# Gene discovery: Polymorphism association with obsessive-compulsive disorder and proteome analysis of *Artemisia annua*.

Een zoektocht naar genen:
Associatie van polymorfismen met obsessieve compulsieve stoornis en proteoom analyse van *Artemisia annua*.

**Apr. Filip Van Nieuwerburgh**

Februari 2006

Ghent University
Faculty of Pharmaceutical Sciences
Laboratory of Pharmaceutical Biotechnology
Promoter: Prof. D. Deforce
Co-promoter: Dr. D. Denys

# TABLE OF CONTENTS

**TABLE OF CONTENTS** ..... 1

**ABBREVIATIONS** ..... 3

**PART I: Polymorphism association with Obsessive-Compulsive Disorder** ..... 7

1. Obsessive Compulsive Disorder: Introduction ..... 9

*1.1. History* ..... 9
*1.2. DSM-IV-TR* ..... 10
*1.3. ICD-10* ..... 12
*1.4. Yale-Brown Obsessive-Compulsive Scale (Y-BOCS)* ..... 13
*1.5. The OCD-cycle* ..... 13
*1.6. Prevalence* ..... 14
*1.7. OCD: many manifestations* ..... 14
*1.8. High Comorbidity: Is OCD a discrete psychiatric disorder?* ..... 15
*1.9. Does OCD represent different clinical subtypes?* ..... 16
*1.10. Treatment* ..... 17
*1.11. (Pharmaco)genetics of OCD* ..... 21

2. Association between serotonergic candidate genes and specific phenotypes of OCD ..... 35

3. Association between the DRD2 TAQI-A2 and COMT-low allele with OCD in males ..... 47

4. Prediction of response to Paroxetine and Venlafaxine by serotonin genes in OCD ..... 59

**PART II: Proteome analysis of Artemisia annua** ..... 73

1. Objective and strategy ..... 75

2. Introduction ..... 79
*2.1. Malaria* ..... 79
*2.2. Artemisia annua L.* ..... 83
*2.3. Gene discovery: Which genes influence artemisinin production?* ..... 89

3. A proteome analysis of Artemisia annua L. ..... 109

4. Quantitation of artemisinin and its precursors in A. annua by HPLC Q-TOF MS/MS ..... 123

**SUMMARY:** ..... 139

**SAMENVATTING:** ..... 143

**APPENDICES:** ..... 147
Appendix 1: Identified proteins in the trichomes-*versus*-whole leafs proteome study ..... 149
Appendix 2: Identified proteins in the upper-*versus*-lower leafs proteome study ..... 153
Appendix 3: Identified proteins in the one-minute chloroform extract proteome study ..... 155
Appendix 4: BLAST results for the trichomes-*versus*-whole leafs proteome study ..... 157
Appendix 5: BLAST results for the upper-*versus*-lower leafs proteome study ..... 167
Appendix 6: BLAST results for the one-minute chloroform extract proteome study ..... 171

**DANKWOORD:** ..... 178

# Abbreviations

| | |
|---|---|
| 2-DE | two dimensional gel electrophoresis |
| 5-HT | 5-hydroxytryptamine (serotonin) |
| 5-HTT | serotonin transporter |
| 5-HTTLPR | serotonin transporter gene promoter length polymorphism |
| 6-FAM | 6-carboxy-fluorescein |
| ACT | artemisinin based combination therapy |
| ADHD | attention deficit hyperactivity disorder |
| ADS | amorpha-4,11-diene synthase |
| AFLP | amplified fragment length polymorphism |
| AIDS | acquired immune deficiency syndrome |
| ANOVA | analysis of variance |
| BAP | 6-benzyl aminopurine |
| CBT | cognitive-behavioural therapy |
| cDNA | complementary DNA |
| CE | capillary electrophoresis |
| CHAPS | 3-[(3-Cholamidopropyl)-dimethyl-ammonio]-1-propane sulfonate |
| COMT | catechol-o-methyl transferase |
| COMT H | high-activity COMT allele |
| COMT L | low-activity COMT allele |
| CT | cognitive therapy |
| CV% | coefficient of variation |
| CYP71AV1 | amorpha-4,11-diene monooxygenase (cytochrome P450 enzyme) |
| DAT | dopamine transporter |
| DHAA | dihydroartemisinic acid |
| DIGE | difference gel electrophoresis |
| DMAPP | dimethylallyl diphosphate |
| DNA | deoxyribonucleic acid |
| DRD2 | dopamine D2 receptor |
| DSM-IV-TR | diagnostic and statistical manual of mental disorders - 4$^{th}$ edition – text revision |
| DXP | 1-Deoxy-D-xylulose-5-phosphate |
| ECD | electrochemical detection |
| ELSD | evaporative light scattering detector |
| ERP | exposure response prevention |

| | |
|---|---|
| ESI | electrospray ionisation |
| EST | expressed sequence tag |
| FA | formic acid |
| FDA | food and drug administration |
| FID | flame ionization detector |
| fpf1 | flowering promoter factor 1 |
| FPP | Farnesyl diphosphate |
| FPPS | farnesyl diphosphate synthase |
| GA3 | gibberellic acid |
| GC | gas chromatography |
| GPP | geranyl diphosphate |
| HAM-A | Hamilton anxiety scale |
| HAM-D | Hamilton depression scale |
| HCl | hydrochloric acid |
| HMG-CoA | 3S-Hydroxy-3-methylglutaryl-CoA |
| HPLC | high performance liquid chromatography |
| 5-HTTLPR S-allele | short allele: reduces efficiency of the 5-HTT gene promoter |
| 5-HTTLPR L-allele | long allele of this polymorphism |
| ICD-10 | international statistical classification of diseases and related health problems - 10th revision |
| IPG | immobilized pH gradient |
| IPP | isopentenyl diphosphate |
| JA | jasmonic acid |
| KCl | potassium chloride |
| LLOQ | lower limit of quantitation |
| LOD | limit of detection |
| MAE | microwave-assisted extraction |
| MALDI | matrix assisted laser desorption ionization |
| MAO-A | monoamine oxidase A |
| MDD | major depressive disorder |
| MEP | 2-C-Methyl-Derythritol-4-phosphate |
| M.I.N.I. | Mini-International Neuropsychiatric Interview |
| mRNA | messenger RNA |
| MS | mass spectrometry |
| MS/MS | tandem mass spectrometry |
| MSDB | mass spectrometry protein sequence database |

| | |
|---|---|
| MVA | mevalonate |
| NED | 2,7',8'-benzo-5'-fluoro-2',4,7,-trichloro-5-carboxyfluorescein |
| NMR | nuclear magnetic resonance |
| OCD | obsessive compulsive disorder |
| PCR | polymerase chain reaction |
| pI | isoelectric point |
| PMF | peptide mass fingerprint |
| PSE | pressurized solvent extraction |
| QTOF | quadrupole time-of-flight |
| SAM | S-adenosyl methionine |
| SD | standard deviation |
| SDS | sodium dodecyl sulfate |
| SFE | super critical fluid extraction |
| SPE | solid-phase extraction |
| SPECT | single photon emission computed tomography |
| SQC | Sesquiterpene cyclase |
| SRI | serotonin reuptake inhibitors |
| SSRI | selective serotonin reuptake inhibitors |
| TDZ | thidiazuron |
| TLC | thin layer chromatography |
| TRIS | 2-Amino-2-(hydroxymethyl)-1,3-propandiol |
| WHO | World Health Organisation |
| Y-BOCS | Yale-Brown obsessive-compulsive scale |
| Y-BOCS CL | Y-BOCS symptom checklist |

# PART I:

## Polymorphism association with Obsessive-Compulsive Disorder

# 1. Obsessive Compulsive Disorder: Introduction

## 1.1. History

A condition resembling obsessive compulsive disorder (OCD) has been recognized for more than three centuries. Through history, the view on OCD has been influenced by the intellectual and scientific climate of the period.

Centuries ago, individuals with obsessive blasphemous or sexual thoughts were considered to be possessed. The logical treatment was one designed to exorcize the invading entity from the unfortunate soul who was possessed.

With time, the explanation of the cause of obsessions and compulsions moved from a religious view to a medical one. The first description in psychiatric literature of a patient with obsessive-compulsive disorder is found in the writings of J.E.D. Esquirol in 1838[1]. He describes a female patient with kleptophobia and stresses the two characteristics of the continuous fight against the obsessive thoughts and the accompanying awareness of their ridiculous nature.

Carl Westphal presented his paper “Uber Zwangvorstellungen” in 1877[2]. By linking compulsions to obsessions, Westphal was the first to characterize OCD with a definition very close to its current definition in contemporary diagnostic manuals. His definition comprises four criteria:

1. integrity of intelligence
2. absence of affective causal pathology
3. intrusion of obsessive-compulsive representations against the subjects will and the inability to suppress them
4. recognition of the bizarreness and abnormality of the representations

By the beginning of the 20th century, theories of OCD shifted towards psychoanalytical explanations. With Freud's writings on psychoanalysis of the Rat Man[3], OCD was explained to be resulting from their emotional antecedents. As a result of these theories, treatment of OCD turned away from attempts to relieve the obsessional symptoms themselves and toward treatment of unconscious conflicts which were presumed to underlie the symptoms.

Over the last few years, research on the biology of OCD grew exponentionally with ongoing studies of pharmacologic agents, neurosurgery, genetics, brain imaging, neuropsychological dysfunction and the association of OCD symptoms with other possibly related illnesses.

Today, OCD is viewed mainly as a neuropsychiatric disease, and is classified by the "Diagnostic and Statistical Manual of Mental Disorders - Fourth Edition – Text Revision" (DSM-IV-TR) as an anxiety disorder, and by the "International Statistical Classification of Diseases and Related Health Problems - 10th revision" (ICD-10) as a stand-alone disorder. The current definition of OCD, opposing obsessions (anxiety inducing) to compulsions (anxiety reducing) is largely derived from behavioral theories. From a neurochemical perspective, OCD is believed to be related to the serotonergic system[4]. From a neuro-anatomical perspective, OCD is believed to be related to the hyperactivity of prefrontal-striatal-thalamic circuitry[4].

## 1.2. DSM-IV-TR

The DSM-IV-TR, published by the American Psychiatric Association, is the handbook used most often in diagnosing mental disorders in the United States. The ICD-10 is a commonly-used alternative internationally.

The Diagnostic and Statistical Manual of Mental Disorders, presently in its fourth revised (IV-TR, 2000) edition, systemizes psychiatric diagnosis in five axes:
Axis I: major mental disorders, developmental disorders and learning disabilities
Axis II: underlying pervasive or personality conditions, as well as mental retardation
Axis III: any non-psychiatric medical condition ("somatic")
Axis IV: social functioning and impact of symptoms
Axis V: Global Assessment of Functioning
Common Axis I disorders include depression, anxiety disorders, OCD, bipolar disorder, attention deficit hyperactivity disorder (ADHD), and schizophrenia. Common Axis II disorders include borderline personality disorder, antisocial personality disorder, narcissistic personality disorder, and mild mental retardation.

For the diagnosis of OCD, the DSM-IV-TR requires that a patient has either obsessions or compulsions which are a significant source of distress; are time-consuming, or significantly interfere with the person's normal routine, occupational functioning, or usual social activities or relationships with others.
At some point during the course of the illness, the patient must recognize that the obsessions or compulsions are excessive or unreasonable. This is not a necessary requirement for young children.

The Quick Reference to the diagnostic criteria from DSM-IV-TR (2000) describes these obsessions and compulsions:

**Obsessions are defined by (1), (2), (3), and (4):**
1) Recurrent and persistent thoughts, impulses, or images that are experienced at some time during the disturbance, as intrusive and inappropriate and cause marked anxiety or distress.
2) The thoughts, impulses, or images are not simply excessive worries about real-life problems.
3) The person attempts to ignore or suppress such thoughts or impulses or to neutralize them with some other thought or action.
4) The person recognizes that the obsessions are the product of his or her own mind.

**Compulsions are defined by (1) and (2):**
1) Repetitive behaviors that the person feels driven to perform in response to an obsession, or according to rules that must be applied rigidly.
2) The behaviors or mental acts are aimed at checking. This can include mental compulsions such as praying, counting and repeating words silently. Such repetitive mental actions generally serve to decrease, prevent or reduce distress or some dreaded event or situation, however, these behaviors or mental acts either are not connected in a realistic way with what they are designed to neutralize or prevent, or are clearly excessive.

## 1.3. ICD-10

The classification is created from the World Health Organization (WHO). OCD is described under chapter V "Mental and behavioral disorders", subsection F40-48 "Neurotic, stress-related and somatoform disorders"

**Obsessive-Compulsive Disorder**

The essential feature is recurrent obsessional thoughts or compulsive acts. Obsessional thoughts are ideas, images, or impulses that enter the patient's mind again and again in a stereotyped form. They are almost invariably distressing and the patient often tries, unsuccessfully, to resist them. They are, however, recognized as his or her own thoughts, even though they are involuntary and often repugnant. Compulsive acts or rituals are stereotyped behaviors that are repeated again and again. They are not inherently enjoyable, nor do they result in the completion of inherently useful tasks. Their function is to prevent some objectively unlikely event, often involving harm to or caused by the patient, which he or she fears might otherwise occur. Usually, this behavior is recognized by the patient as pointless or ineffectual and repeated attempts are made to resist. Anxiety is almost invariably present. If compulsive acts are resisted the anxiety gets worse.

The ICD-10 defines several subtypes:

- Predominantly Obsessional Thoughts Or Ruminations

  These may take the form of ideas, mental images, or impulses to act, which are nearly always distressing to the subject. Sometimes the ideas are an indecisive, endless consideration of alternatives, associated with an inability to make trivial but necessary decisions in day-to-day living. The relationship between obsessional ruminations and depression is particularly close and a diagnosis of OCD should be preferred only if ruminations arise or persist in the absence of a depressive episode.
- Predominantly Compulsive Acts (Obsessional Rituals)

  The majority of compulsive acts are concerned with cleaning (particularly hand washing), repeated checking to ensure that a potentially dangerous situation has not been allowed to develop, or orderliness and tidiness. Underlying the overt behavior is a fear, usually of danger either to or caused by the patient, and the ritual is an ineffectual or symbolic attempt to avert that danger.
- Mixed obsessional thoughts and acts

## 1.4. Yale-Brown Obsessive-Compulsive Scale (Y-BOCS)

The Y-BOCS is one of the most widely used OCD measures. It was developed in 1986 for the assessment of symptom severity and treatment outcome in OCD, and has since demonstrated excellent validity and reliability[5]. It is accompanied by the Y-BOCS symptom checklist (Y-BOCS CL), which provides a comprehensive list of 64 obsessions and compulsions, classified in thirteen specific categories

## 1.5. The OCD-cycle

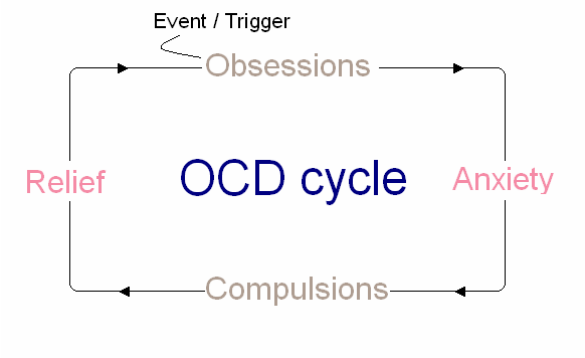


Figure 1: The OCD cycle

Figure 1 is often used as a visual aid to understand why some patients go through hours of meticulous rituals, only to find some temporal relief and why they engage in all kinds of absurd and time-consuming behaviors which end up making them feel unhappy / restricted for having done them.

The combination of having anxiety triggered, not being in control of this and finding relief through rituals that give you a sense of being in control, is making the effects of this OCD-cycle so powerful. This is why treatment (see chapter 1.4.) often includes some kind of behavioral therapy which tries to break the OCD-cycle by slowly but steadily taking away the section that brings the "relief".

## 1.6. Prevalence

OCD usually begins in early adulthood with over half of the patients becoming symptomatic by age 25. It affects men and women in roughly equal number. Even within the last decade, OCD was considered to be extremely rare (approximately 0.05% of the population). One of the most recent prevalence studies (Swiss population) reports a prevalence rate of 3.5 % for OCD (males 1.7%, females 5.4%). The onset of obsessive compulsive symptoms was 18 years (median); and in 70% before age 20. OCD was treated only in one third of cases[6].

The main reason for this "hidden epidemic" is that the patients want to keep it secret and don't want to disclose their "crazy" symptoms.

## 1.7. OCD: many manifestations

A paper by Merlo *et al.*[7] gives an excellent overview of the most common obsession and compulsions. In extreme cases of OCD, these obsessions and compulsions can occupy the patient every hour of the day. The following table was adapted from the Yale-Brown Obsessive-Compulsive Scale and the Children's Yale-Brown Obsessive-Compulsive Scale.

| Common obsessions in OCD | **OBSESSIVE CONCERNS** |
|---|---|
| **Contamination** | Dirt; germs; animals/insects; illnesses; bodily waste; contaminants; household cleaners; "sticky" substances; spreading contamination, germs, illnesses, etc |
| **Aggression** | Harming self or others (even accidentally); causing harm to self or others due to thoughts or behaviors; acting upon aggressive impulses; blurting out inappropriate words/phrases; stealing or breaking things; causing something terrible to happen; frightening/violent images |
| **Sexual** | Forbidden/perverse sexual thoughts, images; disturbing sexual impulses, desires; homosexuality; molestation; sexual acts toward others |
| **Hoarding/saving** | Losing things; throwing away objects that might be important |
| **Magical thinking** | Lucky/unlucky numbers, colors, names, etc |
| **Health/body** | Contracting illness (especially if fatal or rare); appearance;<br>physical abnormalities (real or imagined) |
| **Mortality/religion** | Dying and not going to Heaven; offending God; being sinful; morality/perfection; right/wrong |

| | |
|---|---|
| **Miscellaneous** | Remembering certain things; saying things exactly right; not saying certain words/phrases |
| Common compulsions in OCD | **COMPULSIVE RITUALS** |
| Washing & cleaning | Excessive/ritualized hand washing, showering, bathing, tooth brushing, grooming, toileting; cleaning clothing/personal items; avoiding "contaminated" objects/places |
| Checking | Checking locks, alarms, school supplies, homework, toys, books, etc; checking associated with washing, dressing, undressing, somatic concerns; checking that did/will not harm self or others; checking that nothing terrible did/will happen; checking for mistakes |
| Repeating | Rewriting; rereading; recopying; retying (e.g., shoelaces); erasing; going in/out door or taking items in/out of schoolbag; getting up/down from seat; repeating words/phrases |
| Counting | Counting objects; mental counting (especially up to a "magic" number); counting steps, chewing, hair-brushing, etc |
| Ordering/arranging | Lining up objects in a certain way; arranging things in specific patterns; making objects/piles/groups "even"; making things symmetrical; "balancing" actions (eg, doing thing on the right and on the left) |
| Hoarding & saving | Keeping unimportant/unnecessary items and/or trash; storing items of no particular value; having difficulty throwing things away; sorting through trash to ensure that nothing important has been thrown away |
| Superstitions | Touching/tapping routines to prevent bad things from happening; avoiding stepping on cracks, lines, etc; avoiding "unlucky" objects/places |
| Reassurance-seeking | Asking a parent to repeatedly answer the same questions; asking parents to describe what they are doing/planning to do; forcing family members to do things in a certain way or at a certain time; forcing family members to avoid certain things/activities |
| Miscellaneous | Mental rituals; needing to tell/ask/confess; ritualized eating behaviors; excessive list-making; needing to touch/tap/rub; needing to do things until it feels "just right" hair-pulling; measures to prevent something bad from happening |

Table 1: The many manifestations of OCD

## 1.8. High Comorbidity: Is OCD a discrete psychiatric disorder?

- About half of the OCD patients meet criteria for another psychiatric disease[6]. If OCD rarely presents in isolation of other conditions, the validity to retain OCD as a separate diagnostic category is in doubt.

- Besides a high comorbidity with other psychiatric disorders, OCD itself also encompasses a wide range of mental and behavioral phenomena[7].

One of the largest studies to report on comorbid DSM-IV diagnoses in a sample of OCD patients is by Denys *et al.*[8]. This study examined the cross-sectional prevalence of comorbid DSM-IV axis I, and personality disorders in 420 outpatients with primary OCD. Forty-six percent of the patients were diagnosed with a comorbid disorder. Twenty-seven percent met the criteria for at least one comorbid axis I disorder, 15.6 for a comorbid personality disorder, and 20.4 for both a comorbid axis I disorder and a personality disorder. Associated axis I comorbidity did not affect clinical severity of OCD, but was related to higher levels of depression and anxiety, whereas axis II comorbidity impaired the overall functioning to a higher extend. As follows, the treatment planning and prognosis of OCD patients should be adapted according to their comorbid status.

## 1.9. Does OCD represent different clinical subtypes?

OCD is a complex and heterogeneous disorder that encompasses a wide range of mental behavioral phenomena[7]. Studying clinical subtypes may enhance the likelihood of delineating pathogenic mechanisms, neurobiological underpinnings or genetic transmissions. By consequence, the identification of homogeneous subgroups has been a focus of interest and extensive investigation. Distinct clinical subtypes have formerly been built upon demographic and clinical characteristics (sex, age of onset, course of the disease), nature of OCD symptoms (predominance of obsessions or compulsions, washers/checkers) or comorbid conditions (tic disorder, schizophrenia)[9].

Lately strategies for identifying clinical subtypes have focused on using factor analytical techniques on generally recognized OCD-scales. One of the largest factor analyses on OCD is by Denys *et al.*[9]. This study aimed to identify symptom dimensions in OCD, in order to reveal distinct clinical phenotypes. A factor analysis of the of the Y-BOCS checklist on item level was performed on data from 335 outpatients with primary OCD. The relationship of demographic and clinical characteristics with the resulting factor scores was examined. A principal components analysis identified five consistent symptom dimensions:

1. Contamination and cleaning
2. Aggressive, sexual and religious obsessions
3. Somatic obsessions and checking
4. Symmetry and counting / arranging compulsions
5. High risk assessment and checking

Significant differences in sex distribution, age of onset, Y-BOCS scores, and familial prevalence of OCD in relation to symptom dimensions, which provide further evidence for distinct clinical phenotypes in OCD.

## 1.10. Treatment

Currently the most effective treatment for OCD is a combination of behavior therapy and pharmacotherapy[10-15]. This combined approach can be expected to improve the condition of most patients substantially, and occasionally completely, within a few months[16]. As a last resort, surgery to certain parts of the brain (psychosurgery) may be considered in certain severe cases resistant to all other treatments.

### 1.10.1. Cognitive-Behavior Therapy

Cognitive-behavioral therapy (CBT) refers to two distinct treatments: behavior therapy and cognitive therapy (CT). These treatments are increasingly offered in combination[13-17].

Usually, behavior therapy focuses on Exposure Response Prevention (ERP)[13] in which the patient is exposed to the feared situation or object and has to resist the urge to perform compulsions. Behaviorists regard compulsions as a form of learned avoidance, which are reinforced because they were perceived to reduce fear and anxiety. ERP begins with real life exposure to triggers that initiate the obsessive thoughts and accompanying anxiety. The typical ritual response is prevented and the patient remains in distress. With prolonged exposure to the triggering situation, unrelieved by compulsions, the patient habituates and his or her discomfort dissipates. Behavior therapy produces the largest changes in rituals, such as compulsive cleaning and checking, whereas changes in obsessive thoughts are less predictable. Behavior therapy is now regarded as the treatment of choice (in combination with pharmacotherapy) when behavioral rituals predominate[17].

CT helps the patients to identify and re-evaluate their obsessions and compulsions. For example, a patient who fears to shake hands may believe he will get infected. This interpretation of this fear can be challenged and re-evaluated so that shaking hands is no longer considered dangerous. CT can also help to re-evaluate the consequences of not performing compulsions, working towards eliminating them.

### 1.10.2. Pharmacotherapy

#### 1.10.2.1. Serotonin hypothesis

All Serotonin Reuptake Inhibitors (SRIs) have shown to be effective in OCD, and their exclusive efficacy has given grounds to the "serotonin hypothesis" that serotonin plays an important role in the pathogenesis of OCD[10, 18, 19]. The efficacy of clomipramine (like other tricyclics, clomipramine inhibits noradrenalin and serotonin uptake into central nerve terminals, possibly by blocking the membrane-pump of neurons, thereby increasing the concentration of transmitter monoamines at receptor sites) and Selective Serotonin Reuptake Inhibitors (SSRIs), such as paroxetine, fluvoxamine, fluoxetine, sertraline and citalopram in alleviating OCD symptoms has been firmly established in double-blind, placebo-controlled trials[16, 20-32]. However, this serotonergic hypothesis is not sufficient. It is clear that the dopaminergic mechanism is also implicated in the pathogenesis of OCD.

There has been some debate in the literature about which treatment is more effective, clomipramine or the SSRIs, but most clinicians agree that the efficacy is comparable. However, the safety and tolerability of the SSRIs clearly make them first-line therapy[33-40]. There is no evidence yet that one SSRI is more effective than another, but individual patients may respond better to one drug over another[41].

#### 1.10.2.2. SRIs with addition of antipsychotic drugs

Although SRIs are the mainstay of treatment for OCD, at least 40% of the patients do not respond to an initial medication trial[42]. Moreover, patients who fail to respond to a treatment with a SRI have 25% less chance to achieve response with another SRI compared to a patient who previously responded to a SRI[43]. Clearly, an improved understanding of determinants of response to pharmacotherapy may help to develop more efficient treatment strategies in OCD.

In case of non-response or partial response to SRIs, evidence has accumulated that the addition of antipsychotic drugs to SRIs might lead to symptom improvement. To date, risperidone, olanzapine, and quetiapine were shown to be effective as add-on to SRIs in a number of case reports and open studies[44-48]. Double-blind, placebo-controlled studies have confirmed the efficacy of risperidone[49], olanzapine[50] and haloperidol[51] for SRI-refractory OCD. Abovementioned studies also included OCD patients with comorbid disorders. The beneficial effect of the addition of the atypical antipsychotic quetiapine for treatment of SRI refractory OCD patients without DSM-IV axis I comorbidity is described by a double-blind, randomized, placebo controlled trial by Denys *et al.*[52].

#### 1.10.2.3. Venlafaxine

Drugs with potential activity against the symptoms of OCD are continually being investigated, and one that has received recent attention is venlafaxine, which is a structurally unique antidepressant agent. The drug and its active metabolite, O-desmethylvenlafaxine, are potent inhibitors of serotonin and noradrenalin reuptake and weakly inhibit dopamine reuptake. Serotonin reuptake inhibition is the most potent action of venlafaxine, so it occurs at lower doses. Higher doses cause the additional action of norepinephrine reuptake inhibition, whereas the highest doses are required for dopamine reuptake inhibition[41].

A number of meta-analyses have reported that treatment with clomipramine is more effective than treatment with SSRIs, an effect that has been attributed to its combined action at the serotonin and noradrenalin reuptake sites[53]. Since venlafaxine is also a potent inhibitor of the neuronal reuptake of both serotonin and noradrenalin, venlafaxine might be superior to SSRIs and clomipramine.

A double-blind comparison of venlafaxine and paroxetine in OCD by Denys *et al.*[54] shows that venlafaxine was equally effective as paroxetine in treating patients with OCD. Venlafaxine may be a useful therapy for OCD, but is not superior to SSRIs.
A single-blind comparison of venlafaxine and paroxetine in OCD by Albert *et al.*[55] shows a 36% response for venlafaxine vs. 50% response for clomipramine.

#### 1.10.2.4. The pathophysiology of OCD: a role for dopamine?

Adding an atypical antipsychotic drug to SRIs has proven to be beneficial in a number of neuropsychiatric disorders, such as major depression, schizophrenia and OCD. Combination therapies appear to be effective mostly in treatment-resistant cases and produce a more rapid response, possibly because of pharmacological interaction. (See chapter 1.4.2.2.)

Several studies provide *in vivo* evidence for abnormalities in the binding potential of the dopamine D2 receptor and the dopamine transporter (DAT). These studies suggest the direct involvement of the dopamine system in the pathophysiology of OCD.

- Denys *et al.* quantitated the striatal dopaminergic D2 receptor binding by [123I] iodobenzamide single photon emission computerized tomography in ten medication-free OCD patients and ten healthy controls, matched for age and gender. The dopamine D2 receptor binding in the left caudate nucleus (one of the basal ganglia nuclei involved with control of voluntary movement in the brain) was found to be significantly lower in the patients with OCD than in healthy controls[56].
- Van der Wee et al reported differences between OCD patients and healthy subjects in the [I-123]beta-CIT binding pattern for DAT in the left caudate and left putamen[57].
- Kim *et al.* showed that compared normal control adults, patients with OCD showed a significant increase of specific/non-specific DAT binding ratio in the right basal ganglia and a tendency towards an increase of specific/non-specific DAT binding ratio in the left basal ganglia[58].
- Hesse *et al.* showed a significantly reduced availability (corrected for age) of striatal DAT and of thalamic/hypothalamic, midbrain and brainstem serotonin transporter in OCD patients[59].

Several studies investigated the serotonin and/or dopamine levels in the brain after treatment with SRI, APD or a combination of them:

- Zhang *et al.*[60] showed that the combination of olanzapine with fluoxetine may synergistically increase dopamine and noradrenergic levels in the rat prefrontal cortex.
- Denys *et al.*[61] investigated the effects of combining the atypical antipsychotic drug quetiapine with an the SSRI fluvoxamine on extracellular serotonin and dopamine levels in the rat dorsal striatum (area involved in antipsychotic induced extrapyramidal symptoms), prefrontal cortex, nucleus accumbens (terminal dopaminergic areas

implicated in the pathophysiology of schizophrenia) and thalamus (may be of interest as it is implicated in the pathophysiology of OCD), by means of microdialysis coupled to high performance liquid chromatography (HPLC) with electrochemical detection.

- Although neither quetiapine nor fluvoxamine in monotherapy affected dopamine levels in the prefrontal cortex and thalamus, the combination produced a significant increase of dopamine levels in both these brain areas.
- It is worth noting that the administration of fluvoxamine plus quetiapine, did not result in augmented serotonin levels

### 1.10.3. Psychosurgery

Researchers have been studying repetitive transcranial magnetic stimulation[62], deep brain stimulation[63, 64] and neurosurgical approaches such as gamma-knife- and thermo-capsulotomy[65] and frameless stereotactic subcaudate tractotomy[66] to learn if these procedures are effective in treating treatment-resistant OCD. Repetitive transcranial magnetic stimulation has possibilities not only as a therapy but also as an instrument that can help researchers describe the neurocircuitries involved in OCD[67].

## 1.11. (Pharmaco)genetics of OCD

### 1.11.1. Genetics of OCD

Available evidence suggesting a genetic basis for OCD etiology includes: (a) twin studies with a concordance rate of approximately 50–60% in monozygotic twins compared to 10% in dizygotic twins[68, 69], and (b) family studies showing a significant aggregation of illness within families compared to population prevalence with an age corrected morbid risk as high as 35% in first-degree relatives[68, 70-76]. Furthermore, genome scan studies on OCD have identified a candidate region in 9p24 at marker D9S288[77, 78].

Because in all twin data reported to date, the concordance for monozygotic twins has always been <1.0, obsessive and compulsive behaviors are also influenced by nongenetic / environmental factors. Known environmental stressors are abuse, changes in living situation, illness, occupational changes or problems, school-related problems etc.

