Proteome Analysis of Glandular Trichome from *Artemisia annua* L.

WU, Ting

A Thesis Submitted in Partial Fulfillment

of the Requirements for the Degree of

Master of Philosophy

in

Biology

The Chinese University of Hong Kong

August 2011

Thesis/Assessment Committee

Professor Diane D. J. Guo (Supervisor)

Professor Peter C. K. Cheung (Chairman)

Professor S. M. Ngai (Committee Member)

Professor Z. W. Cai (External Examiner)

Statement

All experimental work reported in this thesis was performed by the author, unless stated otherwise.

WU Ting

Acknowledgements

I would like to give my deepest thanks to people who have supported me during my M.Phil' studies. At First, I am grateful to my supervisor, Prof. Diane D. J. Guo, for giving me this precious opportunity to work in her lab and learn lots of knowledge and techniques about plant molecular biology. I am thankful for her patient instruction, professional advice and constant concern which give me great encouragement during my study.

Secondly, I would like to acknowledge Prof. C. K. Cheung and S. M. Ngai for serving as my thesis committee members. They always concerned about my research progress and gave many valuable suggestions and guidance on my seminar, experiments and thesis. In addition, I would like to thank my external examiner, Prof. Z. W. Cai for his precious comments on my thesis.

Special thanks give to Prof. Jianhua Zhang form Hong Kong Baptist University for providing place for growth of *Artemisia annua.* I thank Prof. Dongtao Ren from China Agricultural University for technical assistance with 2DE and Prof. Paul K S Lam from City University of Hong Kong for providing MALDI-TOF/TOF MS.

I also thank Yejun Wang, who conducted microarray and 454 pyrosequencing raw data processing and established EST database for *Artemisia annua* L. Besides, I would like to thank all my labmates in G94, including Qing Zhang, Mingan Sun, Wei Wang, Han Cheng, Baohui Cheng, Yan Qi, Man Yuan, Dr X.

Q. Wang, Dongming Ma and Yiuman Chan for their company, kind help and technical support.

Lastly, my greatest appreciation would go to my beloved family for their love and support in all my endeavors.

Abstract

Malaria is a mosquito-borne infectious disease which causes one to three millions of deaths in the world each year. Artemisinin, as the most prominent anti-malarial drug, is mainly isolated from the aerial parts of *Artemisia annua* L., an aroma plant of the Asteraceae family. However, low artemisinin content (only 0.01% to 1.1% of dry weight) hindered its wide application. In *A. annua*, glandular secreting trichomes (GSTs) are the major sites for the synthesis, storage and secretion of diverse classes of plant secondary metabolites, including arteminisin. They are also considered as the first guard to protect plant against biotic and abiotic stress. Recently, our group has conducted a global survey of glandular trichome transcriptome in *A. annua* using high throughput pyrosequencing technology. Such effort has provided a useful resource of Expressed Sequence Tags (ESTs) for gene discovery. However, abundance of expressed transcripts may not represent the protein activity and a comprehensive proteomic study of *A. annua* glandular trichome is necessary for further understanding trichome function and arteminisin metabolism using a system biology approach.

The main objective of the my thesis research is to conduct proteomics study of *A. annua* glandular trichomes in order to unravel candidate proteins involved in artemisinin biosynthesis and transport, as well as in the development of glandular trichomes. 2D gel electrophoresis and mass spectrometry were employed for separation and analysis the total proteins in glandular trichomes. The proteomic

2D gel pattern of glandular trichome was furthermore compared with that of leaves. Over 700 spots on 2D gel of glandular trichomes were resolved, of which ~102 non-redundant proteins were confidently identified by searching National Center for Biotechnology Information (NCBI) database and our in-house Expressed Sequence Tag (EST) database. The most abundant proteins were those involved in photosynthesis, such as ribulose-1,5- bisphosphate carboxylase. Except for these proteins, over 70% of the rest 91 proteins were highly expressed in glandular trichomes. They participate in electron transport chain, proteolysis, transcription and translation, metabolism and detoxification or defense and stress response etc. Among the 22 proteins that were only identified from our Artemisia EST database, 14 had no homolog with other species even with quite loose blast parameters, indicating they might be both species-specifically and tissue-specifically expressed. Moreover, two enzymes (artemisinic aldehyde delta-11(13) reductase and 4-diphosphocytidyl-2-C-methyl- D-erythritol synthase) were identified to be enriched in glandular trichomes that directly participate in artemisinin biosynthetic pathways. In conclusion, our proteomics study provides for the first time a general picture of proteins specifically expressed in Artemisia glandular trichomes. Enriched but inter-species un-conserved genes identified in this study, together with our previous transcriptome study suggested that more GST-specific genes are likely involved in Artemisia metabolism and trichome. Some of these candidates will provide potential targets for genetic manipulation of glandular trichomes for enhanced artemisinin production in *A.annua.*

摘　要

瘧疾,作為一種急性寄生蟲傳染病，在全球範圍內每年可導致 100 到 300 萬人死亡。青蒿素作為最有效的抗瘧疾藥，主要來源於菊科植物青蒿的地上部分。然而，由於植物中青蒿素的含量太低（只有 0.01% 至 1.1%干重），嚴重限制了青蒿素的廣泛應用。在青蒿中，腺毛是合成、儲藏和分泌不同種類次級代謝產物的主要器官，包括青蒿。並且腺毛被認為是植物抵禦各種生物和非生物脅迫的第一道防線。目前，本課題組利用高通量的 pyrosequecing 測序法研究了青蒿分泌性腺毛中的轉錄本。這項研究成果為發現與青蒿素合成相關的基因提供了寶貴的資源。但是，轉錄水平調控並不能全面代表蛋白之表達水平。因此，對青蒿腺毛的全面的蛋白質組學研究是必需的，並且，通過此系統的生物學研究可以幫助我們進一步的了解腺毛的功能和青蒿素代謝途徑。

本課題的主要目的是通過對青蒿腺毛的蛋白質組的研究，從而發現一些參與青蒿素合成，轉運，以及腺毛發育的蛋白。雙向電泳和質譜技術被用來分離和鑒定腺毛中的總蛋白。同時我們還比較了腺毛蛋白和葉片蛋白膠圖的差異。在腺毛膠圖上，約 700 多個蛋白質點被分離和質譜分析。通過檢索 NCBI 和 EST 兩個數據庫，近 102 個非冗餘的蛋白質被鑒定。其中豐度最高的蛋白是參與植物光合作用的蛋白，如 1,5-二磷酸核酮糖羧化酶。除此之外，在其餘的 91 個蛋白質中，在腺毛中上調的超過了 70%。這些蛋白質參與了多種生物過程包括呼吸鏈、酶解、轉錄和翻譯、代謝、以及防禦和抗逆反應等。在

EST 數據庫單独鑒定的 22 個蛋白質中，14 個缺少同源蛋白信息，說明這些蛋白質可能是物種或者組織特異表達的蛋白質。另外，兩個參與青蒿素合成的酶（artemisinic aldehyde delta-11(13) reductase 和 4-diphosphocytidyl-2-C-methyl-D-erythritol synthase）在腺毛中的表達量也明顯地上調。總之，本課題的研究第一次提供了青蒿腺毛中蛋白質的信息。結合本課題組轉錄組的研究和此次研究中表達量增高並且物種間非保守的基因，說明許多腺毛特異表達的基因很可能參與青蒿的代謝和腺毛中。其中一些蛋白質可以作為重要的研究對象，為基因調控提高青蒿素含量提供理論依據。

Table of Contents

List of Abbreviations

2-DE	two-dimensional electrophoresis
A. annua	*Artemisia annua* L.
ABC transporters	ATP-binding cassette transporters
ACTs	artemisinin-based combination therapies
ADS	amorpha-4,11-diene synthase
AMDS	Amorpha-4,11-diene synthase
APS	adenosine 5′ phosphosulfate
ATP	adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
BSA	Bovine Serum Albumin
CHAPS	3-(3-cholamidopropyl) dimethylammonio-1-propane sulfonate
cDNA	complementary DNA
CYP	CYP71AV1
DDT	dichloro-diphenyl-trichloroethane
DMAPP	dimethylallyl diphosphate
DNA	Deoxyribonucleic acid
DOXP	1-deoxy-D-xylulose 5-phosphate
DTT	dithiothreitol
DW	dry weight

EST	Expressed Sequence Tag
FA	formic acid
FDS	Synthase Farnesyl Diphosphate
FPP	farnesyl diphosphate
FPS	farnesyl diphosphate synthase
GGPP	geranylgeranyl diphosphate
GO	Gene Ontology
GPP	geranyl diphosphate
GST	glandular secreting trichomes
IPG	immobilized pH gradient
IPP	isopentenyl diphosphate
HMG-CoA	3S-Hydroxy-3-methylglutaryl-CoA
Kb	Kilo base pair
kDa	Kilodalton
MALDI	matrix assisted laser desorption ionization
Mb	megabase
MEP	2-C-Methyl-D-erythritol-4-phosphate
mRNA	messenger RNA
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MSDB	mass spectrometry protein sequence database
MW	molecular weight
MVA	mevalonic acid

NAD(P)H	Nicotinamide Adenine Dinucleotide Phosphate
NCBI	National Center for Biotechnology Information, USA
NT	non-glandular trichomes
PAGE	polyacrylamide gel electrophoresis
P450	Cytochrome P450
pI	isoelectric point
PMF	peptide mass fingerprint
PPi	pyrophosphate
PVP	polyvinylpyrrolidone
RNA	Ribonucleic acid
SDS	sodium dodecyl sulfate
SEM	scanning electron microscope
SES	sesquiterpene synthase
TCA	trichloroacetic acid
TEM	transmission electron microscopy
TFA	trifluoracetic acid
TOF	time-of-flight
WHO	World Health Organization

Chapter 1 Literature Review

1.1 The disease of malaria

Malaria is a devastating disease which has been described since the beginning of history. Several characteristic symptoms of malaria has been mentioned in the Chinese recipes (Nei Ching, The Canon of Medicine) dating back to 2700 B.C (http://malaria.emedtv.com). In 1740, the name "mal aria" (meaning "bad air" in Italian) was first used by H. Walpole to describe the disease. Later in 1880, the parasites in human blood were identified and mosquito's vector was discovered consequently (http://www.medicinenet.com). Malaria is caused by protozoal parasites of the genus *Plasmodium* transmitted by female *Anopheles* mosquitoes. Four species of *Plasmodium* can infect human including *P. falciparum*, *P. vivax*, *P. ovale*, *P. malaria*, of which *P. falciparum* is the most severe species (Singh *et al.*, 2004). Symptoms of malaria including high fevers, chills, sweating, headache and vomiting etc. usually appear between 10 to 15 days after mosquito bite. If not treated in time, it is lethal with serious organ failures or blood problems.

Currently, there are 350-500 million clinical cases of malaria worldwide, resulting in more than 1 million deaths annually (Greenwood *et al.*, 2005; Skarbinski *et al.*, 2006). Most of severe cases occur in sub-Saharan Africa, but the disease also exists in large areas of Asia, Africa, and Central and South America (Fig. 1).

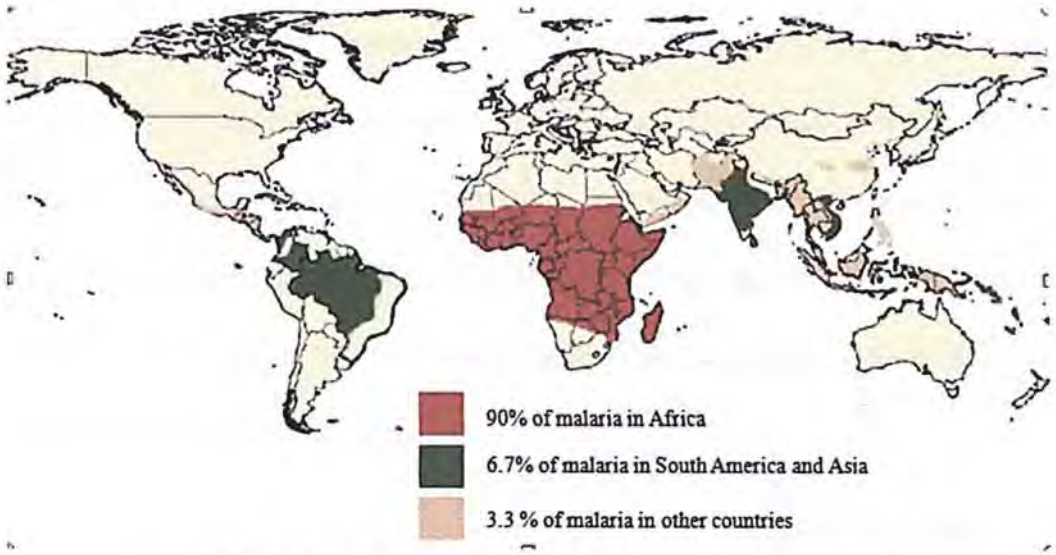

Figure 1 The distribution of malaria in the world Modified from (http://vcrc.res.in/malaria3.html)

1.1.1 Pathogenesis

The pathogenesis of malaria is very complicated and includes several steps (Fig. 2). Firstly, *Plasmodium* parasite grows and matures in the *Anopheles* mosquito's gut, and then passes to its saliva glands. When the infected mosquito bites a person, sporozoites are injected into the human blood. The sporozoites then travel to the liver cells and multiply asexually to produce 10000-30000 merozoites (Greenwood *et al.*, 2008; Miller *et al.*, 2002). This stage of infection lasts about 9-16 days. Following rupture of infected host cells, the merozoites return to the blood and invade the red blood cells where they continue the asexual multiplication. They progressively break down the red blood cells and attack other fresh red blood cells. Meanwhile, merozoites will develop into gametocytes in some red blood cells (Barnwell *et al.*, 1998; Miller *et al.* 2002). During this period, toxins are released which cause characteristic symptoms of malaria including fever, chills and anemia (Aravind *et al.*,

2003). The life cycle of *Plasmodium* parasites is completed when the gametocytes are ingested by mosquitoes. The following mosquito stage initiates firstly in the mosquito's gut where fertilization and sexual recombination of the parasites occur. The oocysts formed there produce the majority of sporozoites. The new released sporozoites then migrate into the mosquito salivary glands again for the next round of infection (Greenwood *et al.*, 2008).

