf area (D) and glandular secretory trichome density (E) recorded for 12-week old plants. Green bars represent Artemis (HAP-chemotype) and brown bars represent NCV (LAP-chemotype). Error bars – SEM (n=15), letters represent one-way ANOVA Tukey's range test results; Groups not sharing letters indicate statistically significant differences

Figure 1.JPEG

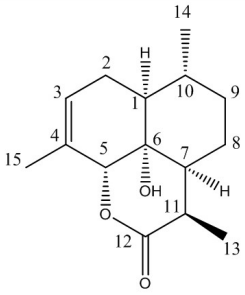
A



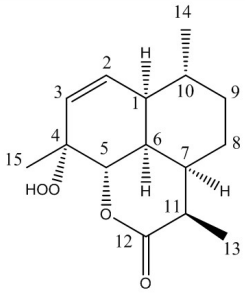
5β-hydroperoxy-eudesma-4(15),11-diene **(4)**



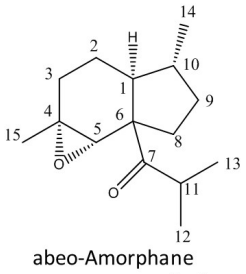
11-Hydroxy-arteannuin I **(18)**



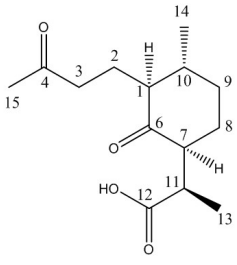
6α-Hydroxy-arteannuin J **(19)**



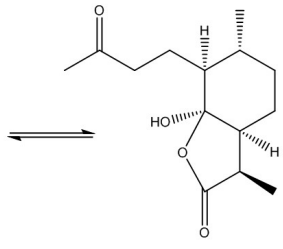
Arteannuin P **(20)**



abeo-Amorphane sesquiterpene **(27)**



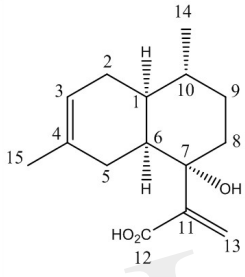
25 (keto form)



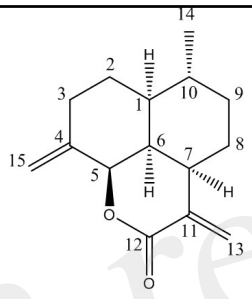
26 (ketal form)

Arteannuin Q **(25) /(26)**

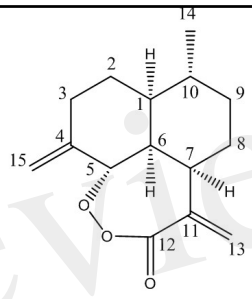
B



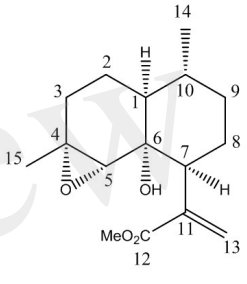
7α-Hydroxy-isoartemisinic acid **(52)**



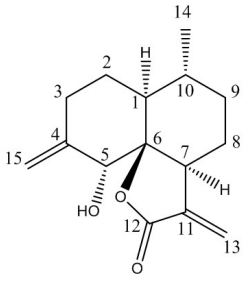
Arteannuin R **(54)**



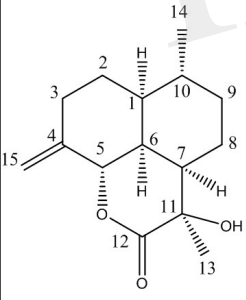
Arteannuin S **(55)**



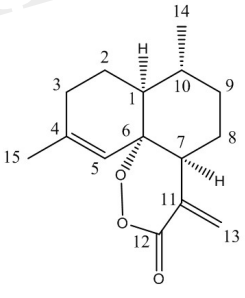
4α, 5α-Epoxy-6α-hydroxyartemisinic acid methyl ester **(57)**