### 1.11.2. Pharmacogenetics of OCD

Research, by which the treatment of patients is tuned on the basis of their genotype, is called pharmacogenetics. Strong evidence suggests that genetic variation plays an important role in inter-individual differences in medication response and toxicity. Gene variants of drug-metabolizing enzymes can dramatically change the pharmacokinetics of a specific drug. Polymorphisms in genes (receptors, enzymes …) involved in the pharmacodynamics of a specific drug can also influence the clinical response.

A high proportion of OCD patients treated with SSRIs respond inadequately. To circumvent this, it is often necessary to switch between different SSRIs to find a more suitable alternative. This trial and error approach, where patients often do not display a full therapeutic response until several weeks after initiation, is unfavorable. This makes SSRI treatment of OCD a classic example where pharmacogenetics could bring a solution. Reviews by Veenstra-VanderWeele *et al.*[79] and Mancama *et al.*[80] give an overview of the role of pharmacogenetics in individualizing treatment with SSRIs and possible genetics influences on therapeutic response to drugs affecting the serotonin system.

- Drug metabolizing enzymes: Reuptake inhibitors of serotonin (citalopram, fluvoxamine, fluoxetine, paroxetine, sertraline) and noradrenaline (reboxetine) or both (venlafaxine), are almost totally biotransformed before excretion[81]. Metabolism generally proceeds through sequential or parallel oxidative pathways. They interact strongly with the cytochrome P450 enzymes, being both substrates and, in some cases, inhibitors (possible drug-drug interactions e.g. with tricyclic antidepressants). Particularly well characterized is the significant effect of the CYP2D6 and CYP2C19 polymorphisms on the pharmacokinetics of almost all tricyclic and many other antidepressants. The activities of these two enzymes are both bimodally distributed in the Caucasian population, allowing classification of individuals into extensive, intermediate and poor metabolizers. Some studies recommend dose adjustments for these antidepressants based on the CYP2D6 and/or CYP2C19 genotype of the patient[80-82].

- Pharmacodynamics of SSRIs: Pharmacogenetic studies on the pharmacodynamics of SSRI response have primarily been focused on 5-HTT polymorphisms. Some studies suggest the possible involvement of 5-$HT_{1A}$, 5-$HT_{1B}$, 5-$HT_{2A}$, 5-$HT_3$ polymorphisms. The gene encoding tryptophan hydroxylase has also been proposed as a candidate, given its pivotal role in serotonin formation[80]. Chapter 4 describes a pharmacogenetic study designed to determine whether polymorphisms of the serotonin transporter (5-HTT) and the receptor genes 5-$HT_{1B}$ and 5-$HT_{2A}$ affect the efficacy of SRI treatment in OCD.

### 1.11.3. Association studies

An excellent review by Hemmings *et al.* describes the current status of association studies in obsessive-compulsive disorder[83]. During the last 2 decades, a large number of association studies have been dedicated to disentangle the genetic components that may be involved in the etiology of OCD. The preliminary and frequently inconsistent nature of the data represented in the majority of OCD psychiatric genetic-association studies may seem discouraging. Failure to confirm previously identified susceptibility loci could result from a number of reasons, including the potential for population admixture, the clinical heterogeneity of OCD, small sample sizes (and subsequent lack of power) or epistasis.

Epistasis takes place when the action of one gene is modified by one or more other genes. The effect can occur directly at the genomic level, where one gene could code for a protein preventing transcription of the other gene. Alternatively, the effect can occur at the phenotypic level. For example, the gene causing albinism would hide the gene controlling the color of a person's hair. Epistasis in the statistical sense means a deviation from the additive effects of two or more loci and occurs at the population level. OCD has a complex multifactorial etiology. For this reason, it is not surprising that studies of the effects of single genes on OCD have often failed to replicate the original findings. This failure is because the impact of single alleles on the risk of OCD is dependent on genetic variations at other loci (i.e. gene-gene interactions) and on environmental factors (i.e. gene-environment interactions). Thus, studies that do not consider the appropriate genetic and/or environmental contexts may not identify important susceptibility loci. The identification and characterization of such gene-gene and gene-environment interactions have been limited by a lack of powerful statistical methods and/or a lack of large enough sample sizes.[84]

A valid way to overcome the clinical heterogeneity of the patients and possibly epistasis would be to stratify the patient sample according to clinically defined sub-types, such as obsession and compulsion subtypes, age at onset of the disorder, and severity of the disorder. Unfortunately, the number of subjects decreases after stratification, thereby limiting the power of the studies. Of course, epistasis could also be tackled by analyzing as much polymorphisms as possible in the same study.

In light of the putative role of the serotonergic, dopaminergic and possibly (nor)adrenergic system in OCD, following polymorphisms were analyzed in a sample of >100 OCD patients and a control sample of >100 ethnically matched Caucasian subjects by means of a case-control study (See chapter 2 and 3).

#### 1.11.3.1. Taq IA polymorphism in the non-coding region flanking the 3' end of the dopamine D2 receptor (DRD2) gene

This polymorphism creates a restriction site for the Taq I restriction enzyme. The TaqIA1 (uncut by Taq I restriction enzyme) allele has been associated with measures of low striatal dopamine receptor D2 density[85, 86].

The non-coding region flanking the 3' end of the DRD2 gene (A region) contains nucleotide sequences that do not code for the structure of the receptor, but may have an important role in regulating the expression of the gene by its influence on the mRNA stability. On the other hand, the relationship between the Taq IA polymorphism and the phenotype data could be due to linkage between the DRD2 Taq IA polymorphism and another adjacent functional polymorphism.

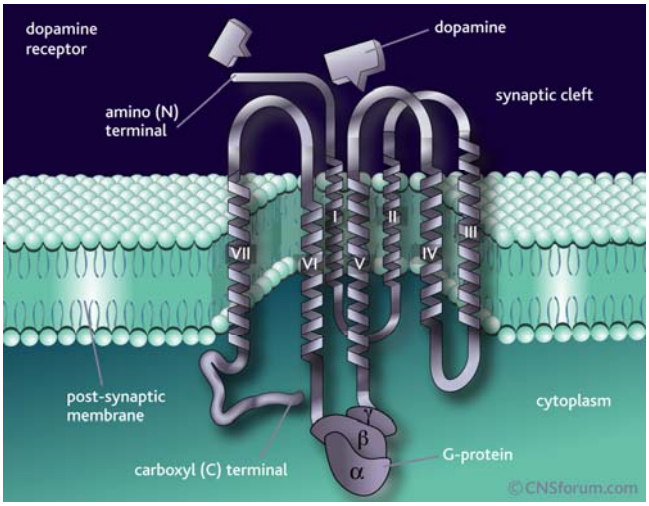


Figure 2: There are four broad 'superfamilies' of receptors: (1) the channel-linked (ionotropic) receptors; (2) the G-protein coupled (metabotropic) receptors; (3) the kinase-linked receptors; and (4) receptors that regulate gene transcription. Dopamine receptors belong to the G-protein coupled superfamily. They are membrane receptors that have 7

transmembrane spanning α-helices. Dopamine binding to the 'binding groove' on the extracellular portion of the receptor activates the G-proteins, which initiate secondary messenger signaling pathways. The downstream effect will be either inhibitory or stimulatory, depending on the types of G-protein linked to the receptor – dopamine D1, D5 receptors (D1 – like) are linked to stimulatory G-proteins, whereas dopamine D2, D3, D4 (D2 – like) are linked to inhibitory G-proteins. The D2 receptor exists as an autoreceptor and as a post-synaptic receptor. Typical and atypical antipsychotics are dopamine D2-like antagonists[87].

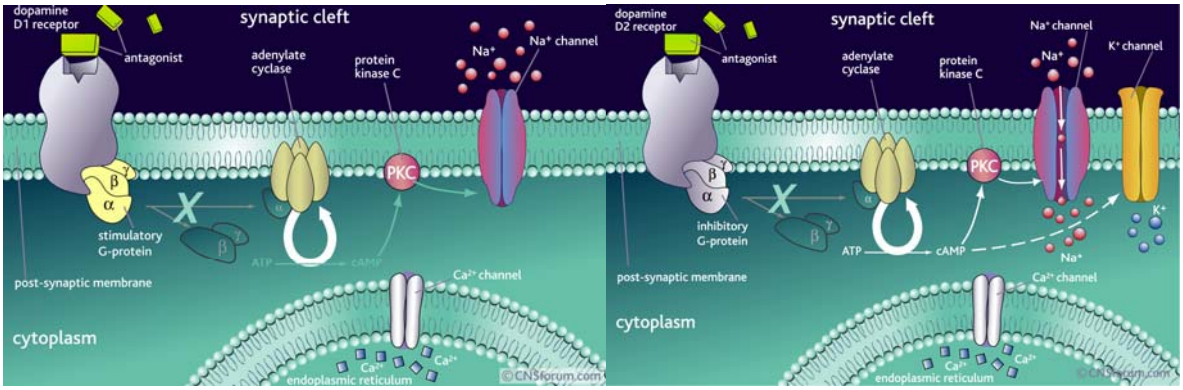


Figure 3: Signal transduction mechanism of the dopamine receptor

### 1.11.3.2. Catechol-O-Methyl Transferase (COMT) NlaIII high/low polymorphism

This polymorphism is a G to A transition at codon 158 (creating a restriction site for the restriction enzyme *Nla*III) of the COMT gene that results in a valine-to-methionine substitution in the enzyme. A valine results in a heat stable, high activity COMT variant, whereas a methionine results in a heat labile low activity variant[88].

COMT is involved in the breakdown of the catecholamine neurotransmitters, dopamine, epinephrine and norepinephrine. The enzyme introduces a methyl group to the catecholamine which is donated by S-adenosyl methionine (SAM).

### 1.11.3.3. 1438 A/G polymorphism within the promoter region of the postsynaptic 5-$HT_{2A}$ receptor

This polymorphism has been associated with several behavioral disorders. The A to G transition at nucleotide 1438 creates a restriction site for the restriction enzyme *Msp*I.

The 5-hydroxytryptamine (5-HT)-1, 2, 4, 5, 6 and 7 receptors belong to the G-protein coupled superfamily. They are membrane receptors that have 7 transmembrane spanning α-helices. 5-HT binding to the 'binding groove' on the extracellular portion of the receptor activates the G-proteins, which initiate secondary messenger signaling pathways. The downstream effect is either inhibitory or stimulatory, depending on the type of G-protein linked to the receptor – 5-HT1 receptors are linked to inhibitory G-proteins, whereas 5-HT2, 4, 6 and 7 are linked to stimulatory G-proteins. The $5\text{-HT}_{2A}$ receptor is the most important serotonergic receptor with respect to behavioral effects.

#### 1.11.3.4. $5\text{-HT}_{1D\beta}$: G861C polymorphism silent G-to-C substitution at nucleotide 861 of the coding region of the $5\text{-HT}_{1DB}$ autoreceptor gene

The $5\text{-HT}_{1D\beta}$ autoreceptor regulates the presynaptic inhibition of the serotonin release.

This polymorphism is a G to C transition at nucleotide 861 (creating a restriction site for the restriction enzyme *Hinc* II) of the coding region of the $5\text{-HT}_{1D\beta}$ gene. Because both alleles encode valine at this position (condons are GTC and GTG), the structure of the receptor is not changed by the polymorphism. However, there could be a linkage between the G861C polymorphism and another adjacent functional polymorphism.

#### 1.11.3.5. 5-HTTLPR: serotonin transporter gene (5-HTT) promoter 44-bp deletion

The serotonin transporter protein causes reuptake of serotonin from the synapse. In this way, it influences the amount of serotonin present in the synapse, and thus the serotonin effects on the receiving neuron's receptor.

The short variant of this polymorphism reduces the transcriptional efficiency of the 5-HTT gene promoter, resulting in decreased 5-HTT expression and 5-HT reuptake[89].

# References


1. Esquirol J: **Des maladies mentales considerées sous les rapports médical, hygiéniques et médico-legal**. *Paris: Ballière* 1838.

2. Westphal C: **Über zwangvorstellungen**. *Archive fur psychiatrie und nervenheilkunde* 1878:734-760.

3. Freud S: **Notes upon a case of obsessional neurosis (1909)**. *Complete Psychological Works, standard edition, vol 10* 1955:153-318.

4. Navarro JF, Puigcerver A: **Biological basis of obsessive-compulsive disorder**. *Psicologia Conductual* 1998, **6**(1):79-101.

5. Taylor S: **Assessment of Obsessions and Compulsions - Reliability, Validity and Sensitivity to Treatment Effects**. *Clinical Psychology Review* 1995, **15**(4):261-296.

6. Angst J, Gamma A, Endrass J, Goodwin R, Ajdacic V, Eich D, Rossler W: **Obsessive-compulsive severity spectrum in the community: prevalence, comorbidity, and course**. *European Archives of Psychiatry and Clinical Neuroscience* 2004, **254**(3):156-164.

7. Merlo LJ, Storch EA: **Obsessive-compulsive disorder: Tools for recognizing its many expressions**. *Journal of Family Practice* 2006, **55**(3):217-222.

8. Denys D, Tenney N, van Megen HJGM, de Geus F, Westenberg HGM: **Axis I and II comorbidity in a large sample of patients with obsessive-compulsive disorder**. *Journal of Affective Disorders* 2004, **80**(2-3):155-162.

9. Denys D, de Geus F, van Megen HJGM, Westenberg HGM: **Use of factor analysis to detect potential phenotypes in obsessive-compulsive disorder**. *Psychiatry Research* 2004, **128**(3):273-280.

10. Denys D: **Pharmacotherapy of obsessive-compulsive disorder and obsessive-compulsive spectrum disorders**. *Psychiatric Clinics of North America* 2006, **29**(2):553-+.

11. Kampman M, Keijsers GPJ, Hoogduin CAL, Verbraak MJPM: **Addition of cognitive-behaviour therapy for obsessive-compulsive disorder patients non-responding to fluoxetine**. *Acta Psychiatrica Scandinavica* 2002, **106**(4):314-319.

12. Greist JH, Jefferson JW: **Pharmacotherapy for obsessive-compulsive disorder**. *British Journal of Psychiatry* 1998, **173**:64-70.

13. Marks I: **Behaviour therapy for obsessive-compulsive disorder: A decade of progress**. *Canadian Journal of Psychiatry-Revue Canadienne De Psychiatrie* 1997, **42**(10):1021-1027.

14. O'Connor KP, Aardema F, Robillard S, Guay S, Pelissier MC, Todorov C, Borgeat F, Leblanc V, Grenier S, Doucet P: **Cognitive behaviour therapy and medication in**

**the treatment of obsessive-compulsive disorder**. *Acta Psychiatrica Scandinavica* 2006, **113**(5):408-419.

15. O'Connor K, Todorov C, Robillard S, Borgeat F, Brault M: **Cognitive-behaviour therapy and medication in the treatment of obsessive-compulsive disorder: A controlled study**. *Canadian Journal of Psychiatry-Revue Canadienne De Psychiatrie* 1999, **44**(1):64-71.

16. Foa EB, Liebowitz MR, Kozak MJ, Davies S, Campeas R, Franklin ME, Huppert JD, Kjernisted K, Rowan V, Schmidt AB *et al*: **Randomized, placebo-controlled trial of exposure and ritual prevention, clomipramine, and their combination in the treatment of obsessive-compulsive disorder**. *American Journal of Psychiatry* 2005, **162**(1):151-161.

17. Mataix-Cols D, Marks IM, Greist JH, Kobak KA, Baer L: **Obsessive-compulsive symptom dimensions as predictors of compliance with and response to behaviour therapy: Results from a controlled trial**. *Psychotherapy and Psychosomatics* 2002, **71**(5):255-262.

18. Fineberg NA, Gale TM: **Evidence-based pharmacotherapy of obsessive-compulsive disorder**. *International Journal of Neuropsychopharmacology* 2005, **8**(1):107-129.

19. Dougherty DD, Rauch SL, Jenike MA: **Pharmacotherapy for obsessive-compulsive disorder**. *Journal of Clinical Psychology* 2004, **60**(11):1195-1202.

20. Geller DA, Wagner KD, Emslie G, Murphy T, Carpenter DJ, Wetherhold E, Perera P, Machin A, Gardiner C: **Paroxetine treatment in children and adolescents with obsessive-compulsive disorder: A randomized, multicenter, double-blind, placebo-controlled trial**. *Journal of the American Academy of Child and Adolescent Psychiatry* 2004, **43**(11):1387-1396.

21. Kamijima K, Murasaki M, Asai M, Higuchi T, Nakajima T, Taga C, Matsunaga H: **Paroxetine in the treatment of obsessive-compulsive disorder: randomized, double-blind, placebo-controlled study in Japanese patients**. *Psychiatry and Clinical Neurosciences* 2004, **58**(4):427-433.

22. Hollander E, Koran LM, Goodman WK, Greist JH, Ninan PT, Yang HC, Li D, Barbato LM: **A double-blind, placebo-controlled study of the efficacy and safety of controlled-release fluvoxamine in patients with obsessive-compulsive disorder**. *Journal of Clinical Psychiatry* 2003, **64**(6):640-647.

23. Geller DA, Hoog SL, Heiligenstein JH, Ricardi R, Tamura R, Kluszynski S, Jacobson JG: **Fluoxetine treatment for obsessive-compulsive disorder in children and adolescents: A placebo-controlled clinical trial**. *Biological Psychiatry* 2001, **49**(8):98S-99S.

24. Geller DA, Hoog SL, Heiligenstein JH, Ricardi RK, Tamura R, Kluszynski S, Jacobson JG, Team FPOS: **Fluoxetine treatment for obsessive-compulsive disorder in children and adolescents: A placebo-controlled clinical trial**. *Journal of the American Academy of Child and Adolescent Psychiatry* 2001, **40**(7):773-779.

25. Kronig MH, Apter J, Asnis G, Bystritsky A, Curtis G, Ferguson J, Landbloom R, Munjack D, Riesenberg R, Robinson D *et al*: **Placebo-controlled, multicenter study of sertraline treatment for obsessive-compulsive disorder**. *Journal of Clinical Psychopharmacology* 1999, **19**(2):172-176.

26. Ushijima S, Kamijima K, Asai M, Murasaki M, Nakajima T, Kudo Y, Tashiro N, Kurihara M, Miura S: **Clinical evaluation of sertraline hydrochloride, a selective serotonin reuptake inhibitor in the treatment of obsessive-compulsive disorder - A double-blind placebo controlled trial**. *Japanese Journal of Neuropsychopharmacology* 1997, **19**(6):603-623.

27. Goodman WK, Kozak MJ, Liebowitz M, White KL: **Treatment of obsessive-compulsive disorder with fluvoxamine: A multicentre, double-blind, placebo-controlled trial**. *International Clinical Psychopharmacology* 1996, **11**(1):21-29.

28. Greist JH, Jefferson JW, Kobak KA, Chouinard G, Duboff E, Halaris A, Kim SW, Koran L, Liebowtiz MR, Lydiard B *et al*: **A 1-Year Double-Blind Placebo-Controlled Fixed-Dose Study of Sertraline in the Treatment of Obsessive-Compulsive Disorder**. *International Clinical Psychopharmacology* 1995, **10**(2):57-65.

29. Chouinard G: **Sertraline in the Treatment of Obsessive-Compulsive Disorder - 2 Double-Blind, Placebo-Controlled Studies**. *International Clinical Psychopharmacology* 1992, **7**:37-41.

30. Montgomery SA, Montgomery DB, Fineberg N: **Early Response with Clomipramine in Obsessive-Compulsive Disorder - a Placebo Controlled-Study**. *Progress in Neuro-Psychopharmacology & Biological Psychiatry* 1990, **14**(5):719-727.

31. Greist JH, Jefferson JW, Rosenfeld R, Gutzmann LD, March JS, Barklage NE: **Clomipramine and Obsessive-Compulsive Disorder - a Placebo-Controlled Double-Blind-Study of 32 Patients**. *Journal of Clinical Psychiatry* 1990, **51**(7):292-297.

32. Jenike MA, Baer L, Summergrad P, Weilburg JB, Holland A, Seymour R: **Obsessive-Compulsive Disorder - a Double-Blind, Placebo-Controlled Trial of Clomipramine in 27 Patients**. *American Journal of Psychiatry* 1989, **146**(10):1328-1330.

33. Piccinelli M, Pini S, Bellantuono C, Wilkinson G: **Efficacy of Drug-Treatment in Obsessive-Compulsive Disorder - a Metaanalytic Review**. *British Journal of Psychiatry* 1995, **166**:424-443.

34. Mundo E, Rouillon F, Figuera ML, Stigler M: **Fluvoxamine in obsessive-compulsive disorder: similar efficacy but superior tolerability in comparison with clomipramine**. *Human Psychopharmacology-Clinical and Experimental* 2001, **16**(6):461-468.

35. Mundo E, Maina G, Uslenghi C, Grp MS: **Multicentre, double-blind, comparison of fluvoxamine and clomipramine in the treatment of obsessive-compulsive disorder**. *International Clinical Psychopharmacology* 2000, **15**(2):69-76.

36. Milanfranchi A, Ravagli S, Lensi F, Marazziti D, Cassano GB: **A double-blind study of fluvoxamine and clomipramine in the treatment of obsessive-compulsive disorder**. *International Clinical Psychopharmacology* 1997, **12**(3):131-136.

37. Bisserbe JC, Lane RM, Flament MF, vanMoffaert M, Magerman J, Ansart E, Danan A, Bensoussan M, Faure M, Hantz D *et al*: **A double-blind comparison of sertraline and clomipramine in outpatients with obsessive-compulsive disorder**. *European Psychiatry* 1997, **12**(2):82-93.

38. LopezIbor JJ, Saiz J, Cottraux J, Note I, Vinas R, Bourgeois M, Hernandez M, GomezPerez JC: **Double-blind comparison of fluoxetine versus clomipramine in the treatment of obsessive compulsive disorder**. *European Neuropsychopharmacology* 1996, **6**(2):111-118.

39. Koran LM, McElroy SL, Davidson JRT, Rasmussen SA, Hollander E, Jenike MA: **Fluvoxamine versus clomipramine for obsessive-compulsive disorder: A double-blind comparison**. *Journal of Clinical Psychopharmacology* 1996, **16**(2):121-129.

40. Freeman CPL, Trimble MR, Deakin JFW, Stokes TM, Ashford JJ: **Fluvoxamine Versus Clomipramine in the Treatment of Obsessive-Compulsive Disorder - a Multicenter, Randomized, Double-Blind, Parallel-Group Comparison**. *Journal of Clinical Psychiatry* 1994, **55**(7):301-305.

41. Phelps NJ, Cates ME: **The role of venlafaxine in the treatment of obsessive-compulsive disorder**. *Annals of Pharmacotherapy* 2005, **39**(1):136-140.

42. Goodman WK, Ward HE, Murphy TK: **Biologic approaches to treatment-refractory obsessive-compulsive disorder**. *Psychiatric Annals* 1998, **28**(11):641-+.

43. Denys D, Burger H, van Megen H, de Geus F, Westenberg H: **A score for predicting response to pharmacotherapy in obsessive-compulsive disorder**. *International Clinical Psychopharmacology* 2003, **18**(6):315-322.

44. Horcajadas FA, Soto JA, Garcia-Cantalapiedra MJ, Calvin JLR, Morales J, Salgado M: **Effectiveness and tolerability of addition of risperidone in obsessive-compulsive disorder with poor response to serotonin reuptake inhibitors**. *Actas Espanolas De Psiquiatria* 2006, **34**(3):147-152.

45. Yoshimura R, Kaneko S, Shinkai K, Nakamura J: **Successful treatment for obsessive-compulsive disorder with addition of low-dose risperidone to fluvoxamine: Implications for plasma levels of catecholamine metabolites and serum brain-derived neurotrophic factor levels**. *Psychiatry and Clinical Neurosciences* 2006, **60**(3):389-393.

46. Arias E, Soto JA, Garcia MJ, Rodriguez-Calvin JL, Morales J, Salgado M: **Efficacy and tolerance of risperidone addition in serotonin reuptake inhibitors (SRI) treatment for refractory obsessive-compulsive disorder**. *European Neuropsychopharmacology* 2002, **12**:S341-S341.

47. Denys D, van Megen H, Westenberg H: **Quetiapine addition to serotonin reuptake inhibitor treatment in patients with treatment-refractory obsessive-compulsive disorder: An open-label study**. *Journal of Clinical Psychiatry* 2002, **63**(8):700-703.

48. Weiss EL, Potenza MN, McDougle CJ, Epperson CN: **Olanzapine addition in obsessive-compulsive disorder refractory to selective serotonin reuptake inhibitors: An open-label case series**. *Journal of Clinical Psychiatry* 1999, **60**(8):524-527.

49. McDougle CJ, Epperson CN, Pelton GH, Wasylink S, Price LH: **A double-blind, placebo-controlled study of risperidone addition in serotonin reuptake inhibitor-refractory obsessive-compulsive disorder**. *Archives of General Psychiatry* 2000, **57**(8):794-801.

50. Shapira NA, Ward HE, Mandoki M, Murphy TK, Yang MC, Blier P, Goodman WK: **A double-blind, placebo-controlled trial of olanzapine addition in fluoxetine-refractory obsessive-compulsive disorder**. *Biological Psychiatry* 2004, **55**(5):553-555.

51. McDougle CJ, Goodman WK, Leckman JF, Lee NC, Heninger GR, Price LH: **Haloperidol Addition in Fluvoxamine-Refractory Obsessive-Compulsive Disorder - a Double-Blind, Placebo-Controlled Study in Patients with and without Tics**. *Archives of General Psychiatry* 1994, **51**(4):302-308.

52. Denys D, de Geus F, van Megen HJGM, Westenberg HGM: **A double-blind, randomized, placebo-controlled trial of quetiapine addition in patients with obsessive-compulsive disorder refractory to serotonin Reuptake inhibitors**. *Journal of Clinical Psychiatry* 2004, **65**(8):1040-1048.

53. Ackerman DL, Greenland S: **Multivariate meta-analysis of controlled drug studies for obsessive-compulsive disorder**. *Journal of Clinical Psychopharmacology* 2002, **22**(3):309-317.

54. Denys D, van der Wee N, van Megen HJGM, Westenberg HGM: **A double blind comparison of venlafaxine and paroxetine in obsessive-compulsive disorder**. *Journal of Clinical Psychopharmacology* 2003, **23**(6):568-575.

55. Albert U, Aguglia E, Maina G, Bogetto F: **Venlafaxine versus clomipramine in the treatment of obsessive-compulsive disorder: A preliminary single-blind, 12-week, controlled study**. *Journal of Clinical Psychiatry* 2002, **63**(11):1004-1009.

56. Denys D, van der Wee N, Janssen J, De Geus F, Westenberg HGM: **Low level of dopaminergic D-2 receptor binding in obsessive-compulsive disorder**. *Biological Psychiatry* 2004, **55**(10):1041-1045.

57. van der Wee NJ, Stevens H, Hardeman JA, Mandl RC, Denys DA, van Megen HJ, Kahn RS, Westenberg HM: **Enhanced dopamine transporter density in psychotropic-naive patients with obsessive-compulsive disorder shown by [I-123]beta-CIT SPECT**. *American Journal of Psychiatry* 2004, **161**(12):2201-2206.

58. Kim CH, Koo MS, Cheon KA, Ryu YH, Lee JD, Lee HS: **Dopamine transporter density of basal ganglia assessed with [I-123]IPT SPET in obsessive-compulsive disorder**. *European Journal of Nuclear Medicine and Molecular Imaging* 2003, **30**(12):1637-1643.

59. Hesse S, Mullller U, Lincke T, Barthel H, Villmann T, Angermeyer MC, Sabri O, Stengler-Wenzke K: **Serotonin and dopamine transporter imaging in patients with obsessive-compulsive disorder**. *Psychiatry Research-Neuroimaging* 2005, **140**(1):63-72.

60. Zhang W, Perry KW, Wong DT, Potts BD, Bao JQ, Tollefson GD, Bymaster FP: **Synergistic effects of olanzapine and other antipsychotic agents in combination with fluoxetine on norepinephrine and dopamine release in rat prefrontal cortex**. *Neuropsychopharmacology* 2000, **23**(3):250-262.

61. Denys D, Klompmakers AA, Westenberg HGM: **Synergistic dopamine increase in the rat prefrontal cortex with the combination of quetiapine and fluvoxamine**. *Psychopharmacology* 2004, **176**(2):195-203.

62. Mantovani A, Lisanby SH, Pieraccini F, Ulivelli M, Castrogiovanni P, Rossi S: **Repetitive transcranial magnetic stimulation (rTMS) in the treatment of obsessive-compulsive disorder (OCD) and Tourette's syndrome (TS)**. *International Journal of Neuropsychopharmacology* 2006, **9**(1):95-100.

63. Abelson JL, Curtis GC, Sagher O, Albucher RC, Harrigan M, Taylor SF, Martis B, Giordani B: **Deep brain stimulation for refractory obsessive-compulsive disorder**. *Biological Psychiatry* 2005, **57**(5):510-516.

64. Gabriels L, Cosyns P, Nuttin B, Demeulemeester H, Gybels J: **Deep brain stimulation for treatment-refractory obsessive-compulsive disorder: psychopathological and neuropsychological outcome in three cases**. *Acta Psychiatrica Scandinavica* 2003, **107**(4):275-282.

65. Oliver B, Gascon J, Aparicio A, Ayats E, Rodriguez R, de Leon JLM, Garcia-Bach M, Soler PA: **Bilateral anterior capsulotomy for refractory obsessive-compulsive disorders**. *Stereotactic and Functional Neurosurgery* 2003, **81**(1-4):90-95.

66. Woerdeman PA, Willems PWA, Noordmans HJ, van der Sprenkel JWB, van Rijen PC: **Frameless stereotactic subcaudate tractotomy for intractable obsessive-compulsive disorder**. *Acta Neurochirurgica* 2006, **148**(6):633-637.

67. Pallanti S, Hollander E, Goodman WK: **A qualitative analysis of nonresponse: Management of treatment-refractory obsessive-compulsive disorder**. *Journal of Clinical Psychiatry* 2004, **65**:6-10.

68. Eapen V, Pauls D, Robertson M: **The role of clinical phenotypes in understanding the genetics of obsessive-compulsive disorder**. *Journal of Psychosomatic Research* 2006, **61**:359-364.

69. Rasmussen SA, Tsuang MT: **Clinical Characteristics and Family History in Dsm-III Obsessive-Compulsive Disorder**. *American Journal of Psychiatry* 1986, **143**(3):317-322.

70. Lenane MC, Swedo SE, Leonard H, Pauls DL, Sceery W, Rapoport JL: **Psychiatric-Disorders in 1St-Degree Relatives of Children and Adolescents with Obsessive-Compulsive Disorder**. *Journal of the American Academy of Child and Adolescent Psychiatry* 1990, **29**(3):407-412.

71. Riddle MA, Scahill L, King R, Hardin MT, Towbin KE, Ort SI, Leckman JF, Cohen DJ: **Obsessive-Compulsive Disorder in Children and Adolescents - Phenomenology and Family History**. *Journal of the American Academy of Child and Adolescent Psychiatry* 1990, **29**(5):766-772.

72. Black DW, Noyes R, Goldstein RB, Blum N: **A Family Study of Obsessive-Compulsive Disorder**. *Archives of General Psychiatry* 1992, **49**(5):362-368.

73. Pauls DL, Alsobrook JP, Goodman W, Rasmussen S, Leckman JF: **A Family Study of Obsessive-Compulsive Disorder**. *American Journal of Psychiatry* 1995, **152**(1):76-84.

74. Nestadt G, Samuels J, Riddle M, Bienvenu OJ, Liang KY, LaBuda M, Walkup J, Grados M, Hoehn-Saric R: **A family study of obsessive-compulsive disorder**. *Archives of General Psychiatry* 2000, **57**(4):358-363.