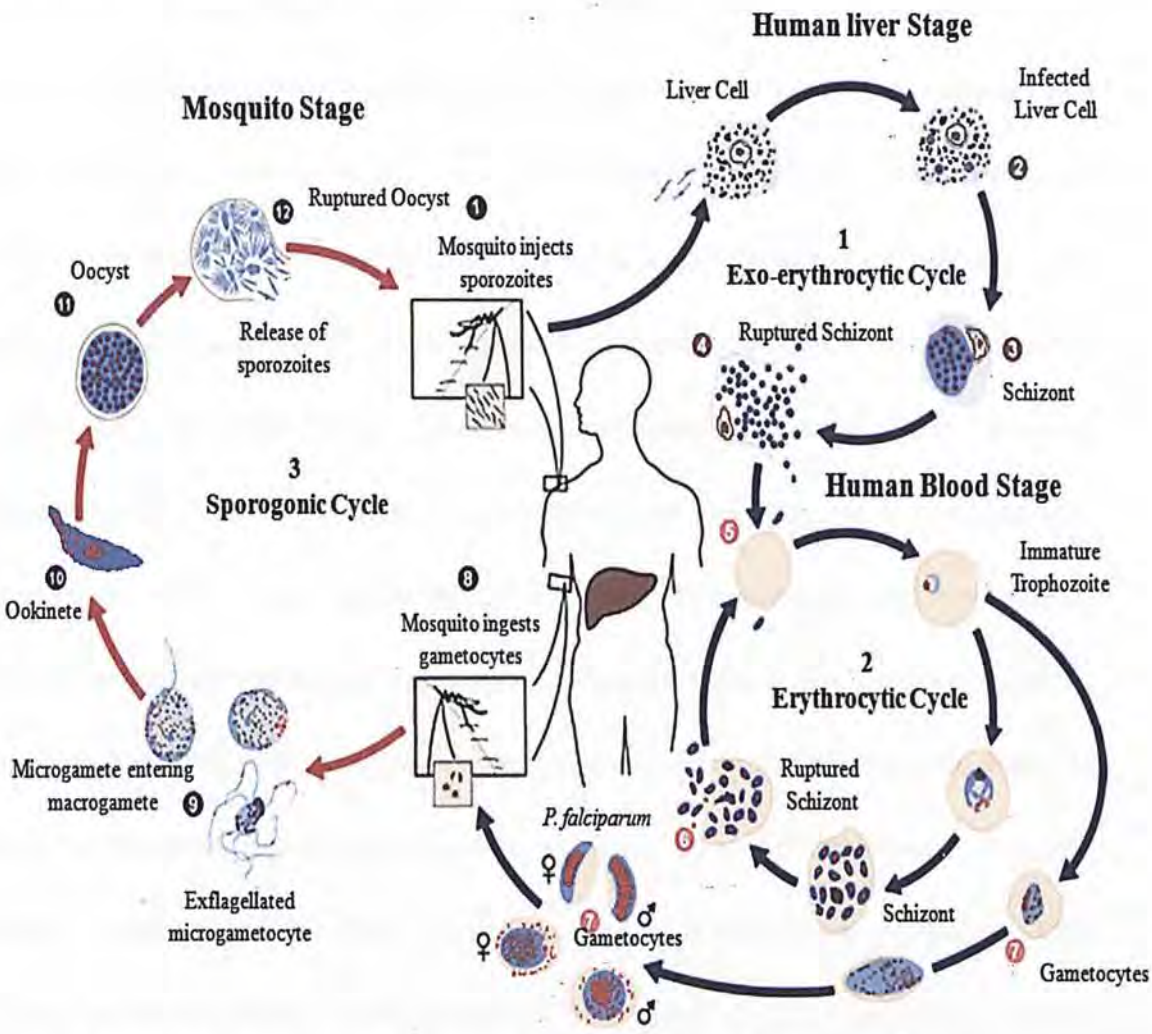

Figure 2 Malaria cycle (modified from http://www.dpd.cdc.gov/dpdx)

1.1.2 The treatment of malaria

In order to prevent the spread of malaria and to protect individuals in endemic areas, tremendous efforts on disease control have been made in the past decades. At the beginning, the fight against malaria depended on two strategies: extermination of mosquito vectors with pesticides such as dichlorodiphenyltrichloroethane (DDT); prevention and treatment of malaria with antimalarial drugs such as quinine and chloroquine (Brown, 2010). Despite of the progress made in some areas such as India and Sri Lanka, the emergence of DDT-resistant *Anopheles* mosquitoes and chloroquine-resistant *Plasmodium* parasites made it more and more difficult to fight against malaria (Ronn *et al.*, 1996). Thus, alternative classes of drugs are urgently needed. From the early 1970s, a group of sesquiterpene lactones, Artemisinin and its semisynthetic derivatives such as artemether and arteether were consequently discovered. They were demonstrated to have a remarkable activity against multidrug-resistant strains of *Plasmodium falciparum* without any reported resistance (Brown, 2010; White *et al.*, 1996). Artemisinin is a constituent of *Artemisia annua* L. which is currently regarded as the only economic source of artemisinin. In 2001, WHO recommended the use of artemisinin-based combination therapies (ACTs) instead of conventional monotherapies in countries where multidrug-resistant *Plasmodium falciparum* presents (Davis *et al.*, 2005; Hallett *et al.*, 2004). The main objective of this strategy is to forestall the emergence of artemisinin resistance, because the chance of a mutant parasite emerging which is resistant to two or more drugs at the same time is lower. Till now, ACTs are still adopted and used as the most

effective method in many countries.

1.2 The Plant of *Artemisia annua* L.

1.2.1. Horticulture

Artemisia annua L. (Fig. 3), commonly known as Sweet Annie, Sweet Sagewort, Sweet Wormwood, Annual Wormwood or Qing Hao, is an annual aromatic herb of the Asteraceae family (Ferreira *et al.*, 1997). It is originated from China and grows wildly in many countries such as Argentina, France, the United States (Klayman, 1989; Klayman, 1993). As a vigorous plant, *A. annua* can adapt to different conditions but grows best in the moist areas. The plant usually has a single stem with alternative branches. Grown in nature, it could reach about 2m tall. Its alternate, fern-like leaves range from 2.5 to 5 cm in length. The nodding flowers (generally 2 to 3 mm in diameter) usually appear greenish or yellowish and are displayed in loose panicles. The pollination of *A. annua* naturally relies on insect and wind (Ferreira *et al.*, 1996; Ferreira *et al.*, 2009).

The 10-celled biseriate glandular trichomes and the 5-celled T-shaped non-glandular trichomes appear on leaves, stems, and flowers. Glandular trichomes contain various types of essential oils including more than 40 volatile compounds and several non-volatile sesquiterpenes, of which artemisinin and its derivatives are the ones possess potential antimalarial properties (Woerdenbag *et al.*, 1994; Charles *et al.*, 1991).

Figure 3 Plant of *Artemisia annua* L.

1.2.2. Historical Importance

The pharmacological use of *A. annua* has a very long history in China. The herb Qinghao was first described in the book "Wu Shi Er Bing Fang" (《五十二病方》, Prescriptions for Fifty-two Diseases)found in the tomb of the Mawangdui Han dynasty (168 B.C.) for the treatment of haemorroids. This plant was mentioned later in the "Zhou Hou Bei Ji Fang" (《肘後備急方》, Handbook of Prescriptions for Emergency) written by Ge Hong in 340 A.D., where it was first recorded for the treatment of fevers. In 1596, Li Shizhen, a famous Chinese herbalist, recommended the use of *A. annua* to treat symptoms associated with fever and malaria in "Ben Cao Gang Mu" (《本草綱目》,Compendium of Materia Medica) (Kuhn *et al.*, 2008; Klayman, 1985).

In 1967, the Chinese government embarked on a systematic investigation of traditional Chinese medicinal herbs in search for effective antimalaria drugs. In 1971, Chinese scientists tested a cold diethyl ether extract from *A. annua* which showed encouraging bioactivity against the malaria parasite *Plasmodium berghei* in mice

(Klayman, 1985; Brown, 2010). Further efforts in 1972 led to the isolation of a new active compound named qinghaosu [QHS, qing hau sau, arteannuin, artemisinin] (Fig. 4) and its chemical structure was determined later by mass spectrometry and X-ray diffraction (China Cooperative Research Group on qinghaosu and its derivatives as antimalarials, 1982). Artemisinin is a sesquiterpene lactone endoperoxide which could be isolated from all aerial tissues of *A. annua*, but highly expressed in flowers (Ferreira *et al.* 1995a). Modern investigations have shown that artemisinin is not only effective against malaria, but also has potential uses in other diseases therapy such as schistosomiasis, cancer, tryptanosomiasis and hepatitis (Xiao, 2005; Efferth, 2006; Kaiser *et al.*, 2002; Paeshuyse *et al.*, 2006).

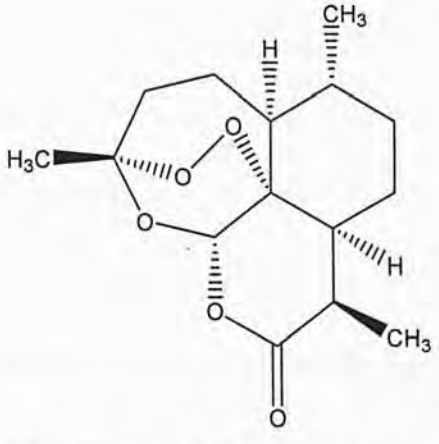

Figure 4 The Chemical Structure of Artemisinin

1.3 Artemisinin

1.3.1 The content and distribution of artemisinin

Artemisinin has been reported to localize in the glandular trichomes on the aerial parts of *A. annua*. The highest yield of artemisinin is found in leaves and flowers while low yield is in stems, which is in accordance with the distribution of the glandular

trichomes. However, neither artemisinin nor its precursor artemisitene was detected in the roots and the pollens (Ferreira *et al.*, 1995a). The concentration of artemisinin in *A. annua* is very low, which is in the range from 0.01% to 1.1% of the dry mass. The concentration mainly depends on the plant materials and growth conditions, and it is also influenced by climate or environment factors, such as temperature and salt (Bhakuni *et al.*, 2001; Ferreira *et al.*, 1997). The variance in artemisinin content is estimated to have a positive correlation with the plant developmental stages (Geldre *et al.*, 1997). According to several reports, there are two stages in which the highest concentration of artemisinin is detected: one is before flowering, while the other is in the full flowering (Abdin *et al.*, 2003; ElSohly *et al.*,1990; Ferreira *et al.*, 1995b). However, studies of the impact of *fpf1* (flowering promoting factor1) gene on flowering time of *A. annua* and the relationship between flowering and artemisinin biosynthesis showed that flowering was not necessary for increasing the artemisinin content, rather the artemisinin content reaches its peak during the period between the late stage of vegetation and the emergence of flowering buds (Wang *et al.*, 2004). The inconsistent distribution of artemisinin throughout the plant is reported by several research groups. In some clones, artemisinins are accumulated more on the top of the plant but are equally distributed in other clones (Charles *et al.*, 1990; Laughlin, 1995).

1.3.2 The biosynthesis of artemisnin

Artemisinin is an endoperoxide sesquiterpene lactone belonging to terpenoids, the largest and varied class of plant secondary metabolites. Terpenoids serve different

functions in plant, such as defense and hormone signaling. Terpenoids are found to be constructed from two precursors: isopentenyl diphosphate (IPP) and its isomer-dimethylallyl diphosphate (DMAPP) (Akhila *et al.*, 1987; Akhila *et al.*, 1990). Two pathways employed in terpenoids biosynthesis are the mevalonate pathway (MVA) in the cytosol and the MVA-independent pathway (also known as MEP/DOXP pathway) in the plastid (Fig. 5). Different sets of isoprenoids are produced in the two pathways: sesquiterpenoids and triterpenoids etc. are produced by the MVA pathway, whereas monoterpenes or diterpenes are derived from the MEP pathway. However, certain extent of crosstalk between these two pathways has been identified (Adam *et al.*, 1998; Laule *et al.*, 2003).

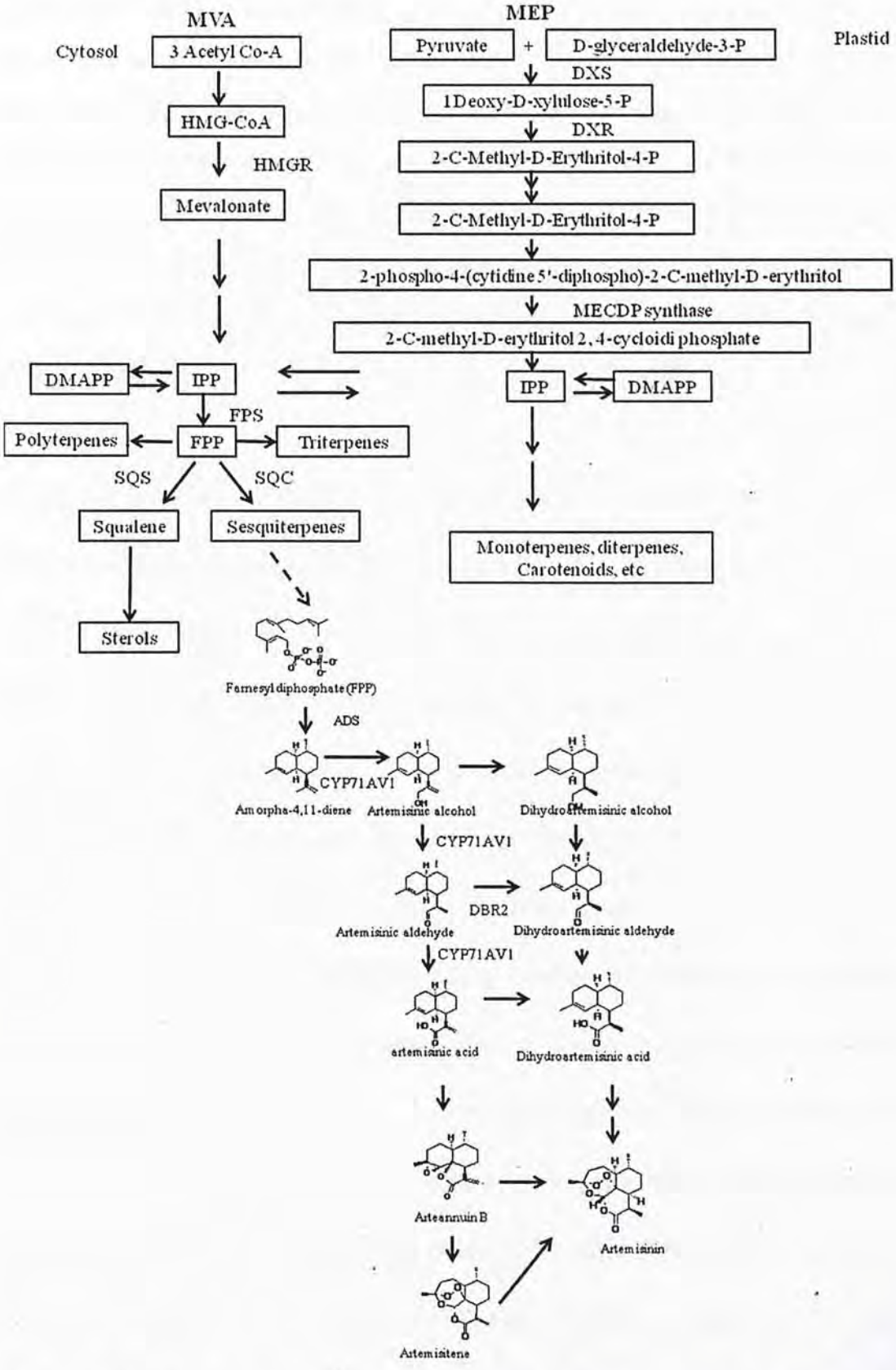

Figure 5 The simplied pathway of terpenoids biosynthsis and the theoretical pathways for the biosynthesis of artemsinin Modified from (Pamela *et al.*, 2006;

Zhang *et al.*, 2008; Liu *et al.*, 2006; Chang *et al.*, 2007; Covello *et al.*, 2007) MVA, 3R-Mevalonic acid; MEP, 2-C-Methy-D-erythritol 4-phosphate; DMAPP, dimethylallyl diphosphate; IPP, isopentenyl diphosphate; FPS, farnesyl diphosphate synthase; HMGR, 3-hydroxy-3-methylglutaryl coenzyme A reductase; DXS, 1-deoxy-D-xylulose-5-phosphate synthase; DXR, 1-deoxy-D-xylulose-5-phosphate reductoisomerase; SQC, sesquiterpene cyclase; SQS, squalene synthase; ADS: amorpha-4,11-diene synthase; CYP71AV1: an enzyme of cytochrome P450 family, which can catalyze multiple oxidations of amorpha-4,11-diene to artemisinic acid; DBR2: artemisinic aldehyde Δ11(13) reductase, a sesquiterpenoid double bond

Despite of the tremendous efforts in the past decade, the biosynthetic routes of artemisinin have been only partially elucidated (Fig. 5). For many years, artemisinin biosynthesis has been estimated to take place in the glandular trichomes. A recent study demonstrated that the two outer apical cells of the glandular trichomes were the exact location where artemisinin biosynthesis occurs (Olsson *et al.*, 2009). The first step of artemisinin biosynthesis starts from the conversion of IPP and DMAPP to farnesyl diphosphate (FPP) by the enzyme farnesyl diphosphate synthase (FPS). FPS catalyzes the "head-to-tail" chain elongation via the reaction of one DMAPP with two IPP and its gene has now been cloned from *A. annua* (Matsushita *et al.*, 1996). Despite both MVA and MEP pathway lead to the formation of IPP and DMAPP, it has been recently demonstrated that the isoprenoid fragment in the FPP precursor is mainly from MEP pathway (Schramek *et al.*, 2010). After FPP synthesis, amorpha-4, 11-diene synthase (ADS) catalyzes cyclization of FPP into (1S, 6R, 7R, 10R)-amorpha-4,11-diene, which was proved by the use of deuterium labeled FPP in 2006 (Kim *et al.*, 2006, Picaud *et al.*, 2006).