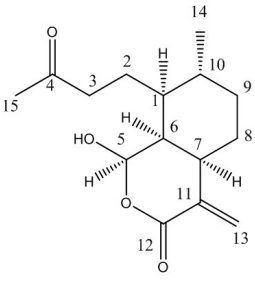
Dehydroarteannuin L **(59)**



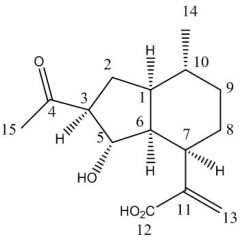
epi-11-Hydroxy-arteannuin I **(64)**



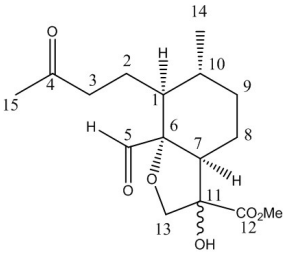
Artemisinic acid 6α-peroxy ester (**65**)



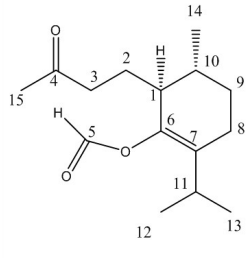
Arteannuin T **(69)**



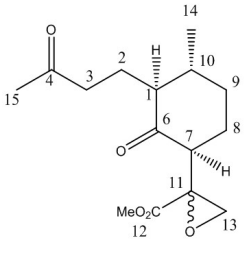
Arteannuin U **(70)**



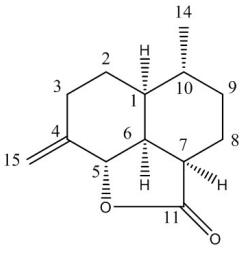
Arteannuin V **(72)**



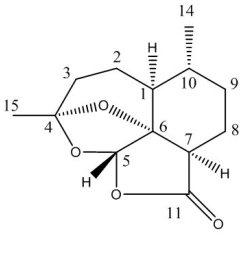
Arteannuin W **(73)**



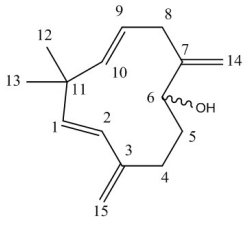
Arteannuin Y **(74)**



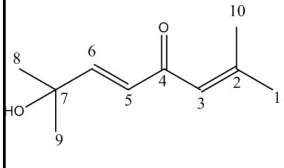
Isoarteannuin A (**77**)



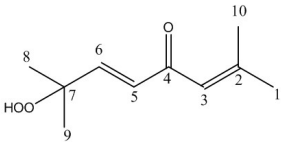
Arteannuin Z **(78)**



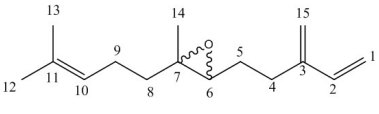
6-Hydroxy-γ-humulene **(48)**



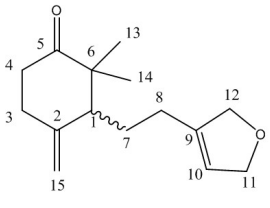
(E)-7-Hydroxy-2,7-dimethylocta-2,5-dien-4-one **(43)**



(E)-7-Hydroperoxy-2,7-dimethylocta-2,5-dien-4-one **(44)**

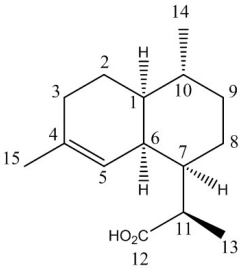


6,7-Epoxy-6,7-dihydro-β-farnesene (**45**)

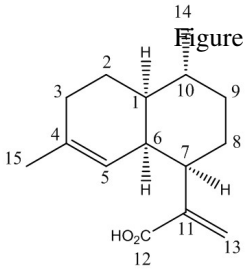


3-(2-(2,5-Dihydrofuran-3-yl)ethyl)-2,2-dimethyl-4-methylenecyclohexan-1-one **(79)**

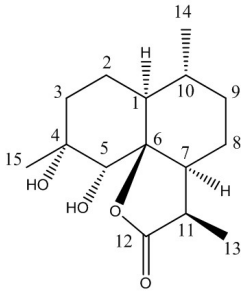
Figure 2.JPEG



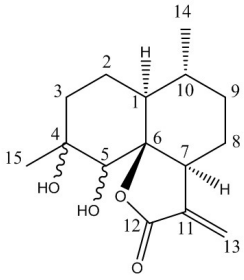
Dihydroartemisinic acid (**8**)



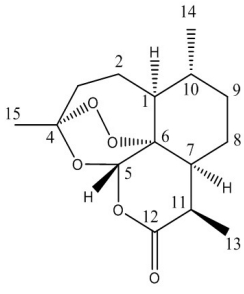
Artemisinic acid (**9**)