75. do Rosario-Campos MC, Leckman JF, Curi M, Quatrano S, Katsovitch L, Miguel EC, Pauls DL: **A family study of early-onset obsessive-compulsive disorder**. *American Journal of Medical Genetics Part B-Neuropsychiatric Genetics* 2005, **136B**(1):92-97.

76. Chabane N, Delorme R, Millet B, Mouren MC, Leboyer M, Pauls D: **Early-onset obsessive-compulsive disorder: a subgroup with a specific clinical and familial pattern?** *Journal of Child Psychology and Psychiatry* 2005, **46**(8):881-887.

77. Hanna GL, Veenstra-VanderWeele J, Cox NJ, Boehnke M, Himle JA, Curtis GC, Leventhal BL, Cook EH: **Genome-wide linkage analysis of families with obsessive-compulsive disorder ascertained through pediatric probands**. *American Journal of Medical Genetics* 2002, **114**(5):541-552.

78. Willour VL, Shugart YY, Samuels J, Grados M, Cullen B, Bienvenu OJ, Wang Y, Liang KY, Valle D, Hoehn-Saric R *et al*: **Replication study supports evidence for linkage to 9p24 in obsessive-compulsive disorder**. *American Journal of Human Genetics* 2004, **75**(3):508-513.

79. Veenstra-VanderWeele J, Anderson GM, Cook EH: **Pharmacogenetics and the serotonin system: initial studies and future directions**. *European Journal of Pharmacology* 2000, **410**(2-3):165-181.

80. Mancama D, Kerwin RW: **Role of pharmacogenomics in individualising treatment with SSRIs**. *Cns Drugs* 2003, **17**(3):143-151.

81. Caccia S: **Metabolism of the newer antidepressants - An overview of the pharmacological and pharmacokinetic implications**. *Clinical Pharmacokinetics* 1998, **34**(4):281-302.

82. Kirchheiner J, Brosen K, Dahl ML, Gram LF, Kasper S, Roots I, Sjoqvist F, Spina E, Brockmoller J: **CYP2D6 and CYP2C19 genotype-based dose recommendations for antidepressants: a first step towards subpopulation-specific dosages**. *Acta Psychiatrica Scandinavica* 2001, **104**(3):173-192.

83. Hemmings SMJ, Stein DJ: **The current status of association studies in obsessive-compulsive disorder**. *Psychiatric Clinics of North America* 2006, **29**(2):411-+.

84. Moore JH, Williams SM: **New strategies for identifying gene-gene interactions in hypertension**. *Annals of Medicine* 2002, **34**(2):88-95.

85. Jonsson EG, Nothen MM, Grunhage F, Farde L, Nakashima Y, Propping P, Sedvall GC: **Polymorphisms in the dopamine D2 receptor gene and their relationships to striatal dopamine receptor density of healthy volunteers**. *Molecular Psychiatry* 1999, **4**(3):290-296.

86. Pohjalainen T, Rinne JO, Nagren K, Lehikoinen P, Anttila K, Syvalahti EKG, Hietala J: **The A1 allele of the human D-2 dopamine receptor gene predicts low D-2 receptor availability in healthy volunteers**. *Molecular Psychiatry* 1998, **3**(3):256-260.

87. Strange PG: **Antipsychotic drugs: Importance of dopamine receptors for mechanisms of therapeutic actions and side effects**. *Pharmacological Reviews* 2001, **53**(1):119-133.

88. Lachman HM, Papolos DF, Saito T, Yu YM, Szumlanski CL, Weinshilboum RM: **Human catechol-O-methyltransferase pharmacogenetics: Description of a functional polymorphism and its potential application to neuropsychiatric disorders**. *Pharmacogenetics* 1996, **6**(3):243-250.

89. Heils A, Teufel A, Petri S, Stober G, Riederer P, Bengel D, Lesch KP: **Allelic variation of human serotonin transporter gene expression**. *Journal of Neurochemistry* 1996, **66**(6):2621-2624.

# 2. Association between serotonergic candidate genes and specific phenotypes of obsessive compulsive disorder

JOURNAL OF AFFECTIVE DISORDERS (2006) **91**(1): 39-44 MAR

Damiaan Denys [1], Filip Van Nieuwerburgh[2], Dieter Deforce[2] Herman G.M. Westenberg [1]

1) Rudolf Magnus Institute of Neuroscience, Department of Psychiatry University Medical Center Utrecht, Utrecht, the Netherlands.

2) Laboratory of Pharmaceutical Biotechnology, Ghent University, Harelbekestraat 72, 9000 Ghent, Belgium.

## ABSTRACT

**Background**: The successful use of serotonin reuptake inhibitors (SRIs) in obsessive-compulsive disorder (OCD) has led to the hypothesis that serotonin plays a pivotal role in the pathogenesis of OCD. The purpose of the present study was to investigate the role of the serotonin transporter (5-HTT) and serotonin 5-$HT_{1D}$ and 5-$HT_{2A}$ receptor genes in OCD.

**Method**: The distribution of polymorphic variants was analyzed in 156 OCD cases and 134 control individuals by means of case-control association studies. Potential relevant OCD phenotypes founded on age of onset, positive family history for OCD, clinical subtypes, comorbidity and symptom severity were stratified according to 5-HTT, 5-$HT_{1D}$ and 5-$HT_{2A}$ genotypes.

**Results**: Patients did not show significant differences in genotype distribution and allele frequency for polymorphisms investigated relative to controls. However, taking in account OCD phenotypes, we found indication towards an association of the 5-HTTLPR S-allele with female OCD patients, and the 5-$HT_{2A}$ G-allele and GG genotype with patients with a positive family history of OCD and an early onset of disease.

**Conclusions**: Our data yields interesting preliminary results as regards the genetic underpinnings of OCD phenotypes that warrant further discussion and investigation.

## 2.1. Introduction

OCD is a common and severe, but still under-recognized psychiatric disorder. Family and twin studies have provided evidence for the involvement of a genetic factor in OCD[1-3]. Although very little is known about the disorder's pathogenesis, serotonergic and dopaminergic pathways may be implicated[4]. A role for serotonin in the pathophysiology of OCD is supported by pharmacologic challenge studies and the unique efficacy of serotonin reuptake inhibitors. Although negative reports have been published, some studies have shown associations between OCD with the serotonin transporter (5-HTT), the 5-$HT_{1D}$, 5-$HT_{2A}$ and 5-$HT_{2C}$ genes[5]. In light of the putative role of the serotonin system in OCD, we tested the frequency of alleles and the distribution of genotypes of the polymorphism in the promoter region 5-HTT, the silent G-to-C substitution at nucleotide 861 of the coding region of the 5-$HT_{1D}$ receptor gene and the -1438 A/G polymorphism of the 5-$HT_{2A}$ receptor gene in a OCD sample of 156 patients and of 134 ethnically matched Caucasian subjects from the Netherlands as control population. As of yet, the association between the 5HTTLPR and the -1438 5HT2A polymorphism and OCD has been inconsistent, possibly reflecting sample–related differences, diagnostic diversity and heterogeneity of the disorder and the 5-$HT_{1B}$ G861C polymorphism has not been investigated through a case-control study. Since OCD is clinically heterogeneous and this heterogeneity is likely to be due to etiologic heterogeneity we considered OCD phenotypes based on qualitative traits such as gender, positive family history of OCD, and age of onset of OCD symptoms.

## 2.2. Material and methods

### 2.2.1. Study sample

The patient sample comprised 156 unrelated patients with OCD from consecutive referrals to the anxiety research unit of the department of psychiatry at the University Medical Centre Utrecht, who gave written informed consent for participation in this study that had been approved by the University of Utrecht Medical Ethical Review committee (Utrecht, The Netherlands). All patients were diagnosed with OCD according to DSM-IV criteria and the Mini-International Neuropsychiatric Interview (M.I.N.I.), a clinical and structured interview, was used to confirm the diagnosis[6]. Severity of obsessive-compulsive symptoms was rated with the Y-BOCS, depression with the Hamilton depression scale (HAM-D), and anxiety with

the Hamilton anxiety scale (HAM-A)[7-9]. Information on family history was obtained by direct interviews with the patients and the presence of vocal and/or motor tics was assessed during the clinical interview. The control sample was composed of 134 ethnically matched and unrelated Caucasian subjects from the Netherlands, selected among healthy volunteers.

| | Total sample (n=156) |
|---|---|
| Gender (Male/female) | 56/100 |
| Age on admission | 36.6 ± 11.5 |
| Positive family history | 43 |
| Mean age of onset | 17.7 ± 8.3 |
| ≤15 years age of onset | 60 |
| > 20 years age of onset | 45 |
| Duration of illness | 18.7 ± 11.6 |
| Symptom dimensions | |
| Contamination fear and washing | 23 |
| Aggressive, sexual and religious obsessions | 14 |
| Somatic obsessions and checking | 17 |
| Symmetry and exactness | 54 |
| High risk assessment and checking | 47 |
| Y-BOCS | 24.9 ± 5.7 |
| HAM-D | 9.5 ± 5.8 |
| HAM-A | 1.6 ± 6.7 |
| Comorbid depressive disorder | 24 |
| Comorbid anxiety disorder | 13 |

Table 1. Demographic and clinical characteristics of the patients sample

### 2.2.2. Genotyping

Blood samples were collected from each subject and frozen at – 80°. DNA was extracted from 10 ml samples of peripheral blood according to standard procedures. The total number of subjects genotyped for the genes in this study was 310. All subjects were genotyped at the University of Ghent (Belgium) based on a coded identification number. The 5-HTT, 5-$HT_{1D}$ and 5-$HT_{2A}$ genotyping was performed following a standardized procedure.

#### 2.2.2.1. 5-HTT

For the detection of the 44 bp insertion/deletion 5-HTTLPR polymorphism, the oligo-nucleotide primers 5'-6FAM-GGCGTTGCCGCTCTGAATGC–3' and 5'–AGGGACTGAGC TGGACAACCAC-3' were used to amplify a 484/528 bp fragment comprising the 5-HTT-linked polymorphic region. The polymerase chain reaction (PCR) reaction was performed according following conditions: 94°C for 1 min, 60 °C for 1min, 72°C for 1min40sec per

cycle, for a total of 35 cycles. The PCR products were analysed by capillary electrophoresis on an Applied Biosystems 3100 Genetic Analyser.

#### 2.2.2.2. 5-$HT_{1D}$

For detection of the 5-$HT_{1B}$ or (5-$HT_{1D\beta}$) G861C polymorphism, the oligonucleotide primers 5' -GAAACAGACGCCCAACAGGAC - 3' and 5' -CCAGAAACCGCGAAAGAAGAT - 3' were used to amplify a 548 bp region comprising the G861C polymorphism site. The PCR reaction was performed under the following conditions: 90°C for 1 min, 55 °C for 2 min, 72°C for 3 min per cycle, for a total of 32 cycles. Digestion of 10µl of PCR product was accomplished by incubation for 4 hours with 10 units of Hinc II restriction enzyme at 37°C. Digestion with Hinc II yields either two fragments (452 bp and 96 bp) for the G-allele or three fragments (310 bp, 142 bp and 96 bp) for the C-allele. The fragments resulting from the digestion were resolved on a 1.5% agarose gel and visualized by ethidium bromide staining.

#### 2.2.2.3. 5-$HT_{2A}$

For the detection of the 5-$HT_{2A}$ 1438G/A polymorphism within the promoter region of the 5-$HT_{2A}$ receptor gene, the oligonucleotide primers 5'-6FAM-AAGCTGCAAGGTAGCAACA GC–3' and 5'-NED-AACCAACTTATTTCCTACCAC-3' were used to amplify a 468 bp region comprising the $5HT_{2A}$ 1438G/A polymorphism site. The PCR reaction was performed under the following conditions: 95°C for 1 min, 47 °C for 1min, 72°C for 1min20sec per cycle, for a total of 40 cycles. Digestion of 10µl of PCR product was accomplished by overnight incubation with 10 units of Msp I restriction enzyme at 37°C. After incubation with Msp I, the 1438A allele remains intact while the 1438G allele is cut into a 223 bp piece (6FAM-labelled) and a 243 bp piece (NED-labeled). The fragments resulting from the digestion were analyzed by capillary electrophoresis on an Applied Biosystems 3100 Genetic Analyser.

### 2.2.3. Data analysis

The following statistical procedure was pursued. The genotypic pattern of distribution and the allele frequencies of the 5-HTT, 5-$HT_{1D}$ and 5-$HT_{2A}$ polymorphisms were analyzed in the whole sample comparing patients with controls. Second, a similar analysis was performed in a stratified sample according to gender comparing female patients with female controls. Third, patients with a positive family history for OCD were compared with controls and patients without a positive family history. Association tests were preformed by means of chi-square

tests for the comparison of allele and genotype frequencies between patients and comparisons subjects. Considering a partial Bonferroni's correction, the p value for statistical significance would be 0.022 with an alpha of 0.05, 5 tests, 2 degrees of freedom, and a correlation correction factor of 0.5. The association between the distribution of the genotypes and allele frequencies with the subjects and expected frequencies to assess Hardy-Weinberg equilibrium were ascertained by cross-tabulation and $\chi^2$ analysis. The data are presented as mean ± standard deviation (SD), and performed at 5% level of significance. All statistical analyses were conducted with the SPSS statistical package version 11.5.

## 2.3. Results

The patient sample was slightly skewed towards the female population (63%) with a mean ± SD age at admission of 36.6 ± 11.5 years for both sexes. (Table 1) The mean age at onset of obsessive-compulsive symptoms in our sample was 17.7 ± 8.3 years, with a length of illness of 18.7 ± 11.6 years at entry. Males had a significantly earlier onset of illness than females. (15.7 ± 8.0 years and 19.0 ± 8.3 years, respectively) ($\chi^2$ = 5.85, df=1, p=0.016) Twenty-four out of 60 patients with an early onset of illness had a positive family history for OCD. The mean Y-BOCS score for the whole sample was 24.9 ± 5.7, with a mean Y-BOCS obsession score of 12.5 ± 3.9 and a mean Y-BOCS compulsion score of 12.4 ± 3.8. Twenty-seven percent (43 patients) of the sample had a first-degree relative with OCD and nine patients (6 males/3 females) reported comorbid tics at some time in their life.

The genotypic pattern of distribution and the allele frequencies of the 5-HTT, 5-$HT_{1D}$ and 5-$HT_{2A}$ polymorphisms are shown in table 2. The incidences of the polymorphisms of the 5-HTT, 5-$HT_{1D}$ and 5-$HT_{2A}$ genes were similar in patients and controls, and no difference in frequencies of any of the alleles was observed between patients and controls. Both groups were in Hardy-Weinberg equilibrium at each locus investigated.

When the sample was stratified by gender, there was a statistically significant higher frequency of the 5-HTTLPR S-allele in female patients (L= 0.53 and S=0.47) compared to female controls (L= 0.67 and S=0.33) ($\chi^2$ = 6.0, df =1, p=0.014) (OR 1.81, 95% CI 1.12-2.93). Patients with a positive family history of OCD (n=43) had a statistically significant predominance of the 5-$HT_{2A}$ GG genotype (AA= 16%, AG= 33%, GG= 51%) versus patients without genetic load for OCD (n=111) (AA= 21%, AG= 49%, GG= 30%) ($\chi^2$ = 6.2, df =2, p=0.043) (OR 2.42, 95% CI 1.35-4.33), and a higher frequency of the 5-$HT_{2A}$ G-allele (A=

0.32 and G=0.68) versus patients without genetic load for OCD (A= 0.45 and G=0.55) ($\chi^2$ = 4.2, df =1, p=0.039) (OR 1.72, 95% CI 1.02-2.91), however the p-value failed to be statistical significant after Bonferroni's correction. The subpopulation with a positive family history had also a statistically significant predominance of the 5-$HT_{2A}$ GG genotype (AA= 16%, AG= 33%, GG= 51%) versus controls (AA= 15%, AG= 57%, GG= 28%) ($\chi^2$ = 8.427, df =2, p=0.015) (OR 2.67, 95% CI 1.48-4.81). A comparable predominance of the 5-$HT_{2A}$ GG genotype that just failed to be statistically significant (AA= 23%, AG= 37%, GG= 40%) was observed in the subpopulation of patients with an early onset of disease relative to controls (AA= 15%, AG= 57%, GG= 28%) ($\chi^2$ = 6.5, df =2, p=0.038) (OR 1.71, 95% CI 0.94-3.09).

| | n | allele frequencies | | p-value | genotypes | | | p-value |
|---|---|---|---|---|---|---|---|---|
| **5-$HT_{1D}$** | | **C** | **G** | | **CC** | **CG** | **GG** | |
| Controls | 117 | 0.30 | 0.70 | 0.647 | 7 (6.0%) | 55 (47.0%) | 55 (47.0%) | 0.866 |
| Patients | 141 | 0.28 | 0.72 | | 8 (5.7%) | 62 (44.0%) | 71 (50.4%) | |
| **5-$HT_{2A}$** | | **A** | **G** | | **AA** | **AG** | **GG** | |
| Controls | 116 | 0.43 | 0.57 | 0.776 | 17 (14.7%) | 66 (56.9%) | 33 (28.4%) | 0.143 |
| Patients | 154 | 0.42 | 0.58 | | 30 (19.5%) | 69 (44.8%) | 55 (35.7%) | |
| **5-HTT** | | **L** | **S** | | **LL** | **L/S** | **SS** | |
| Controls | 134 | 0.58 | 0.42 | 0.607 | 48 (35.8%) | 60 (44.8%) | 26 (19.4%) | 0.787 |
| Patients | 156 | 0.56 | 0.44 | | 50 (32.1%) | 75 (48.1%) | 31 (19.9%) | |

Table 2. Allele frequencies and genotype distribution of the 5-$HT_{1D,}$ 5-$HT_{2A}$ receptor and 5-HTT polymorphisms. Two-sided *P* values were calculated from 3x2 and 2x2 contingency tables for the genotype and allele distributions using the $\chi^2$-test.

## 2.4. Discussion

The findings of this study provide no evidence for a general association between 5-HTT, 5-$HT_{1D}$ and 5-$HT_{2A}$, polymorphisms and OCD. However, assuming the heterogeneous character of OCD, we found some hints towards an association of the 5-HTTLPR S-allele with female OCD patients, and the 5-$HT_{2A}$ G-allele and GG genotype with patients with a positive family and an early onset of disease.

The 5-HTTLPR gene has been studied extensively in OCD by means of case-control as well as family based studies. Except for Mc Dougle et al and Bengel et al who both reported an

association with the L allele, the majority of studies have been negative[10-13]. The association of the S-allele with female OCD patients in our study contrasts with previous reports, but is in line with the serotonin hypothesis of OCD. The S-allele has been related with relatively lower 5-HTT expression and lower serotonin transporter binding capacity than the L-allele[14, 15]. The S-allele also has been related with neuroticism and with an increased activation response in the right amygdale in response to fearful stimuli compared to the L-allele[16, 17]. OCD patients show increased activation response to fearful stimuli, and single photon emission computed tomography (SPECT) studies recently have provided evidence for reduced 5-HTT densities in patients with OCD relative to controls[18]. OCD is thus expected to be associated predominantly with the S-allele of the 5-HTTLPR. It is not clear why the association of 5-HTT S- allele that we have found is restricted to female patients. Either the apparent specific association is a spurious finding as result of a type two error due to small sample sizes, or it might be explained by the heterogeneous nature of OCD. Gender-related differences have been observed regularly in obsessive-compulsive symptomatology. In general, women have been noted to exhibit more aggressive and contamination obsessions and cleaning rituals, while men tend to report more frequently primary obsessive slowness, sexual, exactness and symmetry obsessions and odd rituals[19]. Comorbidity studies have shown that depressive and eating disorders are more prevalent in female OCD patients[19]. A recent cluster analysis showed that impulsive OCD behavior such as compulsive shopping, kleptomania and eating disorders was linked with female gender, emotional abuse and susceptibility to early traumatic experiences. Sexually dimorphic associations in OCD have been found also with COMT, monoamine oxidase A (MAO-A), and 5-$HT_{2A}$ genes[20]. COMT and MAO-A genes were associated with male OCD patients whereas 5-$HT_{2A}$ genes were associated with female OCD patients. This suggests that genetic mechanisms operant in OCD may be gender specific. The association of 5-HTTLP S- allele with female patients might represent a particular OCD phenotype rooted in a specific OCD genotype.

The -1438G/A polymorphism of the 5-$HT_{2A}$ receptor has been linked before with OCD. Enoch et al reported an association between the A-allele and female OCD patients, and Walitza et al between the A-allele and children and adults with OCD[21, 22]. Our results of an association with the 5-$HT_{2A}$ G-allele contradict both studies. Though differences in ages of onset may not account for these divergent results since both other studies pertain to a group of early onset patients, the discrepancy probably reflects dissimilarities in OCD phenotypes. The 5-$HT_{2A}$ receptor is of particular relevance to OCD since 5-$HT_{2A}$ receptor-binding characteristics such as may discriminate between affected and unaffected OCD subjects in

families with OCD[23]. In addition, an increase in 5-$HT_{2A}$ receptor binding has been found in the caudate nuclei of 15 untreated patients with OCD[24]. At present, it is difficult to interpret the significance of a link wit the 5-$HT_{2A}$ promoter region since it is not known whether the -1438 G and A alleles have functional implications[25]. Finally, it should be noted that the S-allele of 5HTTLPR and the G-allele of the -1438 5HT2A polymorphism may be in linkage disequilibrium with the particular OCD phenotype.

In conclusion, we have found that OCD in female patients was associated with the 5-HTTLPR S-allele and, that the subpopulation of patients with a positive family history of OCD and with an early onset of disease had a predominance of the 5-$HT_{2A}$ GG genotype. Our results should be interpreted with caution given the limited sample sizes and multiple testing. Nevertheless, they strengthen the argument that different OCD phenotypes may have different genetic susceptibilities, and add further evidence to the complex pattern of inheritance in OCD.

# References


1. Pauls DL, Alsobrook JP, Goodman W, Rasmussen S, Leckman JF: **A Family Study of Obsessive-Compulsive Disorder**. *American Journal of Psychiatry* 1995, **152**(1):76-84.

2. Nestadt G, Lan T, Samuels J, Riddle M, Bienvenu OJ, Liang KY, Hoehn-Saric R, Cullen B, Grados M, Beaty TH *et al*: **Complex segregation analysis provides compelling evidence for a major gene underlying obsessive-compulsive disorder and for heterogeneity by sex**. *American Journal of Human Genetics* 2000, **67**(6):1611-1616.

3. Hudziak JJ, van Beijsterveldt CEM, Althoff RR, Stanger C, Rettew DC, Nelson EC, Todd RD, Bartels M, Boomsma DI: **Genetic and environmental contributions to the child behavior checklist Obsessive-Compulsive Scale - A cross-cultural twin study**. *Archives of General Psychiatry* 2004, **61**(6):608-616.

4. Denys D, Zohar J, Westenberg HGM: **The role of dopamine in obsessive-compulsive disorder: Preclinical and clinical evidence**. *Journal of Clinical Psychiatry* 2004, **65**:11-17.

5. Grados MA, Walkup J, Walford S: **Genetics of obsessive-compulsive disorders: new findings and challenges**. *Brain & Development* 2003, **25**:S55-S61.

6. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC: **The Mini-International Neuropsychiatric Interview (MINI): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10**. *Journal of Clinical Psychiatry* 1998, **59**:22-33.

7. Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, Heninger GR, Charney DS: **The Yale-Brown Obsessive Compulsive Scale.1. Development, Use, and Reliability**. *Archives of General Psychiatry* 1989, **46**(11):1006-1011.

8. Hamilton M: **The assessment of anxiety states by rating**. *Br J Med Psychol* 1959, **32**:50-55.

9. Hamilton M: **A rating scale for depression**. *J Neurol Neurosurg Psychiatry* 1960, **23**:56-62.

10. Bengel D, Greenberg BD, Cora-Locatelli G, Altemus M, Heils A, Li Q, Murphy DL: **Association of the serotonin transporter promoter regulatory region polymorphism and obsessive-compulsive disorder**. *Molecular Psychiatry* 1999, **4**(5):463-466.

11. McDougle CJ, Epperson CN, Price LH, Gelernter J: **Evidence for linkage disequilibrium between serotonin transporter protein gene (SLC6A4) and obsessive compulsive disorder**. *Molecular Psychiatry* 1998, **3**(3):270-273.

12. Cavallini MC, Di Bella D, Siliprandi F, Malchiodi F, Bellodi L: **Exploratory factor analysis of obsessive-compulsive patients and association with 5-HTTLPR polymorphism**. *American Journal of Medical Genetics* 2002, **114**(3):347-353.

13. Di Bella D, Erzegovesi S, Cavallini MC, d'Annucci A, Bellodi L: **Obsessive-compulsive disorder, treatment response and the 5HTT gene.** *American Journal of Medical Genetics* 2000, **96**(4):536-536.

14. Williams RB, Marchuk DA, Gadde KM, Barefoot JC, Grichnik K, Helms MJ, Kuhn CM, Lewis JG, Schanberg SM, Stafford-Smith M *et al*: **Serotonin-related gene polymorphisms and central nervous system serotonin function**. *Neuropsychopharmacology* 2003, **28**(3):533-541.

15. Smith GS, Lotrich FE, Malhotra AK, Lee AT, Ma YL, Kramer E, Gregersen PK, Eidelberg D, Pollock BG: **Effects of serotonin transporter promoter polymorphisms on serotonin function**. *Neuropsychopharmacology* 2004, **29**(12):2226-2234.

16. Munafo MR, Roberts K, Johnstone EC, Walton RT, Yudkin PL: **Association of serotonin transporter gene polymorphism with nicotine dependence: no evidence for an interaction with trait neuroticism**. *Personality and Individual Differences* 2005, **38**(4):843-850.

17. Hariri AR, Mattay VS, Tessitore A, Kolachana B, Fera F, Goldman D, Egan MF, Weinberger DR: **Serotonin transporter genetic variation and the response of the human amygdala**. *Science* 2002, **297**(5580):400-403.

18. Stengler-Wenzke K, Muller U, Angermeyer MC, Sabri O, Hesse S: **Reduced serotonin transporter-availability in obsessive-compulsive disorder (OCD)**. *European Archives of Psychiatry and Clinical Neuroscience* 2004, **254**(4):252-255.

19. Zohar J, Gross-Isseroff R, Hermesh H, Weizman A: **Is there sexual dimorphism in obsessive-compulsive disorder?** *Neuroscience and Biobehavioral Reviews* 1999, **23**(6):845-849.

20. Lochner C, Hemmings SMJ, Kinnear CJ, Niehaus DJH, Nel DG, Corfield VA, Moolman-Smook JC, Seedat S, Stein DJ: **Cluster analysis of obsessive-compulsive spectrum disorders in patients with obsessive-compulsive disorder: clinical and genetic correlates**. *Comprehensive Psychiatry* 2005, **46**(1):14-19.

21. Enoch MA, Greenberg BD, Murphy DL, Goldman D: **Sexually dimorphic relationship of a 5-HT2A promoter polymorphism with obsessive-compulsive disorder**. *Biological Psychiatry* 2001, **49**(4):385-388.

22. Walitza S, Wewetzer C, Warnke A, Gerlach M, Geller F, Gerber G, Gorg T, Herpertz-Dahlmann B, Schulz E, Remschmidt H *et al*: **5-HT2A promoter polymorphism -1438G/A in children and adolescents with obsessive-compulsive disorders**. *Molecular Psychiatry* 2002, **7**(10):1054-1057.

23. Delorme R, Betancur C, Callebert J, Chabane N, Laplanche JL, Mouren-Simeoni MC, Launay JM, Leboyer M: **Platelet serotonergic markers as endophenotypes for obsessive-compulsive disorder**. *Neuropsychopharmacology* 2005, **30**(8):1539-1547.

24. Adams BL, Warneke LB, Mcewan AJB, Fraser BA: **Single-Photon Emission Computerized-Tomography in Obsessive-Compulsive Disorder - a Preliminary-Study**. *Journal of Psychiatry & Neuroscience* 1993, **18**(3):109-112.

25. Veenstra-VanderWeele J, Anderson GM, Cook EH: **Pharmacogenetics and the serotonin system: initial studies and future directions**. *European Journal of Pharmacology* 2000, **410**(2-3):165-181.

# 3. Association between the Dopamine D2 receptor TAQI-A2 allele and low activity COMT allele with obsessive-compulsive disorder in males

EUROPEAN NEUROPSYCHOPHARMACOLOGY (2006) **16**,446-450

Damiaan Denys [1], M.D., Filip Van Nieuwerburgh[2], Dieter Deforce[2] Herman Westenberg [1]

1) Rudolf Magnus Institute of Neuroscience, Department of Psychiatry University Medical Center Utrecht, Utrecht, the Netherlands.

2) Laboratory of Pharmaceutical Biotechnology, Ghent University, Harelbekestraat 72, 9000 Ghent, Belgium.

## ABSTRACT

**Background:** Mounting evidence suggests the involvement of the dopamine system in the pathophysiology of obsessive–compulsive disorder.
**Method:** The relationship of the dopamine D2 receptor (DRD2) TAQI-A, and catechol-0-methyl-transferase (COMT) *Nla* III High/Low activity polymorphism to obsessive–compulsive disorder (OCD) was examined in a sample of 150 patients and 150 controls.
**Results:** OCD patients did not show significant differences in genotype distribution and allele frequency for polymorphisms investigated relative to controls. However, when the sample was stratified by gender, there was a mildly significantly predominance of the DRD2 A2A2 genotype ($p = 0.049$), and a higher frequency of the DRD2 A2 allele ($p = 0.020$) and low-activity COMT allele ($p = 0.035$) in male OCD patients compared to male controls. In addition, we observed an association of the DRD2 A2A2 genotype in patients with an early onset of disease (≤ 15 years) ($p= 0.033$).
**Conclusions:** Our findings replicate previous reports and provide support for a potential role of the COMT and DRD2 locus in subgroup of male, early onset patients with OCD.

## 3.1. Introduction

OCD is a common and severe, but still under-recognized psychiatric disorder. Family and twin studies have provided evidence for the involvement of a genetic factor in OCD[1, 2]. Although very little is known about the disorder's pathogenesis, both serotonergic and dopaminergic pathways may be implicated. A role for dopamine in the pathophysiology of OCD is supported by the observation that pharmacological agents enhancing dopamine release such as methylphenidate, cocaine, and bromocriptine may induce obsessive-compulsive symptoms[3-5]. In addition, evidence has accumulated that augmentation strategies with antipsychotics are beneficial for treatment-refractory OCD patients[6-8]. Recently, SPECT studies provided evidence for higher dopamine transporter densities and lower dopamine D2 binding ratios in patients with OCD relative to controls[9-11]. The combined results of these receptor-binding studies provide circumstantial in vivo evidence for an increased dopaminergic activity in OCD patients.

Catechol-O-methyl transferase (COMT) is an enzyme that has a crucial role in the elimination of dopamine. Since COMT is involved in the inactivation of dopamine, and higher dopamine levels may be implicated in OCD, the COMT gene is an attractive candidate for OCD. The G→A transition in codon 158 of the COMT gene results in a valine to methionine substitution and is associated with a three- to fourfold decrease in enzyme activity (valine = high-activity, methionine = low-activity)[12]. It has already been reported by Karayiourgou *et al.*[13] that the low-activity COMT (COMT L) allele occurs significantly more frequently in male OCD patients, but opposing results have also been obtained[14-16]. In addition, one might theorize that lower densities of the D2 receptor in OCD patients are caused by genetic factors. The A1 allele of TaqI A polymorphism in the D2 receptor (DRD2) gene locus has been suggested to be associated with reduced DRD2 receptor densities[17, 18]. Data on the DRD2 TaqI A polymorphism in OCD are limited, but Nicolini *et al.* found a higher frequency of the DRD2 TaqI A2 allele in a small subgroup of OCD patients (n=12) with tics, when compared to controls[19].