The following biosynthetic pathway is the conversion of amorpha-4, 11-diene to putative intermediates: dihydroartemisinc acid and/or artemisinic acid. In 2005, based on the isolation and the identification of several intermediate compounds (artemisinic alcohol, artemisinic aldehyde etc.) from leaves and glandular trichomes of *A.* annua and enzyme assay, Bertea *et al.* proposed a pathway as follows (Fig. 5): amorpha-4, 11-diene is hydroxylated firstly to artemisinic alcohol, followed by the oxidation of artemisinic alcohol to artemisinic aldehyde, and then Δ11(13) double bond of artemisinic aldehyde is reduced to dihydroartemisinic aldehyde, which is finally oxidized to dihydroartemisinic acid (Bertea *et al.*, 2005). This hypothesis was later proved by several research groups. In 2006, two research groups cloned the gene coding an enzyme named cytochrome P450 (Cyp71av1) involved in three sequential oxidation reactions which converts amorpha-4, 11-diene to artemisinic acid (Teoh *et al.*, 2006; Ro *et al.*, 2006). More recently, a recombinant enzyme, artemisinic aldehyde Δ11(13) reductase (DBR2) has been characterized. This enzyme is specific for the reduction of the Δ 11(13) double bond in artemisinic aldehyde and it converts artemisinic aldehyde into dihydroartemisinic aldehyde. This result supports Bertea's hypothesis that the reduction of the Δ11(13) double bond must occur in aldehyde (Zhang *et al.*, 2008). However, the final steps in the artemisinin biosynthetic pathway (from dihydroartemisinic acid to artemisinin) have not been elucidated till now.

Furthermore, the in vivo evidences for other proposed biosynthetic pathways have been reported by other groups. By using the isotope labeled mevalonate, Akhila *et al.*

suggested the sequential pathway starting from farnesyl pyrophosphate (FPP) followed by germacrene skeleton, dihydrocostunolide, cadinanolide, arteannuin B and ending in artemisinin (Akhila *et al.*, 1987). Conversion of artemisinic acid to arteannuin B or artemisinin, and arteannuin B to artemisinin has been reported in a cell free system (Sangwan *et al.*, 1993; Nair *et al.*, 1993).

1.4 Trichomes

1.4.1. Structure and Function of Trichomes

Trichomes are hair cells or leaf hairs originated from the outgrowth of specialized epidermal cells on leaves or other organs of plants. They play an important role in plant taxonomy due to its variance in size, shape, morphology, cell number and composition (Anthony *et al.*, 2008). Normally, trichomes are divided into two general types: non-glandular trichomes (also known as simple trichomes) and glandular trichomes (also called glandular secreting trichomes, GSTs). Non-glandular trichomes are mainly found on aerial surfaces of angiosperms, gymnosperms and bryophytes, whereas glandular trichomes present on almost 30 % of the vascular plants (Wagner *et al.*, 2004). However, it is not uncommon for both types to co-exist on the same individual plant or even on the same organ. The typical representations of non-glandular trichomes are Arabidopsis and cotton fiber trichomes, whereas glandular trichomes in basil and peppermint have been studied extensively (Schilmiller *et al.*, 2008; Turner *et al.*, 2000). Generally, non-glandular trichomes are single-celled with none or less branches. However, glandular trichomes are much

more complex. Some are small structures consisting of a few cells, while others are much larger with differentiated basal, stalk and apical secreting cells.

Trichomes perform a wide range of functions. One of the most remarkable functions of trichomes is that they synthesize, store and secrete various types of metabolites: terpenoids, essential oils, phenylpropanoid derivatives, acyl sugars, methylketones and flavonoids etc (Anthony *et al.*, 2008). Many of these trichome-specified compounds have important commercial values in pharmaceutical, food and cosmetic manufacture. Besides, trichomes also possess significant physical functions. Their presence can affect respiration and transpiration rates by modulating air flowing over the leaf surface and prevent photodamage by blocking light reaching on the leaf surface (Johnson, 1975). On the other hand, non-glandular trichomes and glandular trichomes are proposed to serve different functions. Both of them can form at different development stages of organs and affect plant physiological and ecological aspects. Generally, non-glandular trichomes play important roles in water absorption, seed dispersal, deterring herbivores and so on, whereas glandular trichomes function in the biosynthesis and secretion of phytochemicals to affect host disease and pest attraction (Wagner *et al.*, 2004).

1.4.2 Trichome Investigations in *A. annua*

Two types of trichomes co-exist on the aerial surface of *A. annua* L.: 10-celled biseriate glandular secreting trichomes (GSTs) and non-glandular 5-cell T-shaped

non-glandular trichomes (Fig 6). The morphology and development of trichomes on leaves and floral buds of *A. annua* has been studied by light and electron microscopy (Duke *et al.*, 1993; Ferreira *et al.*, 1995a). The non-glandular trichomes are generally abundant on leaves, stems but not on florets. They don't produce essential oils and their functions are proposed to be similar to those of Arabidopisis. However, glandular trichomes are found throughout the aerial parts of the plant and they are densely distributed on the flower buds. Glandular trichomes of *A. annua* develop very early in the life of the leaf (Duke *et al.*, 1993). The 10 cells of glandular trichomes are divided into one basal, one stalk and three secretory pairs, all of which are distinct from each other upon TEM analysis. Under light microscope, only the two sub-apical cell pairs of secretory cells are green while other cells are colorless (Covello *et al.*, 2007). When glandular trichomes become mature, the apical cell pair contains leucoplasts or proplastids but no thylakoids, whereas the two sub-apical cell pairs contains chloroplasts which show autofluorescence (Olsson *et al.*, 2009).

Glandular trichomes are thought to be the site of biosynthesis and accumulation of artemisinin-related sesquiterpenoid compounds. It is supported by the evidence that neither artemisitene nor artemisinin could be detected from an *A. annua* glandless biotype (Woerdenbag *et al.*, 1994). Duke *et al.* reported that 97% of the artemisinin and 100% of artemisitene could be extracted by a 5-sec leaf dip in chloroform without visible damage to other leaf epidermal cells (Duke *et al.*, 1994). In a recent study, the exact location where artemisinin biosynthesis occurs was found to be the two outer

apical cells (Olsson *et al.*, 2009). The researchers separated the apical and sub-apical cell pairs of glandular trichomes using laser microdissection pressure catapulting method and investigated the expression of several transcripts using cDNA from those separated cell pairs. Three enzymes involved in artemisinin biosynthesis (ADS, CYP71AV1 and DBR2) were expressed only in the apical cells. The researchers speculated that 5 cell pairs of glandular trichomes played different roles in the metabolism of artemisinin. Besides, majority of essential oils (e.g. a-pinene, artemisia ketone, 1,8-cineole and camphor) are also thought to be excreted and stored in the subcuticular space of glandular trichomes. In addition, the content of artemisinin was reported to be 4 to 11 times higher in the inflorescence compared to leaves (Ferreira *et al.*, 1995b) and the correlation of the presence and development of glandular trichomes in the inflorescences with artemisinin production has also been extensively studied (Ferreira *et al.*, 1995a). Kapoor and his coworkers inoculated *A. annua* with two arbuscular mycorrhizal (AM) fungi, *Glomus macrocarpum* and *Glomus fasciculatum* and found a significant increase of artemisinin concentration associated with an increase of glandular trichome density on leaves (Kapoor *et al.*, 2007).

Currently, the method of isolation glandular trichomes based on the mechanized abrasion enables researchers to get further insight into the function of glandular trichomes. Several genes encoding enzymes involved in the isoprenoid biosynthesis have been cloned and functionally characterized, these include germacrene A synthase, farnesyl diphosphate synthase, cytochrome P450 CYP71AV1 and artemisinic

aldehyde Δ11(13) reductase (Gershenzon *et al.*, 1992; Teoh *et al.*, 2006; Zhang *et al.*, 2008). With gas chromatography–mass spectrometry (GC-MS), liquid chromatography–mass spectrometry (LC-MS) or nuclear magnetic resonance (NMR), many intermediate compounds (e.g. artemisinic alcohol and dihydroartemisinic alcohol) participating in artemisinin production has been identified (Bertea *et al.*, 2005). These cell extracts from glandular trichomes showed no significant difference to those obtained from leaves. Besides, recent global surveys of glandular trichomes transcriptome in *A. annua* were conducted using high throughput pyrosequencing technology (Wang *et al.*, 2009; Graham *et al.*, 2010). Such efforts have provided a useful resource of ESTs for gene discovery.

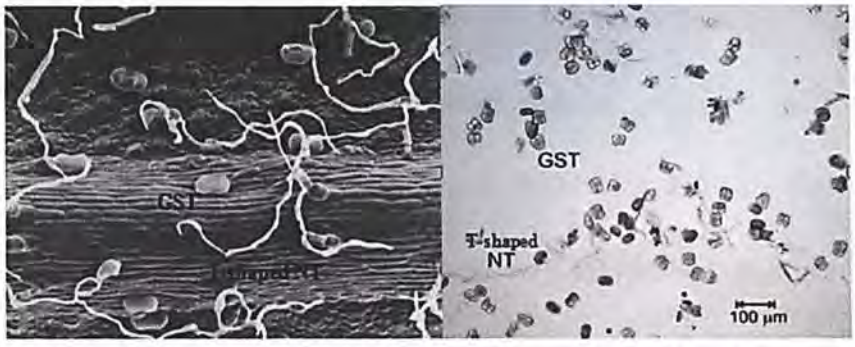

Figure 6 Scanning electron micrograph of an *A. annua* leaf (Left) and Light micrograph of non glandular trichomes and glandular trichomes from a floral bud (Right) Modified from (Duke *et al.*, 1993; Covello *et al.*, 2007). NT: non-glanduar trichome; GST: glandular secreting trichome

1.5 Proteomics

1.5.1 The basics of proteomics

Proteomics has made rapid progress during the last decades. Complement of the analysis of the transcriptome and the metabolome, proteomics stands for the systematic analysis of the proteins, particularly their structures and functions. Its

objective includes protein profiling, protein quantification, posttranslational modifications (PTMs), protein/protein interactions, subcellular localization and activities.

New transcriptomics technologies provide important information regarding complete sets of RNA transcripts which are produced by the genome at given time. However, due to the complex and dynamic regulatory routes from PTMs to protein accretion, it is difficult to predict the abundance of cognate proteins from the mRNAs level. The proteome approach is necessary to reveal the functional information of the biochemical processes underlying phenotypes which is not accessible by other means (Zivy *et al.*, 2000). Many reviews have addressed the problems, pitfalls, and the developments of plant proteomes (Newton *et al.*, 2004; Hisashi *et al.*, 2004; Baginsky *et al.*, 2009; Wijk, 2001). They provide a well-established platform for plant researchers.

1.5.2 Two-dimensional gel electrophoresis

Two-dimensional gel electrophoresis (2-DE) has long been recognized as a powerful and standard proteomics technique for the analysis of complex protein mixtures extracted from biological samples such as cells, tissues. 2-DE separates proteins in two dimensions according to two independent properties. Protein mixtures are first separated according to their isoelectric points (pI) by isoelectric focusing (IEF). Then proteins are separated further according to their molecular weights (MW) by

SDS-polyacrylamide gel electrophoresis (SDS-PAGE). In this way, thousands of different proteins can be separated at one time and the relative amount of each protein can be determined. Currently, the broad pI range (typically pI 3–10) and sensitive staining method (e.g. silver staining) has increased the resolution and the resolving power. However, some challenges still exist for 2DE technique, particular with respect to the detection of low-abundance proteins and membrane proteins. Thus, further researches and improvement are needed.

1.5.3 Mass spectrometry

Mass spectrometry is an essential analytical tool for protein identification. A mass spectrometry instrument generally consists of ion source, mass spectrometer and ion detector. Two forms of mass spectrometry, based on "soft" ionization methods, are commonly used. The first one is matrix-assisted laser desorption ionization (MALDI)-time-of-flight (TOF) mass spectrometry, which is used for high-throughput protein identification based on peptide mass fingerprinting (PMF). The second one is electrospray ionization (ESI) tandem mass spectrometry (MS/MS), usually combined high performance liquid chromatography (HPLC) separation and used to produce an amino acid sequence tag (Russell *et al.*, 2004). The standard procedure of protein identification based on MS analyses is as follows: the protein spots of interest are first excised from the 2D gel and digested with a site-specific protease (commonly trypsin); finally the resulting mixture of peptides was analyzed by MS. Then the experimental PMF or MS/MS spectrum is matched against a calculated PMF or MS/MS spectrum in the public database (e.g. Mascot, SEQUEST). At last, a score is calculated

automatically for each match and a threshold is given in the final report. To be significant, the score has to be higher than the threshold ($P<0.5$).

1.5.4 Gel-free Proteomics

Although 2-DE has been widely used and regarded by far as the best technique for analysis of post-translational modifications of proteins, it has several limitations including the low throughput of samples, incapability in detecting proteins with relatively low abundance, extreme pI or hydrophobicity, and small or large molecular weights. Thus, alternative "gel free" proteomics techniques have been developed to improve resolution of complex protein samples. Recently, the use of these techniques has been increased in the plant proteome study such as rice (Koller *et al.*, 2002; Jorrín-Novo *et al.*, 2009). One of the well known examples of gel-free techniques is multidimensional protein identification technology (MudPIT), which was first used by Washburn's group to analyze membrane spanning proteins (Washburn *et al.*, 2001). Unlike 2DE, MudPIT requires that all proteins of a sample are firstly digested before the separation. The resulting mixture of peptides is then separated on a strong cation exchange (SCX) column in the first dimension followed by a reverse phase chromatography (RP) in the second dimension. Finally, the separated peptides are analyzed by mass spectrometer. Based on this technique, total 1484 proteins were detected and identified from *Saccharomyces cerevisiae* in Washburn's study, including 131 proteins with transmembrane domains (Washburn *et al.*, 2001). Besides, MudPIT has also been reported to be used in quantitative and differential comparative proteomic study (Gustavsson *et al.*, 2005; Hu *et al.*, 2006). Therefore, MudPIT has

been proven to be an effective tool in increasing protein coverage of a sample, analysis of protein modifications in protein mixtures, and the identification of membrane proteins. However, the large number of spectra produced by MudPit analysis lays heavy burdens on the data storage and analysis. In addition, only a few fractions of peptides obtained from digestion of hundreds of proteins in a sample will be selected for MS/MS analysis. Thus, in order to circumvent these random sampling errors, multiple analyses are usually required.

Other gel-free proteomic approaches such as Isotope coded affinity tag (ICAT) and chip-based techniques have been summarized by several groups (Zhu *et al.*, 2003; Baggerman *et al.*, 2005).

1.6 Objectives

Although tremendous efforts have been dedicated to study of *A. annua* and several accomplishments have been obtained these years, it is still difficult for researchers to get an explicit picture about the biosynthesis of artemisinin and the function of glandular trichomes. Thus, in order to discover the cellular mechanism operating in glandular trichomes of *A. annua* and enzymes involved in the artemisinin biosynthesis, we initiated the proteome study of glandular trichomes of *A. annua*. There are three objectives in this study:

1. Conduction of proteomics study of Artemisia glandular trichomes and setting up proteins database.

2. Comparative study of proteins from leaves and glandular trichomes to find the trichome-specific proteins.
3. Through proteomic study, it helps us to discover candidate genes which provide potential targets for genetic manipulations for enhanced artemisinin production in *A.annua.*

Chapter 2 Materials and Methods

2.1 Chemicals

All chemicals used in this study are of AR available from Sigma-Aldrich Chemical Co. (USA), Bio-Rad Co. (USA), USB Corporation (USA), BDH Chemicals Ltd. (England), or Amresco (USA).