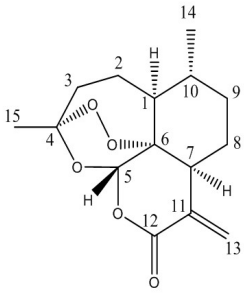
Arteannuin M (**15**)



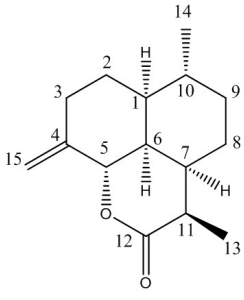
Dehydroarteannuin M (**61**)



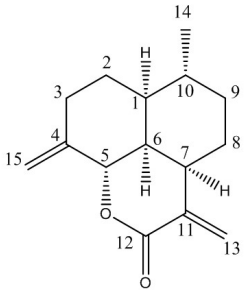
Artemisinin **(22)**



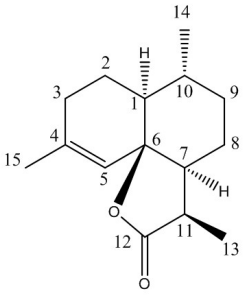
Artemisitene **(66)**



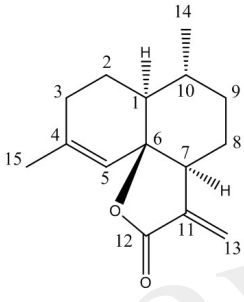
Arteannuin I (**16**)



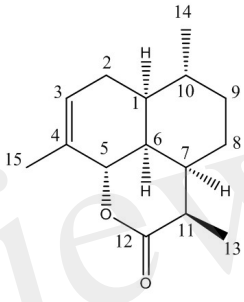
Annulide (**62**)



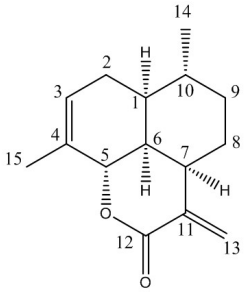
Dihydro-*epi*-deoxyarteannuin B **(12)**



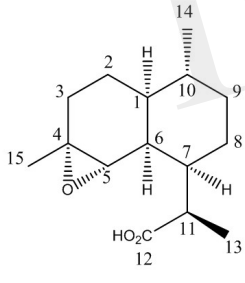
epi-Deoxyarteannuin B **(13)**



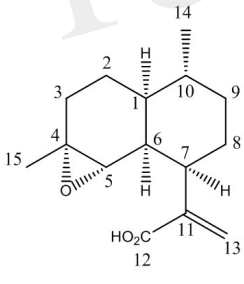
Arteannuin J **(17)**



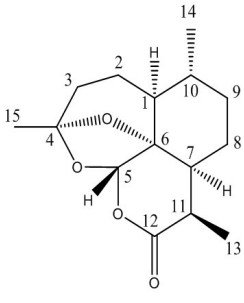
Isoannulide (**63**)



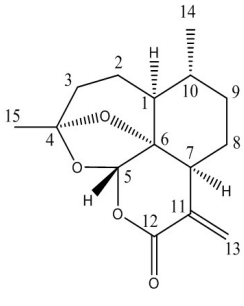
α-Epoxy-dihydroartemisinic acid **(10)**



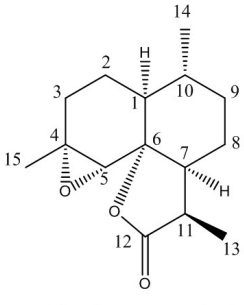
α-Epoxy-artemisinic acid **(56)**



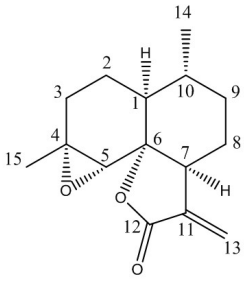
Deoxyartemisinin **(23)**



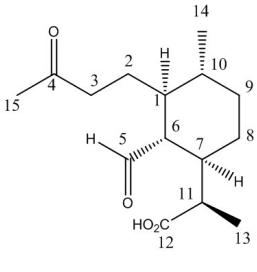
*Deoxyartemisitene **(67)**



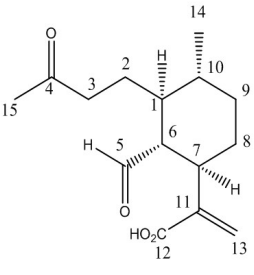
Dihydroarteannuin B (**14**)