In light of the putative role of the dopamine system, and in particular the dopamine D2 receptor in OCD, we tested the frequency of alleles and the distribution of genotypes of the DRD2 receptor TaqI A and of COMT genes in an OCD sample of 150 patients. As a control population, we tested 150 ethnically matched Caucasian subjects. As there is evidence for gender specificity of the D2 receptor and COMT gene and because previous findings have suggested gender differences in the clinical manifestation of OCD, males and females were

analyzed separately[15, 20-22]. We firstly hypothesized that OCD patients would have higher frequencies of the COMT L allele, resulting in higher synaptic dopamine levels. Secondly, we hypothesized that OCD patients would have higher frequencies of the DRD2 receptor TaqI A1 allele, resulting in lower synaptic DRD2 density.

## 3.2. Material and methods

### 3.2.1. Study sample

The patient sample comprised 159 unrelated patients with OCD from consecutive referrals to the anxiety research unit of the department of psychiatry at the University Medical Centre Utrecht, who gave written informed consent for participation in this study that had been approved by the University of Utrecht Medical Ethical Review committee (Utrecht, The Netherlands). All patients were diagnosed with OCD according to DSM-IV criteria and the M.I.N.I., a clinical and structured interview, was used to confirm the diagnosis[23]. Severity of obsessive-compulsive symptoms was rated with the Y-BOCS, depression with the HAM-D, and anxiety with the HAM-A[24-26]. Information on family history was obtained by direct interviews with the patients and the presence of vocal and/or motor tics was assessed during the clinical interview. The control sample was composed of 151 ethnically matched and unrelated Caucasian subjects, selected among healthy volunteers.

### 3.2.2. Genotyping and data analysis

Blood samples were collected from each subject and frozen at – 80°. DNA was extracted from 10 ml samples of peripheral blood according to standard procedures. The total number of subjects genotyped for the genes in this study was 310. All subjects were genotyped at the University of Ghent (Belgium) based on a coded identification number. The COMT and DRD2 genotyping was performed following a standard protocol.

### 3.2.3. COMT

For detection of the *Nla*III polymorphism in codon 158, the following oligonucleotide primers were used (5' - TCACCATCGAGATCAACCCC - 3' and 5' -ACAACGGGTCAGGCATGCA - 3') to amplify a 96 bp region comprising the Val158Met polymorphism site. The PCR reaction was performed under the following conditions: 94°C for 30sec, 64 °C for 1min, 72°C for 1min per cycle, for a total of 35 cycles. Digestion of 9µl of PCR product was accomplished by incubation for 3 to 4

hours with 5 units of NlaIII restriction enzyme at 37°C. Digestion with NlaIII yields either two fragments (13 bp and 83 bp) for the Val-allele (COMT H) or three fragments (13 bp, 18 bp and 65 bp) for the Met-allele (COMT L). The fragments were resolved on a 2.5% agarose gel and visualized by ethidium bromide staining.(Karayiorgou *et al.* 1997)

### 3.2.4. DRD2

For the detection of the polymorphism in the TaqA site in the DRD2 gene the oligonucleotide primers (5'-CCGTCGACGGCTGGCCAAGTTGTCTA–3' and 5'–CCGTCGACCCTTCCTGAGTGTCATCA -3') were used to amplify a 310 bp region comprising the TaqA site[27]. The PCR reaction was performed under the following conditions: 94°C for 1 min, 50 °C for 1min, 72°C for 1,5min per cycle, for a total of 35 cycles. Digestion of 10µl of PCR product was accomplished by overnight incubation with 5 units of TaqI restriction enzyme at 65°C. After incubation with TaqI, the A1 allele remains intact while the A2 allele is cut into a 130 bp piece and a 180 bp piece. The fragments resulting from the digestion were resolved on a 1,5% agarose gel and visualized by staining with ethidium bromide.

The association between the distribution of the genotypes and allele frequencies with the subjects, and expected frequencies to assess Hardy-Weinberg equilibrium, were ascertained by cross-tabulation and $\chi^2$ analysis.

## 3.3. Results

The patient sample was slightly skewed towards the female population (63%) with a mean ± SD age at admission of 36.0 ± 11.0 years for both sexes. The mean age at onset of obsessive-compulsive symptoms in our sample was 17.7 ± 8.3 years, with a length of illness of 18.0 ± 11.0 years at entry. Males had a significantly earlier onset of illness than females. (15.7 ± 8.0 years and 19.0 ± 8.3 years, respectively) ($\chi^2$ = 5.85, df=1, p=0.016) The mean Y-BOCS score for the whole sample was 24.9 ± 5.7, with a mean HAM-D score of 9.5 ± 5.8 and a mean HAM-A score of 11.6 ± 6.7. Twenty-seven percent (43 patients) of the sample had a first-degree relative with OCD and nine patients (6 males/3 females) reported comorbid tics at some time in their life.

The genotypic pattern of distribution and the allele frequencies of the DRD2 and COMT polymorphisms are shown in table 1 and 2. The representations of the polymorphism of the DRD2 receptor and the COMT gene were similar in patients and controls. No difference in frequencies of any of the alleles was observed between patients and controls. Both groups were in Hardy-Weinberg equilibrium at each locus investigated.

When the sample was stratified by gender, there was a statistically significant predominance of the DRD2 A2A2 genotype in the male patient group ($\chi^2$ = 6.0, df=2, p=0.049) and a higher frequency of the DRD2 A2 allele in male patients compared to male controls ($\chi^2$ = 5.4, df=2, p=0.020). In addition, a significant association was observed between the frequency of the COMT L allele and male patients ($\chi^2$ = 4.4, df=2, p=0.035). Although the frequency of the COMT LL genotype was higher in male patients (37.5%) compared to male controls (21%), the difference failed to reach statistical significance ($\chi^2$ = 4.6, df=2, p=0.10).

| | | n | allele frequencies | | p-value | genotypes | | | p-value |
|---|---|---|---|---|---|---|---|---|---|
| | | | A1 | A2 | | A1A1 | A1A2 | A2A2 | |
| Total sample | Controls | 135 | 0.20 | 0.80 | 0.218 | 5 (3.7%) | 45 (33.3%) | 85 (63.0%) | 0.452 |
| | Patients | 141 | 0.16 | 0.84 | | 3 (3.7%) | 40 (28.4%) | 98 (69.5%) | |
| Males | Controls | 67 | 0.22 | 0.78 | 0.020 | 2 (3.0%) | 26 (38.8%) | 39 (58.2%) | 0.049 |
| | Patients | 51 | 0.11 | 0.89 | | - | 11 (21.6%) | 40 (78.4%) | |
| Females | Controls | 66 | 0.17 | 0.83 | 0.672 | 3 (4.5%) | 17 (25.8%) | 46 (69.7%) | 0.692 |
| | Patients | 88 | 0.19 | 0.81 | | 3 (3.4%) | 28 (31.8%) | 57 (64.8%) | |

Table 1. Allele frequencies and genotype distribution of the DRD2 TAQIA polymorphism

| | | n | allele frequencies | | p-value | genotypes | | | p-value |
|---|---|---|---|---|---|---|---|---|---|
| | | | L | H | | LL | LH | HH | |
| Total sample | Controls | 152 | 0.51 | 0.49 | 0.643 | 39 (25.7%) | 77 (50.7%) | 36 (23.7%) | 0.779 |
| | Patients | 158 | 0.53 | 0.47 | | 46 (29.1%) | 75 (47.5%) | 37 (23.4%) | |
| Males | Controls | 79 | 0.47 | 0.53 | 0.035 | 17 (21.5%) | 40 (50.6%) | 22 (27.8%) | 0.100 |
| | Patients | 56 | 0.60 | 0.40 | | 21 (37.5%) | 25 (44.6%) | 10 (17.9%) | |
| Females | Controls | 71 | 0.56 | 0.44 | 0.227 | 21 (29.6%) | 37 (52.1%) | 13 (18.3%) | 0.440 |
| | Patients | 99 | 0.49 | 0.51 | | 24 (24.2%) | 49 (49.5%) | 26 (26.3%) | |

Table 2. Allele frequencies and genotype distribution of the COMT polymorphism

## 3.4. Discussion

The findings of this study provide evidence for an association between the DRD2 TaqI A2 allele and the low-activity COMT allele on the one hand, OCD on the other, in male OCD patients.

Two previous studies have examined the association between the DRD2 TaqI A system and OCD. Nicolini *et al.* found no association in 67 patients with OCD, but observed a higher frequency of the A2 allele in a subgroup of OCD patients with tics (n=12)[19]. Billet *et al.*[28] did not find an association in a sample of 100 OCD patients either. Since the A1 allele of the DRD2 TaqI A system has been found to be associated with a variety of addictive, impulsive and compulsive disorders, the association of the A2 allele with our OCD-sample was

unexpected[29]. In addition, we presumed a higher frequency of the A1 allele, as it has been suggested that the A1 allele is associated with a mutation that decreases the D2 receptor expression. This latter suggestion has been recently confirmed by Pohjalainen *et al.*[30], but was contradicted by Laruelle *et al.*[31] At the moment, it is still unclear whether or not the Al allele is associated with lower D2 expression[17]. Since other reports of associations between the A2 allele and similar neuropsychiatric disorders are scarce, the association of the A2 allele in our sample is difficult to interpret. The A2 allele has previously been related to hyperactive and impulsive symptoms in attention deficit hyperactivity disorder (ADHD) in a sample of 166 children and to compulsive smoking habits in 793 subjects[32, 33]. Interestingly, the ADHD sample comprised 81% males, and the association with compulsive smoking was only significant for males. This suggests that the A2A2 genotype of the DRD2 TaqI A system, regardless of the diagnosis, is associated with a broad spectrum of impulsive/compulsive symptoms in a gender specific manner. On the other hand, without further evidence of its functional significance, and in the absence of other studies reporting similar associations, our finding of the association of the A2A2 genotype and OCD in male patients warrants replication in other samples, as well as family-based designs.

The COMT locus has been reported to be associated with OCD in several previous studies. Karayiourgou *et al.*[13, 15] found evidence for an association between the low-activity COMT allele and OCD in male OCD patients, in a case-control study and a family-based study, whereas Alsobrook *et al.*[34] found evidence pointing to an association between the low-activity COMT allele and OCD in female OCD patients. Niehaus *et al.*[35] reported a preponderance of COMT high/low heterozygotes in an Afrikaner population of 54 OCD patients, but did not observe gender differences. Schindler *et al.*[36] did not find an association between any particular allele and OCD, but found a tendency for an association with homozygosity at the COMT locus. Ohara *et al.*[16] did not find any association in a small sample of 24 Japanese patients and neither did Erdal *et al.*[14] in a sample of 59 Turkish patients.' A recent meta-analysis of the COMT gene in 144 OCD patients and 337 controls showed insufficient evidence to support an association[37].

On the other hand, since a higher prevalence of the low-activity COMT allele in OCD patients has been established in different independent samples, the finding of an association between COMT and OCD remains interesting. Especially because the results are compatible with the assumption that increased dopamine levels are associated with obsessive-compulsive symptoms. It is possible that subjects with a low-activity COMT genotype have a longer lasting and more effective dopamine release, which makes them more vulnerable to the

development of obsessive-compulsive symptoms. On the other hand, this is hard to reconcile with the observation that the association is gender specific. In this regard, it is remarkable that our results are strikingly similar to both reports by Karayiourgou *et al.*[13, 15], providing further evidence to the previously reported gender-selective association between COMT polymorphism and male patients. It is important to emphasize that in our control sample neither the COMT genotype distribution ($\chi^2$ = 4.6, df=2, p=0.1), nor the allele frequency ($\chi^2$ = 4.6, df=2, p=0.1) differed significantly between males and females. The significance of a gender specific association may not be easily explained. Karayiourgou *et al.*[13, 15] proposed that females have evolved mechanisms to compensate for their lower levels of COMT activity and are therefore less vulnerable to developing OCD in association with a low-activity COMT genotype. On the other hand, as has been noted by Schindler *et al.*[36], the specific association in males may be a sampling phenomenon, since males typically demonstrate an earlier onset of OCD than females. In our sample, males (15.7 ± 8.0 years) had a significantly earlier age of onset than females (19.0 ± 8.3 years) and age of onset was significantly correlated with gender (r=0.18, p=0.033). Since only 10 out of 54 male patients had an onset of disease later than 21 years, the bias of age of onset cannot be excluded.

To eliminate the possible confounding factor of age of onset, we dichotomized the patient population into an early-onset-group (≤ 15 years (n=60)), and a late-onset-group (≥ 21 years (n=45)). We found that the low-activity COMT genotype was significantly associated with the early-onset-group (36.7%), relative to late-onset-group (22.7%) ($\chi^2$ = 6.83, df=2, p=0.033), though allele frequencies did not significantly differ ($\chi^2$ = 0.13, df=1, p =0.71; data not shown). Fifty-five percent of the early-onset-group was female, which suggests that age of onset might be an independent factor in the association with the low-activity COMT genotype. Therefore, it is conceivable that both male gender and early age of onset represent different subgroups in OCD, which are independently related to the COMT gene.

To summarize, this study suggests that DRD2 and COMT genes may be etiologically relevant in OCD, in a gender specific manner, and that early-onset-patients represent a genetically different subgroup. Further analysis of these phenotypic subtypes in larger samples is warranted to confirm our data.

# References


1. Nestadt G, Lan T, Samuels J, Riddle M, Bienvenu OJ, Liang KY, Hoehn-Saric R, Cullen B, Grados M, Beaty TH *et al*: **Complex segregation analysis provides compelling evidence for a major gene underlying obsessive-compulsive disorder and for heterogeneity by sex**. *American Journal of Human Genetics* 2000, **67**(6):1611-1616.

2. Pauls DL, Alsobrook JP, Goodman W, Rasmussen S, Leckman JF: **A Family Study of Obsessive-Compulsive Disorder**. *American Journal of Psychiatry* 1995, **152**(1):76-84.

3. Crum RM, Anthony JC: **Cocaine Use and Other Suspected Risk-Factors for Obsessive-Compulsive Disorder - a Prospective-Study with Data from the Epidemiologic Catchment-Area Surveys**. *Drug and Alcohol Dependence* 1993, **31**(3):281-295.

4. Jenike MA, Baer L, Summergrad P, Minichiello WE, Holland A, Seymour R: **Sertraline in Obsessive-Compulsive Disorder - a Double-Blind Comparison with Placebo**. *American Journal of Psychiatry* 1990, **147**(7):923-928.

5. Satel SL, Mcdougle CJ: **Obsessions and Compulsions Associated with Cocaine Abuse**. *American Journal of Psychiatry* 1991, **148**(7):947-947.

6. Denys D, van Megen H, Westenberg H: **Quetiapine addition to serotonin reuptake inhibitor treatment in patients with treatment-refractory obsessive-compulsive disorder: An open-label study**. *Journal of Clinical Psychiatry* 2002, **63**(8):700-703.

7. McDougle CJ, Goodman WK, Leckman JF, Lee NC, Heninger GR, Price LH: **Haloperidol Addition in Fluvoxamine-Refractory Obsessive-Compulsive Disorder - a Double-Blind, Placebo-Controlled Study in Patients with and without Tics**. *Archives of General Psychiatry* 1994, **51**(4):302-308.

8. McDougle CJ, Epperson CN, Pelton GH, Wasylink S, Price LH: **A double-blind, placebo-controlled study of risperidone addition in serotonin reuptake inhibitor-refractory obsessive-compulsive disorder**. *Archives of General Psychiatry* 2000, **57**(8):794-801.

9. Denys D, van der Wee N, Janssen J, De Geus F, Westenberg HGM: **Low level of dopaminergic D-2 receptor binding in obsessive-compulsive disorder**. *Biological Psychiatry* 2004, **55**(10):1041-1045.

10. Kim CH, Koo MS, Cheon KA, Ryu YH, Lee JD, Lee HS: **Dopamine transporter density of basal ganglia assessed with [I-123]IPT SPET in obsessive-compulsive disorder**. *European Journal of Nuclear Medicine and Molecular Imaging* 2003, **30**(12):1637-1643.

11. van der Wee N, Stevens H, Hardeman H, Denys D, Megen HJv, Kahn RS, Westenberg HG: **Enhanced densities of dopamine but not of serotonin**

**transporters in psychotropic-naïve patients with obsessive-compulsive disorder**. *J Nucl Med* 2001, **42 (5)**:238.

12. Lotta T, Vidgren J, Tilgmann C, Ulmanen I, Melen K, Julkunen I, Taskinen J: **Kinetics of Human Soluble and Membrane-Bound Catechol O-Methyltransferase - a Revised Mechanism and Description of the Thermolabile Variant of the Enzyme**. *Biochemistry* 1995, **34**(13):4202-4210.

13. Karayiorgou M, Altemus M, Galke BL, Goldman D, Murphy DL, Ott J, Gogos JA: **Genotype determining low catechol-O-methyltransferase activity as a risk factor for obsessive-compulsive disorder**. *Proceedings of the National Academy of Sciences of the United States of America* 1997, **94**(9):4572-4575.

14. Erdal ME, Tot S, Yazici K, Yazici A, Herken H, Erdem P, Derici E, Camdeviren H: **Lack of association of catechol-O-methyltransferase gene polymorphism in obsessive-compulsive disorder**. *Depression and Anxiety* 2003, **18**(1):41-45.

15. Karayiorgou M, Sobin C, Blundell ML, Galke BL, Malinova L, Goldberg P, Ott J, Gogos JA: **Family-based association studies support a sexually dimorphic effect of COMT and MAOA on genetic susceptibility to obsessive-compulsive disorder**. *Biological Psychiatry* 1999, **45**(9):1178-1189.

16. Ohara K, Nagai M, Suzuki Y, Ochiai M, Ohara K: **No association between anxiety disorders and catechol-O-methyltransferase polymorphism**. *Psychiatry Research* 1998, **80**(2):145-148.

17. Hitzemann RJ: **The regulation of D-2 dopamine receptor expression**. *Molecular Psychiatry* 1998, **3**(3):198-203.

18. Jonsson EG, Nothen MM, Grunhage F, Farde L, Nakashima Y, Propping P, Sedvall GC: **Polymorphisms in the dopamine D2 receptor gene and their relationships to striatal dopamine receptor density of healthy volunteers**. *Molecular Psychiatry* 1999, **4**(3):290-296.

19. Nicolini H, Cruz C, Camarena B, Orozco B, Kennedy JL, King N, Weissbecker K, delaFuente JR, Sidenberg D: **DRD2, DRD3 and 5HT2A receptor genes polymorphisms in Obsessive-Compulsive Disorder**. *Molecular Psychiatry* 1996, **1**(6):461-465.

20. Castle DJ, Deale A, Marks IM: **Gender Differences in Obsessive-Compulsive Disorder**. *Australian and New Zealand Journal of Psychiatry* 1995, **29**(1):114-117.

21. Kaasinen V, Nagren K, Hietala J, Farde L, Rinne JO: **Sex differences in extrastriatal dopamine D-2-like receptors in the human brain**. *American Journal of Psychiatry* 2001, **158**(2):308-311.

22. Lensi P, Cassano GB, Correddu G, Ravagli S, Kunovac JL, Akiskal HS: **Obsessive-compulsive disorder - Familial-developmental history, symptomatology, comorbidity and course with special reference to gender-related differences**. *British Journal of Psychiatry* 1996, **169**(1):101-107.

23. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC: **The Mini-International Neuropsychiatric Interview (MINI): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10**. *Journal of Clinical Psychiatry* 1998, **59**:22-33.

24. Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, Heninger GR, Charney DS: **The Yale-Brown Obsessive Compulsive Scale.1. Development, Use, and Reliability**. *Archives of General Psychiatry* 1989, **46**(11):1006-1011.

25. Hamilton M: **The assessment of anxiety states by rating**. *Br J Med Psychol* 1959, **32**:50-55.

26. Hamilton M: **A rating scale for depression**. *J Neurol Neurosurg Psychiatry* 1960, **23**:56-62.

27. Grandy DK, Zhang Y, Civelli O: **Pcr Detection of the Taqa Rflp at the Drd2 Locus**. *Human Molecular Genetics* 1993, **2**(12):2197-2197.

28. Billett EA, Richter MA, Sam F, Swinson RP, Dai XY, King N, Badri F, Sasaki T, Buchanan JA, Kennedy JL: **Investigation of dopamine system genes in obsessive-compulsive disorder**. *Psychiatric Genetics* 1998, **8**(3):163-169.

29. Comings DE, Blum K: **Reward deficiency syndrome: genetic aspects of behavioral disorders**. *Cognition, Emotion and Autonomic Responses: The Integrative Role of the Prefrontal Cortex and Limbic Structures* 2000, **126**:325-341.

30. Pohjalainen T, Rinne JO, Nagren K, Lehikoinen P, Anttila K, Syvalahti EKG, Hietala J: **The A1 allele of the human D-2 dopamine receptor gene predicts low D-2 receptor availability in healthy volunteers**. *Molecular Psychiatry* 1998, **3**(3):256-260.

31. Laruelle M, Gelernter J, Innis RB: **D-2 receptors binding potential is not affected by Taq1 polymorphism at the D-2 receptor gene**. *Molecular Psychiatry* 1998, **3**(3):261-265.

32. Rowe DC, Van den Oord EJCG, Stever C, Giedinghagen LN, Gard JMC, Cleveland HH, Gilson M, Terris ST, Mohr JH, Sherman S *et al*: **The DRD2 Taql polymorphism and symptoms of attention deficit hyperactivity disorder**. *Molecular Psychiatry* 1999, **4**(6):580-586.

33. Hamajima N, Ito H, Matsuo K, Saito T, Tajima K, Ando M, Yoshida K, Takahashi T: **Association between smoking habits and dopamine receptor $D_2$ *Taq*I A A2 allele in Japanese males: a confirmation study**. *J Epidemiol* 2002, **12 (4)**:297-304.

34. Alsobrook JP, Zohar AH, Leboyer M, Chabane N, Ebstein RP, Pauls DL: **Association between the COMT locus and obsessive-compulsive disorder in females but not males**. *American Journal of Medical Genetics* 2002, **114**(1):116-120.

35. Niehaus DJH, Kinnear CJ, Corfield VA, du Toit PL, van Kradenburg J, Moolman-Smook JC, Weyers JB, Potgieter A, Seedat S, Emsley RA *et al*: **Association between**

**a catechol-o-methyltransferase polymorphism and obsessive-compulsive disorder in the Afrikaner population**. *Journal of Affective Disorders* 2001, **65**(1):61-65.

36. Schindler KM, Richter MA, Kennedy JL, Pato MT, Pato CN: **Association between homozygosity at the COMT gene locus and obsessive compulsive disorder**. *American Journal of Medical Genetics* 2000, **96**(6):721-724.

37. Azzam A, Mathews C, Reus V: **A meta-analysis of the association between the catecholamine-O-methyl-transferase gene and obsessive-compulsive disorder**. *American Journal of Medical Genetics* 2002, **114**(7):834-835.

# 4. Prediction of response to Paroxetine and Venlafaxine by serotonin related genes in obsessive compulsive disorder

JOURNAL OF CLINICAL PSYCHIATRY, in press

Damiaan Denys[1], Filip Van Nieuwerburgh[2], Dieter Deforce[2] and Herman GM Westenberg[1]

1) Rudolf Magnus Institute of Neuroscience, Department of Psychiatry University Medical Center Utrecht, Utrecht, the Netherlands.

2) Laboratory of Pharmaceutical Biotechnology, Ghent University, Harelbekestraat 72, 9000 Ghent, Belgium.

## ABSTRACT

**Objective:** Serotonin reuptake inhibitors (SRIs) are the most effective pharmacological treatment currently available for patients with obsessive-compulsive disorder (OCD). Still, up to 40 to 60% of OCD patients do not respond to SRI treatment. The purpose of the present study was to determine whether polymorphisms of the serotonin transporter (5-HTT), 5-$HT_{1B}$ and 5-$HT_{2A}$ receptor genes affect the efficacy of SRI treatment in OCD.
**Method:** 91 outpatients with primary OCD according to DSM-IV criteria consented to the study and were randomly assigned in a 12-week, double-blind trial to receive dosages titrated upward to 300 mg/day of venlafaxine, or 60 mg/day of paroxetine. Primary efficacy was assessed by the change from baseline on the Yale-Brown obsessive-compulsive scale (Y-BOCS), and response was defined as a ≥ 25 % reduction on the Y-BOCS. Responders and non-responders were stratified according to 5-HTT, 5-$HT_{1B}$ and 5-$HT_{2A}$ genotypes, and differentiated in paroxetine or venlafaxine treated groups.
**Results**: In the whole group, 64% of responders carried the S/L genotype of the 5-HTTLPR polymorphism ($\chi^2$ = 7.17, df=2, p=0.028). In the paroxetine treated patients, the majority of responders carried the G/G genotype of the 5-$HT_{2A}$ polymorphism ($\chi^2$ = 8.66, df=2, p=0.013), whereas in the venlafaxine treated patients, the majority of responders carried the S/L genotype of the 5-HTTLPR polymorphism ($\chi^2$ = 9.71, df=2, p=0.008).
**Conclusions:** The results of this study suggest that response in paroxetine treated OCD patients is associated with the G/G genotype of the 5-$HT_{2A}$ polymorphism and in venlafaxine treated OCD patients with the S/L genotype of the 5-HTTLPR polymorphism.

## 4.1. Introduction

Obsessive-compulsive disorder (OCD) is a common and severe, but still under-recognized psychiatric disorder. Although serotonin reuptake inhibitors (SRIs) are the most effective pharmacological treatment for OCD, up to 40 to 60% of OCD patients do not respond to treatment[1]. Even after a switch to a second SRI-treatment, 30 to 40% of OCD patients fail to respond[2]. Clearly, an improved understanding of determinants of response to SRIs would be immensely valuable to develop more efficient treatment strategies in OCD.

Among a number of factors that have been proposed to determine treatment outcome with SRIs, genetic differences between patients may play a significant role[3]. In major depression, for example, it has been reported repeatedly that the short form (S-allele) of the 44-bp deletion/insertion functional polymorphism within the promoter region of the serotonin transporter gene (5-HTTLPR) is associated with impaired efficacy of SRIs[4]. In OCD, three studies have investigated the role of the 5-HTTLPR and treatment response. Mc Dougle et al found an association of the L-allele with poorer response to SRIs, whereas Billet et al and Di Bella et al failed to find a relation between response and 5-HTT genotypes[5-7]. Other receptors that might be involved in the therapeutic efficacy of SRIs are the terminal 5-$HT_{1B}$ autoreceptor and the postsynaptic 5-$HT_{2A}$ receptor. As of yet, neither polymorphisms of 5-$HT_{1B}$ or 5-$HT_{2A}$ receptor genes have been investigated with regard to treatment response of SRIs in OCD.

In this study we tested the hypothesis that variations of the 5-HTT (L-allele of the 5-HTTLPR), 5-$HT_{1B}$ and 5-$HT_{2A}$ gene expressions are linked to treatment response with SRIs in OCD. We report the results of 91 patients who participated in 12-week, double-blind trial with paroxetine and venlafaxine and were assessed for the 44 bp insertion/deletion 5-HTTLPR, the 5-$HT_{1B}$ (5-$HT_{1D\beta}$) G861C, and the 5-$HT_{2A}$ 1438G/A polymorphism.

## 4.2. Material and methods

### 4.2.1. Study sample

Ninety-one outpatients gave written informed consent for participation in this study that had been approved by the University of Utrecht Medical Ethical Review committee (Utrecht, The Netherlands). All patients were diagnosed with OCD according to DSM-IV criteria and the M.I.N.I., a clinical and structured interview, was used to confirm the diagnosis[8]. Severity of

obsessive-compulsive symptoms was rated with the Y-BOCS, depressive symptoms with the HAM-D, and anxiety with the HAM-A[9-11]. Only patients with a score of at least 18 on the Y-BOCS, or at least 12, if only obsessions or only compulsions were present, were included. Patients with a major depressive disorder or patients with a total score of 15 or more on the 17-item Hamilton Depression Rating Scale (HAM-D) on admission were excluded. Information on family history of OCD and other psychiatric disorders was obtained by direct interviews with the patients and the presence of vocal and/or motor tics was assessed during the clinical interview.

| | Non Responders (n=32) | Responders (n=56) |
|---|---|---|
| Gender (Male/female) | 14/18 | 20/36 |
| Age on admission | 31.7 ± 12.0 | 34.1 ± 11.3 |
| Positive family history | 11 | 19 |
| Mean age of onset | 14.7 ± 9.3 | 17.2 ± 7.4 |
| ≤15 years age of onset | 12 | 33 |
| > 20 years age of onset | 20 | 23 |
| Y-BOCS baseline | 26.8 ± 5.8 | 25.2 ± 5.2 |
| Y-BOCS endpoint | 24.8 ± 5.7 | 13.2 ± 5.4 |
| Y-BOCS mean % decrease | 6.8 ± 11.0 | 48.6 ± 18.0 |
| HAM-D | 5.6 ± 10.7 | 7.8 ± 10.8 |
| HAM-A | 7.4 ± 6.7 | 9.8 ± 7.5 |
| Paroxetine (n=40) | 9 | 31 |
| Venlafaxine (n=44) | 20 | 24 |

Table 1. Demographic and clinical characteristics of the patients sample

### 4.2.2. Study design

Patients were randomly assigned to receive either paroxetine or venlafaxine XR for twelve weeks in a single-center, double blind controlled, and parallel-group study design. Paroxetine treatment was initiated at a dose of 15 mg/day, and gradually increased to 60 mg/day using a fixed dosing schedule. Venlafaxine treatment was initiated at a dose of 75 mg/day and gradually increased to 300 mg/day. Psychotropic drugs or psychotherapy were not allowed. Obsessive-compulsive symptoms were measured with the Y-BOCS, and response to treatment was prospectively defined as a ≥ 25% decrease in Y-BOCS score. Three out of ninety-one patients dropped out during the study because of lack of motivation or side effects. A detailed description of the study has been published earlier[12, 13].

### 4.2.3. Genotyping

Blood samples were collected from each subject and frozen at – 80°. DNA was extracted from 10 ml samples of peripheral blood according to standard procedures. The total number of subjects genotyped for the genes in this study was 88. In seven cases, the genotyping of the 5-$HT_{1B}$ polymorphism failed, and in one case the genotyping of the 5-HTT polymorphism. All subjects were genotyped at the University of Ghent (Belgium) based on a coded identification number. The 5-HTT, 5-$HT_{1B}$ and 5-$HT_{2A}$ genotyping was performed following a standardized protocol.

#### 4.2.3.1. 5-HTT

For the detection of the 44 bp insertion/deletion 5-HTTLPR polymorphism, the oligonucleotide primers 5'-6FAM-GGCGTTGCCGCTCTGAATGC–3' and 5'–AGGGACTGAGCTGGACAACCAC-3' were used to amplify a 484/528 bp fragment comprising the 5-HTT-linked polymorphic region. The PCR reaction was performed according following conditions: 94°C for 1 min, 60 °C for 1min, 72°C for 1min40sec per cycle, for a total of 35 cycles. The PCR products were analysed by capillary electrophoresis on an Applied Biosystems 3100 Genetic Analyser.