2.2 Plant materials

The seeds of *A. annua* were purchased from You Yang, Si Chuang Province, China. The seeds were geminated under controlled condition (average 30°C) in the greenhouse of the Biology Department, Hong Kong Baptist University. Leaf samples of 16-to-18 week old plants were harvested and used for trichomes isolation and microscopy. Leaf samples were harvested from the same plants. Three independent groups of plant materials were included in the study.

2.3 Isolation of glandular trichomes

Glandular trichomes were isolated from leaves of *A. annua* plants as described previously (Teoh *et al.*, 2006) with minor modifications. The isolated glandular trichomes were then subjected to additional sucrose gradients purification. Leaves were collected carefully and imbibed with ice-cold distilled water for 1h. Subsequently, 20–30 g of leaves was placed into a 350 ml chamber of Bead Beater (BioSpec Products, Inc., Bartlesville, OK, USA). The 350 ml chamber was then filled with 80-100 g of glass beads (0.5 mm diameter) (BioSpec Products, Inc., Bartlesville,

OK, USA), XAD-4 resin (1 g/g plant material) (Amberlite® XAD® 4, SUPELCO, Sigma-Aldrich Chemical Co., Bellefonte, PA, USA), and isolation buffer (25 mM MOPSO, pH 6.6, 200 mM sorbitol, 10 mM sucrose, 5 mM thiourea, 2 mM dithiothreitol, 5 mM $MgCl_2$, 0.5 mM sodium phosphate, 0.6% (w/v) methylcellulose and 1% (w/v) polyvinylpyrrolidone (PVP) (Mr 40000). The trichomes were gently abraded from the leaves on ice for 3x1min with 1 min intervals. After abrasion, the crude cellular extract was separated by sequentially filtering through a 300 μm and a 105 μm nylon mesh (Small Parts Inc., Miami Lake, FL, USA). The residual plant materials were rinsed twice with rinse buffer (Isolation buffer without PVP and methylcellulose). The resulting solution of mixed cells was then concentrated by centrifugation (10 min, 4°C, 6000g) using an eppendorf 5810R centrifuge (Eppendorf, Germany) and was resuspended in 15-20 ml pre-cold rinse buffer. 3-5 ml of the concentrated mixture was gentled layered on the top of 30ml 40% sucrose solution along the wall of 50 ml centrifuge tube (Neptube, Germany). The tube was then put into the swing-bucket rotor and was centrifuged at 4°C for 10min at 750rcf (with acceleration levels 5 and deceleration level 0). Consequently, upper layer was carefully transferred to the new tubes and was concentrated. After being washed for 3 times with ice-cold distilled water, the isolated trichomes were ready for protein extraction and microscopic analysis. Leaf samples and trichomes were all frozen in liquid nitrogen and stored at -80°C.

2.4 Protein Extraction

The extraction procedure was performed as described in previous studies with minor modification (Görg A. *et al.*, 2000; Amme *et al.*, 2005). The frozen plant material was ground into fine powder with a mortar under liquid nitrogen. To precipitate the proteins, the obtained powder was incubated with 10 times (m/v) of precipitation solution containing 10% w/v trichloroacetic acid and 0.07% w/v 2-mercaptoethanol in ice-cold acetone and kept at -20°C overnight. The material was then centrifuged at 14000g for 30 min at 4°C. After centrifugation, the protein pellets obtained were washed 2~3 times with ice-cold acetone containing 0.07% w/v 2-ME and freeze-dried. Dried samples were re-suspended in lysis buffer (7M urea, 2M thiourea, 4% CHAPS, 100mM DTT and 2% Carrier Ampholyte solution) and incubated at room temperature for 2hrs with 10 min sonication per hour. The insoluble material was removed by centrifugation at 14000g for 60 min at 4°C. The supernatant was then purified by centrifugation through 0.45 mm filter units (Ultrafree-MC, Millipore, USA). Protein concentration was measured using a quick start bradford dye reagent (Biorad, USA) with BSA (4 $mg.ml^{-1}$) as standard.

2.5 Two-dimensional gel electrophoresis

350µl of protein solution containing 0.002% v/v bromophenol blue was used to rehydrate IPG strips of 18 cm with a linear pH gradient of 4-7 (Readystrip, Biorad, USA). The solution was then incubated for 20hrs at room temperature. IEF was carried out on an Ettan IPGphor unit (GE Healthcare, USA) using the following

settings: 100 V for 3 h, 300V for 3 h, 1000 V for 1 h, 3000 V for 24000 Vhr and 8000 V for 64000 Vhr at 20°C with a maximum current setting of 50 mA/strip. After the isoelectric focusing, the IPG strips were first equilibrated for 10 min in equilibration buffer (50 mM Tris, pH 8.8, 6 M urea, 20% v/v glycerol, 2% w/v SDS) containing 2% w/v DTT, and equilibrated for another 10 min with the same buffer containing 2.5% w/v iodoacetamide. After equilibration, the IPG strips were placed on the top of 14% polyacrylamide gel and subjected to electrophoresis in a vertical Protean Lixi Cell electrophoresis system (Biorad, USA) at 80V for 1hr and 200V for approximate 7hrs. After electrophoresis, the gels were washed with Milli-Q water and then fixed for 1 hr with a fixation buffer (40% menthol, 10% acetic acid). After fixation, the gels were stained overnight with "Blue silver" staining solution (10% phosphoric acid, 10% ammonium sulfate, 20% anhydrous methanol, 0.12% Brilliant Blue G-250). Finally the gels were distained using 500ml Milli-Q water for 30 min and the operation was repeated 3 times.

2.6 Image analysis

After staining and distaining, the gels were scanned under visible light at 300dpi using ImageScanner (Amershan Biosciences) with Labscan software. Analysis was carried out using the computer program, ImageMaster 2D Platinum 5.0 (Amershan Biosciences). Spots were detected automatically while manual spot editing and deleting was performed if necessary. Distinct spots were selected throughout the gel for alignment and matching of the spots after the mass-spectrometry analysis. The gel

with the protein spots from the glandular trichomes were used as the reference gel.

2.7 In gel digestion and protein identification by mass spectrometry

Spots of interest were excised from the 2-DE gel obtained from glandular trichomes after image analysis. The spots transferred to 1.5ml microcentrifuge tubes were distained with 50mM ammonium bicarbonate, dehydrated with ACN and then vacuum dried completely by a SpeedVac (LABCONCO). For the rehydration and digestion of proteins, 5-7μl 25mM NH_4HCO_3 containing 40ng/μl sequence grade modified trypsin (Promega, Madison, USA) were added to each sample and samples were incubated for 30 minutes on ice. In-gel digestion with trypsin was continued over night at 37°C. After sonication for 10min, the supernatant was removed to a new tube. Afterwards, the reaction was stopped by adding 10μl of extraction buffer (50%, 2.5% trifluoracetic acid), followed by sonication for another 10 min. The extracts were combined and completely dried under vacuum and dissolved in 5μl 0.1% trifluoracetic acid. Digested sample solution (1.5μl) was spotted on MALDI-TOF disposable target plates (4800, Applied Biosystems, Foster City, CA, USA), then covered with 0.5μl matrix (α-cyano-4-hydroxycinnamic acid). Peptide mass determinations were carried out using a MALDI-TOF/TOF mass spectrometer (Applied Biosystems 4800 Proteomics Analyzer, Applied Biosystems) in the m/z range 900 to 4000. Proteins were identified by searching against the NCBI database and our in-house *A. annua* EST database using the following parameters:

monoisotopic mass accuracy, 50 ppm; missed cleavages, 1; allowed variable modifications, oxidation (Met). Proteins with a Mascot score > 69 were selected.

Chapter 3 Results and Discussion

3.1 The isolation of glandular trichomes

Trichome isolation is a critical step in constructing the *A. annua* glandular trichomes proteins database for identification of trichome-specific proteins. Previous study demonstrated that the mechanical abrasion method was suitable for the separation of trichomes from *A. annua* flower buds (Teoh *et al.*, 2006). Our group also succeeded in establishing an EST database for glandular trichomes of *A. annua* using similar method (Wang *et al.*, 2009). However, it takes more than 6 months to harvest the flower buds and the trichomes exist only few days before senescence occurred. This has created a bottleneck for the large-scale proteome study. Thus, leaves were selected as candidates in our study.

Two important factors affect the successful isolation of glandular trichomes from leaf materials: harvest time and isolation method. According to our previous study, the density of glandular trichomes on the surface of leaf increases dramatically from the 3rd month and remained stable till the inflorescences stage (Fig. 7). This result is consisted with previous reports (Ferreira *et al.*, 1995a; Ferreira *et al.*, 1996). Thus, 4-month old plants were subjected to the proteome study. For the isolation method, reproducible results were reported by applying Bead Beater abrasion to isolate the trichomes from mint and tobacco leaves (Lange *et al.*, 2000; Steffen, *et al.*, 2005). This method is suitable for these plants for two reasons: the leaves are large and thick; the trichomes lay flat on the surfaces of leaves. Thus, the beads are able to access the

trichomes easily. However, the leaves of *A. annua* are generally smaller and deeply dissected. The glandular trichomes are slightly embedded into the leaf surface and hence the glass beads would have difficulty in contacting the trichome's base. In a test experiment, we found that the quality of isolated glandular trichomes was low usually contaminated with a higher percentage of mesophyll or cell debris (Fig. 8) using Bead Beater abrasion method. 2DE patterns of proteins from leaves and trichomes had no significant difference (Fig.9) and over 50% of identified proteins were Rubisco (data not shown). Therefore, an additional sucrose gradient was attempted for further purification after the mechanical abrasion and filtration. Sucrose with different concentration and different centrifuge speed were tested. A 40% sucrose solution with centrifugation speed at least 750rcf yielded better results: more than 90% of the preparations were glandular trichomes with the remaining 10% being cell debris and no T-shaped trichomes were detected (Fig. 8). The quality of the preparation was appropriate for protein extraction.

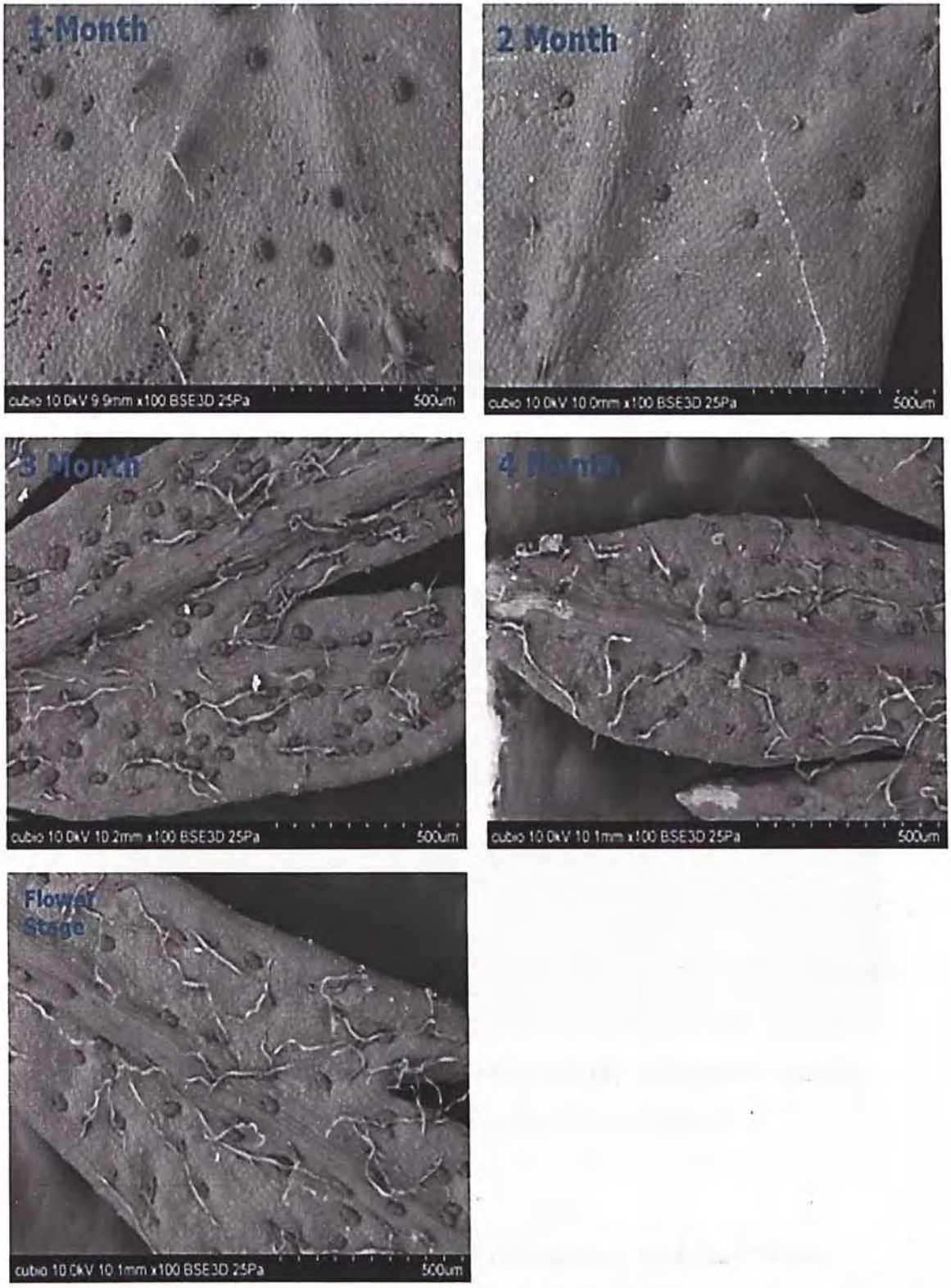

Figure 7 The trichomes distribution on leaf surface of *A. annua* L. in different developmental stages

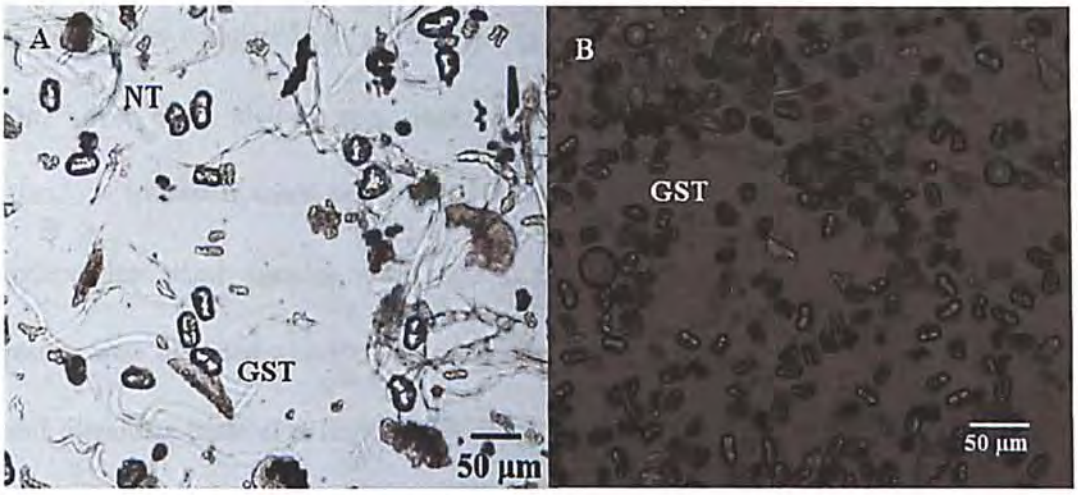

Figure 8 The isolated glandular trichomes of *A. annua* recovered from 30 µm mesh (A) and sucrose gradient purification (B) (NT, non-glandular trichomes; GST glandular trichomes)