Arteannuin B (**60**)



4,5-*seco*-4,5-Keto,aldehyde-amorphan-12-oic acid (**24**)



4,5-*seco*-4,5-Keto, aldehyde-amorphan-11,13-ene-12-oic acid (**68**)

Artemis NCV Artemis NCV

Figure 3.JPEG

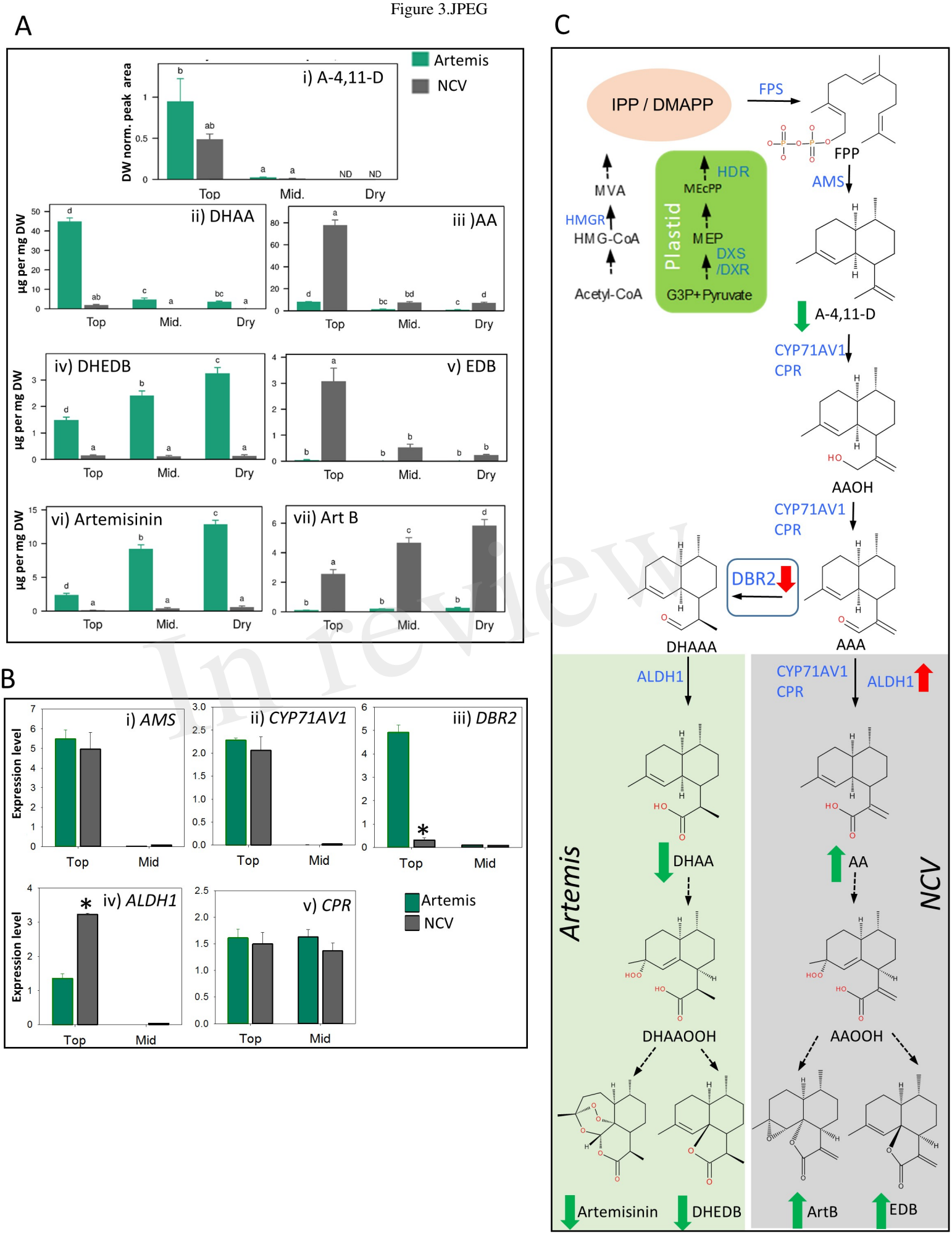


Figure 4.JPEG

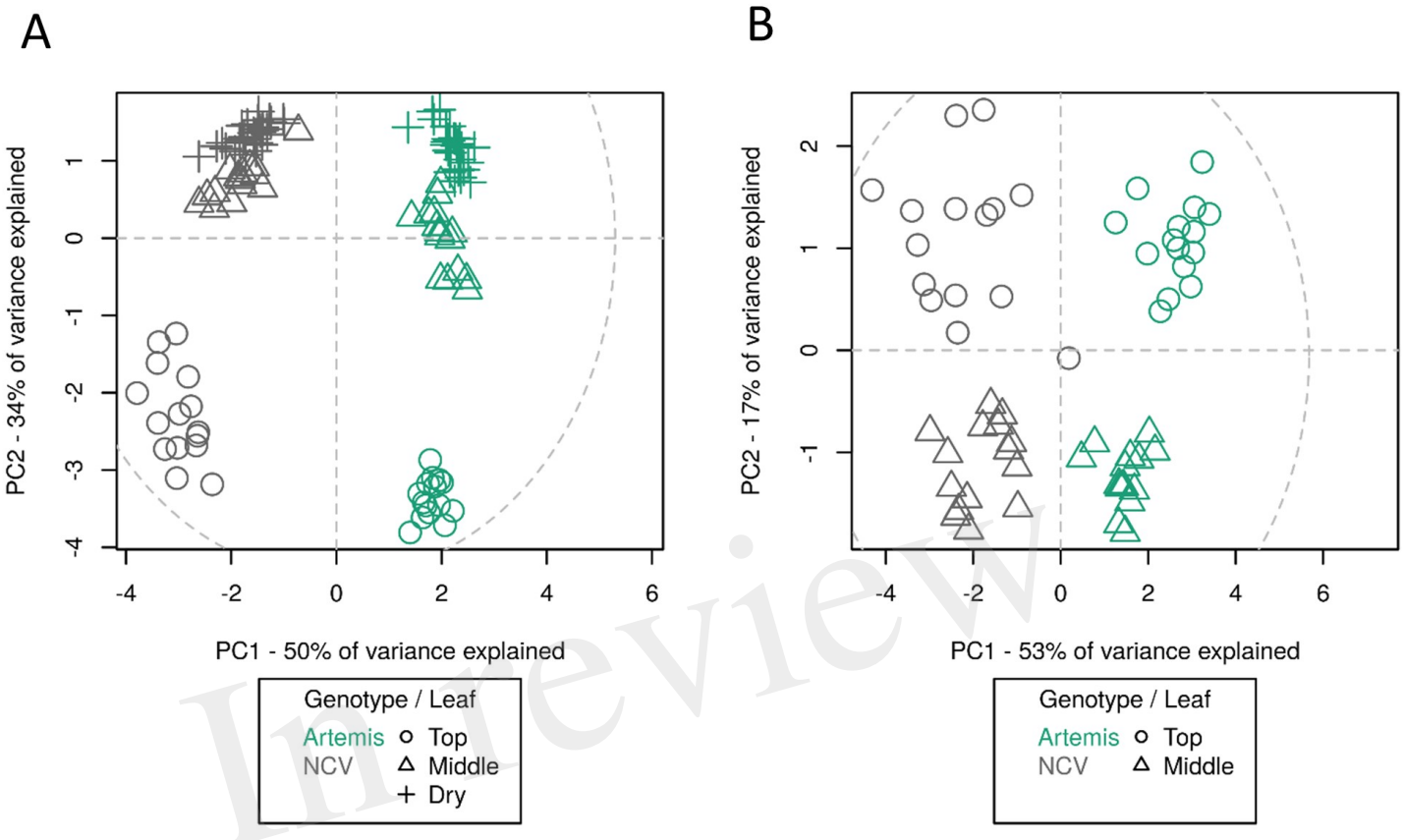
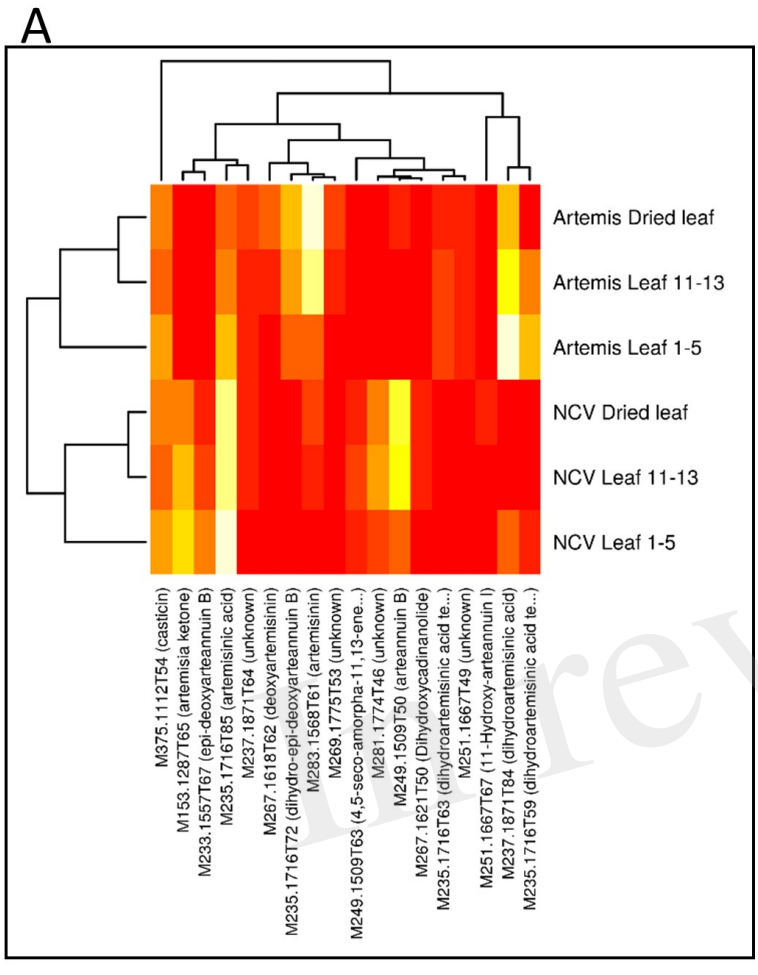