#### 4.2.3.2. 5-$HT_{1B}$

For detection of the 5-$HT_{1B}$ or (5-$HT_{1D\beta}$) G861C polymorphism, the oligonucleotide primers 5' - GAAACAGACGCCCAACAGGAC - 3' and 5' -CCAGAAACCGCGAAAGAAGAT - 3' were used to amplify a 548 bp region comprising the G861C polymorphism site. The PCR reaction was performed under the following conditions: 90°C for 1 min, 55 °C for 2 min, 72°C for 3 min per cycle, for a total of 32 cycles. Digestion of 10µl of PCR product was accomplished by incubation for 4 hours with 10 units of Hinc II restriction enzyme at 37°C. Digestion with Hinc II yields either two fragments (452 bp and 96 bp) for the G-allele or three fragments (310 bp, 142 bp and 96 bp) for the C-allele. The fragments resulting from the digestion were resolved on a 1.5% agarose gel and visualized by ethidium bromide staining.

#### 4.2.3.3. 5-$HT_{2A}$

For the detection of the 5-$HT_{2A}$ 1438G/A polymorphism within the promoter region of the 5-$HT_{2A}$ receptor gene, the oligonucleotide primers 5'-6FAM-AAGCTGCAAGGTAGCAAC AGC–3' and 5'-NED-AACCAACTTATTTCCTACCAC-3' were used to amplify a 468 bp region comprising the $5HT_{2A}$ 1438G/A polymorphism site. The PCR reaction was performed under the following conditions: 95°C for 1 min, 47 °C for 1min, 72°C for 1min20sec per cycle, for a total of 40 cycles. Digestion of 10µl of PCR product was accomplished by overnight incubation with 10 units of Msp I restriction enzyme at 37°C. After incubation with Msp I, the 1438A allele remains intact while the 1438G allele is cut into a 223 bp piece (6FAM-labelled) and a 243 bp piece (NED-labelled). The fragments resulting from the digestion were analysed by capillary electrophoresis on an Applied Biosystems 3100 Genetic Analyser.

### 4.2.4. Data analysis

The following statistical procedure was pursued. Firstly, the genotypic pattern of distribution and the allele frequencies of the 5-HTT, 5-$HT_{1D}$ and 5-$HT_{2A}$ polymorphisms were analyzed in the whole sample (N=88). Secondly, an analogous analysis was performed in the paroxetine treated patients (N=40), and in the venlafaxine treated patients (n=44) separately. Medication use and dose was uncertain in four patients and they were excluded from the treatment groups. The association between the distribution of the genotypes and allele frequencies with the responders and non-responders were assessed by cross-tabulation and $\chi^2$ analyses. One way analysis of variance (ANOVAs) were calculated to determine whether significant differences were present between genotypes in mean decrease of Y-BOCS scores. Considering a partial Bonferroni's correction, the p value for statistical significance would be 0.020 with an alpha of 0.05, 6 tests, 2 degrees of freedom, and a correlation correction factor of 0.5. The data are presented as mean ± SD, and performed at 5% level of significance. All statistical analyses were conducted with the SPSS statistical package version 11.5.

## 4.3. Results

Demographic variables and outcome measures are presented in table 1. The patient sample was slightly skewed towards the female population (63%). Fifty-six out of 88 patients (63%) were rated as responders, 31 out of 40 patients in the paroxetine group and 24 out of 44 patients in the venlafaxine group. Four patients were not assigned to a particular treatment group (see methods section). There were no statistically significant differences between responders and non responders as regards gender, age, age of onset, family history, and baseline Y-BOCS, HAM-A, or HAM-D measures.

| | n | allele frequencies | | p-value | genotypes | | | p-value |
|---|---|---|---|---|---|---|---|---|
| **5-HT$_{1B}$** | | **C** | **G** | | **CC** | **CG** | **GG** | |
| Non responders | 30 | 0.23 | 0.74 | 0.273 | 1 (3.3%) | 12 (40.0%) | 17 (56.7%) | 0.510 |
| Responders | 51 | 0.31 | 0.69 | | 4 (7.8%) | 24 (47.1%) | 23 (45.1%) | |
| | | | | | | | | |
| **5-HT$_{2A}$** | | **A** | **G** | | **AA** | **AG** | **GG** | |
| Non responders | 32 | 0.48 | 0.52 | 0.418 | 5 (15.6%) | 21 (65.6%) | 6 (18.8%) | 0.144 |
| Responders | 56 | 0.42 | 0.58 | | 11 (19.6%) | 25 (44.6%) | 20 (35.7%) | |
| | | | | | | | | |
| **5-HTT** | | **L** | **S** | | **LL** | **L/S** | **SS** | |
| Non responders | 32 | 0.55 | 0.45 | 0.551 | 12 (37.5%) | 11 (34.4%) | 9 (28.1%) | **0.028** |
| Responders | 55 | 0.50 | 0.50 | | 10 (18.2%) | 35 (63.6%) | 10 (18.2%) | |

Table 2. Allele frequencies and genotype distribution of the 5-HT$_{1B}$-, 5-HT$_{2A}$ receptor, and 5-HT transporter polymorphisms in the whole sample (n=88)

In the whole sample (Table 2), a difference in genotype distribution of the 5-HTTLPR polymorphism was found between responders and non–responders. Sixty-four percent of the responders carried the S/L genotype of the 5-HTTLPR polymorphism compared to 18 % carrying the S/S genotype and 18 % carrying the L/L genotype. The difference just failed to be statistically significant after Bonferroni's correction ($\chi^2$ = 7.17, df=2, p=0.028). When the mean Y-BOCS decrease was stratified by 5-HTTLPR genotype, a superior response was observed in the S/L
genotype (37% decrease) versus the S/S genotype (28%) and the L/L genotype (29%), but the ANOVA failed to reach statistical significance ($F_{2,84}$= 1.2, p = 0.30). Allele frequencies of the 5-HTTLPR polymorphism between responders and non–responders were not statistically significant different ($\chi^2$ = 0.05, df=1, p=0.71), and there were no significant differences between responders and non–responders in allele or genotype frequencies for the 5-HT$_{1B}$ and 5-HT$_{2A}$ polymorphisms in the whole sample.

| | n | allele frequencies | | p-value | genotypes | | | p-value |
|---|---|---|---|---|---|---|---|---|
| **5-HT$_{1B}$** | | **C** | **G** | | **CC** | **CG** | **GG** | |
| Non responders | 8 | 0.19 | 0.81 | 0.513 | 0 (0.0%) | 3 (37.5%) | 5 (62.50%) | 0.625 |
| | 28 | 0.27 | 0.73 | | 3 (10.7%) | 9 (32.1%) | 16 (57.1%) | |
| **5-HT$_{2A}$** | | **A** | **G** | | **AA** | **AG** | **GG** | |
| Non responders | 9 | 0.67 | 0.33 | **0.004** | 3 (33.3%) | 5 (66.7%) | 0 (0.0%) | **0.013** |
| Responders | 31 | 0.29 | 0.71 | | 4 (12.9%) | 10 (32.3%) | 17 (54.8%) | |
| **5-HTT** | | **L** | **S** | | **LL** | **L/S** | **SS** | |
| Non responders | 9 | 0.56 | 0.44 | 0.772 | 3 (33.3%) | 4 (44.4%) | 2 (22.2%) | 0.787 |
| Responders | 30 | 0.52 | 0.48 | | 7 (23.3%) | 17 (56.7%) | 6 (20.0%) | |

Table 3. Allele frequencies and genotype distribution of the 5-HT$_{1B}$-, 5-HT$_{2A}$ receptor, and 5-HT transporter polymorphisms in the paroxetine treated group (n=40).

In the paroxetine treated patients (Table 3), the majority of responders carried the G/G genotype of the 5-HT$_{2A}$ polymorphism ($\chi^2$ = 8.66, df=2, p=0.013). The association of a superior response with the G/G genotype was confirmed in the ANOVA when the mean Y-BOCS decrease was broken down according to the genotypes. Patients carrying the G/G genotype of the 5-HT$_{2A}$ polymorphism had a mean decrease of 51% on the Y-BOCS compared to 34% with the A/A genotype and 29% with the A/G genotype ($F_{,2.39}$= 4.95, p = 0.012). In general, responders carried predominantly the G-allele compared to non-responders ($\chi^2$ = 8.43, df=1, p=0.004) (OR 4.89 95% CI 1.59-15.02).

| | n | allele frequencies | | p-value | genotypes | | | p-value |
|---|---|---|---|---|---|---|---|---|
| **5-HT$_{1B}$** | | **C** | **G** | | **CC** | **CG** | **GG** | |
| Non responders | 19 | 0.24 | 0.76 | 0.214 | 1 (5.3%) | 7 (36.8%) | 11 (57.9%) | 0.221 |
| Responders | 22 | 0.36 | 0.64 | | 1 (4.5%) | 14 (63.6%) | 7 (31.8%) | |
| **5-HT$_{2A}$** | | **A** | **G** | | **AA** | **AG** | **GG** | |
| Non responders | 20 | 0.40 | 0.60 | 0.087 | 2 (10.0%) | 12 (60.0%) | 6 (30.0%) | 0.165 |
| Responders | 24 | 0.58 | 0.42 | | 7 (29.2%) | 14 (58.3%) | 3 (12.5%) | |
| **5-HTT** | | **L** | **S** | | **LL** | **L/S** | **SS** | |
| Non responders | 21 | 0.55 | 0.45 | 0.393 | 8 (40.0%) | 6 (30.0%) | 6 (30.0%) | **0.008** |
| Responders | 23 | 0.46 | 0.54 | | 2 (8.3%) | 18 (75.0%) | 4 (16.7%) | |

Table 4. Allele frequencies and genotype distribution of the 5-HT$_{1B}$-, 5-HT$_{2A}$ receptor, and 5-HT transporter polymorphisms in the venlafaxine treated group (n=44).

In the venlafaxine treated patients (Table 4), the majority of responders carried the S/L genotype of the 5-HTTTLPR polymorphism ($\chi^2$ = 9.72, df=2, p=0.008). The ANOVA showed a difference in favor of the S/L genotype with a mean Y-BOCS decrease of 38 % compared to 24% in patients with the S/S

genotype and 15% in patients with the LL genotype, who had the worst outcome, but failed to be statistically significant after correction ($F_{2, 43}$= 3.27, p = 0.04).

Since the number of responder appeared to be correlated to the G/G genotype of the 5-$HT_{2A}$ polymorphism in the paroxetine treated patients, and to the S/L genotype of the 5-HTTLPR polymorphism in the venlafaxine treated patients, we analyzed thereupon responder rates in patients who had either one of the genotypes in the full sample. More than 81 % of the responders (45 out of 55) carried either the G/G genotype of the 5-$HT_{2A}$ polymorphism or the S/L genotype of the 5-HTTLPR polymorphism ($\chi^2$ = 8.1, df=1, p=0.004). All of patients (n=9) who carried both the G/G genotype of the 5-$HT_{2A}$ polymorphism and to the S/L genotype of the 5-HTTLPR polymorphism were responders. There was a statically significant difference between the mean Y-BOCS decrease of 49% in these patients compared to the remainder of patients ($\chi^2$ = 16.0, df=8, p=0.01).

## 4.4. Discussion

The main finding of this study is that OCD patients with the S/L genotype of the 5-HTTLPR polymorphism have a more favorable response following paroxetine and venlafaxine treatment. This effect was more pronounced for the venlafaxine treated patients, while response to paroxetine mainly was associated with the G/G genotype of the 5-$HT_{2A}$ polymorphism. The small group of patients (n=9) who both carried the S/L genotype of the 5-HTTLPR polymorphism and the G/G genotype of the 5-$HT_{2A}$ polymorphism responded all to treatment.

Three previous studies have investigated the role of the 5-HTTLPR and treatment response in OCD. Mc Dougle et al found in a sample of 33 patients a trend for an association of the L-allele with poorer response to SRIs (clomipramine, fluvoxamine, fluoxetine, sertraline and paroxetine)[5]. Billet et al examined retrospectively 72 patients after a 10-week trial with SRIs and found no association, and Di Bella et al failed to find a relation between response and 5-HTTLPR genotypes in a sample of 99 patients following a standardized fluvoxamine treatment of 12 weeks[6, 7]. Our results are in line with the findings of Mc Dougle and do not suggest a better outcome with SRIs in carriers of the L/L genotype of the 5-HTTLPR, which is in flat contradiction with the majority of reports in mood disorders in which the presence of the L variant of the 5-HTTLPR has been related to a more favorable and faster response with SRIs[4]. On the other hand, in Asian populations, an association in the opposite direction was found with a better response for carriers of the S allele. In sum, our findings are in line with those of Mc Dougle et al but disagree with Billet et al and Di Bella et al and the majority of the studies in major depressive disorder (MDD) patients[6, 7]. It is possible that this discrepancy is due to pathophysiologic and neurobiological dissimilarities between OCD and MDD. It has been suggested that SRIs exert their beneficial effects with their typical delay of 6 to 8 weeks in OCD by down regulating 5-$HT_{1B}$ receptors in the orbito-frontal whereas in MDD a faster response is observed

probably due 5-HT auto-receptor desensitization in other brain areas such as the hippocampus and hypothalamus[14]. This supposition is appealing, but still needs to be confirmed.

It is unclear exactly why the S/L genotype of the 5-HTTLPR would confer a favorable potential for a better response with SRIs in OCD. One might comprehend the connection of the L/L genotype with a superior response since it has been related to higher 5-HTT densities and hence an increased efficacy of SRIs. On the other hand, the L/L genotype of the 5-HTTLPR has been associated to placebo response as well, thereby questioning the rationale of the direct link between the 5-HTTLPR and therapeutic efficacy of SSRIs[4]. Furthermore, it still needs to be clarified whether or not the 5-HTTLPR determines the number of 5-HTT in the human brain *in vivo*[15, 16]. Some studies have reported that L/L homozygous individuals had higher 5-HTT availability compared to S/L or S/S homozygous individuals in the raphe area, but others failed to find an association in the diencephalon, brainstam, and the thalamus[17-20]. Equally, post mortem studies did not detect any significant influence of 5-HTTLPR on 5-HTT density in the hippocampus or frontal cortex[21, 22]. Thus it would be premature to relate superior response of the S/L genotype carriers in OCD to lower 5-HTT densities since it still remains to be elucidated whether the 5-HTTLPR genotypes relate to 5–HTT function and hence different psychopharmacological mechanisms of SRIs.

Except for Tot et al, who failed to find an association between the -1438G/A and T102C polymorphism of the 5-$HT_{2A}$ receptor in 52 patients following a 12 week trial with fluvoxamine, fluoxetine or sertraline, no further study has investigated the 5-$HT_{2A}$ receptor gene with regard to treatment response in OCD[23]. This is surprising since sensitization of the 5-$HT_{2A}$ receptor has been hypothesized to be a common mechanism of SRIs treatment[24, 25]. For example, Meyer et al have reported increased densities of the 5-$HT_{2A}$ receptor after paroxetine treatment. Massou et al, on the other hand, have found the opposite[26, 27]. A recent study in 54 Japanese patients with MDD failed to find a major role for the -1438G/A promoter polymorphism in therapeutic response to fluvoxamine, and similarly, Choi et al found no significant association between the 5-$HT_{2A}$ G-1438A genotype and treatment response[28, 29]. Thus far, it is unclear whether the -1438A/G promoter polymorphism results in functional effects[30]. Spurlock et al found no effect of the -1438A/G promoter polymorphism on basal or cAMP- and protein kinase C induced gene transcription in HeLa cells, and found no difference in lymphocyte 5-$HT_{2A}$ receptor mRNA expression between 1438A/A and G/G homozygotes[31]. Turecki et al, in a small postmortem study, reported higher prefrontal 5-$HT_{2A}$ receptor binding in subjects with the -1438A allele, but Bray et al failed to find a significant effect on 5-$HT_{2A}$ receptor mRNA expression in post mortem brain tissue[32, 33].

It is puzzling why response in paroxetine treated patients is related to the 5-$HT_{2A}$ receptor genotype and response in venlafaxine treated patients to the 5-HTTLPR. It has been reported that chronic treatment with paroxetine produces a significant desensitization in post synaptic 5-$HT_{2A}$ receptor function[26, 34]. On the other, the 5-HTT and 5-$HT_{2A}$ receptor are intimately linked, for example the

constitutive lack of the 5-HTT alters the density of the 5-$HT_{2A}$ receptor in a brain region specific manner, with an increase in the hypothalamus and decrease in the striatum[25, 35]. Thus, the apparent specific association of paroxetine and venlafaxine might be a spurious finding as result of a type two error due to the small sample sizes. Further investigation in larger samples might clarify this issue.

In summary, this study suggests a better outcome in OCD after treatment with SRIs for patients carrying the S/L genotype of the 5-HTTLPR polymorphism. This effect was more pronounced for the venlafaxine treated patients, whereas response to paroxetine was associated with the G/G genotype of the 5-$HT_{2A}$ polymorphism. The small group of patients who both carried the S/L genotype of the 5-HTTLPR polymorphism and the G/G genotype of the 5-$HT_{2A}$ polymorphism responded all to treatment. Our results indicate that 5-$HT_{2A}$ and 5-HTTLPR polymorphisms may be markers for treatment outcome in OCD.

# References


1. Hollander E, Bienstock CA, Koran LM, Pallanti S, Marazziti D, Rasmussen SA, Ravizza L, Benkelfat C, Saxena S, Greenberg BD *et al*: **Refractory obsessive-compulsive disorder: State-of-the-art treatment**. *Journal of Clinical Psychiatry* 2002, **63**:20-29.

2. March J: **Treatment of obsessive-compulsive disorder. The Expert Consensus Panel for Obsessive-compulsive disorder.** *J Clin Psychiatry* 1997, **58 Suppl. 4**:2-72.

3. Denys D, Burger H, van Megen H, de Geus F, Westenberg H: **A score for predicting response to pharmacotherapy in obsessive-compulsive disorder**. *International Clinical Psychopharmacology* 2003, **18**(6):315-322.

4. Smits KM, Smits LJM, Schouten JSAG, Stelma FF, Nelemans P, Prins MH: **Influence of SERTPR and STin2 in the serotonin transporter gene on the effect of selective serotonin reuptake inhibitors in depression: a systematic review**. *Molecular Psychiatry* 2004, **9**(5):433-441.

5. McDougle CJ, Epperson CN, Price LH, Gelernter J: **Evidence for linkage disequilibrium between serotonin transporter protein gene (SLC6A4) and obsessive compulsive disorder**. *Molecular Psychiatry* 1998, **3**(3):270-273.

6. Billett EA, Richter MA, King N, Heils A, Lesch KP, Kennedy JL: **Obsessive compulsive disorder, response to serotonin reuptake inhibitors and the serotonin transporter gene**. *Molecular Psychiatry* 1997, **2**(5):403-406.

7. Di Bella D, Erzegovesi S, Cavallini MC, d'Annucci A, Bellodi L: **Obsessive-compulsive disorder, treatment response and the 5HTT gene.** *American Journal of Medical Genetics* 2000, **96**(4):536-536.

8. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC: **The Mini-International Neuropsychiatric Interview (MINI): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10**. *Journal of Clinical Psychiatry* 1998, **59**:22-33.

9. Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, Heninger GR, Charney DS: **The Yale-Brown Obsessive Compulsive Scale.1. Development, Use, and Reliability**. *Archives of General Psychiatry* 1989, **46**(11):1006-1011.

10. Hamilton M: **The assessment of anxiety states by rating**. *Br J Med Psychol* 1959, **32**:50-55.

11. Hamilton M: **A rating scale for depression**. *J Neurol Neurosurg Psychiatry* 1960, **23**:56-62.

12. Denys D, Van Megen HJGM, Westenberg HGM: **A double blind comparison of venlafaxine and paroxetine in obsessive compulsive disorder**. *European Neuropsychopharmacology* 2002, **12**:S343-S343.

13. Denys D, Van Megen HJGM, Westenberg HGM: **A double blind switch study of venlafaxine and paroxetine in obsessive compulsive disorder**. *European Neuropsychopharmacology* 2002, **12**:S343-S344.

14. Elmansari M, Bouchard C, Blier P: **Alteration of Serotonin Release in the Guinea-Pig Orbitofrontal Cortex by Selective Serotonin Reuptake Inhibitors - Relevance to Treatment of Obsessive-Compulsive Disorder**. *Neuropsychopharmacology* 1995, **13**(2):117-127.

15. Williams RB, Marchuk DA, Gadde KM, Barefoot JC, Grichnik K, Helms MJ, Kuhn CM, Lewis JG, Schanberg SM, Stafford-Smith M *et al*: **Serotonin-related gene polymorphisms and central nervous system serotonin function**. *Neuropsychopharmacology* 2003, **28**(3):533-541.

16. Smith GS, Lotrich FE, Malhotra AK, Lee AT, Ma YL, Kramer E, Gregersen PK, Eidelberg D, Pollock BG: **Effects of serotonin transporter promoter polymorphisms on serotonin function**. *Neuropsychopharmacology* 2004, **29**(12):2226-2234.

17. Heinz A, Jones DW, Mazzanti C, Goldman D, Ragan P, Hommer D, Linnoila M, Weinberger DR: **A relationship between serotonin transporter genotype and in vivo protein expression and alcohol neurotoxicity**. *Biological Psychiatry* 2000, **47**(7):643-649.

18. Van Dyck CH, Malison RT, Staley JK, Jacobsen LK, Seibyl JP, Laruelle M, Baldwin RM, Innis RB, Gelernter J: **Central serotonin transporter availability measured with [I-123]beta-CIT SPECT in relation to serotonin transporter genotype**. *American Journal of Psychiatry* 2004, **161**(3):525-531.

19. Shioe K, Ichimya T, Suhara T, Takano A, Sudo Y, Yasuno F, Hirano M, Shinohara M, Kagami A, Okubo Y *et al*: **No association between genotype of the promoter region of serotonin transporter gene and serotonin transporter binding in human brain measured by PET**. *Synapse* 2003, **48**(4):184-188.

20. Willeit M, Stastny J, Pirker W, Praschak-Rieder N, Neumeister A, Asenbaum S, Tauscher J, Fuchs K, Sieghart W, Hornik K *et al*: **No evidence for in vivo regulation of midbrain serotonin transporter availability by serotonin transporter promoter gene polymorphism**. *Biological Psychiatry* 2001, **50**(1):8-12.

21. Naylor L, Dean B, Pereira A, Mackinnon A, Kouzmenko A, Copolov D: **No association between the serotonin transporter-linked promoter region polymorphism and either schizophrenia or density of the serotonin transporter in human hippocampus**. *Molecular Medicine* 1998, **4**(10):671-674.

22. Mann JJ, Huang JY, Underwood MD, Kassir SA, Oppenheim S, Kelly TM, Dwork AJ, Arango V: **A serotonin transporter gene promoter polymorphism (5-HTTLPR) and prefrontal cortical finding in major depression and suicide**. *Archives of General Psychiatry* 2000, **57**(8):729-738.

23. Tot S, Erdal ME, Yazici K, Yazici AE, Metin O: **T102C and-1438 G/A polymorphisms of the 5-HT2A receptor gene in Turkish patients with obsessive-compulsive disorder**. *European Psychiatry* 2003, **18**(5):249-254.

24. Tilakaratne N, Yang ZL, Friedman E: **Chronic Fluoxetine or Desmethylimipramine Treatment Alters 5-Ht2 Receptor-Mediated C-Fos Gene-Expression**. *European Journal of Pharmacology-Molecular Pharmacology Section* 1995, **290**(3):263-266.

25. Li Q, Muma NA, Battaglia G, Van der Kar LD: **Fluoxetine gradually increases [I-125]DOI-labelled 5-HT2A/2C receptors in the hypothalamus without changing the levels of G(q)- and G(11)-proteins**. *Brain Research* 1997, **775**(1-2):225-228.

26. Meyer JH, Kapur S, Eisfeld B, Brown GM, Houle S, DaSilva J, Wilson AA, Rafi-Tari S, Mayberg HS, Kennedy SH: **The effect of paroxetine on 5-HT2A receptors in depression: An [F-18]setoperone PET imaging study**. *American Journal of Psychiatry* 2001, **158**(1):78-85.

27. Massou JM, Trichard C, AttarLevy D, Feline A, Corruble E, Beaufils B, Martinot JL: **Frontal 5-HT2A receptors studied in depressive patients during chronic treatment by selective serotonin reuptake inhibitors**. *Psychopharmacology* 1997, **133**(1):99-101.

28. Sato K, Yoshida K, Takahashi H, Ito K, Kamata M, Higuchi H, Shimizu T, Itoh K, Inoue K, Tezuka T *et al*: **Association between-1438G/A promoter polymorphism in the 5-HT2A receptor gene and fluvoxamine response in Japanese patients with major depressive disorder**. *Neuropsychobiology* 2002, **46**(3):136-140.

29. Choi MJ, Lee HJ, Lee HJ, Ham BJ, Cha JH, Ryu SH, Lee MS: **Association between major depressive disorder and the-1438A/G polymorphism of the serotonin 2A receptor gene**. *Neuropsychobiology* 2004, **49**(1):38-41.

30. Veenstra-VanderWeele J, Anderson GM, Cook EH: **Pharmacogenetics and the serotonin system: initial studies and future directions**. *European Journal of Pharmacology* 2000, **410**(2-3):165-181.

31. Spurlock G, Heils A, Holmans P, Williams J, D'Souza UM, Cardno A, Murphy KC, Jones L, Buckland PR, McGuffin P *et al*: **A family based association study of T102C polymorphism in 5HT2A and schizophrenia plus identification of new polymorphisms in the promoter**. *Molecular Psychiatry* 1998, **3**(1):42-49.

32. Turecki G, Briere R, Dewar K, Antonetti T, Lesage AD, Seguin M, Chawky N, Vanier C, Alda M, Joober R *et al*: **Prediction of level of serotonin 2A receptor binding by serotonin receptor 2A genetic variation in postmortem brain samples from subjects who did or did not commit suicide**. *American Journal of Psychiatry* 1999, **156**(9):1456-1458.

33. Nakamura T, Matsushita S, Nishiguchi N, Kimura M, Yoshino A, Higuchi S: **Association of a polymorphism of the 5HT2A receptor gene promoter region with alcohol dependence**. *Molecular Psychiatry* 1999, **4**(1):85-88.

34. Chen Z, Waimey K, Van de Kar LD, Carrasco GA, Landry M, Battaglia G: **Prenatal cocaine exposure potentiates paroxetine-induced desensitization of 5-HT2A receptor function in adult male rat offspring**. *Neuropharmacology* 2004, **46**(7):942-953.

35. Rioux A, Fabre V, Lesch KP, Moessner R, Murphy DL, Lanfumey L, Hamon M, Martres MP: **Adaptive changes of serotonin 5-HT2A receptors in mice lacking the serotonin transporter**. *Neuroscience Letters* 1999, **262**(2):113-116.

# PART II:

## Proteome analysis

## of

## *Artemisia annua*

# 1. Objective and strategy

The objective of this study is to find genes of the plant *Artemisia annua* L. that are involved in the production of the antimalarial artemisinin. These genes do not necessarily have to code for the enzymes involved in the biosynthetic pathway, but can also be involved in the mechanisms influencing the amount of produced artemisinin. (e.g. genes involved in the formation of trichomes)

Three strategies were followed to accomplish this challenge: a proteome analysis, a quantitative cDNA amplified fragment length polymorphism analysis (cDNA AFLP) and the construction of three full length Expressed Sequence Tag (EST) cDNA libraries. The results of these 3 strategies can be compared and complemented with each other.

- We investigated the proteome of *A. annua* by identifying proteins that were differentially expressed between trichomes and whole leaf tissue. We also identified differentially expressed proteins between lower leafs and upper leafs of *A. annua* plants. In a third proteome analysis, we explored the proteins present in a one-minute chloroform extract of *A. annua* plants.

  Trichomes are considered to be the artemisinin factories of *A. annua* (chapter 2.2.2.). The upper leafs produce more artemisinin than the lower leafs (chapter 2.2.2.). The proteins that are upregulated in these samples might thus be involved in artemisinin production. With the chloroform extraction, the content of the capitate glands of the trichomes is extracted (chapter 4). These glands possibly contain proteins involved in artemisinin production.

  Proteins of interest were identified by tandem mass spectrometry. The fragmentation spectra were searched against the public Mass Spectrometry protein sequence DataBase (MSDB) and against the three EST libraries. Because only a few genes and proteins of A. annua have been characterized (chapter 2.3.1.2.), identification of proteins by searching public protein databases, is only possible due to homology with known proteins of other plants.

  For a general introduction on proteomics: see chapter 2.3.3. For the results of this study: see chapter 3

- Quantitative cDNA AFLP analysis is a genome-wide messenger RNA (mRNA) expression analysis. The level of mRNA expression was compared between samples of *A. annua* leafs, taken at different time points during a 72h time period after exposure to jasmonic acid (JA). Transcripts that are overexpressed by JA are possibly involved in the production of artemisinin, because the production of artemisinin is stimulated by JA (chapter 2.3.1.3).

  This analysis was conducted by the department of Plant Systems Biology of the "Vlaams interuniversitair instituut voor Biotechnologie" (VIB), Ghent University, Technologiepark 927, B-9052 Ghent, Belgium.

  For a general introduction on quantitative cDNA AFLP see chapter 2.3.2. For the results of this study: see chapter 3.

- Construction of three EST cDNA libraries: one from poly-A RNA from flower buds of *A. annua* and one from poly-A RNA from the trichomes on the flower buds. A subtracted cDNA library using poly-A RNA from the flower buds and the trichomes was also constructed. ESTs that are present in the trichomes and not in the flower buds are possibly involved in the production of artemisinin because trichomes are considered to be the artemisinin factories of *A. annua* (chapter 2.2.2.).

  Searching for homologies between the EST sequences and sequences of known plant genes revealed a cDNA clone encoding a cytochrome P450 enzyme. This enzyme (CYP71AV1) was expressed and characterized in *Saccharomyces cerevisiae* and was found to catalyze three steps: oxidation of amorpha-4,11-diene to artemisinic alcohol, artemisinic alcohol to artemisinic aldehyde and artemisinic aldehyde to artemisinic acid (see chapter 2.3.1.1.).

  Proteins pointed out by the proteomics analysis and sequences obtained from the cDNA AFLP analysis were compared with the cDNA EST libraries. The most promising genes are of course those that are found to be differential in all three techniques.

  The cDNA library construction was performed by the Plant Biotechnology Institute, 110 Gymnasium Place, Saskatoon, Canada and is described in a paper by Teoh *et al.*[1].

To investigate which plant hormones trigger *A. annua* to produce more artemisinin, to check the correlation between the artemisinin content of *A. annua* leafs and the number of trichomes on these leafs, and to be able to check transformed plants or shoot cultures for enhanced artemisinin production, a quantitation method for artemisinin and its bioprecursors had to be developed. This resulted in the finding that a chloroform extract resulting from a one-minute immersion of fresh plant material in chloroform, could be used for high performance liquid chromatography – electrospray ionization – quadrupole time-of-flight tandem mass spectrometry analysis (HPLC – ESI – QTOF MS/MS) without additional sample preparation steps. This research is described in chapter 4.

# 2. Introduction

## 2.1. Malaria

### 2.1.1. Malaria: a devastating disease with a long history

Malaria has been described since the dawn of history. The symptoms of shivering, fever, and spleen enlargement are described in Egyptian (Ebers papyrus, 1570 B.C.) and Chinese (Nei Ching, The Canon of Medicine, 1700 B.C.) writings. The *Corpus Hippocraticum* (fifth century B.C.) describes the recurrence of fevers at regular intervals and the connection of the disease to marches. In the seventeenth century, Italians believed that breathing bad air (mal airia) arising from swamps was responsible for the disease, and in the first half of the nineteenth century the term malaria entered the English literature. The French physician Charles Louis Alphonse Laveran first identified the parasite under the microscope in 1880. In 1899, the mystery of malaria transmission was solved independently, by Ronald Ross, an English physician working in India and Giovanni Battista Grassi, an Italian physician, who proved that the disease was spread by the bite of female mosquitoes of the genus *Anopheles*[2].