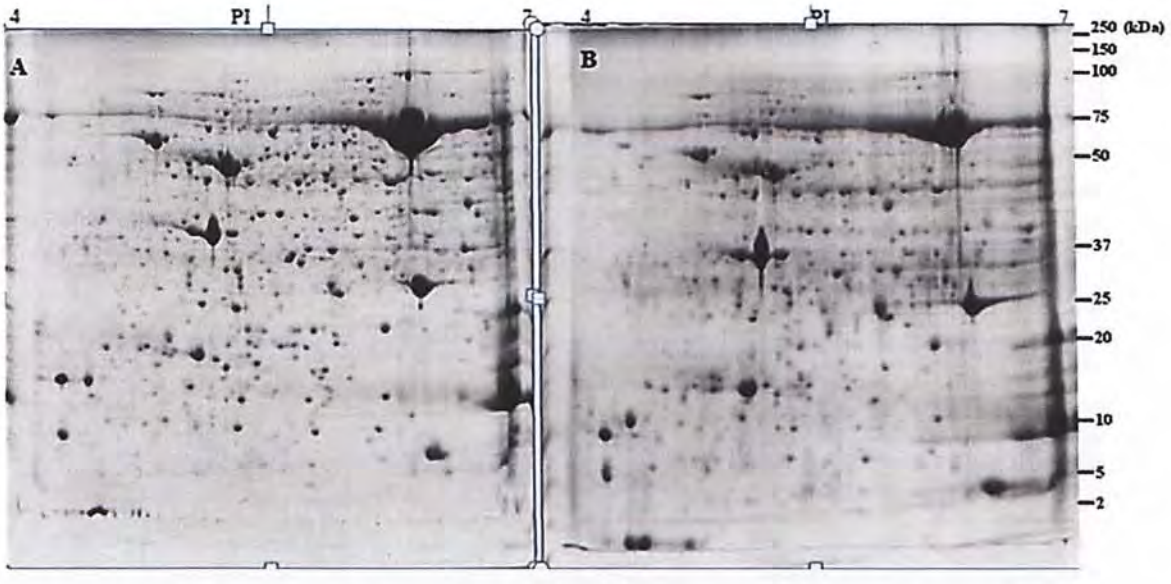

Figure 9 2-DE patterns of leaves (A) and trichomes (B) of Artemisia proteins Trichomes were isolated using the Bead Beater method. kDa:kiloDalton

3.2 2DE patterns of *A. annua* leaf trichomes and leaf tissue

In order to set up a glandular trichomes protein database and detect highly expressed trichome-specific proteins, the 2-D gel patterns of the leaves and glandular trichomes

proteins were compared. As showed in Fig. 11, distinct gel patterns were observed. On "silver blue" stained gels, over 1000 distinct protein spots from leaves and glandular trichomes could be separated between the pI ranges of 4-7. Three biological independent plant samples were used to extract proteins and highly reproducible results were obtained with 95% proteins at the same relative position on gels of leaves and glandular trichomes respectively (Fig. 10). Therefore, protein spots (including those highly expressed ones) from gels of glandular trichomes were excised for MS analysis.

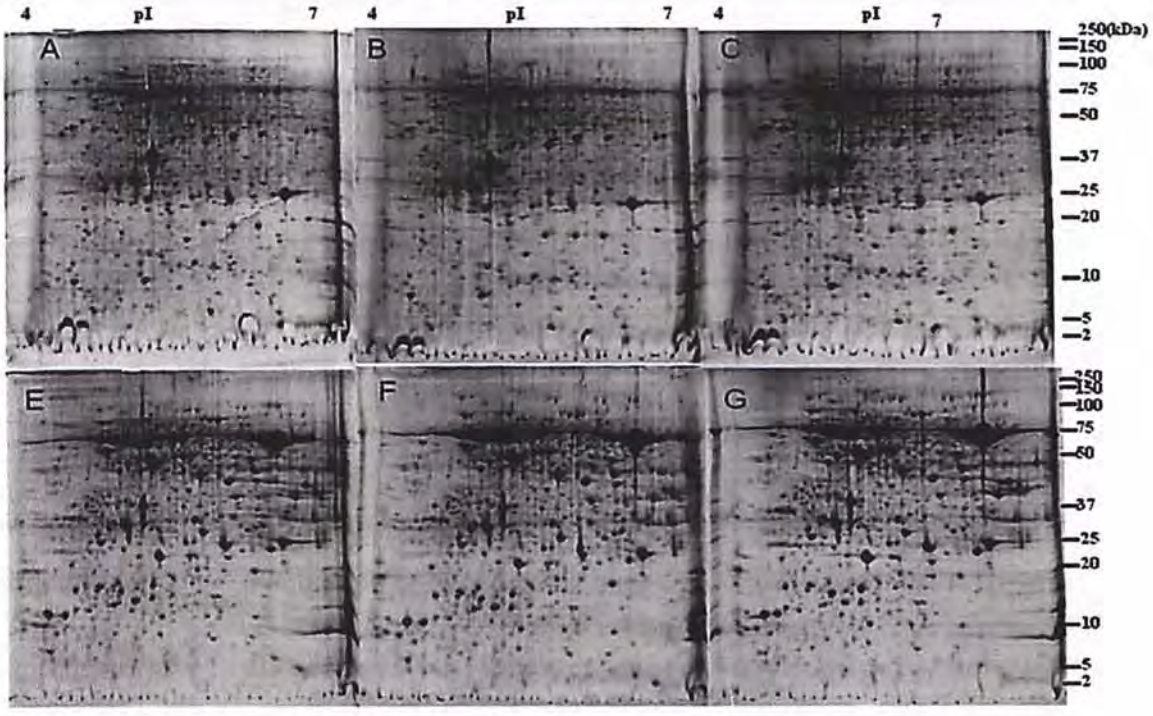

Figure 10 2DE patterns of proteins from trichomes (A, B, C) and leaves (E, F, G) (Three biological repeats were obtained in the study). kDa:kiloDalton

3.3 Identification of proteins in glandular trichcomes

This is the first report providing a global view of proteins expressed in glandular trichomes of *A. annua*. Of over 1000 spots on 2D gel of glandular trichomes, 738 spots were excised. We searched both NCBI database using green plant taxonomy and our in-house EST database. Out of these 738 spots, 206 spots (28%) were identified in the NCBI database, whereas 162 (22%) spots were found in EST database. These spots represented ~102 non-redundant proteins. More than 70% of the protein spots analyzed could not be assigned mainly due to the limited *A. annua* genome sequence and ESTs information, and partially due to the low abundance of those proteins or the poor quality data generated by MS analysis. On the other hand, a large number of spots with different pI/Mr values were attributed to the same protein. The reason has been reported as follows: post-translational modification, isozyme variation, protein degradation, alternative splicing and allelic variation of the same protein (Mathesius *et al.*, 2001).

Among ~102 non-redundant proteins, the most abundant proteins were those involved in photosynthesis (oxygen-evolving enhancer protein, ribulose-1,5-bisphos- phate carboxylase, chloroplast light-harvesting chlorophyll a/b-binding protein, photosystem II reaction center psbP protein, oxygen evolving protein of photosystem, Manganese-stabilizing protein, photosystem I subunit VII, photosynthetic electron transfer-like protein, Ribulose bisphosphate carboxylase/oxygenase activase, ribulose-1,5-bisphosphate carboxylase/oxygenase large/small subunit), which was

consistent with previous reports that the glandular trichomes of *A. annua* had two pairs of green secretary cells where photosynthesis existed (Bertea *et al.*, 2006; Covello *et al.*, 2007). These proteins were omitted in following analysis. Out of the rest 91 proteins, 69 were positively matched to the NCBI databases and 54 proteins were positively matched our in-house EST databases (Fig 12). Among the 22 proteins that were only identified from our Artemisia EST database, 14 had no homolog with other species even with quite loose blast parameters, indicating they might be both species-specifically and tissue-specifically expressed.

Changes in spot intensity between glandular trichomes and the whole leaves were also qualified by software analysis. Over 70% of the 91 non-redundant proteins were highly expressed in glandular trichomes. Among them, one of the most important one was artemisinic aldehyde delta-11(13) reductase (Dbr2), an enzyme directly involved in artemisinin biosynthesis in *A. annua*. Furthermore, in our previous study on isolation of trichomes by Beat Beater method, 4-diphosphocytidyl-2-C-methyl-D-erythritol synthase (MECDP-synthase) implicated in the MEP pathway were also identified. Both of them are enriched in the glandular trichomes. The identification of these two enzymes could be considered as a proof-of-concept for the proteome study. The list of proteins highly expressed in trichomes was showed in **Figure 11** and **Table 1**.

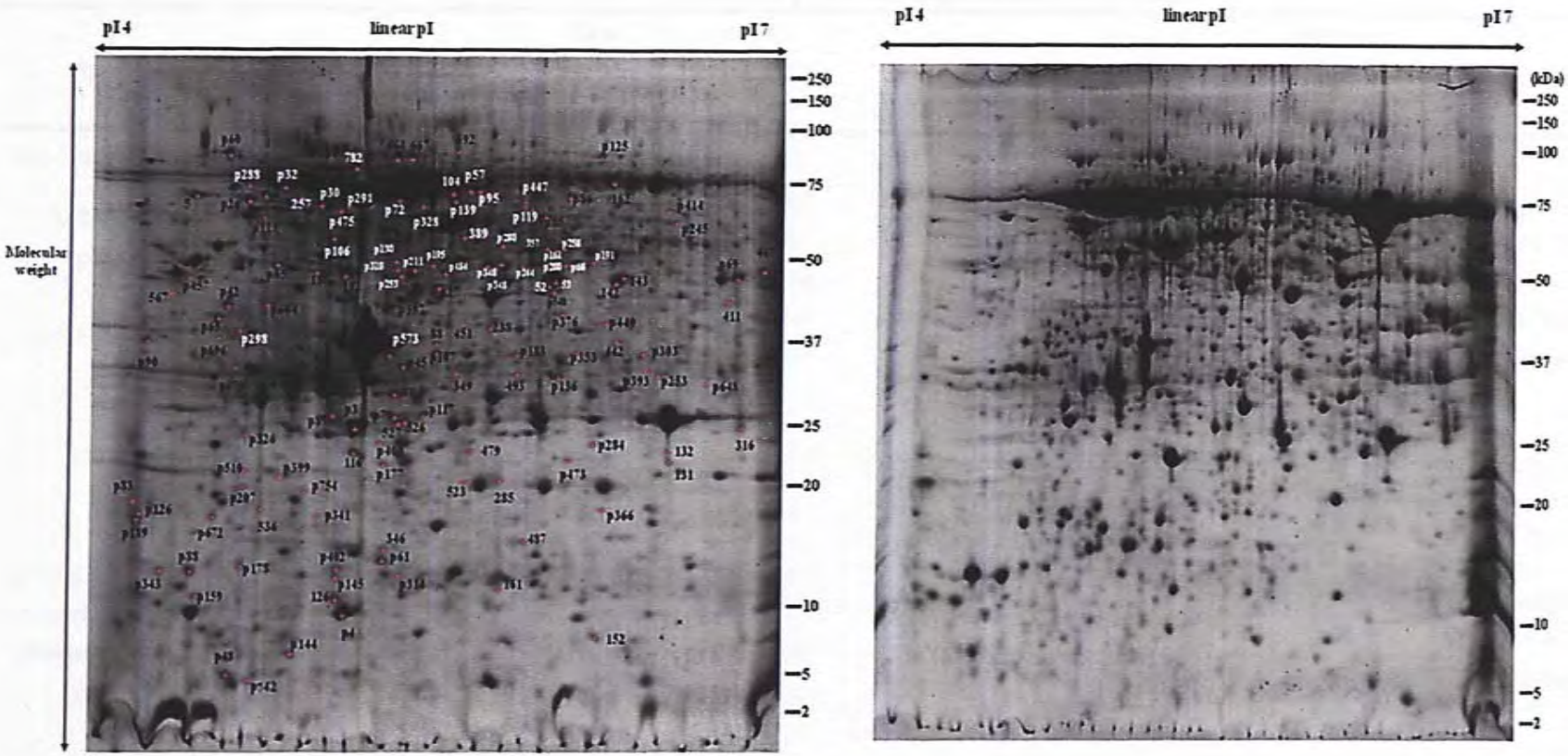

Figure 11 2DE patterns of trichomes (left) and leaves (right) proteins 91 non-redundant proteins were shown on the gel of glandular trichomes. Over 70% of them were highly expressed in the glandular trichomes. kDa:kiloDalton

Table 1 identified proteins in glandular trichomes

Spot ID	Changes[a]	pI/Mr	Best protein score[b]	Homolog_GI	Protein Name	EST best protein score[b]	Best matching EST ID
Electron transport chain							
p684|527|p95|p136|p79	up	5.20/55482	472	gi|94502485 gi|114522 gi|115502358	ATP synthase CF1 alpha subunit [Helianthus annuus]	108	092121_2539_3868_3r
p314|102|88|161	up	5.42 /13772	443	gi|81176256 gi|22795943	ATP synthase CF1 epsilon subunit [Lactuca sativa]	115	132965_1015_3960_2p_3 Contig8503_3p_21
p402|p145	up	5.42 /13772	358	gi|6635347	ATPase epsilon chain [Lactuca sativa]	94	Contig16137_3p_17
p125	up	5.86 /69612	239	gi|15240075	SDH1-1; ATP binding / succinate dehydrogenase [Arabidopsis thaliana]		
p195|p280|p65|p696|p72|p87|p298|p675|p604|p144|p42|p316|p257|p288	up	5.11/52197	384	gi|5758861 gi|7708518 gi|153012228 gi|241912832	ATP synthase CF1 beta subunit [Medicago truncatula]	391	Contig8503_1p_4 Contig2065_3p_1 Contig14249_3r
p3	up	5.04/33964	134	gi|224159128	predicted protein [Populus trichocarpa]		
p57	up	7.87/213765	350	gi|218194450	hypothetical protein OsI_15081 [Oryza sativa Indica Group]		
p440|238|162	up	6.02/55452	440	gi|114404 gi|20146574	ATP synthase subunit alpha, mitochondrial	181	141790_3502_1279_3r_1 Contig35597_2p_6

(Continued)

Spot ID	Changes[a]	pI/Mr	Best protein score[b]	Homolog_GI	Protein Name	EST best protein score[b]	Best matching EST ID
p400|p177|p284|p341|116	up	5.31/19707	176	gi|192910736	mitochondrial F0 ATP synthase D chain [Elaeis guineensis]	338	Contig3641_2p_4
104		5.63/20241	157	gi|210076131	mitochondrial ATP synthase beta subunit [Dimocarpus longan]	132	048139_1764_1521_1r
p161|p258	up	5.95/40049	175	gi|83777062	ATP synthase gamma subunit [Phaseolus vulgaris]	233	Contig13579_3p_4
782|p102	up	5.14/55538	169	gi|225465149	PREDICTED: hypothetical protein [Vitis vinifera]		
p70	up	8.98/26768[c]	104	gi|416681	ATP synthase delta chain, chloroplastic	110	Contig4565_2p_3
667	up	5.29/68792	331	gi|224109966	V-type (H+)-ATPase V1, A subunit	144	Contig5858_1r
52|53|142|143		8.37/40419	227	gi|61969078 gi|222423270	putative ferredoxin-NADP reductase [Lycopersicon peruvianum]	136	026370_1594_0779_2p_1 214374_0319_3906_3r_1 Contig23006_1p
p510	up	5.10/23523	138	gi|209573110	NAD(P)H-quinone oxidoreductase subunit M		
p178	up	9.19/64123	111	gi|255548890	hypothetical protein RCOM_0923760 [Ricinus communis]		
p90*	up	NA[d]	86	gi|168061725	cytochrome C oxidase polypeptide vib	138	Contig22469_3p_3

(Continued)