Figure 5.JPEG



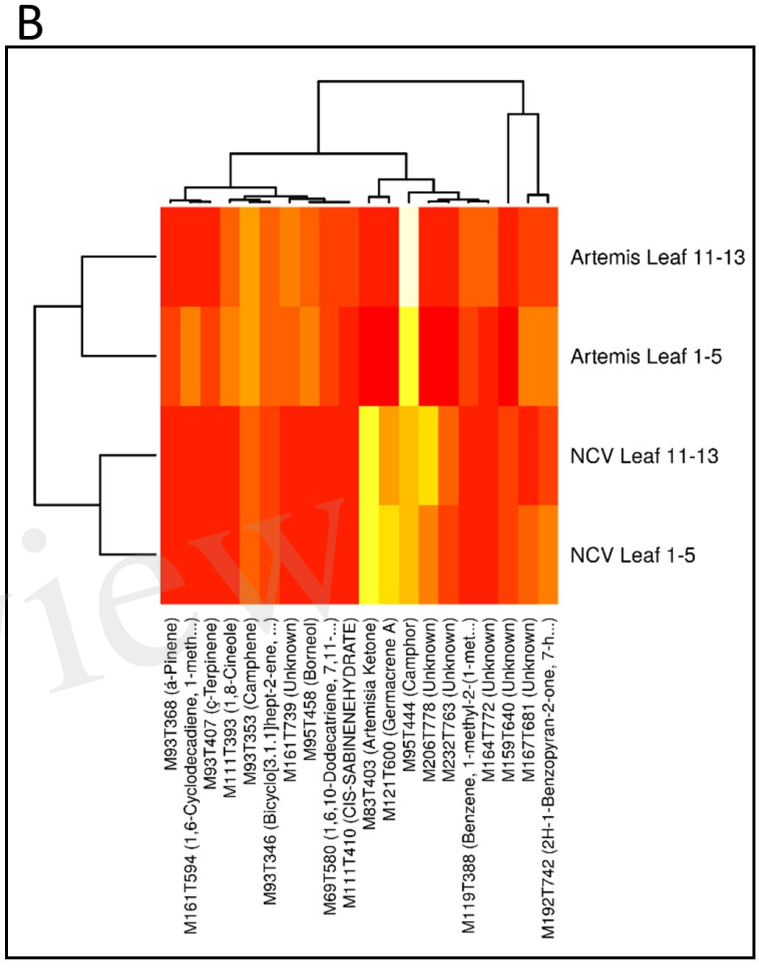


Figure 6.JPEG