Today, malaria is one of the world's most devastating parasitic diseases. There are at least 300 million acute cases of malaria each year globally, resulting in more than a million deaths. Most of these cases occur in Africa, but large areas of Asia, Central, and South America have high incidences of the disease[3, 4].

### 2.1.2. Pathogenesis

Human malaria is caused by four major Plasmodium species: *falciparum*, *vivax*, *malariae* and *ovale*. *P. falciparum* causes the large majority of the clinical cases and mortalities[5].

The development and spreading of the parasite, goes in 4 phases[6, 7]:

1. A female anopheline mosquito carrying malaria (vector) takes a blood meal and injects haploid sporozoites (the infectious form of the parasite released from the anopheline mosquitoes salivary glands) into the blood stream.

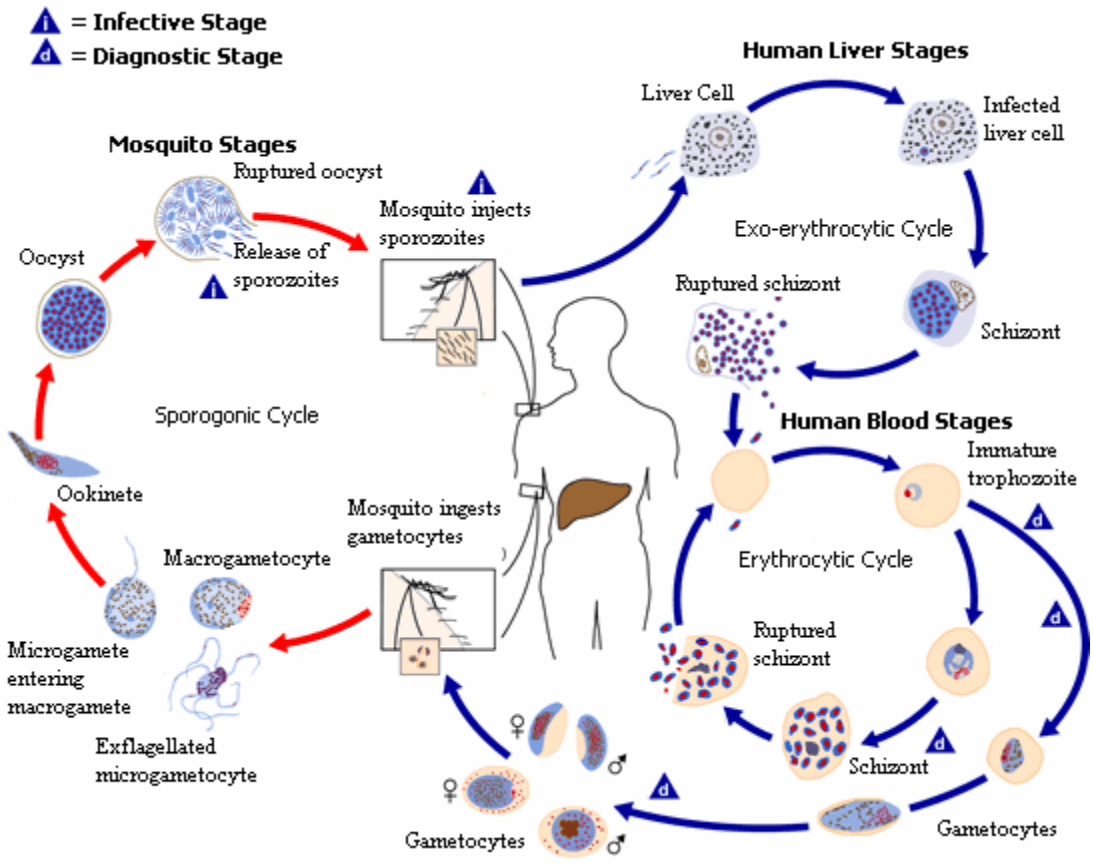


**Figure 1**: Malaria cycle (Centers of Disease Control - http://www.dpd.cdc.gov/dpdx)

2. The sporozoites enter the hepatocytes (liver cells) and initiate the exo-erythrocytic cycle in the liver. In the hepatocytes, the sporozoites undergo multiple asexual fissions, or schizogony, to produce thousands of infective, haploid merozoites.
3. The infected hepatocytes rupture and the merozoites invade erythrocytes where they continue the asexual cycle (erythrocytic cycle). Periodically, the infected red blood cells lyse (causing fever, anemia, coma and possible death), and the merozoites invade fresh erythrocytes. In some infected erythrocytes, merozoites develop into gametocytes.
4. When these gametocytes are ingested by mosquitoes, they initiate sexual development in the midgut. The female gametocytes exit the mosquito stomach after fertilization and form oocysts which produce sporozoites. These sporozoites migrate to the mosquito salivary glands and are passed into humans when the mosquito feeds.

The parasite is relatively protected from the immune system, because for most of his human life cycle, it resides in the liver and blood cells. Infected erythrocytes are however destroyed in the spleen. To avoid passage through the spleen, *P. falciparum* produces adhesive proteins on the surface of the infected erythrocytes, causing them to stick to the walls of small blood vessels. The parasite constantly switches between a broad repertoire of surface protein variants, making it difficult for the immune system to effectively develop antibodies against these proteins. High endothelial venules can get obstructed by infected erythrocytes, causing placental and cerebral malaria. In patients with cerebral malaria, infected erythrocytes can eventually breach the blood-brain barrier, which leads to coma and death[6, 7].

Other mammals as well as bird and reptiles also suffer from malaria. Only *P. malariae* can cause malaria both in humans and other higher primates. Other animal forms of malaria do not infect humans. *P. falciparum*, *vivax* and *ovale* are exclusive to humans.

### 2.1.3. Treatment and resistance to treatment

For many years, chloroquine was the antimalarial drug of choice in most parts of the world[8]. Multi-drug resistance of the Plasmodium strains to the cheapest and most widely used antimalarials such as chloroquine, mefloquine and sulfadoxine-pyrimethamine is one of the biggest challenges in the fight against malaria[3]. In parts of Southeast Asia, *P. falciparum* is now resistant to almost all antimalarial drugs and strains of chloroquine resistant *P. vivax* have emerged. In Africa, chloroquine resistance is widespread and resistance to sulphadoxine/pyrimethamine is being detected with increasing frequency[3, 9-12].

*A. annua* L. (sweet wormwood), a herb of the Asteraceae family has been used for centuries for the treatment of fever and malaria. Artemisinin is the main component responsible for this therapeutic effect[13]. Based on artemisinin, several semisynthetic derivatives such as artemether, arteether and artesunate have been produced (see chapter 2.2.4.). The World Health Organization (WHO) recommends that all countries experiencing resistance to conventional monotherapies should use combination therapies, preferably those containing artemisinin derivatives (ACTs or artemisinin based combination therapies)[14, 15].

In addition to the therapeutic benefits, artemisinin based combination therapies also bring a significant reduction in parasite transmission to *Anopheles* mosquitoes, particularly parasites carrying drug resistance genes. Artemisinin and it derivatives act against immature

gametocytes during the period of sequestration (7 days) that precedes emergence into the peripheral circulation as mature infectious gametocytes. In this way, they minimize the carriage of gametocytes, the parasite's transmissible stage[16-18].

Clinically relevant artemisinin resistance has not been demonstrated, but it is likely to occur since artemisinin resistance has been obtained in laboratory models[19].

## 2.2. *Artemisia annua* L.

### 2.2.1. Botany

*A. annua* L. is one of nearly 400 species of the Asteraceae. The herb is native to Asia but now grows in nature in many other countries in Europe and North America. The generic name *A.* refers to Artemis, goddess of maternity, because in antiquity plants of this genus were used to control birth and regulate women's menstrual disorders. The specific name *annua* reflects the annual cycle of the plant.

Figure 1: *A. annua* before (left) and during flowering (right). The left picture is provided by the Wisconsin State Herbarium. The right picture is provided by www.delawarewildflowers.org.

*A. annua*, commonly known in the United states by the names sweet or annual wormwood, can reach 2.0 meters in height. The plant is usually single-stemmed with alternate branches and alternate, deeply insected leaves ranging from 2.5 to 5.0 cm in length. Tiny yellow nodding flowers (capitula) only 2 to 3 mm across are displayed in lose panicles containing numerous bisexual florets in the centre and pistillate marginal florets. The plant is naturally cross pollinated by insect and wind action, which is unusual in the Asteraceae, and senesces after seeds are mature[20].

Non-glandular T-shaped trichomes and 10-celled biseriate glandular trichomes occur on leaves, stems and inflorescences[20]. The morphology and origin of the glandular trichomes

has been described for leaves[21] and capitula[22] using light and/or scanning electron microscopy. These glandular trichomes are known to contain essential oils. The essential oil (common to the Asteraceae) of *A. annua* contains at least 40 volatile compounds and several non-volatile sesquiterpenes, one of which is artemisinin[23].

As artemisinin cannot be synthesized chemically in an economically feasible way, *A. annua* is the only practical source of this valuable drug[24]. Unfortunately, *A. annua* contains only very small amounts of artemisinin ranging from 0.001 to 1.54% of dry weight[25]. Several research programs have been set up trying to increase the concentration of artemisinin in *A. annua* by optimizing the growing and harvesting conditions, by selecting high yielding cultivars or by creating transgenic plants[24, 26].

### 2.2.2. Glandular trichomes as sites of artemisinin accumulation

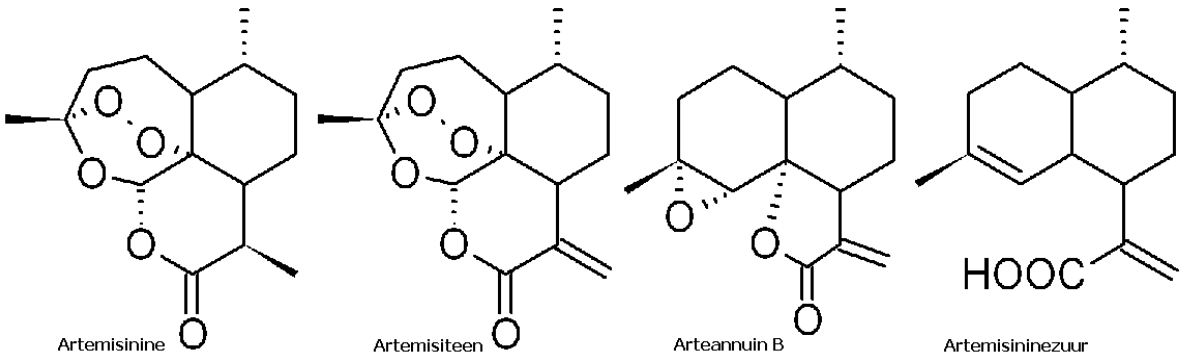


Figure 2: Artemisnin and structural analogs

There is strong circumstantial evidence that artemisinin is produced in the glandular trichomes. Duke *et al.*[27] reported that a 5 seconds dip in chloroform extracted 97% of the artemisinin and 100% of artemisitene from *A. annua*. Light microscopy and transmission electron microscopy revealed that the 5 seconds dip results in collapse of the subcuticular cavity of the glands on the leaf surface without visible damage to the leaf epidermal cells. An *A. annua* biotype without glands contained neither artemisinin nor artemisitene. These results indicate that artemisinin and artemisitene present in foliar tissue are localized entirely in the subcuticular space of glands of *A. annua*. The fact that artemisinin is not detected in parts of the plant that do not bear glandular trichomes also support this hypothesis.

Bertea *et al.*[28] incubated intact gland cell clusters in the presence of [$^3$H]farnesyl diphosphate and products of conversion were analyzed by radio-GC (gas chromatography coupled to detection of radioactive isotopes) and GC-MS (GC coupled to mass spectrometry). Radio-GC analysis showed a large radio-labeled amorpha-4,11-diene peak, indicating that the trichome cells contain amorpha-4,11-diene synthase and are able to perform the first step in the biosynthetic pathway of artemisinin[29].

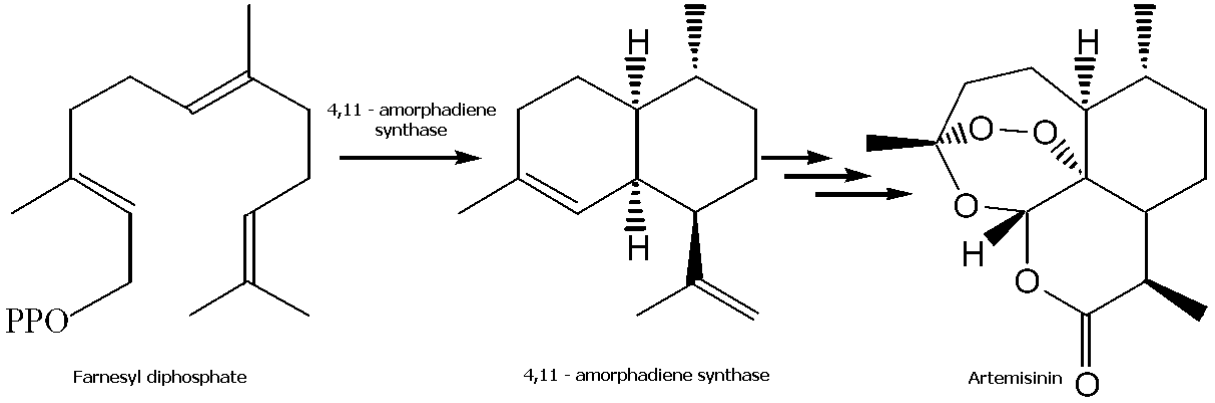


Figure 3: Cyclization of farnesyl diphosphate by 4,11-amorphadienesynthase

We used the quantitation method described in chapter 4 to test if the amount of artemisinin in a leaf correlates with the amount of glandular trichomes on the leaf. Samples were taken from the lower, middle and upper leaves of an approximately 2 months old, non-flowering *A. annua* plant. Microscopic inspection showed that the lower leafs had a much lower amount of trichomes per mm$^2$ than the middle leafs. The upper leafs had the highest amount of trichomes per mm$^2$. Table 1 shows the quantities of artemisinin and some its structural analogs (arteannuin B, artemisitene and artemisinic acid), present in the lower, middle and upper leafs of the same *A. annua* plant. As the amount of analyte present per gram of fresh leaf, is lowest in the lower leafs and highest in the upper leafs, there is a clear correlation with the number of trichomes present on these leafs. These results are completely in accordance with the results recently published by Zhang *et al.*[30], who report the difference between lower-, middle and upper leafs and who report a correlation coefficient of 0.987 between artemisinin content and the density of capitate glands on the surface of different plant tissues.

| Mean quantities (µg/g) | arteannuin B | artemisitene | artemisinin | artemisinic acid |
|---|---|---|---|---|
| Lower leafs | < LLOQ | < LLOQ | 16.58±2.01 | 38.00±0.47 |
| Middle leafs | 84.83±15.86 | < LLOQ | 41.26±3.47 | 41.75±2.45 |
| Upper leafs | 222.10±16.39 | < LLOQ | 94.50±0.96 | 52.63±1.75 |

Table 1: Quantities of artemisinin and some its structural analogs, present in the lower, middle and upper leafs of the same *A. annua* plant. The results are shown as the amount (µg) of analyte present in one gram of fresh plant material. Amounts below the lower limit of quantitation (LLOQ) are not shown.

### 2.2.3. *A. annua*, the only economically feasible source of artemisinin

The first published laboratory procedure for isolation of artemisinin is by Klayman *et al*. (1984)[31]. Air dried leaves were extracted with petroleum ether (bp=30 to 60°C), which was subsequently removed under vacuum. The residue was dissolved in chloroform to which acetonitrile was added to precipitate inert plant components such as sugar and waxes. The concentrated extract was then chromatographed on a column of silica gel. Fractions with a high artemisinin content crystallized readily, recrystallization was achieved with cyclohexane or 50% ethanol.

The most abundant sesquiterpene in *A. annua* is artemisinic acid, which can be 8-10 times more abundant than artemisinin[25, 32, 33] Vonwiller *et al.*[34] developed an efficient method to extract both artemisinic acid and artemisinin from the same material. Artemisinic acid can then be semi-synthetically converted to artemisinin, which greatly increases the yield of artemisinin[35].

None of the reported methods for the total stereospecific synthesis of artemisinin are economically feasible due to the complexity (up to 13 synthesis steps) and very low yields[32, 36-40]. Recently Ro *et al.*[41] engineered a *Saccharomyces cerevisiae* yeast to produce high titers of artemisinic acid, opening perspectives of cost-effective semi-synthesis of artemisinin.

### 2.2.4. Artemisinin and derivatives

Artemisinin, an endoperoxide-containing sesquiterpene lactone is a secondary metabolite produced by *A. annua*. Secondary metabolites give the plant a selective advantage. A huge variety of sesquiterpene lactones are known to act as insect deterrents, vertebrate poisons, etc. Endoperoxide-containing sesquiterpene lactones are however very uncommon.

#### 2.2.4.1. First-generation antimalarial endoperoxides

Artemisinin is surprisingly stable for an endoperoxide. It can be heated up to 50°C above its melting point (200°C) for 2.5 min.[42]. *In vitro* tests showed that artemisinin and various derivatives are effective against *P. falciparum* at nanomolar concentrations and showed little cross-resistance with other antimalarial agents[43]. Adverse effects are rare in patients treated with artemisinin derivatives[19, 44]. In a prospective study of over 3500 patients in Thailand, there was no evidence for serious adverse events[45]. Artemisinin derivatives also appear to be safe for pregnant women[46, 47]. In several animal studies however, artemisinin derivatives have clearly shown to cause neurotoxicity at high doses[19].

Artemisinin has been formulated as tablets, capsules and suppositories. Since artemisinin is poorly soluble in water or oil and has a low bioavailability, water-soluble derivatives (artesunate and artelinate) and oil-soluble derivatives (artemether and arteether) have been synthesized, making it possible to prepare parenteral formulations[43]. All of these drugs have comparable efficacy. The choice of derivative should be based upon availability, cost and quality of the preparation. They are all well-tolerated in both adults and children, with no evidence to date of serious clinical toxicity[44].

Several hundred other semisynthetic artemisinin derivatives have been prepared and tested: ethers or esters of dihydroartemisinin, 9-alkyl derivatives, 10-deoxoartemisinins, lactol ring-contracted derivatives etc. (reviewed by Meshnick *et al.* [43]).

#### 2.2.4.2. Second-generation antimalarial endoperoxides

The complex ring structure of artemisinin is not necessary for antimalarial activity; only the endoperoxide bridge is required. A group of simplified analogs, the trioxanes, also has antimalarial activity both *in vitro* and *in vivo*. These analogs all have the same endoperoxide-containing six-membered ring found in artemisinin but are much simpler to synthesize and some of them have increased stability[43].

A somewhat different group of endoperoxides are based on yingzhaosu, another endoperoxide containing compound discovered by Chinese scientists. One of these compounds, arteflene, is quite effective in infected patients with *P. falciparum*. Unfortunately, it was not superior to the semisynthetic artemisinin derivatives and it is not being developed further[43].

#### 2.2.4.3. Mechanism of action

Artemisinin and its derivatives are toxic to malaria parasites at nanomolar concentrations, whereas micromolar concentrations are required for toxicity to mammalian cells. One reason for this selectivity is the enhanced uptake and concentration of the drug by *P. falciparum* infected erythrocytes to more than a 100 fold higher than uninfected erythrocytes. Artemisinin derivatives are hydrophobic and partition into biological membranes of the parasite[43].

The endoperoxide bridge of artemisinin and its derivatives is necessary for antimalarial activity. Since peroxides are a known source of reactive oxygen species such as hydroxyl radicals and superoxide, this observation suggested that free radicals might be involved in the mechanism of action.

There is no consensus on how the endoperoxide bridge breaks open to form free radicals and how these free radicals kill the parasites. Accumulating biological and chemical evidence supports the hypothesis that the formation of free radicals is heme-mediated. As the malaria parasite is rich in heme-iron, derived from the proteolysis of host cell hemoglobin, this could explain why artemisinin is selectively toxic to parasites. Once formed, the artemisinin derived free radicals appear to damage specific intracellular targets, possibly via alkylation. Possible mechanisms and targets are reviewed by Meshnick[19].

Research by Eckstein-Ludweg *et al*.[48] shows compelling evidence that artemisinins act by inhibiting the "sarco/endoplasmic reticulum $Ca^{2+}$ - ATPase" orthologue (*Pf*ATP*ase*6) of *Plasmodium falciparum* after activation in a non-heme-mediated $Fe^{2+}$ dependent manner. The findings of Jambou *et al*.[49] support *Pf*ATP*ase*6 as the target for artemisinins: A S769N *Pf*ATP*ase*6 mutation in *P. falciparum* isolates from French Guiana was associated with raised (> 30 nmol/L) artemether $IC_{50}$s ($p>0.0001$, Mann-Withney).

## 2.3. Gene discovery: Which genes influence artemisinin production?

### 2.3.1. Genes and biosynthesis of artemisinin

#### 2.3.1.1. Biosynthesis of artemisinin

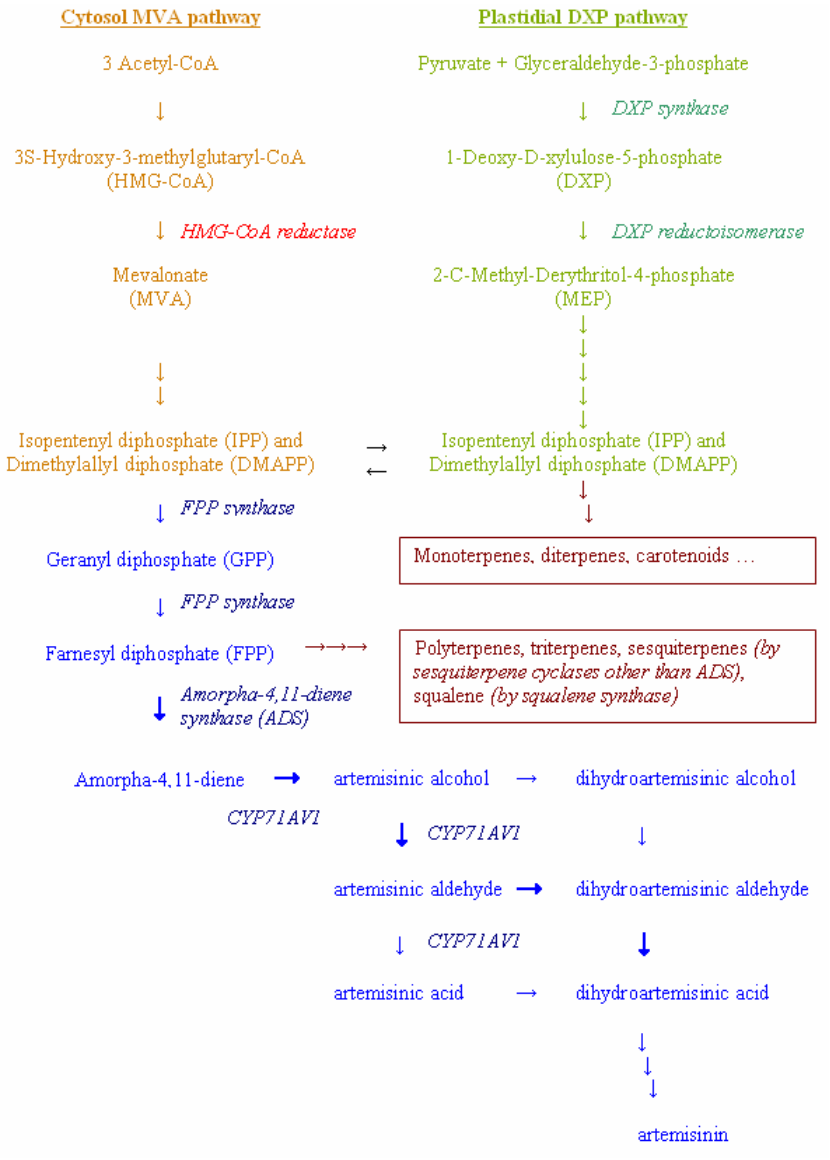


Figure 4: Simplified terpenoid biosynthetic scheme[50-52], supplemented with the artemisinin pathway proposed by Bouwmeester and co-workers (big blue arrows)[29, 53]. The oxidations catalyzed by CYP71AV1, the P450 enzyme characterized by Teoh *et al.*[1] is also indicated.

Artemisinin, an endoperoxide-containing sesquiterpene lactone is a secondary metabolite belonging to the category of the terpenes. Terpenes are the largest class of plant secondary metabolites with several thousands of representatives. These compounds are crucial to normal plant function and a major source for scientifically and commercially important chemicals, including pharmaceuticals (paclitaxel, artemisinin), flavors (menthol), food colors (carotenoids) and pesticides. Despite their great diversity, terpenoids are derived from two common precursors, isopentenyl diphosphate (IPP) and its isomer (catalyzed by isopentenyl diphosphate isomerase), dimethylallyl diphosphate (DMAPP). It has been established that higher plants have two independent biosynthetic pathways leading to the formation of IPP: the cytosolic mevalonate (MVA) pathway and the plastid-localized MVA-independent pathway. (Plastids are organelles responsible for photosynthesis, for storage of products like starch and for the synthesis of many classes of molecules such as fatty acids and terpenes) Terpene synthases, like farnesyl diphosphate synthase (FPPS), then convert IPP and DMAPP to linear prenyl diphosphates of different chain length, who are used to form the different terpenes[50, 52]. Matsushita *et al.*[54] were the first to clone the *A. annua* FPPS gene. They determined that FPPS from *A. annua* was very similar to that of other plants. FPPS catalyzes sequential condensation reactions:

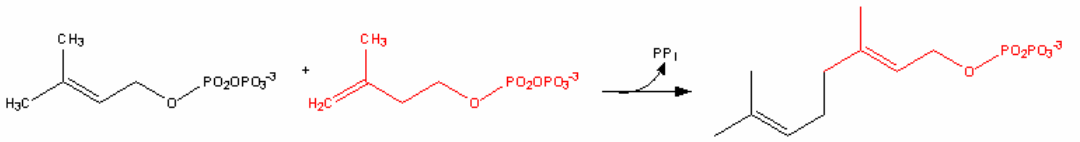


Figure 5: DMAPP reacts with 3 IPP to form geranyl diphosphate

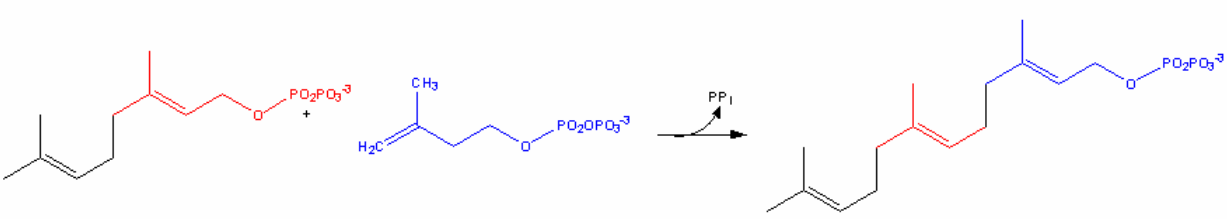


Figure 6: Geranyl diphosphate reacts with 3 IPP to form farnesyl diphosphate (FPP)

The post-FPP biosynthetic pathway of artemisinin is not yet completely elucidated. It is clear that the first dedicated step in the biosynthesis of artemisinin is the cyclization of FPP to amopha-4,11-diene by amorpha-4,11-diene synthase (ADS)[29]. ADS has now been cloned by several groups[55-57]. The amino acid sequence varies from 31 to 52% identity with other known angiosperm sesquiterpene cyclases[55, 56]. *In vitro* enzymatic assays demonstrated production of amorpha-4,11-diene from FPP[55].

Several authors have demonstrated that artemisinic acid and/or dihydroartemisinic acid are further intermediates in the formation of artemisinin[32, 58, 59]. With GC-MS, radio-GC, nuclear magnetic resonance (NMR) and enzyme assays, Bertea *et al.*[53] identified the intermediates and enzymes involved in the conversion of amorpha-4,11-diene to dihydroartemisinic acid. They propose the hydroxylation of amorpha-4,11-diene to artemisinic alcohol, followed by oxidation to artemisinic aldehyde, and reduction of the C11–C13 double bond to dihydroartemisinic aldehyde and oxidation to dihydroartemisinic acid (see figure 2). Teoh *et al.*[1] constructed a cDNA library with trichome-specific ESTs and found a cDNA clone encoding a cytochrome P450 (CYP71AV1). The enzyme was expressed and characterized in a *Saccharomyces cerevisiae* yeast and was found to catalyze three steps: oxidation of amorpha-4,11-diene to artemisinic alcohol, artemisinic alcohol to artemisinic aldehyde and artemisinic aldehyde to artemisinic acid.

The subsequent route to artemisinin and the genes involved herein, have not been characterized. Some papers propose complete biosynthetic pathways, others focus on an individual biosynthetic step. They are reviewed by Van Geldre *et al.*[60] and Liu *et al.*[51]. The following papers describe the *in vivo* evidence for possible biosynthetic steps or pathways:

- Akhila *et al.*[61] studied the *A. annua* biosynthesis starting from isotopically labeled mevalonate. They suggested the following pathway: farnesyl pyrophosphate (FPP) → germacrene skeleton → dihydrocostunolide → cadinanolide → arteannuin B → artemisinin.
- Sangwan *et al.*[62] reported the transformation of artemisinic acid to arteannuin B and artemisinin both *in vivo* and in a cell free system.
- Nair *et al.*[63] reported the conversion of arteannuin B to artemisinin by cell-free leaf homogenate of *A. annua*.
- Wallaart *et al.*[58, 64] isolated dihydroartemisinic acid (DHAA) and DHAA hydroperoxide in *A. annua*. DHAA can be chemically converted to artemisinin by photooxidation[35, 65] under conditions that may also be present in the living plant. DHAA hydroperoxide, which can very easily oxidize to artemisinin, is known as an intermediate of the photochemical oxidation of DHAA leading to artemisinin. The presence of DHAA and DHAA hydroperoxide in the plant and the conditions under which DHAA can be converted into DHAA hydroperoxide, provide evidence for a nonenzymatic, photochemical conversion of DHAA into artemisinin in *A. annua*.

### 2.3.1.2. Genes influencing biosynthesis in *A. annua*

The genes related to artemisinin biosynthesis are reviewed by Weathers *et al.*[50] and are summarized in Table 3.

| Enzyme | Function | GenBank | Ref. |
|---|---|---|---|
| DXP synthase | See figure 2 | AF182286 | [66, 67] |
| DXP reductoisomerase | See figure 2 | AF182287 | [66, 67] |
| HMG-CoA reductase | See figure 2 | AF142473 | [66, 67] |
| FPP synthase | See figure 2 | AF112881 | [26, 66, 67] |
| Sesquiterpene cyclases (SQC): | See figure 2 | | |
| *Epicedrol synthase* | See figure 2 | AJ001539 | [68, 69] |
| *ADS* | See figure 2 | AJ251751 | [55, 69] |
| *B-caryophyllene synthase* | See figure 2 | AF472361 | [70] |
| *B-farnesene synthase* | See figure 2 | AY835398 | [71] |
| *Putative SQC casc125* | See figure 2 | AJ271792 | [72] |
| *Putative SQC casc34* | See figure 2 | AJ271793 | [72] |
| *Putative SQC ses* | See figure 2 | AAD3983 | [73] |
| Squalene synthase | See figure 2 | AY445506 | [74] |
| Squalene synthase fragment | See figure 2 | AF182286 | [67] |
| CYP71AV1 | See figure 2 | DQ31567 | [1] |
| Peroxidase 1 | Stimulates: artemisinic acid→artemisinin | AY208699 | [75] |
| B-pinene synthase | GPP to B-pinene (monoterpene) | AF276072 | [76] |
| (3R)-linalool synthase | GPP to (3R)-linalool (terpene alcohol) | AF154125 | [77] |
| Isopentenyl transferase | Biosynthesis cytokinin phytohormones | M91610 | [78] |

Table 2: Genes related to artemisinin biosynthesis in *A. annua* L.[50]

### 2.3.1.3. Phytohormones

Artemisinin production can be influenced by exposure to exogenous phytohormones. The cytokinin 2-isopentenyladenine increased artemisinin levels in transformed roots of *A. annua*[79]. Artemisinin content is also increased in gibberellic acid (a phytohormone that can induce flowering) treated shoot cultures, and flowering plants[50].