Spot ID	Changes[a]	pI/Mr	Best protein score[b]	Homolog_GI	Protein Name	EST best protein score[b]	Best matching EST ID
Transcription and Translation							
692	up	4.63/37643	323	gi|224172764	predicted protein [Populus trichocarpa]		
p672	up	6.34/20309	77	gi|133088	50S ribosomal protein L12	72	Contig33808_1p_2
p183	up	7.05/29063	156	gi|82621186	transcription factor APFI-like [Solanum tuberosum]	94	Contig31551_1p_1
5		5.15/2643	103	gi|399305	Calreticulin	87	Contig33337_2r_2
p69	up	5.15/172051	81	gi|196128985	RNA polymerase beta' chain [Chlorella vulgaris]		
188		5.10/40050	72	gi|147837944	hypothetical protein VITISV_044014 [Vitis vinifera]		
p291	up	5.41/12112	110	gi|160895751	nitrogen regulatory protein P-II GLUTAMINE		
p83	up	5.30/17121	88	gi|85376261	eukaryotic translation initiation factor eIF5A	98	127634_0717_1720_3r
Metabolism							
p60	up	5.20/72856	145	gi|168044813	predicted protein [Physcomitrella patens subsp. patens]		
479		NA			mangrin [Bruguiera sexangulamangrin]	89	Contig6064_1r_2
p88	up	5.04/134619	76	gi|1346119	Glycine cleavage system H protein		
p106	up	4.78/22695	97	gi|168006632	predicted protein (PRK)	90	Contig40843_3p_1

(Continued)

Spot ID	Changes[a]	pI/Mr	Best protein score[b]	Homolog_GI	Protein Name	EST best protein score[b]	Best matching EST ID
487		9.5/123954	286	gi|229473559	nucleoside diphosphate kinase II	133	097220_0761_3711_3r_1
p242	up	6.43/16189	235	gi|1346675	Nucleoside diphosphate kinase B	161	316948_2128_2485_1p_1
p348|p207|p211|p484|p27	up	6.92 /42547	174	gi|108864048 gi|4827251	plastidic aldolase NPALDP1 [Nicotiana paniculata]	115	Contig6205_2p 344418_1450_0111_3r_1 Contig20966_1p_2
467	up	8.30/36662	80	gi|115371630	glyceraldehyde-3 phosphate dehydrogenase		
p353	up	6.67/28814	80	gi|114326546	phosphoglycerate mutase 1		
p310	up	5.90/48165	73	gi|194702114	unknown [Zea mays]		
p26	up	6.06/24549	82	gi|224080171	phosphopyruvate hydratase		
p414	up	6.59/43596	107	gi|159477247	4-aminobutyrate aminotransferase		
p376	up	7.60/27963	72	gi|62320917	carbonic anhydrase	138	140570_0931_3601_3r_1
526		5.30/36934	70	gi|226460634	predicted protein [Micromonas pusilla CCMP1545]		
p85	up	5.10/114085	72	gi|116783747	Nitrilase/cyanide hydratase		
p245	up	6.13 /42420	232	gi|197310860	artemisinic aldehyde delta-11(13) reductase [Artemisia annua]	95	Contig11872_1p_2

(Continued)

Spot ID	Changes[a]	pI/Mr	Best protein score[b]	Homolog_GI	Protein Name	EST best protein score[b]	Best matching EST ID
Proteolysis							
p191	up	NA			cell division protein FtsH-like protein	77	072367_3859_2372_2p
357		5.40/147942	74	gi|242066450	Peptidases_S8_Tripeptidyl_Aminopeptidase_II		
664|p130|p328	up	5.48/75235	431	gi|84468324	putative zinc dependent protease	254	Contig3343_2p_4
p200	up	5.83/76015	206	gi|1483215	chloroplast FtsH protease	109	Contig28749_2r_1
p399	up	5.80/74312	317	gi|147809607	cell division protein	98	126915_1464_1042_1r_1
p139|p86	up	6.00/54643	221	gi|266567 gi|222424629	Mitochondrial-processing peptidase subunit alpha	227	234760_0781_2015_2p Contig36731_2p_1
442		5.59/27213	134	gi|224139394	predicted protein, proteasome [Populus trichocarpa]	144	Contig10396_1p_1
p457	up	4.25/15650	102	gi|239937266	cysteine protease [Lactuca sativa]	109	Contig13910_3r_8
Detoxificaiton Defence and Stress Response							
p126|p189		NA			dehydration stress-induced protein	165	Contig31528_2r_1
p573|p107*	up	9.61/31275	226	gi|38679339	harpin binding protein 1 [Zea mays]	126	Contig9087_1p
285		9.01/25296	145	gi|242049924	hypothetical protein SORBIDRAFT_02g030550 [Sorghum bicolor]Rhodanese Homology Domain (RHOD)	179	Contig3308_2p_1

(Continued)

Spot ID	Changes[a]	pI/Mr	Best protein score[b]	Homolog_GI	Protein Name	EST best protein score[b]	Best matching EST ID
346		5.31/13235	183	gi|45533923	glycine-rich RNA-binding protein RGP-1c [Nicotiana sylvestris]	429	Contig17268_2p_2
p473	up	NA			Peroxiredoxin	99	Contig38348_2p_5
p244	up	6.67/95677	77	gi|218196275	ABC_NikE_OppD_transporters		
349		NA			ferritin [Conyza canadensis]	92	Contig7767_3p_1
Others							
p61	up	5..10/78212	70	gi|168041407	Retropepsins; pepsin-like aspartate proteases [Physcomitrella patens subsp. patens]		
p32	up	5.20/17463	80	gi|212720942	hypothetical protein LOC100193724 [Zea mays]		
p397	up	5.04/14713	109	gi|132270	Rubber elongation factor protein		
132		NA			Ran binding protein in the microtubule-organising 2443886	71	209657_0077_0711_1r
411		6.43/28995	83	gi|255587064	ef-hand calcium binding protein, putative [Ricinus communis]	73	Contig1107_2p_2
p326	up	8.88/54682	215	gi|20743	pine globulin-1 [Pinus strobus]		
p48	up	NA			putative seed specific protein Bn15D17A [Oryza sativa Japonica Group]	74	Contig2796_2r_5

(Continued)

Spot ID	Changes[a]	pI/Mr	Best protein score[b]	Homolog_GI	Protein Name	EST best protein score[b]	Best matching EST ID
p45	up	5.20/55158	79	gi|115459778	Os04g0550400 [Oryza sativa (japonica cultivar-group)]		
126		NA			F5I10.22 gene product [Arabidopsis thaliana]	82	087095_0622_0875_2p_7
Unknown							
p366	up	10.36/29623	78	gi|125573095	hypothetical protein OsJ_04535 [Oryza sativa Japonica Group]		
p742	up	5.00/158452	75	gi|226458260	predicted protein [Micromonas pusilla CCMP1545]		
561		4.45/53322	75	gi|115488812	Os12g0517000 [Oryza sativa (japonica cultivar-group)]		
p68	up	6.10/7078	74	gi|10716604	hypothetical protein [Oryza sativa]		
81		5.10/53043	74	gi|6143899	unknown protein[Arabidopsis thaliana]		
p475	up	11.44/53043	73	gi|22324448	hypothetical protein [Oryza sativa Japonica Group]		
p303	up	9.66/15246	72	gi|168030864	predicted protein [Physcomitrella patens subsp. patens]		
p253	up	6.20/148525	71	gi|240254562	unknown protein [Arabidopsis thaliana]		
p115	up	4.95/12757	71	gi|224154921	predicted protein [Populus trichocarpa]		

(Continued)

Spot ID	Changes[a]	pI/Mr	Best protein score[b]	Homolog_GI	Protein Name	EST best protein score[b]	Best matching EST ID
85		9.47/35948	74	gi|168010293	predicted protein [Physcomitrella patens subsp. patens]		
p159	up	5.32/17505	87	gi|255637280	unknown [Glycine max]		
ESTs with no homology to known genes							
p119	up	NA	100	Contig31239_3 p_1			
536		NA	160	Contig40779_2 p_1			
523		NA	116	Contig20688_1 r_1			
p343	up	NA	141	Contig16569_1 p_11			
316		NA	134	Contig17501_2 r_2			
493		NA	73	378644_0715_ 0306_1r_1			
p754	up	NA	70	166761_0271_ 2109_3r_1			
389		NA	86	237117_1464_ 0273_3p			

(Continued)

Spot ID	Changes[a]	pI/Mr	Best protein score[b]	Homolog_GI	Protein Name	EST best protein score	Best matching EST ID
152		NA	78	277247_0225_3091_1r			
p4	up	NA	76	395769_2662_2273_2r_6			
451		NA	75	136083_0743_1553_1p_1			
p447	up	NA	74	233564_2556_0114_3p_4			
p393	up	NA	73	201199_3918_1922_3r_1			
131		NA	136	Contig3481_1p_2			

*The best protein result in NCBI database is different from that in EST database; the protein with higher protein score was chosen. [a]up-reglated proteins in trichomes (difference ratio>2 fold); [b] the highest score was chosen even if multiple protein spots were attributed to the same proteins. [c]calculated pI/Mr values are different from experiment ones; reasons are given in the discussion. [d]pI/Mr values were not available in EST database for the proteins predicted in the EST database are short.

The 91 identified proteins were divided into several different classes based on the UniProt database (http://www.uniprot.org): electron transport chain, proteolysis, transcription and translation, metabolism, detoxification or defense and stress response, others and unknown function proteins (Table 1; Fig. 13). These proteins were assigned mainly according to their predominant function even if they might play various roles in different subcellular compartments or at different developmental stages. This result was consistent with our previous transcriptome study of the glandular trichomes on flower buds (Wang *et al.*, 2009).

Figure 12 Venn diagram for the identified proteins in glandular trichomes

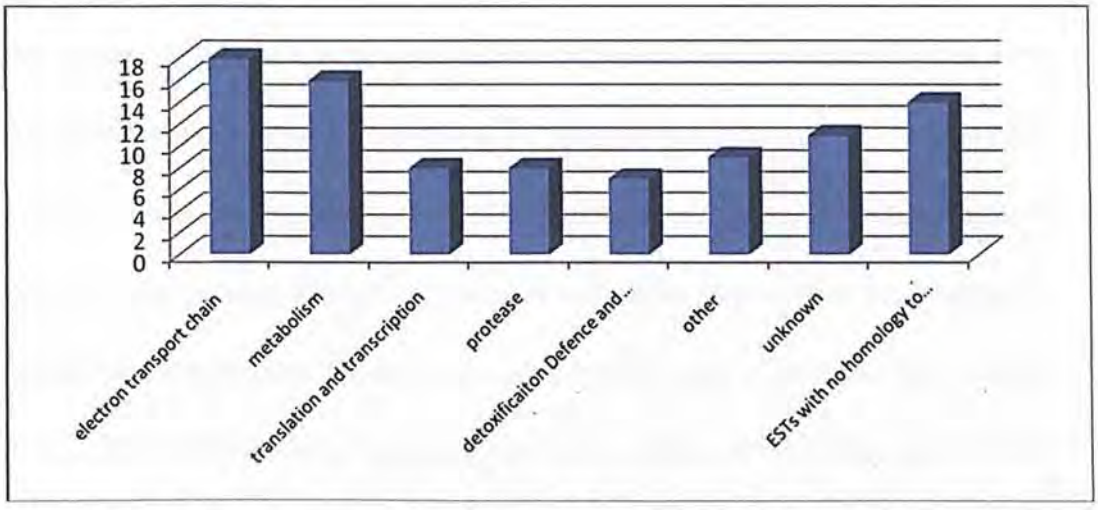

Figure 13 Categorization and Distribution of proteins in glandular trichomes

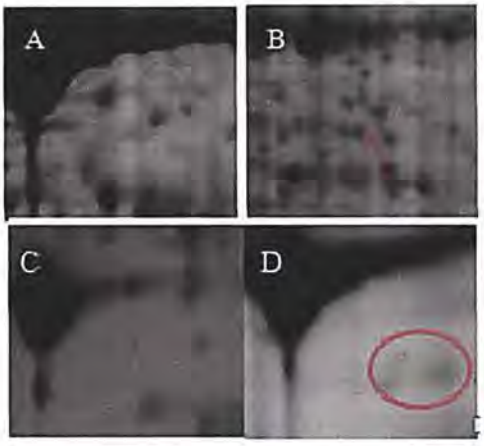

Figure 14 Enlarged windows of 2D gels for isolated trichomes (A, C) and corresponding leaves (B, D). The labeled protein spots showed a higher protein accumulation in the isolated trichomes compared with the corresponding leaf tissue. B: DBR2; D: MECDP-synthase.

3.3.1 Proteins involved in electron transport chain

This class mainly comprised several enzymes in ATPase family including subunits of the ATP-synthase complex (alpha, beta, gamma, epsilon subunits etc) and V-type (H^+)-ATPase V1. The ATPase family is known as one of the most important enzymes that create energy, determinate and maintain ion concentration balance in plant cells. Among them, H^+-ATPase is regarded as the basis for salt tolerance by regulating Na^+ and K^+ homeostasis in cytosol (Zhu, 2003; Ndimba *et al.*, 2005). The enhancement of these proteins in several plants exposed to cold stress implies that more energy is needed to reinforce plant self-defense system to cold stress (Cui *et al.*, 2005; Goulas *et al.*, 2006). Furthermore, we also found some other enzymes: ferredoxin-NADP reductase, NAD(P)H-quinone oxidoreductase, cytochrome c oxidase subunit Vib, ATP binding / succinate dehydrogenase. All of them have been identified and well

characterized in plant mitochondria or chloroplasts and are thought to function in protecting plant from oxidative stress by catalyzing electron transport in the respiratory chain (Millar *et al.*, 2001; Rodriguez *et al.*, 2007; Møller, 2001). In addition, NADP reductase and NAD(P)H-quinone oxidoreductase also participate in various metabolic pathways, including terpenoid biosynthesis, steroid metabolism. The increased abundance of some of these proteins in glandular trichomes of *A. annua* is in consistent with previous study that high energy is demanded for powering other cellular reactions in glandular trichomes, such as isoprenoid or other carbon fixation, fatty acid metabolism (Wagner, 1991).

3.3.2 Proteins involved in metabolism

Another large class representing proteins involved in metabolic process was also identified. It includes two enzymes related to artemisinin biosynthesis, enzymes implicate in glycolysis, and other enzymes with different catalytic functions (e.g. carbon fixation, glycine degradation).

3.3.2.1 Artemisinin biosynthesis

Artemisinic aldehyde delta-11(13) reductase (Dbr2) was first characterized in vitro by Zhang and his coworkers (Zhang *et al.*, 2008). It belongs to an enzyme family that acts on α, β-unsaturated carbonyls and presents the subfamily associated with terpene double bond reduction, the FMN-linked oxidoreductases. It is a trichome-specified enzyme involved in the sesquiterpenoid Δ11(13) double bond reduction leading the

conversion of artemisinic aldehyde to dihydroartemisinic acid. Another important enzyme involved in terpenoid biosynthesis is MECDP-synthase. It belongs to the family of lyases, specifically phosphorus-oxygen lyases which catalyzes the conversion of 2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol to 2-C-methyl-D-erythritol 2, 4-cyclodiphosphate (MECDP) accompanying cytidine monophosphate (CMP) generation. MECDP-synthase participates in the artemisinin biosynthesis in MEP pathway, leading to the formation of artemisinin precursor IPP. The gene encoding this enzyme has been cloned form Arabidopsis (Rohdich *et al.*, 2000). Two x-ray crystal structures of MECDP synthase were also characterized (Richard *et al.*, 2002). The first structure contains a bound Mn^{2+} cation, while the second one is a complex with CMP, MECDP, and Mn^{2+}. These models could help us understand the catalytic mechanism of MECDP synthase. In our study, both of these two enzymes could only be detected on the 2D gel of the glandular trichomes, providing the direct evidence that glandular trichomes are the major site of artemisinin biosythethsis (Fig. 14).

3.3.2.2 Glycolysis

Four enzymes implicated in glycolysis are fructose-1, 6-bisphosphate adolase, glyceraldehyde-3 phosphate dehydrogenase, phosphoglycerate mutase, phosphorpy-ruvate hydratase. Glycolytic enzymes are considered as important suppliers of essential metabolic flexibility that facilitates plant development and acclimation to environmental stress. They convert glucose into pyruvate and provide essential energy

to cells. At the same time, a large number of carbons in the plant glycolytic pathways are utilized in the biosynthesis of numerous compounds, such as isoprenoids, amino acids, nucleic acids, and fatty acids, which are especially important to the autotrophic tissues (Plaxton, 1996). In *A. annua*, the intermediate product glyceraldehyde 3-phosphate and final product pyruvate participate in the formation of IPP and DMAPP, the precursors for the synthesis of artemisinin and other monoterpenes or isoprenoids (Phillips *et al.* 2008). Changes of glycolytic enzymes have been reported in plant and non-plant species against environment stresses including nutrient limitation, salt, drought, and radioactivity (Sobhanian *et al.*, 2011; Moshwene *et al.*, 2003; Bae *et al.*, 2003). However, the mechanism has not been fully elucidated.

3.3.2.3 Other metabolic enzymes

Other enzymes with different catalytic functions were also observed in this study. These include glycine cleavage system H protein, phosphoribulokinase, nucleoside diphosphate kinase, carbonic anhydrase, 4-aminobutyrate aminotransferase etc. These enzymes participate in carbon fixation (phosphoribulokinase), glycine degradation (glycine cleavage system H protein), or in maintaining equilibrium between the concentrations of different nucleoside triphosphates (nucleoside diphosphate kinase). The enhanced expression of these proteins in the glandular trichomes further suggests that glandular trichome is the site with active metabolic activities.

3.3. 3 Proteins involved in transcription and translation

Proteins involved in gene transcription (mRNA synthesis) and mRNA translation (protein synthesis) are also identified in glandular trichomes, and the abundance of these proteins was found to be increased compared to those in the leaves. Changes in the expression of these protein synthesis related enzymes can be attributed to active role of trichomes in the metabolites biosynthesis and protection of plant against biotic or abiotic stress, which requires large number of new proteins and RNA synthesis. Among them, eukaryotic translation initiation factor eIF5A is a highly conserved protein which functions in the transport of a subset of mRNAs out of the nucleus to the ribosome. It is mainly active in translation elongation, mRNA turnover and decay, regulation of cell division, cell growth and cell death. However, its function in high plant is unclear. Recent reports found that eIF5A was up regulated in Arabidopsis root under Fe deficiency condition. They speculated that eIF5A might play an important role in regulating translation under stress conditions and in adapting plants to prevailing environmental conditions (Lan *et al.*, 2011).

3.3.4 Proteins involved in proteolysis

Several enzymes involved in proteolysis are also identified in this study. These include zinc dependent protease, chloroplast FtsH protease, cell division protein FtsH-like protein, cell division protein ftsH, predicted Zn-dependent peptidases, 20S proteasome alpha subunit A, cysteine protease. FtsH, a member of AAA^+ ATPases, is membrane bound ATP dependent metalloprotease which functions in assembly,

operation and disassembly of protein complexes. Changes of FstH following cold acclimation in some plants suggests that the FtsH ATPases may directly involve in chloroplast or plasma–membrane, and post translational modification of cold responsive proteins. It could be a basis for cold tolerance in plants response to cold stress (Kjellsen *et al.*, 2010; Balbuena *et al.*, 2010). In addition, FstH has been reported to participate in cell division in *Bacillus subtilis* (Wehrl *et al.*, 2000). The different expression of several FstH proteins between glandular trichomes and leaves reinforce the roles that glandular trichomes play in plant defense against severe environment condition.

3.3.5 Detoxification, stress defense related proteins

A variety of proteins associated with detoxification, stress defense were also identified. This group of proteins includes dehydration stress-induced protein, rhodanese homology domain, peroxiredoxin, glycine-rich RNA-binding protein, ferritin, ABC_NikE_OppD_ transporters. One of the most important proteins in this group is ABC_NikE_OppD_ transporters. As commonly known, ABC is one of the largest transporter families found from prokaryotes to humans. They coupled ATP hydrolysis to transport a wide variety of substrates such as ions, amino acids, peptides, sugars, lipids, and sterols, across cell membranes. Plant ABC transporters are first reported to be involved in detoxification. However, they have recently been implicated to function in ion regulation and plant growth process etc (Martinoia *et al.*, 2002). The ABC_NikE_OppD_transporters identified in our study is associated with

the transport of dipeptides, oligopeptides (OppD), and nickel, which may be essential in protecting plants against toxin accumulation.

Besides, several proteins whose biological functions are unique or unknown are also listed in the Table 1. Their roles in glandular trichomes need further investigation.

3.4 Prospective

The protein candidates identified in this study can be used to discover genes participated in the artemisinin production. Several methods can be used to validate these candidate genes: fluorescent mRNA in situ hybridization to localize the mRNA in tissues of *A. annua*; real time PCR to validate the overexpression of the mRNA in glandular trichomes; western blot study to detect specific protein candidates; or creation of transformed *A. annua* plants.

3.5 Conclusion

Plant trichomes are very important structures found on the aerial parts of most plant species. They not only synthesize and store a wide range of plant metabolites with high commercial values, but also play an important role in protecting plant against abiotic and biotic stresses. Glandular trichomes of *A. annua* have long been regarded as the essential artemisinin factory in the fight against malaria. Thus, the proteome study on glandular trichomes accomplished in this study provides important information in understanding the molecular functions of glandular trichomes. Over

700 spots in glandular trichomes were separated on the 2DE gels, of which ~102 non-redundant proteins were confidently identified using MALDI-TOF MS by searching NCBI database and our in-house EST database. Over 70% of these proteins were highly expressed in the glandular trichomes. Consistent with our previous EST study, functional clustering of these identified proteins demonstrated that most trichomes enriched proteins participate in metabolic process, transport, and protein or mRNA renewing. Among the 22 proteins identified according to our Artemisia EST database, 14 have no homology with other species even with less stringent blast search parameters, indicating they might be both species-specific and tissue-specific genes. Moreover, two enzymes that directly participate in artemisinin biosynthetic pathways were also identified to be enriched in glandular trichomes.

However, every technology has its limitations. Although 2DE combined mass spectrometry analysis has been widely applied in the plant proteomics study, several pitfalls remain to affect our results. Firstly, the detergent for isoelectric focusing is not suitable to dissolve hydrophobic proteins (e.g. transmembrane proteins). Thus, the gene products presenting in particular subcellular compartments of glandular trichomes, which could provide important information about the functions of glandular trichomes may be lost. In addition, a large amount of chemical compounds in the glandular trichomes and leaves such as pigments, hydrolytic enzymes may be easily degraded or precipitated proteins, resulting in incomplete protein profiling or several spots on 2DE gels attributed to one protein. Secondly, 2DE is usually

time-consuming and demands high amounts of proteins. Compared to the approximately 20% of protein content in microbial cells or mammal tissues, the protein content in plant tissues is low (usually 2%) (Weiss *et al.*, 2007). Thus, enrichment of proteins is necessary. For glandular trichomes of *A. annua*, the production rate is quite low (~500mg wet weight every 400g leaves). Thus, large amounts of plants are needed for proteome study. Furthermore, not all proteins could be separated or visible on 2DE gels. Abundant proteins usually mask low copy number proteins of specialized functions only in certain cell types. Thus, it is not unusual to observe Rubisco or other photosynthesis-related proteins with highest abundance in this study. Besides, high quality of digested peptides is critical for MS analysis.

In conclusion, our proteomics study provides for the first time a general picture of proteins specifically expressed in *Artemisia annua* glandular trichomes. Enriched but inter-species un-conserved genes identified in this study, together with our previous transcriptome study suggested that more trichome-specific genes are likely involved in Artemisia metabolism and trichome. These candidate proteins provide potential targets for further experiments in order to elucidate the trichome function and artemisinin metabolism in the important medicinal plant *A.annua.*

References

Abdin, M.Z., Israr, M., Rehman, R.U., Jain, S.K. (2003). Artemisinin, a novel antimalarial drug: biochemical and molecular approaches for enhanced production. *Planta Medica* **69**, 289-299.

Adam, K., Zapp, J. (1998). Biosynthesis of the isoprene units of chamomile sesquiterpenes. *Phytochemistry* **48**, 953–959.

Akhila, A., Thakur, R.S., Popli, S.P. (1987). Biosynthesis of artemisinin in *Artemisia annua*. *Photochemistry* **16**, 1927-1930.

Akhila, A., Kumkum, R., Thakur, R.S. (1990). Biosynthesis of artemisinic acid in *Artemisia annua*. *Photochemistry* **29**, 2129-2132.

Amme, S., A.; Rutten, T., Melzer M., Sonsmann, G., Vissers, J.P., Schlesier, B., Mock, H.P. (2005). A proteome approach defines protective functions of tobacco leaf trichomes. *Proteomics* **5**, 2508–2518.

Anthony, L. S., Robert, L., Eran, P. (2008). Harnessing plant trichome biochemistry for the production of useful compounds. *The Plant Journal* **54**, 702–711.

Aravind, L., Iyer, L.M., Wellems, T.E., and Miller, L.H. (2003). *Plasmodium* biology: genomic gleanings. *Cell* **115,** 771-785.

Bae, M.S., Cho, E.J., Choi, E.Y., Park, O.K. (2003). Analysis of the Arabidopsis nuclear proteome and its response to cold stress. *The Plant Journal* **36**, 652-63.

Baggerman, G., Vierstraete, E., De, L.A, Schoofs, L. (2005). Gel-based versus gel-free proteomics: A review. *Combinatorial Chemistry & High Throughput*

Screening **8**, 669-677.

Baginsky, S. (2009). Plant proteomics: concepts, applications, and novel strategies for data interpretation. *Mass Spectrometry Reviews* **28**, 93– 120.

Balbuena, T.S., Salas, J.N.J., Force, E.M., Garcés, R.G., Thelen, J.J. (2011). Proteome Analysis of Cold Acclimation in Sunflower. *Journal of Proteome Research* **10**, 2330-2346.

Barnwell, J.W., Galinski, M.R. (1998). Invasion of vertebrate cells: Erythrocytes. In *Malaria: Parasite Biology, Pathogenesis, and Protection*, (Sherman, I.W., ed), Washington, D. C.: ASM Press, pp. 93-120.

Bertea, C.M., Freije, J.R., van der Woude, H., Verstappen, F.W., Perk, L., Marquez, V., De Kraker, J.W., Posthumus, M.A., Jansen, B.J., de Groot, A., Franssen M.C., Bouwmeester H.J. (2005). Identification of intermediates and enzymes involved in the early steps of artemisinin biosynthesis in *Artemisia annua*. *Planta Medica* **71**, 40-47.

Bertea, C.M., Voster, A., Verstappen, F.W., Maffei, M., Beekwilder, J., and Bouwmeester, H.J. (2006). Isoprenoid biosynthesis in *Artemisia annua*: cloning and heterologous expression of a germacrene A synthase from a glandular trichome cDNA library. *Archives of Biochemistry and Biophysics* **448**, 3-12.

Bhakuni, R.S., Jain, D.C., Sharma, R.P., Kumar, S. (2001). Secondary metabolites of *Artemisia annua* and their biological activity. *Current Science* **80**, 35-48. .

Brown, G.D. (2010). The Biosynthesis of Artemisinin (Qinghaosu) and the Phytochemistry of Artemisia annua L. (Qinghao). *Molecules* **15**, 7603-7698.

Covello, P.S., Teoh, K.H., Polichuk, D.R., Reed, D.W., Nowak, G. (2007). Functional genomics and the biosynthesis of artemisinin. *Phytochemistry* **68**, 1864–1871.

Chang, M.C., Eachus, R.A., Trieu, W., Ro, D.K., Keasling, J.D. (2007). Engineering Escherichia coli for production of functionalized terpenoids using plant P450s. *Nature Chemical Biology* **3**, 274-277.

Charles, D.J., Cebert, E., Simon, J.E. (1991). Characterization of the essential oils of *Artemisia annua* L. *Journal of Essential Oil Research* **3**, 33-39.

Charles, D.J., Simon, J.E., Wood, K.V., Heinstein, P. (1990). Germplasm variation in artemisinin content of *Artemisia annua* using an alternative method of artemisinin analysis from crude plant extracts. *Journal of Natural Products* **53**, 157-160.

China Cooperative Research Group on qinghaosu and its derivatives as antimalarials. (1982). Chemical studies on qinghaosu (artemisinine). *Journal of traditional Chinese medicine* **2**, 3-8.

Covello, P.S., Teoh, K.H., Polichuk, D.R., Reed, D.W., Nowak, G. (2007). Functional genomics and the biosynthesis of artemisinin. *Phytochemistry* **68**, 1864-1871.

Covello, P.S. (2008). The Molecular cloning of artemisinic aldehyde reductase and its role in glandular trichome-dependent biosynthesis of artemisinin in *Artemisia annua*. *Journal of Biological Chemistry* **283**, 21501-8.

Cui, S.X., Huang, F., Wang, J., Ma, X., Cheng, Y. (2005). A proteomic analysis of cold stress responses in rice. *Proteomics* **5**, 3162–3172.

Davis, T.M., Karunajeewa, H.A., Ilett, K.F. (2005). Artemisinin-based combination therapies for uncomplicated malaria. *Medical Journal of Australia* **182**,181-185.

Duke, M.V., Paul, R.N., El-Sohly, H.N., Strutz, G., Duke, S.O. (1994). Localization of artemisinin and artemisitene in foliar tissues of glanded and glandless biotypes of *Artemisia annua* L. *International Journal of Plant Sciences* **55**, 365-372.

Duke, S.O., Paul, R.N. (1993). Development and fine structure of glandular trichomes of *Artemisia annua* L. *International Journal of Plant Sciences* **154,** 107-118.

Efferth, T. (2006). Molecular pharmacology and pharmacogenomics of artemisinin and its derivatives in cancer cells. *Current Drug Targets* **7**, 407-421.

ElSohly, H.N., Croom, E.M., ElFeraly, F.S., ElSherei, M.M. (1990). A large-scale extraction technique of artemisinin from Artemisia annua. *Journal of Natural Products* **53**, 1560–1564.

Ferreira, J.F.S., Janick, J. (1995a). Floral morphology of *Artemisia annua* with special reference to trichomes. *International Journal of Plant Sciences* **156**, 807-815.

Ferreira, J.F., Simon, J.E., Janick, J. (1995b). Developmental studies of *Artemisia annua*: flowering and artemisinin production under greenhouse and field conditions. *Planta Medica* **61**, 167-170.

Ferreira, J.F.S., Janick, J. (1996). Distribution of Artemisinin in *Artemisia annua*. In *Progress in new crops*, (Janick, J., ed), Alexandria: ASHS Press, pp. 579-584.

Ferreira, J.F.S., Simon, J.E., and Janick, J. (1997). *Artemisia annua*: botany, horticulture and pharmacology. *Horticulture Reviews* **19,** 319-371.

Ferreira, J.F.S., Janick, J. (2009). Annual Wormwood (*Artemisia annua* L.) *New Crop FactSHEET*

Frevert, U., Crisanti, A. (1998). Invasion of vertebrate cells: Hepatocytes. In *Malaria: Parasite Biology, Pathogenesis, and Protection*, (Sherman, I.W., ed), Washington, D. C.: ASM Press, pp. 73-91.

Gang, D.R., Wang, J., Dudareva, N., Nam, K.H., Simon, J.E., Lewinsohn, E., Pichersky, E. (2001). An investigation of the storage and biosynthesis of phenylpropenes in sweet basil. *Plant Physiology* **125**, 539-555.

Geldre, E. V., Annemieke, V., Elfride, V.D. (1997). State of the art of the production of the antimalarial compound artemisinin in plants. *Eeckhout Plant Molecular Biology* **33**, 199-209.

Geoffrey, D.B. (2010). The Biosynthesis of Artemisinin (Qinghaosu) and the Phytochemistry of Artemisia annua L. (Qinghao). *Molecules* **15**, 7603-7698.

Gershenzon, J., McCaskill, D., Rajaonarivony, J.I., Mihaliak, C., Karp, F., Croteau, R. (1992). Isolation of secretory cells from plant glandular trichomes and their use in biosynthetic studies of monoterpenes and other gland products. *Analytical Biochemistry* **200**, 130-138.

Görg, A., Obermaier, C., Boguth, G., Harder, A., Scheibe, B., Wildgruber, R., Weiss, W. (2000). The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* **21**(6), 1037-1053.