One of the responses of plants to cytokinins is stimulation of shoot growth. Artemisinin is produced in shoots, so it was reasonable to measure the effect of cytokinins on artemisinin production in *A. annua*. Geng *et al.*[78] transferred the *ipt* gene (considered to be an enzyme at a rate-limiting step in the cytokinin biosynthesis[80]) into *A. annua*. Results showed that

two cytokinins were elevated 2- to 3-fold and artemisinin increased 30-70% compared with the control.

We used the quantitation method described in chapter 3 to investigate which of the 4 plant hormones tested (see below) triggered the highest artemisinin production in *A. annua* plants of only a few weeks old. Table 2 shows the quantities of artemisinin and some its structural analogs (arteannuin B, artemisitene and artemisinic acid), present in plants treated with 6-benzyl aminopurine (BAP), jasmonic acid (JA), gibberellic acid (GA3) and thidiazuron (TDZ). For each plant hormone, 3 ways of administration (local administration, administration by spraying the whole plant and administration by pouring the hormone into the soil of the plant) and a control were analyzed. The only plant hormone that clearly triggered artemisinin production was jasmonic acid.

| **Quantities** (µg/g) | **arteannuin B** | **Artemisitene** | **artemisinin** | **Artemisinic** |
|---|---|---|---|---|
| BAP local administration | 63 | < LLOQ | 16 | 81 |
| BAP total administration | 24 | < LLOQ | < LLOQ | 116 |
| BAP soil administration | 40 | < LLOQ | 15 | 260 |
| BAP control | 88 | < LLOQ | 33 | 891 |
| JA local administration | 180 | < LLOQ | 85 | 711 |
| JA total administration | 111 | < LLOQ | 31 | 179 |
| JA soil administration | 228 | < LLOQ | 107 | 758 |
| JA control | 45 | < LLOQ | < LLOQ | 624 |
| GA3 local administration | 28 | < LLOQ | 12 | 38 |
| GA3 total administration | 12 | < LLOQ | < LLOQ | 58 |
| GA3 soil administration | 48 | < LLOQ | 15 | 639 |
| GA3 control | 49 | < LLOQ | 14 | 645 |
| TDZ local administration | 18 | < LLOQ | < LLOQ | 77 |
| TDZ total administration | 53 | < LLOQ | 16 | 59 |
| TDZ soil administration | 77 | < LLOQ | 27 | 59 |
| TDZ control | 30 | < LLOQ | < LLOQ | 712 |

Table 3: Quantities of artemisinin and some of its structural analogs, present in *A. annua* plants after treatment with BAP, JA, GA3 and TDZ. The results are shown as the amount (µg) of analyte present in one gram of fresh plant material.

#### 2.3.1.4. Flowering

Several reports state that the highest artemisinin content occurs just before or at full flowering[50]. Wang *et al.*[81] investigated whether such a linkage does indeed exist by altering the plant flowering time through transgenics. They transferred the flowering promoter factor (fpf1) from *A. thaliana* into *A. annua* via *A. tumefaciens*. No differences in artemisinin content were detected between the flowering transgenic plants and the non-flowering non-transgenic plants. This work suggest that flowering is not a necessary factor for increased artemisinin content in *A. annua*.

#### 2.3.1.5. Metabolic engineering of artemisinin

Although several highly efficient transformation and regeneration systems for *A. annua* have been published (reviewed be Liu *et al.*[51]), only a few interesting transgenic *A. annua* varieties have been created.

Chen *et al.*[26] used an Agrobacterium tumefaciens-mediated transformation system to transfer cDNA FPPS (placed under a CaMV 35S promoter) into *A. annua*. The artemisinin content in the transgenic plants regenerated from the shoot lines was about 2-3 times higher than in the control. Han *et al.*[82] did a similar experiment.

Martin *et al.*[83] introduced the ADS gene into *E. coli* together with the mevalonate isoprenoid pathway from *Saccharomyces cerevisiae*, to construct a bacterial system producing the artemisinin precursor amorpha-4,11-diene from acetyl-CoA. The transformed *E. coli* was reported to produce around 100 mg of amorpha-4,11-diene per liter of culture in 12h. These *E. coli* strains can serve as platform hosts for the production of essentially any terpenoid for which the genes are known.

Lindahl *et al.*[84] transformed *Saccharomyces cerevisiae* with the gene encoding for ADS. The plasmid and genome-transformed yeasts produced 600 and 100 µg/l of the artemisinin precursor amorpha-4,11-diene, respectively, during a 16-days' batch cultivation.

Ro *et al.*[41] report the engineering of *Saccharomyces cerevisiae* to produce high titers (up to 100 mg/l) of artemisinic acid using an engineered mevalonate pathway, ADS, and a novel cytochrome P450 monooxygenase (CYP71AV1)[1]. The synthesized artemisinic acid is

transported out and retained on the outside of the engineered yeast, meaning that a simple and inexpensive purification process can be used to obtain the desired product.

### 2.3.2. Gene discovery: cDNA AFLP

Quantitative cDNA AFLP analysis is a genome-wide mRNA expression analysis. It is an invaluable tool for identifying and analyzing genes involved in, or controlling various biological responses. Although DNA microarrays provide a convenient tool for genome-wide expression analysis, their use is limited to organisms for which the complete genome sequence or a large cDNA collection is available. For other organisms, including most plant species, DNA fragment analysis based methods, such as cDNA-AFLP, provide a more appropriate tool for genome-wide expression analysis. A review by Breyne and Zabeau[85] discusses and compares the merits and drawbacks of the major technologies for genome-wide expression analysis.

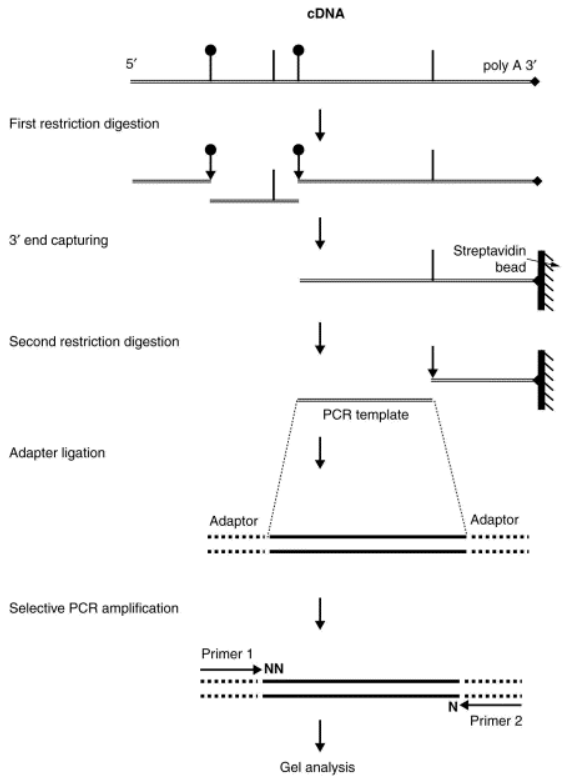


Figure 7: Outline of the cDNA-AFLP procedure[85]

- Messenger RNA (mRNA) is converted into double stranded cDNA using a biotinylated oligo-dT primer.
- After digestion with a first restriction enzyme, the 3′ termini of the cDNA are recovered by binding to streptavidin beads.
- Following digestion with a second enzyme, restriction fragments are released that serve as polymerase chain reaction (PCR) templates.
- After ligation of site-specific adapters, primers that match the adapter sequences and that carry selective nucleotides at their 3′ end are used to amplify subsets of the transcript fragments. Following amplification, fragments of between 50 and 500 bases are separated and displayed on polyacrylamide gels.
- After the fragments have been excised out of the gel, they can be sequenced and compared with genomic databases or EST libraries. The sequence tags can also be used to *de novo* sequence the complete mRNA.

The principal advantage of cDNA-AFLP compared to microarrays are that it allows genome-wide expression analysis in any species without prior sequence knowledge and that both known and unknown genes can be analyzed. Gel analysis of the amplified fragments (see figure 5) reveals which fragments are differentially expressed. These fragments can be sequenced and the resulting sequence tags can be compared to existing genomic databases and EST libraries. If no mach is found, the sequence tag can be used to clone the unknown gene.

For studying biological processes in plant species, cDNA AFLP presents another advantage compared to microarrays. The high levels of redundancy in plant genomes will likely remain a major obstacle for detailed microarray studies. Fragment-based technologies, such as cDNA-AFLP, overcome the problem caused by redundancy, and will continue to be the best alternative for performing in-depth analyses of gene expression in plant species.

### 2.3.3. Gene discovery: Proteomics

Proteomics stands for the large-scale characterization of the entire protein complement of a cell, tissue or organism. Proteomics studies can be subdivided in three main areas of interest: An expression proteomics analysis is designed to identify all the protein species present in a proteome of a cell, tissue or organism at a certain time. In structural proteomics the focus is on identifying the molecular structure. Functional proteomics describes the changes in protein abundance and modifications during a biological process.

New transcriptomics technologies permit simultaneous examination of thousands of transcripts. However, the complex regulatory routes, from post-translational modifications to protein turnover cannot be studied at the cDNA level. The proteome approach is necessary to help answer questions of functional analysis. Several reviews address the application of proteomics to plant biology[86-98]. Some of them address the technical aspects of plant proteomics[99-101].

#### 2.3.3.1. Technology

There are 2 main proteomics technologies: gel-based and gel-free proteomics. Two dimensional gel electrophoresis (2-DE) for the separation and visualization of proteins has been the standard proteomics technique during the past decade. Over the last years, several gel-free proteomics techniques have been developed to either complement 2-DE or to entirely

replace the gel based techniques. A review by Baggerman *et al.*[102] summarizes the most important gel-free and gel-based proteomics techniques and compares their advantages and drawbacks. A review by Tilleman *et al.*[103] discusses the basics of the current proteomics technologies. This chapter offers background information on the gel based technique used in chapter 3.

2-DE is a high resolution method for separating proteins in two dimensions; according to their isoelectric point (pI) in the first dimension and according to their size (molecular weight) in the second dimension. Exploring differences in spot intensity or tracking protein spots that appear or disappear on gels derived from experimental and control conditions is the main goal of functional proteomics.

**Isoelectric focusing:** After the proteins have been extracted from the tissue sample, they are brought into a small gel strip that contains an immobilized pH gradient (IPG). When an electric field is applied over this IPG strip, the proteins migrate along the pH gradient in the strip until they reach their pI. The strip is then applied onto a polyacrylamide gel and the proteins are subsequently size-separated by electrophoresis.

**2-D gel visualization:** In order to visualize the separated proteins, 2-D gels can be stained with a variety of different stains: Coomassie Blue (cheap, low sensitivity: 10ng per protein spot), silver stains (sensitive, 0.5-1ng per protein spot, very small linear range, time-consuming protocol, less compatible with mass spectrometry identification), radio-active labeling (very sensitive, hazardous and expensive) and the current state of the art fluorescent dyes like Sypro Ruby (sensitive, 1ng per protein spot, large linear range, extremely easy to use, compatible with MS identification)[104, 105]. To study post-translational modifications, dyes can be used that selectively stain glycoproteins (e.g. Pro-Q Emerald 488 glycoprotein stain), phosphoproteins and other modified proteins[105]. The 'difference gel electrophoresis' (DIGE) approach is gaining interest. This approach allows to fluorescently label (with cyanine dyes Cy2, Cy3, Cy4) as many as 3 different complex protein populations prior to mixing them together and running them on the same 2-D gel[105].

**Digestion and identification:** After protein spots have been picked from the gels, they are digested using a sequence-specific protease (e.g. trypsin). The resulting mixture of peptides is

desalted and analyzed by MS (producing a “peptide mass fingerprint”) or MS/MS (producing an amino acid sequence tag). MS and MS/MS analysis can complement each other.

A “peptide mass fingerprint” (PMF) is the result of the cleavage of a particular protein using a sequence-specific protease such as trypsin. The masses from the resulting peptides can be measured by MS. Usually matrix assisted laser desorption ionization (MALDI) is used to produce a “peptide mass fingerprint”. MALDI produces singly charged peptide ions leading to a less convoluted spectrogram than the multiple charged peptides ions resulting from electrospray ionization (ESI).

In MS/MS analysis, the double charged peptide ions are one-by-one selected out of a mixture of peptides ions by the first mass analyzer of the mass spectrometer. Between the first and the second analyzer, the peptides are fragmented by collision with an inert gas. During these energetic collisions of the selected peptide and the collision gas, bonds are broken along the peptide backbone. In most applications, this leads to so called b and y ions, which indicate fragmentation at the amide bond with charge retention on the N or C terminus, respectively. The masses of the resulting fragments are analyzed by the second mass analyzer. Each peptide fragment in the spectrum differs from its neighbor by only one amino acid. It is therefore possible to determine the amino acid sequence by considering the mass difference between the neighboring peaks in a series. Usually ESI is used to produce an peptide MS/MS spectrum. In contrast with MALDI, ESI produces peptides ions with 2 (or more) positive charges, making it possible to deduct the amino acid sequence both from the b and the y ions.

The experimental PMF or MS/MS spectrum is matched against a calculated PMF or MS/MS spectrum for all peptides in the database (e.g. Mascot, SEQUEST). A score is calculated for each match: the higher the score, the lower the probability that the match is a random event. To be significant, the score has to be higher than a calculated threshold. The threshold is the score resulting from the probability for which the expectancy for a match is 5%.

#### 2.3.3.2. Combining proteomic en genetic studies in plants

Reviews by Thiellement *et al.*[87, 106] focuses on the various, mainly genetic, applications of the proteomics in plants: characterization of individuals or lines, estimation of genetic variability within and between populations, establishment of genetic distances that can be used in phylogenetic studies, characterization of mutants and localization of the genes

encoding the revealed proteins. Main focus is on proteomics as an invaluable tool for deciphering the function and role of the genes that are or will be sequenced: Innumerable proteins have been described, whose relative abundance depends on the conditions (light, heat, cold, hormones, pathogens, etc.), and on the developmental stages or organs. As far as these proteins can be sequenced (e.g. partially by MS/MS), the corresponding genes that may exist in the databases as ESTs may thus be further characterized and tentative functions may be proposed.

Transcriptomics tools, such as cDNA microarrays and cDNA AFLP are usefully complemented by proteomics, since the amounts of a protein and of its mRNA are not necessarily correlated. When studying conditions with high and low artemisinin production, a protein that is overexpressed under one of the conditions, does not necessarily imply an overexpression of its mRNA. While a proteomics study will point to this protein, a transcriptomics study might not point to the corresponding mRNA.

# References


1. Teoh KH, Polichuk DR, Reed DW, Nowak G, Covello PS: **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* 2006, **580**(5):1411-1416.

2. Capanna E: **Grassi versus Ross: who solved the riddle of malaria?** *International Microbiology* 2006, **9**(1):69-74.

3. Greenwood B, Mutabingwa T: **Malaria in 2002**. *Nature* 2002, **415**(6872):670-672.

4. Snow RW, Craig M, Deichmann U, Marsh K: **Estimating mortality, morbidity and disability due to malaria among Africa's non-pregnant population**. *Bull World Health Organ* 1999, **77**(8):624-640.

5. Bozdech Z, Llinas M, Pulliam BL, Wong ED, Zhu JC, DeRisi JL: **The transcriptome of the intraerythrocytic developmental cycle of Plasmodium falciparum**. *Plos Biology* 2003, **1**(1):85-100.

6. Miller LH, Baruch DI, Marsh K, Doumbo OK: **The pathogenic basis of malaria**. *Nature* 2002, **415**(6872):673-679.

7. Miller LH, Good MF, Milon G: **Malaria Pathogenesis**. *Science* 1994, **264**(5167):1878-1883.

8. White NJ: **The treatment of malaria**. *New England Journal of Medicine* 1996, **335**(11):800-806.

9. Ridley RG: **Medical need, scientific opportunity and the drive for antimalarial drugs**. *Nature* 2002, **415**(6872):686-693.

10. Fryauff DJ, Tuti S, Mardi A, Masbar S, Patipelohi R, Leksana B, Kain KC, Bangs MJ, Richie TL, Baird JK: **Chloroquine-resistant Plasmodium vivax in transmigration settlements of West Kalimantan, Indonesia**. *American Journal of Tropical Medicine and Hygiene* 1998, **59**(4):513-518.

11. Mutabingwa T: **Chlorproguanil-dapsone for treatment of drug-resistant falciparum malaria in Tanzania (vol 358, pg 1218, 2001)**. *Lancet* 2001, **358**(9292):1556-1556.

12. White NJ: **Antimalarial drug resistance**. *Journal of Clinical Investigation* 2004, **113**(8):1084-1092.

13. Klayman DL: **Qinghaosu (artemisinin): an antimalarial drug from China**. *Science* 1985, **228**(4703):1049-1055.

14. Olliaro PL, Taylor WR: **Developing artemisinin based drug combinations for the treatment of drug resistant falciparum malaria: A review**. *J Postgrad Med* 2004, **50**(1):40-44.

15. Davis TM, Karunajeewa HA, Ilett KF: **Artemisinin-based combination therapies for uncomplicated malaria**. *Med J Aust* 2005, **182**(4):181-185.

16. Hallett RL, Sutherland CJ, Alexander N, Ord R, Jawara M, Drakeley CJ, Pinder M, Walraven G, Targett GAT, Alloueche A: **Combination therapy counteracts the enhanced transmission of drug-resistant malaria parasites to mosquitoes**. *Antimicrobial Agents and Chemotherapy* 2004, **48**(10):3940-3943.

17. Drakeley CJ, Jawara M, Targett GAT, Walraven G, Obisike U, Coleman R, Pinder M, Sutherland CJ: **Addition of artesunate to chloroquine for treatment of Plasmodium falciparum malaria in Gambian children causes a significant but short-lived reduction in infectiousness for mosquitoes**. *Tropical Medicine & International Health* 2004, **9**(1):53-61.

18. Targett G, Drakeley C, Jawara M, von Seidlein L, Coleman R, Deen J, Pinder M, Doherty T, Sutherland C, Walraven G *et al*: **Artesunate reduces but does not prevent posttreatment transmission of Plasmodium falciparum to Anopheles gambiae**. *Journal of Infectious Diseases* 2001, **183**(8):1254-1259.

19. Meshnick SR: **Artemisinin: mechanisms of action, resistance and toxicity**. *International Journal for Parasitology* 2002, **32**(13):1655-1660.

20. Ferreira J, Janick J: **Distribution of Artemisinin in *Artemisia annua***. *J Janick (ed), Progress in new crops* 1996:579-584.

21. Duke SO, Paul RN: **Development and Fine-Structure of the Glandular Trichomes of Artemisia-Annua L**. *International Journal of Plant Sciences* 1993, **154**(1):107-118.

22. Ferreira JFS, Janick J: **Floral Morphology of Artemisia-Annua with Special Reference to Trichomes**. *International Journal of Plant Sciences* 1995, **156**(6):807-815.

23. Woerdenbag HJ, Pras N, Chan NG, Bang BT, Bos R, Vanuden W, Y PV, Boi NV, Batterman S, Lugt CB: **Artemisinin, Related Sesquiterpenes, and Essential Oil in Artemisia-Annua during a Vegetation Period in Vietnam**. *Planta Medica* 1994, **60**(3):272-275.

24. Delabays N, Simonnet X, Gaudin M: **The genetics of artemisinin content in Artemisia annua L. and the breeding of high yielding cultivars**. *Curr Med Chem* 2001, **8**(15):1795-1801.

25. Bhakuni RS, Jain DC, Sharma RP, Kumar S: **Secondary metabolites of Artemisia annua and their biological activity**. *Current Science* 2001, **80**(1):35-48.

26. Chen D, Ye H, Li G: **Expression of a chimeric farnesyl diphosphate synthase gene in Artemisia annua L. transgenic plants via Agrobacterium tumefaciens-mediated transformation**. *Plant Science* 2000, **155**(2):179-185.

27. Duke M, Paul R, Elsohly H, Sturtz G, Duke S: **Localization of artemisinin and artemisitene in foliar tissues of glanded and glandless biotypes of *Artemisia annua* L.** *International Journal of Plant Sciences* 1994, **155**(3):365-372.

28. Bertea CM, Voster A, Verstappen FWA, Maffei M, Beekwilder J, Bouwmeester HJ: **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* 2006, **448**(1-2):3-12.

29. Bouwmeester HJ, Wallaart TE, Janssen MHA, van Loo B, Jansen BJM, Posthumus MA, Schmidt CO, De Kraker JW, Konig WA, Franssen MCR: **Amorpha-4,11-diene synthase catalyses the first probable step in artemisinin biosynthesis**. *Phytochemistry* 1999, **52**(5):843-854.

30. Zhang L, Ye HC, Li GF: **Effect of development stage on the artemisinin content and the sequence characterized amplified region (SCAR) marker of high-artemisinin yielding strains of Artemisia annua L.** *Journal of Integrative Plant Biology* 2006, **48**(9):1054-1062.

31. Klayman DL, Lin AJ, Acton N, Scovill JP, Hoch JM, Milhous WK, Theoharides AD: **Isolation of Artemisinin [Qinghaosu) from Artemisia-Annua Growing in the United-States**. *Journal of Natural Products* 1984, **47**(4):715-717.

32. Abdin MZ, Israr M, Rehman RU, Jain SK: **Artemisinin, a novel antimalarial drug: Biochemical and molecular approaches for enhanced production**. *Planta Medica* 2003, **69**(4):289-299.

33. Woerdenbag HJ, Pras N, Bos R, Visser JF, Hendriks H, Malingre TM: **Analysis of Artemisinin and Related Sesquiterpenoids from Artemisia-Annua L by Combined Gas-Chromatography Mass-Spectrometry**. *Phytochemical Analysis* 1991, **2**(5):215-219.

34. Vonwiller SC, Haynes RK, King G, Wang HJ: **An Improved Method for the Isolation of Qinghao (Artemisinic) Acid from Artemisia-Annua**. *Planta Medica* 1993, **59**(6):562-563.

35. Roth RJ, Acton N: **A Simple Conversion of Artemisinic Acid into Artemisinin**. *Journal of Natural Products* 1989, **52**(5):1183-1185.

36. Avery MA, Chong WKM, Jenningswhite C: **Stereoselective Total Synthesis of (+)-Artemisinin, the Antimalarial Constituent of Artemisia-Annua L**. *Journal of the American Chemical Society* 1992, **114**(3):974-979.

37. Ravindranathan T, Kumar MA, Menon RB, Hiremath SV: **Stereoselective Synthesis of Artemisinin**. *Tetrahedron Letters* 1990, **31**(5):755-758.

38. Yadav JS, Babu RS, Sabitha G: **Stereoselective total synthesis of (+)-artemisinin**. *Tetrahedron Letters* 2003, **44**(2):387-389.

39. Zhou WS, Xu XX: **Total Synthesis of the Antimalarial Sesquiterpene Peroxide Qinghaosu and Yingzhaosu-A**. *Accounts of Chemical Research* 1994, **27**(7):211-216.

40. Schmid G, Hofheinz W: **Total Synthesis of Qinghaosu**. *Journal of the American Chemical Society* 1983, **105**(3):624-625.

41. Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J *et al*: **Production of the antimalarial drug precursor artemisinic acid in engineered yeast**. *Nature* 2006, **440**(7086):940-943.

42. Lin AJ, Klayman DL, Hoch JM, Silverton JV, George CF: **Thermal Rearrangement and Decomposition Products of Artemisinin (Qinghaosu)**. *Journal of Organic Chemistry* 1985, **50**(23):4504-4508.

43. Meshnick SR, Taylor TE, Kamchonwongpaisan S: **Artemisinin and the antimalarial endoperoxides: From herbal remedy to targeted chemotherapy**. *Microbiological Reviews* 1996, **60**(2):301-&.

44. Price RN: **Artemisinin drugs: novel antimalarial agents**. *Expert Opinion on Investigational Drugs* 2000, **9**(8):1815-1827.

45. Price R, van Vugt M, Phaipun L, Luxemburger C, Simpson J, McGready R, ter Kuile F, Kham A, Chongsuphajaisiddhi T, White NJ *et al*: **Adverse effects in patients with acute falciparum malaria treated with artemisinin derivatives**. *American Journal of Tropical Medicine and Hygiene* 1999, **60**(4):547-555.

46. McGready R, Cho T, Cho JJ, Simpson JA, Luxemburger C, Dubowitz L, Looareesuwan S, White NJ, Nosten F: **Artemisinin derivatives in the treatment of falciparum malaria in pregnancy**. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1998, **92**(4):430-433.

47. McGready R, Cho T, Keo NK, Thwai KL, Villegas L, Looareesuwan S, White NJ, Nosten F: **Artemisinin antimalarials in pregnancy: A prospective treatment study of 539 episodes of multidrug-resistant Plasmodium falciparum**. *Clinical Infectious Diseases* 2001, **33**(12):2009-2016.

48. Eckstein-Ludwig U, Webb RJ, van Goethem IDA, East JM, Lee AG, Kimura M, O'Neill PM, Bray PG, Ward SA, Krishna S: **Artemisinins target the SERCA of Plasmodium falciparum**. *Nature* 2003, **424**(6951):957-961.

49. Jambou R, Legrand E, Niang M, Khim N, Lim P, Volney B, Ekala MT, Bouchier C, Esterre P, Fandeur T *et al*: **Resistance of Plasmodium falciparum field isolates to in-vitro artemether and point mutations of the SERCA-type PfATPase6**. *Lancet* 2005, **366**(9501):1960-1963.

50. Weathers PJ, Elkholy S, Wobbe KK: **Artemisinin: The biosynthetic pathway and its regulation in Artemisia annua, a terpenoid-rich species**. *In Vitro Cellular & Developmental Biology-Plant* 2006, **42**(4):309-317.

51. Liu CZ, Zhao Y, Wang YC: **Artemisinin: current state and perspectives for biotechnological production of an antimalarial drug**. *Applied Microbiology and Biotechnology* 2006, **72**(1):11-20.

52. Liu Y, Wang H, Ye HC, Li GF: **Advances in the plant isoprenoid biosynthesis pathway and its metabolic engineering**. *Journal of Integrative Plant Biology* 2005, **47**(7):769-782.

53. Bertea CM, Freije JR, van der Woude H, Verstappen FWA, Perk L, Marquez V, De Kraker JW, Posthumus MA, Jansen BJM, de Groot A *et al*: **Identification of intermediates and enzymes involved in the early steps of artemisinin biosynthesis in Artemisia annua**. *Planta Medica* 2005, **71**(1):40-47.

54. Matsushita Y, Kang W, Charlwood BV: **Cloning and analysis of a cDNA encoding farnesyl diphosphate synthase from Artemisia annua**. *Gene* 1996, **172**(2):207-209.

55. Chang YJ, Song SH, Park SH, Kim SU: **Amorpha-4,11-diene synthase of Artemisia annua: cDNA isolation and bacterial expression of a terpene synthase involved in artemisinin biosynthesis**. *Archives of Biochemistry and Biophysics* 2000, **383**(2):178-184.

56. Mercke P, Bengtsson M, Bouwmeester HJ, Posthumus MA, Brodelius PE: **Molecular cloning, expression, and characterization of amorpha-4,11-diene synthase, a key enzyme of artemisinin biosynthesis in Artemisia annua L.** *Archives of Biochemistry and Biophysics* 2000, **381**(2):173-180.

57. Wallaart TE, Bouwmeester HJ, Hille J, Poppinga L, Maijers NCA: **Amorpha-4,11-diene synthase: cloning and functional expression of a key enzyme in the biosynthetic pathway of the novel antimalarial drug artemisinin**. *Planta* 2001, **212**(3):460-465.

58. Wallaart TE, van Uden W, Lubberink HGM, Woerdenbag HJ, Pras N, Quax WJ: **Isolation and identification of dihydroartemisinic acid from Artemisia annua and its possible role in the biosynthesis of artemisinin**. *Journal of Natural Products* 1999, **62**(3):430-433.

59. Brown GD, Sy LK: **In vivo transformations of dihydroartemisinic acid in Artemisia annua plants**. *Tetrahedron* 2004, **60**(5):1139-1159.

60. Van Geldre E, Vergauwe A, VandenEeckhout E: **State of the art of the production of the antimalarial compound artemisinin in plants**. *Plant Molecular Biology* 1997, **33**(2):199-209.

61. Akhila A, Thakur RS, Popli SP: **Biosynthesis of Artemisinin in Artemisia-Annua**. *Phytochemistry* 1987, **26**(7):1927-1930.

62. Sangwan RS, Agarwal K, Luthra R, Thakur RS, Singhsangwan N: **Biotransformation of Arteannuic Acid into Arteannuin-B and Artemisinin in Artemisia-Annua**. *Phytochemistry* 1993, **34**(5):1301-1302.

63. Nair MSR, Basile DV: **Bioconversion of Arteannuin-B to Artemisinin**. *Journal of Natural Products* 1993, **56**(9):1559-1566.

64. Wallaart TE, Pras N, Quax WJ: **Isolation and identification of dihydroartemisinic acid hydroperoxide from Artemisia annua: A novel biosynthetic precursor of artemisinin**. *Journal of Natural Products* 1999, **62**(8):1160-1162.

65. Acton N, Roth RJ: **On the Conversion of Dihydroartemisinic Acid into Artemisinin**. *Journal of Organic Chemistry* 1992, **57**(13):3610-3614.

66. Souret FDF, Weathers PJ, Wobbe KK: **The mevalonate-independent pathway is expressed in transformed roots of artemisia annua and regulated by light and culture age**. *In Vitro Cellular & Developmental Biology-Plant* 2002, **38**(6):581-588.

67. Souret FF, Kim Y, Wysiouzil BE, Wobbe KK, Weathers PJ: **Scale-up of Artemisia annua L. hairy root cultures produces complex patterns of terpenoid gene expression**. *Biotechnology and Bioengineering* 2003, **83**(6):653-667.

68. Hua L, Matsuda SPT: **The molecular cloning of 8-epicedrol synthase from Artemisia annua**. *Archives of Biochemistry and Biophysics* 1999, **369**(2):208-212.

69. Mercke P, Crock J, Croteau R, Brodelius PE: **Cloning, expression, and characterization of epi-cedrol synthase, a sesquiterpene cyclase from Artemisia annua L**. *Archives of Biochemistry and Biophysics* 1999, **369**(2):213-222.

70. Cai Y, Jia JW, Crock J, Lin ZX, Chen XY, Croteau R: **A cDNA clone for beta-caryophyllene synthase from Artemisia annua**. *Phytochemistry* 2002, **61**(5):523-529.

71. Picaud S, Brodelius M, Brodelius PE: **Expression, purification and characterization of recombinant (E)-beta-farnesene synthase from Artemisia annua**. *Phytochemistry* 2005, **66**(9):961-967.

72. Van Geldre E, De Pauw I, Inze D, Van Montagu M, Van den Eeckhout E: **Cloning and molecular analysis of two new sesquiterpene cyclases from Artemisia annua L.** *Plant Science* 2000, **158**(1-2):163-171.