Goulas, E., Schubert, M., Kieselbach, T., Kleczkowski, L.A., Gardestrom, P., Hurry, W. S.V. (2006). The chloroplast lumen and stromal proteomes of Arabidopsis thaliana show differential sensitivity to short- and long-term exposure to low temperature. the *Plant Journal* **47**, 720-34.

Graham, I.A., Besser, K., Blumer, S., Branigan, C.A. (2010). The Genetic Map of

Artemisia annua L. Identifies Loci Affecting Yield of the Antimalarial Drug Artemisinin. *Science* **327**, 328-331.

Greenwood, B.M., Bojang, K., Whitty, C.J., and Targett, G.A. (2005). Malaria. *Lancet* **365,** 1487-1498.

Greenwood, B.M., Fidock, D.A., Kyle, D.E., Kappe, S.H., Alonso, P.L., Collins, F.H., Duffy, P.E. (2008). Malaria: progress, perils, and prospects for eradication. *Journal of Clinical Investigation* **118**, 1266-1276.

Gustavsson, N., Greber, B., Kreitler, T., Himmelbauer, H., Lehrach H., Gobom, J. (2005). A proteomic method for the analysis of changes in protein concentrations in response to systemic perturbations using metabolic incorporation of stable isotopes and mass spectrometry. *Proteomics* **5**, 3563.

Hallett, R.L., Sutherland, C.J., Alexander, N.R., Jawara, M., Drakeley, C.J., Pinder, M., Walraven, G., Targett, G., Alloueche, A. (2004). Combination therapy counteracts theenhanced transmission of drug-resistant malaria parasites to mosquitoes. *Antimicrobial Agents and Chemotherapy* **48**, 3940-3943.

Hisashi, H., Nazrul, I., Hiroshi, K. (2004). Technical aspects of functional proteomics in plants. *Phytochemistry* **65**, 1487-1498.

Hu, J. Qian, J., Borisov, O., Pan, S., Li, Y., Liu, T., Deng, L., Wannemacher, K., Kurnellas, M., Patterson, C., Elkabes, S., Li, H. (2006). Optimized proteomic analysis of a mouse model of cerebellar dysfunction using amine-specific isobaric tags. *Proteomics* **6**, 4321-4334.

Jonathan, G., David, M., Jean, I.M.R., Mihaliak C. (1992). Isolation of secretory cells from plant glandular trichomes and their use in biosynthetic studies of monoterpenes and other gland products. *Analytical Biochemistry* **200**, 130-138.

Johnson, H.B. (1975). Plant pubescence: An ecological perspective. *Botanical Review* **41**, 233-258.

Jorrín-Novo, J.V., Maldonado, A.M., Echevarría-Zomeño, S., Valledor, L., Castillejo, M.A., Curto, M., Valero, J., Sghaier, B., Donoso, G., Redondo, I. (2009). Plant proteomics update (2007-2008): Second-generation proteomic techniques, an appropriate experimental design, and data analysis to fulfill MIAPE standards, increase plant proteome coverage and expand biological knowledge. *Journal of Proteomics* **72**, 285-314.

Kaiser, A., Gottwald, A., Wiersch, C., Maier, W., Seitz, H.M. (2002). The necessity to develop drugs against parasitic diseases. *Pharmazie* **57**, 723-728.

Kapoor, R., Chaudhary, V., and Bhatnagar, A.K. (2007). Effects of arbuscular mycorrhiza and phosphorus application on artemisinin concentration in *Artemisia annua* L. *Mycorrhiza* **17,** 581-587.

Kim, S.H., Heo, K., Chang, Y.J., Park, S.H., Rhee, S.K., Kim, S.U. (2006). Cyclization mechanism of amorpha-4,11-diene synthase, a key enzyme in artemisinin biosynthesis. *Journal of Natural Products* **69** (5), 758-762.

Kjellsen, T. D., Shiryaeva, L., Schröder, W.P. (2010). Proteomics of extreme freezing tolerance in Siberian spruce (Picea obovata). *Journal of Proteomics* **73**, 965-975.

Klayman, D.L. (1985). Qinghaosu (artemisinin): an antimalarial drug from China. *Science* **228,** 1049-1055.

Klayman, D.L. (1989). Weeding out malaria. *Natural History* **10**, 18-26.

Klayman, D.L. (1993). Artemisia annua: From weed to respectable antimalarial plant. *ACS symposium series*. 242-255.

Koller, A., Washburn, M.P., Lange, B.M. (2002). Proteomic survey of metabolic pathways in rice. *Proceedings of the National Academy of Sciences* **99**, 11969–11974.

Kuhn, T., and Wang, Y. (2008). Artemisinin--an innovative cornerstone for anti-malaria therapy. *Progress in Drug Research* **66,** 385-422.

Lan, P., Schmidt, W. (2011). The enigma of eIF5A in the iron deficiency response of Arabidopsis. *Plant Signaling & Behavior* **6**, 10-15.

Laule, O., Furholz, A., Chang, H.S., Zhu, T., Wang, X., Heifetz, P.B., Gruissem, W., and Lange, B.M. (2003). Cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis in *Arabidopsis thaliana. Proceedings of the National Academy of Sciences* **100**, 6866-6871.

Laughlin, J.C. (1995). The influence of distribution of antimalarial constituents in *Artemisia annua* L. on time and method of harvest. *Acta Horticulturae* **390**, 67-73.

Liu, C., Zhao, Y., Wang, Y. (2006). Artemisinin: current state and perspectives for biotechnological production of an antimalarial drug. *Applied Microbiology and Biotechnology* **72**, 11-20.

Markus, L.B., Mark, R.W., Einar, J.S., Christopher, S. (2000). Probing essential oil biosynthesis and secretion by functional evaluation of expressed sequence tags from mint glandular trichomes. *Proceedings of the National Academy of Sciences* **97**,

2934–2939.

Martinoia, E., Klein, M., Geisler, M., Bovet, L., Forestier, C., Kolukisaoglu, U. (2002). Multifunctionality of plant ABC transporters--more than just detoxifiers. *Planta* **214**, 345-55.

Mathesius, U., Keijzers, G., Natera, S.H.A., Winman, J.J., Djordjevic, M.A. (2001). Establishment of a root proteome reference map for the model legume Medicago truncatula using the expressed sequence tag database for peptide mass fingerprinting. *Proteomics* **1**, 1424-1440.

Matsushita, Y., Kang, W., Charlwood, B.V. (1996). Cloning and analysis of a cDNA encoding farnesyl diphosphate synthase from *Artemisia annua*. *Gene* **172**, 207-209.

MEDTV homepage [http://malaria.emedtv.com]

MedicineNet homepage [http://www.medicinenet.com]

Millar, H.A., Sweetlove, L.J., Giege, P. (2001). Analysis of the Arabidopsis Mitochondrial Proteome. *Plant Physiology* **127**, 1711-1727.

Miller, L.H., Baruch, D.I., Marsh, K., Doumbo, O.K. (2002). The pathogenic basis of malaria. *Nature*, **415**, 673-679.

Møller, I.M. (2001). Plant Mitochomdria and Oxidative Stress: Electron Transport, NADPH Turnover, and Metabolism of Reactive Oxygen Species. *Annual Review Plant Physiology and Plant Molecular Biology* **52**, 561–91.

Motshwene, P., Brandt, W., Lindsey, G. (2003). Significant quantities of the glycolytic enzyme phosphoglycerate mutase are present in the cell wall of yeast Saccharomyces cerevisiae. *Biochemical Journal* **369**, 357-362.

Nair, M.S.R., Basile, D.V. (1993). Bioconversion of Arteannuin-B to Artemisinin. *Journal of Natural Products* **56**, 1559-1566.

Newton, R.P., Brenton, A.G., Smith, C.J., Dudley, E. (2004). Plant proteome analysis by mass spectrometry: principles, problems, pitfalls and recent developments *Phytochemistry* **65**, 1449–1485.

Ndimba, B. K., Chivasa, S., Simon, W.J. (2005). Identification of Arabidopsis salt and osmotic stress responsive proteins using two-dimensional difference gel electrophoresis and mass spectrometry. *Proteomics* **5**, 4185–4196.

Olsson, M.E., Linda, M.O., Ann-Louise, L., Anneli L., Maria B., Peter E. (2009). Localization of enzymes of artemisinin biosynthesis to the apical cells of glandular secretory trichomes of *Artemisia annua* L., *Brodelius Phytochemistry* **70**, 1123–1128.

Paeshuyse, J., Coelmont, L., Vliegen, I., Van hemel, J., Vandenkerckhove, J., Peys, E., Sas, B., De Clercq, E., Neyts, J. (2006). Hemin potentiates the anti-hepatitis C virus activity of the antimalarial drug artemisinin. *Biochemical and Biophysical Research Communication* **348**, 139-144.

Pamela, J. W., Shereen, E., Kristin, K. (2006). Wobbe artermisinin: the biosynthetic pathway and its regulation in *Artemisia Annua*, a terpenoid-rich species in vitro cell. *Developmental. Biology—Plant* **42**, 309–317.

Picaud, S., Mercke, P., He, X., Sterner, O., Brodelius, M., Cane, D.E., Brodelius, P.E. (2006). Amorpha-4,11-diene synthase: mechanism and stereochemistry of the enzymatic cyclization of farnesyl diphosphate. *Archives of Biochemistry and Biophysics* **448**, 150-155.

Phillips, M.A., León P., Boronat, A., Rodríguez-Concepción, M. (2008). The plastidial MEP pathway: unified nomenclature and resources. *Trends in Plant Science* **13** (12), 619-623.

Plaxton, W.C. (1996). The Organization and Regulation of Plant Glycolysis. *Annual Review Plant Physiology and Plant Molecular Biology* **47**, 185–214.

Richard, S. B., Ferrer, J.L., Bowman, M.E., Lillo, A.M., Tetzlaff, C.N. (2002). Structure and Mechanism of 2-C-Methyl-d-erythritol 2,4-Cyclodiphos phate Synthase. *Journal of Biological Chemistry* **277**, 8667-8672.

Rohdich, L., Wungsintaweekul, J., Eisenreich, W., Richter, G., Schuhr, C.A. (2000). Biosynthesis of terpenoids: 4-Diphosphocytidyl-2C-methyl-d-erythritol synthase of *Arabidopsis thaliana. Proceedings of the National Academy of Sciences* **97**, 6451–6456.

Roberts, J.M., Parise, M.E., Slutsker, L., and Newman, R.D. (2006). Malaria surveillance—United States, 2004. *Morbidity and Mortality Weekly Report Surbeillance Summaries* **55**, 23-37.

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

Rodriguez, R.E., Lodeyro, A., Poli, H.O., Zurbriggen, M., Peisker, M., Palatnik, J.F. (2007). Transgenic tobacco plants overexpressing chloroplastic ferredoxin-NADP(H) reductase display normal rates of photosynthesis and increased tolerance to oxidative stress. *Plant Physiology* **143** (2), 639-49.

Ronn, A.M., Msangeni, H.A., Mhina, J., Wernsdorfer, W.H., and Bygbjerg, I.C. (1996). High level of resistance of *Plasmodium falciparum* to sulfadoxine-pyrimethamine in children in Tanzania. *Transactions of the Royal Society of Tropical Medicine & Hygiene* **90**, 179-181.

Russell, P.N., Brenton, A.G., Chris, J.S., Edward, D. (2004). Plant proteome analysis by mass spectrometry: principles, problems, pitfalls and recent developments. *Phytochemistry* **65**, 1449–1485.

Sanger, F., Nicklen, S., Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* **74**, 5463-5467.

Sangwan, R.S., Agarwal, K., Luthra, R., Thakur, R.S., Singhsangwan, N. (1993). Biotransformation of Arteannuic Acid into Arteannuin-B and Artemisinin in Artemisia-Annua. *Phytochemistry* **34**, 1301-1302.

Schilmiller, A.L., Last, R.L., Pichersky, E. (2008). Harnessing plant trichome biochemistry for the production of useful compounds. the *Plant Journal* **54**, 702-711.

Schramek, N., Wang, H., Roemisch-Margl, W., Keil, B., Radykewicz, T., Winzenhoerlein, B., Beerhuse, L., Bacher, A., Rodich, F., Gershenzon, J., Liu, B., Eisenreich, W. (2010). Artemisinin biosynthesis in growing plants of *Artemisia annua*. A $^{13}CO_2$ study. *Phytochemistry* **71**, 179-187.

Singh, B., Kim Sung, L., Matusop, A., Radhakrishnan, A., Shamsul, S.S., Cox-Singh, J., Thomas, A., and Conway, D.J. (2004). A large focus of naturally acquired *Plasmodium knowlesi* infections in human beings. *Lancet* **363,** 1017-1024.

Skarbinski, J., Eliades, M.J., Causer, L.M., Barber, A.M., Mali, S., Nguyen-Dinh, P.,

Sobhanian, H., Aghaei, K., Komatsu, S. (2011). Changes in the plant proteome resulting from salt stress: Toward the creation of salt-tolerant crops? *Journal of Proteomics.*

Teoh, K.H., Polichuk, D.R., Reed, D.W., Nowak, G., 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 Letters* **580**, 1411-1416.

Turner, G.W., Gershenzon, J., and Croteau, R.B. (2000). Development of peltate glandular trichomes of peppermint. *Plant Physiology* **124**, 665-680.

Wagner, G.J., Wang, E., and Shepherd, R.W. (2004). New Approaches for Studying and Exploiting an Old Protuberance, the Plant Trichome. *Annals of Botany* **93**, 3-11.

Wagner, G.J. (1991). Secreting Glandular Trichomes: More than Just Hairs. *Plant Physiology* **96**, 675-679.

Wang, H., Ge, L., Ye, H.C., Chong, K., Liu, B.Y., and Li, G.F. (2004). Studies on the effects of fpf1 gene on *Artemisia annua* flowering time and on the linkage between flowering and artemisinin biosynthesis. *Planta Medica* **70**, 347-352.

Wang, W., Wang, Y.J., Zhang, Q., Qi, Y., Guo D.J. (2009). Global characterization of *Artemisia annua* glandular trichome transcriptome using 454 pyrosequencing. *BMC Genomics* **10**, 1471-2164.

Washburn, M.P., Wolters, D., Yates, J.R., III (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnology* **19**, 242-247.

Wehrl, W., Niederweis, M., Schumann, W. (2000). The FtsH Protein Accumulates at the Septum of Bacillus subtilis during Cell Division and Sporulation. *Journal of Bacteriology* **182**, 3870–3873.

Weiss, W., Gorg, A. (2007). Two-dimensonal electrophoresis for plant proteomics. *Plant Proteomics* **355**, 121-142.

White, N.J., D.Sc., M.D. (1996). The Treatment of Malaria. the *New England Journal of Medicine* **335**, 800-806.

Wijk, K.J.V. (2001). Challenges and Prospects of Plant Proteomics. *Plant Physiology* **126**, 501–508.

Woerdenbag, H.J., Pras, N., Chan, N.G., Bang, B.T., Bos, R., van Uden, W., Van, Y.P., Van Boi, N., Batterman, S., and Lugt, C.B. (1994). Artemisinin, Related Sesquiterpenes, and Essential Oil in *Artemisia annua* During a Vegetation Period in Vietnam. *Planta Medica* **60**, 272-275.

Xiao, S.H. (2005). Development of antischistosomal drugs in China, with particular consideration to praziquantel and the artemisinins. *Acta Tropica* **96**, 153-167.

Zhang, Y., Teoh, K.H., Reed, D.W., Maes, L., Goossens, A., Olson, D.J., Ross, A.R., 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. the Journal of Biological Chemistry* **283**, 21501-21508.

Zivy, M., Vienne, D. (2000). Proteomics: a link between geno- mics, genetics and physiology. *Plant Molecular Biology* **44**, 575–580.

Zhu, J.K. (2003). Regulation of ion homeostasis under salt stress. *Current Opinion in*

Plant Biology **6**, 441–44.

Zhu, H., Snyder, M. (2003). Protein chip technology. *Current Opinion in Chemical Biology* **7**:55–63.