73. Liu Y, Ye HC, Li GF: **Cloning, E-coli expression and molecular analysis of a novel sesquiterpene synthase gene from Artemisia annua**. *Acta Botanica Sinica* 2002, **44**(12):1450-1455.

74. Liu Y, Ye HC, Wang H, Li GF: **Molecular cloning, Escherichia coli expression and genomic organization of squalene synthase gene from Artemisia annua**. *Acta Botanica Sinica* 2003, **45**(5):608-613.

75. Zhang YS, Liu BY, Li ZQ, Ye HC, Wang H, Li GF, Han JL: **Molecular cloning of a classical plant peiroxidase firom Artemisia annua and its effect on the biosynthesis of artemisinin in vitro**. *Acta Botanica Sinica* 2004, **46**(11):1338-1346.

76. Lu S, Xu R, Jia JW, Pang JH, Matsuda SPT, Chen XY: **Cloning and functional characterization of a beta-pinene synthase from Artemisia annua that shows a circadian pattern of expression**. *Plant Physiology* 2002, **130**(1):477-486.

77. Jia JW, Crock J, Lu S, Croteau R, Chen XY: **(3R)-linalool synthase from Artemisia annua L.: cDNA isolation, characterization, and wound induction**. *Archives of Biochemistry and Biophysics* 1999, **372**(1):143-149.

78. Geng S, Ma M, Ye HC, Liu BY, Li GF, Chong K: **Effects of ipt gene expression on the physiological and chemical characteristics of Artemisia annua L.** *Plant Science* 2001, **160**(4):691-698.

79. Weathers PJ, Bunk G, McCoy MC: **The effect of phytohormones on growth and artemisinin production in Artemisia annua hairy roots**. *In Vitro Cellular & Developmental Biology-Plant* 2005, **41**(1):47-53.

80. McKenzie MJ, Mett V, Reynolds PHS, Jameson PE: **Controlled cytokinin production in transgenic tobacco using a copper-inducible promoter**. *Plant Physiology* 1998, **116**(3):969-977.

81. Wang H, Ge L, Ye HC, Chong K, Liu BY, Li GF: **Studies on the effects of fpf1 gene on Artemisia annua flowering time and on the linkage between flowering and artemisinin biosynthesis**. *Planta Medica* 2004, **70**(4):347-352.

82. Han JL, Liu BY, Ye HC, Wang H, Li ZQ, Li GF: **Effects of overexpression of the endogenous farnesyl diphosphate synthase on the artemisinin content in Artemisia annua L.** *Journal of Integrative Plant Biology* 2006, **48**(4):482-487.

83. Martin VJJ, Pitera DJ, Withers ST, Newman JD, Keasling JD: **Engineering a mevalonate pathway in Escherichia coli for production of terpenoids**. *Nature Biotechnology* 2003, **21**(7):796-802.

84. Lindahl AL, Olsson ME, Mercke P, Tollbom O, Schelin J, Brodelius M, Brodelius PE: **Production of the artemisinin precursor amorpha-4,11-diene by engineered Saccharomyces cerevisiae**. *Biotechnology Letters* 2006, **28**(8):571-580.

85. Breyne P, Zabeau M: **Genome-wide expression analysis of plant cell cycle modulated genes**. *Current Opinion in Plant Biology* 2001, **4**(2):136-142.

86. van Wijk KJ: **Challenges and prospects of plant proteomics**. *Plant Physiology* 2001, **126**(2):501-508.

87. Thiellement H, Bahrman N, Damerval C, Plomion C, Rossignol M, Santoni V, de Vienne D, Zivy M: **Proteomics for genetic and physiological studies in plants**. *Electrophoresis* 1999, **20**(10):2013-2026.

88. Rampitsch C, Srinivasan M: **The application of proteomics to plant biology: a review**. *Canadian Journal of Botany-Revue Canadienne De Botanique* 2006, **84**(6):883-892.

89. Hochholdinger F, Sauer M, Dembinsky D, Hoecker N, Muthreich N, Saleem M, Liu Y: **Proteomic dissection of plant development**. *Proteomics* 2006, **6**(14):4076-4083.

90. Ephritikhine G, Ferro M, Rolland N: **Plant membrane proteomics**. *Plant Physiology and Biochemistry* 2004, **42**(12):943-962.

91. Lee SJ, Saravanan RS, Damasceno CMB, Yamane H, Kim BD, Rose JKC: **Digging deeper into the plant cell wall proteome**. *Plant Physiology and Biochemistry* 2004, **42**(12):979-988.

92. Millar AH, Heazlewood JL, Kristensen BK, Braun HP, Moller IM: **The plant mitochondrial proteome**. *Trends in Plant Science* 2005, **10**(1):36-43.

93. Agrawal GK, Yonekura M, Iwashashi Y, Iwahashi H, Rakwal R: **System, trends and perspectives of proteomics in dicot plants Part II: Proteomes of the complex developmental stages**. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* 2005, **815**(1-2):125-136.

94. Agrawal GK, Yonekura M, Iwahashi Y, Iwahashi H, Rakwal R: **System, trends and perspectives of proteomics in dicot plants Part III: Unraveling the proteomes influenced by the environment, and at the levels of function and genetic relationships**. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* 2005, **815**(1-2):137-145.

95. Reddy VS, Reddy ASN: **Proteomics of calcium-signaling components in plants**. *Phytochemistry* 2004, **65**(12):1745-1776.

96. Millar AH: **Location, location, location: surveying the intracellular real estate through proteomics in plants**. *Functional Plant Biology* 2004, **31**(6):563-571.

97. Bestel-Corre G, Dumas-Gaudot E, Gianinazzi S: **Proteomics as a tool to monitor plant-microbe endosymbioses in the rhizosphere**. *Mycorrhiza* 2004, **14**(1):1-10.

98. Park OK: **Proteomic studies in plants**. *Journal of Biochemistry and Molecular Biology* 2004, **37**(1):133-138.

99. Rose JKC, Bashir S, Giovannoni JJ, Jahn MM, Saravanan RS: **Tackling the plant proteome: practical approaches, hurdles and experimental tools**. *Plant Journal* 2004, **39**(5):715-733.

100. Newton RP, Brenton AG, Smith CJ, Dudley E: **Plant proteome analysis by mass spectrometry: principles, problems, pitfalls and recent developments**. *Phytochemistry* 2004, **65**(11):1449-1485.

101. Hirano H, Islam N, Kawasaki H: **Technical aspects of functional proteomics in plants**. *Phytochemistry* 2004, **65**(11):1487-1498.

102. Baggerman G, Vierstraete E, De Loof A, Schoofs L: **Gel-based versus gel-free proteomics: A review**. *Combinatorial Chemistry & High Throughput Screening* 2005, **8**(8):669-677.

103. Tilleman K, Deforce D, Elewaut D: **Rheumatology: a close encounter with proteomics**. *Rheumatology* 2005, **44**(10):1217-1226.

104. Patton WF: **A thousand points of light: The application of fluorescence detection technologies to two-dimensional gel electrophoresis and proteomics**. *Electrophoresis* 2000, **21**(6):1123-1144.

105. Patton WF: **Detection technologies in proteome analysis**. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* 2002, **771**(1-2):3-31.

106. Thiellement H, Zivy M, Plomion C: **Combining proteomic and genetic studies in plants**. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* 2002, **782**(1-2):137-149.

# 3. A proteome analysis of *Artemisia annua* L.

Filip Van Nieuwerburgh[1], Sofie Vande Casteele[1], Lies Maes[2], Alain Goosens[2], Pat Covello[3], Dirk Inzé[2] and Dieter Deforce[1]

3) Laboratory for Pharmaceutical Biotechnology, Ghent University, Harelbekestraat 72, B-9000 Ghent, Belgium

4) Department of Plant Systems Biology, Flanders Interuniversity Institute for Biotechnology, Ghent University, Technologiepark 927, B-9052 Ghent, Belgium

5) Plant Biotechnology Institute, 110 Gymnasium Place, Saskatoon, SK S2N 0W9 Canada

**ABSTRACT.**

The proteome of *A. annua* was investigated to discover candidate genes related to artemisinin production. We identified proteins that were differentially expressed between trichomes and whole leaf tissue. We also identified differentially expressed proteins between lower and upper leafs of *A. annua* plants. In a third proteome analysis, we explored the proteins present in a one-minute chloroform extract of *A. annua* plants. Trichomes are considered to be the artemisinin factories of *A. annua*. Upper leafs produce more artemisinin than the lower leafs. Proteins that are upregulated in these samples might thus be involved in artemisinin production. With the chloroform extraction, the content of the capitate glands of the trichomes is extracted. These glands possibly contain proteins involved in artemisinin production.

Proteins of interest were identified by tandem mass spectrometry. The fragmentation spectra were searched against the public Mass Spectrometry protein sequence DataBase (MSDB) and against three Expressed Sequence Tag (EST) libraries (one from flower buds of *A. annua,* one from the trichomes on the flower buds and one being the result of a cDNA subtraction of both plant tissues). To narrow down the results to the most valid gene candidates, the proteome data was compared with a cDNA Amplified Fragment Length Polymorphism (AFLP) analysis that investigated samples of *A. annua* leafs, taken at different points during a 72h time period after exposure to jasmonic acid (JA). We were able to compile a list of EST candidates, which could be useful for further investigation.

## 3.1. Introduction

Multi-drug resistance of the *Plasmodium* strains to the cheapest and most widely used antimalarials such as chloroquine, mefloquine and sulfadoxine-pyrimethamine is one of the biggest challenges in the fight against malaria[1]. The World Health Organization (WHO) recommends that all countries experiencing resistance to conventional monotherapies should use combination therapies, preferably those containing artemisinin derivatives (ACTs – artemisinin-based combination therapies)[2, 3]. As artemisinin cannot be synthesized chemically in an economically feasible way, *A. annua* is the only practical source of this valuable drug[4]. Unfortunately, *A. annua* contains only very small amounts of artemisinin ranging from 0.001 to 1.54% of dry weight[5]. Several research programs have been set up trying to increase the artemisinin content in *A. annua* by optimizing the growing and harvesting conditions, by selecting high yielding cultivars or by creation of transgenic plants[4, 6].

The objective of this study is to discover genes of the plant *A. annua* L. that are involved in the production of artemisinin. These genes do not necessarily have to code for the enzymes involved in the biosynthetic pathway, but can also be involved in the mechanisms influencing the amount of produced artemisinin (e.g. genes involved in the formation of trichomes). Three strategies were followed to accomplish this challenge: a proteome analysis, quantitative cDNA Amplified Fragment Length Polymorphism analysis (cDNA AFLP) and the construction of three full length Expressed Sequence Tag (EST) cDNA libraries. The results of these 3 strategies can be compared and complemented with each other.

- We investigated the proteome of *A. annua* by identifying proteins that were differentially expressed between trichomes and whole leaf tissue. We also identified differentially expressed proteins between lower leafs and upper leafs of *A. annua* plants. In a third proteome analysis, we explored the proteins present in a one-minute chloroform extract of *A. annua* plants. Trichomes are considered to be the artemisinin factories of *A. annua*[7]. The upper leafs produce more artemisinin than the lower leafs[8]. The proteins that are upregulated in these samples might thus be involved in artemisinin production. With the chloroform extraction, the content of the capitate glands of the trichomes is extracted[9]. These glands possibly contain proteins involved in artemisinin production.

Proteins of interest were identified by tandem mass spectrometry. The fragmentation spectra were searched against the public MSDB database and against the three EST libraries. Because only a few genes and proteins of *A. annua* have been characterized, identification of proteins by searching public protein databases, is only possible due to homology with known proteins of other plants.

- Quantitative cDNA AFLP analysis is a genome-wide messenger RNA (mRNA) expression analysis. The level of mRNA expression was compared between samples of *A. annua* leafs, taken at different time points during a 72h time period after exposure to JA. Transcripts that are overexpressed by JA are possibly involved in the production of artemisinin, because the production of artemisinin is stimulated by JA.
This analysis was conducted by the department of Plant Systems Biology of the "Vlaams interuniversitair instituut voor Biotechnologie" (VIB), Ghent University.

- Construction of three full length EST cDNA libraries: one from poly-A RNA from flower buds of *A. annua* and one from poly-A RNA from the trichomes on the flower buds. A subtracted cDNA library using poly-A RNA from the flower buds and the trichomes was also constructed. ESTs that are present in the trichomes and not in the flower buds are possibly involved in the production of artemisinin because trichomes are considered to be the artemisinin factories of *A. annua*[7]. Searching for homologies between the EST sequences and sequences of known plant genes revealed a cDNA clone encoding a cytochrome P450 enzyme. This enzyme (CYP71AV1) was expressed and characterized in *Saccharomyces cerevisiae* and was found to catalyze three steps: oxidation of amorpha-4,11-diene to artemisinic alcohol, artemisinic alcohol to artemisinic aldehyde and artemisinic aldehyde to artemisinic acid. The cDNA library construction was performed by the Plant Biotechnology Institute, Saskatoon, Canada and is described in a paper by Teoh *et al*.[1].

This paper focuses on the abovementioned proteome studies. To narrow down the results to the most valid gene candidates, we compared the proteome data with the cDNA AFLP analysis. Doing so, we were able to compile a list of EST candidates, which could be useful for further investigation.

## 3.2. Experimental

### 3.2.1. Chemicals

Following chemicals of analytical grade were used in the course of the proteome analysis: Acetonitrile MS grade (Biosolve, Valkenswaard, The Netherlands), Ammonium Bicarbonate ($NH_4HCO_3$) (Sigma-Aldrich, Bornem, Belgium), 3-[(3-Cholamidopropyl)-dimethyl-ammonio]-1-propane sulfonate (CHAPS) (MP Biomedicals, Illkirch, France), DL-Dithiothreitol electrophoresis grade (MP Biomedicals, Illkirch, France), Chloroform (Arcos, Geel, Belgium), Formic acid (FA) MS grade (Biosolve, Valkenswaard, The Netherlands), Glycerol (MP Biomedicals, Illkirch, France), Hydrochloric acid 37% (HCl) (Sigma-Aldrich, Bornem, Belgium), Iodoacetamide (MP Biomedicals, Illkirch, France), Potassium chloride (KCl) (Sigma-Aldrich, Bornem, Belgium), Sodium Dodecyl Sulfate (SDS) (MP Biomedicals, Illkirch, France), Sucrose (Sigma-Aldrich, Bornem, Belgium), Thiourea (Arcos, Geel, Belgium), 2-Amino-2-(hydroxymethyl)-1,3-propandiol (TRIS) (Sigma-Aldrich, Bornem, Belgium). Following for synthesis grade chemicals were used: 2-mercaptoethanol, Acetic acid and Methanol (Merck, Whitehouse Station, USA). Purified water of 18.2 MΩ/cm was obtained from a Milli-Q system from Millipore.

### 3.2.2. *A. annua* samples for differential proteome analysis

Twelve 2-DE gels were compared in two individual statistical analyses. Table 1 gives an overview of the analyzed samples and of the groups that were compared in the statistical analyses.

To prepare the samples, plants were grown under controlled conditions (21 °C; 12h day/12h night regime). Seeds were kindly provided by the National Botanic Garden of Belgium (Meise, Belgium). Trichome samples were prepared from one-month-old plants as described by Gershenzon *et al.*[10, 11] and McConkey *et al.*[12]. Leaf samples were harvested from the same plants.

| Gel Number | Characteristics | Number of spots on gel | Group Trichomes *vs.* | Group Upper *vs.* | Remarks |
|---|---|---|---|---|---|
| 0 | Trichomes taken from leafs only, after flowering | N/A | | | Sample too small |
| 1 | Trichomes taken from leafs only, before flowering | 526 | Group 1 | | |
| 2 | Trichomes taken from whole plant, after flowering | 573 | Group 1 | | |
| 3 | Trichomes taken from whole plant, before flowering | 652 | Group 1 | | |
| 4 | Trichomes taken from leafs only, before flowering – sample 2 | 430 | Group 1 | | |
| 5 | Trichomes taken from whole plant, before flowering – sample 2 | 572 | Group 1 | | |
| 6 | Trichomes taken from whole plant, after flowering – sample 2 | 577 | Group 1 | | |
| 7 | Upper leafs, after flowering | 573 | Group 2 | Group 1 | |
| 8 | Upper leafs, during flowering | 615 | Group 2 | Group 1 | |
| 9 | Upper leafs, before flowering | 561 | Group 2 | Group 1 | |
| 10 | Lower leafs, after flowering | 643 | Group 2 | Group 2 | |
| 11 | Lower leafs, during flowering | 575 | Group 2 | Group 2 | |
| 12 | Lower leafs, before flowering | 459 | Group 2 | Group 2 | |

Table 1: Overview of the 2-DE gels compared during statistical analysis

### 3.2.3. *A. annua* samples for proteome analysis of chloroform extract

To investigate the proteins present in a one-minute chloroform extract, 3 2-DE gels were prepared from chloroform extracts from three individual one-month-old plants. The plants were grown under controlled conditions (21 °C; 12h day/12h night regime). For each sample, an entire plant was used.

### 3.2.4. Protein extraction

The protein content of the 6 trichome samples and the 6 whole leaf samples was extracted as described by Bauw and Van Montagu[13]. For each sample, the plant material was grinded to powder in liquid nitrogen. Hundred fifty mg of the powder was transferred to an Eppendorf cup and homogenized during 1 minute using an Ultra-Turrax T8 homogenizer (Ika-Werke, Stanfer, Germany) in protein extraction buffer. The protein extraction buffer consisted of 0.7M sucrose, 0.5M TRIS, 30mM HCl, 0.1M KCl and 1% 2-mercaptoethanol. Immediately before use, a cocktail of phosphatase inhibitors (cocktail 1 and cocktail 2, Sigma-Aldrich, Bornem, Belgium), endonuclease (Sigma-Aldrich, Bornem, Belgium) and protease inhibitors (Complete Mini tablets, Roche diagnostics, Vilvoorde, Belgium) were added to the protein extraction buffer.

For the proteome analysis of a one-minute chloroform extract, an entire plant was cut into large pieces and transferred into a 250 ml flask. The plant material was shaken for one minute in 100 ml of chloroform, after which the chloroform was transferred to another flask. The chloroform was evaporated to dryness under a nitrogen flow, while keeping the temperature at 1 °C.

After all the chloroform was evaporated, the proteins in the residue were extracted with 10 ml ReadyPrep Sequential Extraction reagent 3 (Biorad, Hercules, USA). Before use, 0.7M sucrose, a cocktail of phosphatase inhibitors (cocktail 1 and cocktail 2, (Sigma-Aldrich, Bornem, Belgium), endonuclease (Sigma-Aldrich, Bornem, Belgium) and protease inhibitors (Complete Mini tablets, Roche diagnostics, Vilvoorde, Belgium) were added to the protein extraction buffer. The subsequent extraction steps are described by Bauw and Van Montagu[13].

### 3.2.5. Two-dimensional gel electrophoresis

2-DE was performed according to Görg A. *et al.*[14] with minor adjustments. For each sample, the complete yield of the protein extraction was dissolved in 360µl of rehydration buffer solution containing 7M urea, 2M thiourea, 4% CHAPS, 100mM DTT and 2% Carrier Ampholyte solution (Ampholene pH3.5-10, Amersham Biosciences, Uppsala, Sweden). This solution was used to rehydrate an immobilized pH gradient strip (IPG) with a linear pH gradient from pH 3 to pH 10 (Readystrip, Biorad, Hercules, USA). After the overnight in-gel rehydration, the proteins in the strip are focused at their isoelectric point on the Protean isoelectric focusing system (Biorad, Hercules, USA) at 18°C, starting at 100V for 6h, followed by linear ramping to 250V over 2h, linear ramping to 500V over 2h, linear ramping to 1 kV over 2h, linear ramping to 3kV over 2h, rapid ramping to 10 kV in 3.5h and steady state at 10 kV for 65 kVh. After the isoelectric focusing, the strips were equilibrated for 2x15 minutes. During the first 15 minutes, the strips were gently shaken in equilibration buffer (50mM TRIS - HCl buffer pH 8.8, 6M urea, 30% v/w glycerol, 2% SDS) with addition of 1.5% v/w DTT. During the second 15 minutes, the strips are gently shaken in equilibration buffer with addition of 4% v/w iodoacetamide. After equilibration, the IPG strips were placed on a polyacrylamide gel (10% T, 3.3% C). These gels were subjected to electrophoresis in a vertical Protean II Multi Cell electrophoresis system (Biorad, Hercules, USA) at 16mA/gel for 30 minutes and 32mA/gel for approximately 6 hours at 10°C. After electrophoresis, the gels were fixed for 30 minutes in a 10% methanol – 7% acetic acid solution and stained overnight with Sypro Ruby (Biorad, Hercules, USA).

### 3.2.6. Gel scanning and image analysis

2-DE gels stained with Sypro Ruby were scanned and digitized with the QuantityOne 4.4.1. software on a Versadoc Imaging system (Biorad, Hercules, USA). For each gel, exposure time was adjusted until 1 spot became oversaturated.

### 3.2.7. Data analysis and statistical analysis

The images of the 12 gels (see table 1) were analyzed with PDQuest 2D-analysis Software v7.1 (Biorad, Hercules, USA). All gels were matched to each other creating a match set standard image containing the match information of all spots on all gels. In order to compare protein expression levels across the 12 gels, the intensity of the spots on each gel was normalized to the total intensity of all spots in each gel image. The PDQuest software was used to perform a Mann-Whitney Signed-Rank test (trichomes group *versus* leafs group) and a Student's T-test (upper leafs group *versus* lower leafs group) on the log transformed intensity data of the spots. Spots of which the intensities were statistically higher ($p < 0.05$) in the trichomes group or the upper leafs group (see table 1), were selected for further identification by mass spectrometry.

For the expression proteomics analysis of a one-minute chloroform extract, no statistical analysis was done. Instead, as much different spots as possible were excised from the three different gels. In this experiment we tried to identify as much proteins as possible that are present in the chloroform extract.

### 3.2.8. Protein identification by mass spectrometry

Spots of interest were excised from the 2-DE gels obtained from samples 4 to 6 (see table 1). The excised spots were washed twice with 100µl 50% acetonitrile / 25mM $NH_4HCO_3$, followed by a reduction with 10mM DDT / 25mM $NH_4HCO_3$, an alkylation with 100mM iodoacetamide / 25mM $NH_4HCO_3$, and a final wash step with 100µl 50% acetonitrile / 25mM $NH_4HCO_3$. After the wash buffer was removed, the gel pieces were dehydrated with 100% acetonitrile and rehydrated in 20µl 25mM $NH_4HCO_3$ containing 10ng/µl sequence grade modified trypsin (Promega, Madison, USA) for 30 minutes on ice. In-gel digestion with trypsin was continued over night at 37°C. Peptides were extracted with 100µl 50% acetonitrile / 50% $H_2O$ for 30 minutes, followed by extraction with 100µl 100% acetonitrile and pooled. The extracts were completely dried under vacuum and dissolved in 23µl 0.1% TFA.

### 3.2.9. Identification by ESI mass spectrometry

Using column switching, 20µl of the peptide solution was injected on a nano-LC system (Dionex, Sunnyvale, USA) coupled to a Q-Tof Ultima mass spectrometer (Waters, Milford, USA) fitted with an electrospray ionization source. The sample was loaded and desalted on an Atlantis dC18 Trap column (Waters, Milford, USA) at a flow-rate of 10µl/min and eluted on a C18 Pepmap100 (75 µm i.d. × 15 cm, 3 µm, Dionex, Sunnyvale, USA) at a flow-rate of 150 nl/min. The data was acquired with the Masslynx software 4.1. from Waters operating in the "automatic function switching" mode.

Fragmentation spectra, resulting from tandem mass spectrometry on the detected peptides, were searched against the MSDB database using the Mascot search engine (http://www.matrixscience.com). Using the ProteinLynx software from Waters, the spectra were also searched against the 3 EST libraries described by Teoh *et al.*[1].

### 3.2.10. Comparison of the proteomics and the AFLP data

We compared the differential proteins resulting from the proteome analysis with the differentially expressed sequences from the cDNA AFLP analysis, allowing us to point out cDNA EST clones of which their possible involvement in artemisinin biosynthesis is supported by both analyses. To be able to do this comparison, we compiled a database containing all nucleic acid sequences pointed out by the proteome analysis. This database contains the sequences of the genes of the proteins that were annotated to the spots by the mascot search engine, the EST sequences that were annotated to the spots by the ProteinLynx software and the complete coding sequence of the gene annotated to these EST sequences by the NCBI BLAST algorithm.

All 830 differential AFLP sequences were blasted against this database (NCBI BLASTN algorithm 2.2.15). Blasts with an E-value below 0.0001 were considered informative. Blasts with an E-value between 0.0001 and 0.001, were taken into consideration if multiple AFLP sequences were linked to the same protein spot.

## 3.3. Results and discussion

### 3.3.1. Protein extraction

Proteome analysis of plants tissues poses some challenges because plant tissues have a low protein content and have high levels of components interfering with 2-DE analysis (lipids, sugars, organic acids, etc.). Like other essential oil producing plants, *A. annua* tissues (especially trichomes) are characterized by a high lipid content.

Three protein extraction protocols were tested in an attempt to optimize the protein extraction:

- The ReadyPrep Sequential Extraction kit (Biorad, Hercules, USA).
- A TRIS-HCl extraction proven to be the best of 4 methods tested by Audrius *et al.*[15] for protein extraction from *Citrus* leafs.
- Extraction protocol described by Bauw and Van Montagu[13] including a phenol extraction.

The ReadyPrep Sequential Extraction kit and the TRIS-HCl extraction proved to be unsatisfactory. In search for a method which is more suitable to extract proteins from samples with a high lipid content, the extraction protocol by Bauw and Van Montagu[13] was evaluated. This method has been used by Tilleman *et al.* for protein extraction from brain tissues[16, 17], which are also characterized by a high lipid content. For the extraction of trichomes and leaf tissue, this method yielded between 459 and 643 of analyzable spots per 2-DE gel (see table 1).

An adaptation of the method by Bauw and Van Montagu[13] was used to extract proteins from the waxy and oily residue after evaporation of a one-minute chloroform extract. The protein extraction buffer had to be replaced by the ReadyPrep Sequential Extraction reagent 3. This reagent is optimized for the extraction of lipophilic proteins. Unlike the protein extraction buffer, this reagent was able to dissolve most of the residue. All other steps in the protocol were performed as described[13]. The adapted protocol yielded 150 spots on each of the gels.

### 3.3.2. Proteins annotated by MSMS analysis

| | Number of differential spots | Number of differential spots that are upregulated | Number of excised, digested and analyzed spots | Number of annotated spots |
|---|---|---|---|---|
| Trichomes *versus* whole leaf | 250 | 115 | 101 | 83 |
| Upper *versus* lower leafs | 25 | 12 | 9 | 8 |
| Proteins in one minute chloroform extract | N/A | N/A | 118 | 66 |

Table 2: Numbers of upregulated, analyzed and annotated spots for each proteome study

Amorpha-4,11-diene synthase and CYP71AV1 are the only two proteins that are known to be essential to artemisinin biosynthesis in *A. annua*. The identification of spots to be one of these enzymes, can be considered as a proof-of-concept for the proteome study. In the differential study between trichomes and whole leaf tissue, two upregulated spots are identified as amorpha-4,11-diene synthase. None of the spots identify as CYP71AV1.

A comprehensive list of the annotated spots for the three proteome studies can be found in appendix 1-3 for the trichomes-*versus*-leafs study, the upper-*versus*-lower leafs study and the chloroform extract study respectively. Spots that identified as histones, chaperonins, proteasomes, porins, DNA binding proteins, Ribulose-1,5-bisphosphate carboxylase and ATPases, were omitted from the list. This list shows several candidate ESTs which could be useful for further investigation: transketolases, cytochrome P450 enzymes, dehydrogenases, etc.

### 3.3.3. Comparison of the proteomics and the AFLP data

To narrow down the results to the most valid gene candidates, we compared the proteome and the AFLP data. We were able to compile a short list of EST candidates (see Table 3), that could be useful for further investigation. The BLAST results for the three proteome studies can be found in appendix 4-6 for the trichomes-*versus*-leafs study, the upper-*versus*-lower leafs study and the chloroform extract study respectively.

| | EST candidates | Annotation of EST |
|---|---|---|
| **Trichomes *vs.* whole** | GSTSUB_047_A04 | glucose acyltransferase [Solanum berthaultii] |
| | AAGST_08_G09 | glyceraldehyde 3-phosphate dehydrogenase [Linum usitatissimum] |
| | AAGST_013_H10 | thioredoxin/transketolase fusion protein [synthetic construct] |
| | AAGST_018_B10 | caffeic acid O-methyltransferase II [Nicotiana tabacum] |
| | AAGST_029_C10 | catalase 2 [Helianthus annuus] |
| | AAGST_032_A11 | monodehydroascorbate reductase [Mesembryanthemum crystallinum] |
| | AAFB_013_C07 | plastidic aldolase NPALDP1 [Nicotiana paniculata] |
| **Upper *vs.* lower leafs** | AAFB_015_D05 | glycolate oxidase/ oxidoreductase [Arabidopsis thaliana] |
| | AAFB_015_H07 | Cytochrome b6-f [Pisum sativum] |
| **Chloroform extract** | GSTSUB_009_C06 | N/A |
| | AAGST_010_C02 | caffeic acid O-methyltransferase [Rosa chinensis var. spontanea] |
| | AAGST_010_C05 | hypothetical protein MtrDRAFT_AC151668g11v1 [Medicago |
| | AAGST_020_B07 | putative hypersensitive-induced response protein [Oryza sativa] |
| | AAGST_032_F11 | MADS-box transcription factor CDM41 [Chrysanthemum x morifolium] |
| | AAFB_014_H05 | Chlorophyll a-b binding protein |

Table 3: EST candidates supported by both the proteome and the AFLP analyses

### 3.3.4. Prospectives

The proposed EST candidates can be used to discover genes involved in the production of artemisinin. Several strategies can be followed to validate the candidate genes:

- Fluorescent mRNA in situ hybridization: A fluorescent labeled RNA probe complementary to the mRNA of a candidate gene can be used to localize the mRNA within tissue sections of *A. annua*. If the probes only hybridize in the trichomes and not in other cell types, this would be an extra indication for the involvement of this gene in artemisinin production.
- The overexpression in trichomes of the mRNA of a candidate gene could be validated using real time PCR.
- Candidate genes, proposed to be involved in the biosynthesis of artemisinin, could be expressed in *Saccharomyces cerevisiae*[18]. Microsomes of this yeast can be assayed with possible substrates to determine the catalytic function of the recombinant protein.
- *A. annua* plants could be transformed with the candidate genes. Transformed *A. annua* plants producing more artemisinin is the ultimate validation of the involvement of the candidate gene.

## 3.4. Conclusions

Using a proteome approach, we identified proteins which are upregulated in trichomes compared to whole leaf tissue and in upper leafs compared to lower leafs. We were also able to identify several proteins which are present in a one-minute chloroform extract of whole *A. annua* plants. The fragmentation spectra of the proteins resulting from the proteomics analysis were searched against the public MSDB database and against three EST libraries (one from flower buds of *A. annua,* one from the trichomes on the flower buds and a subtracted library). This search resulted in several candidate ESTs which could be useful for further investigation in the quest for artemisinin related genes: transketolases, cytochrome P450 enzymes, dehydrogenases, etc.

To narrow down the results to the most valid gene candidates, the proteome data was compared with a cDNA AFLP analysis on samples of A. annua leafs, taken at different points during a 72h time period after exposure to jasmonic acid. This approach allowed us to compile a list of EST candidates of which their possible involvement in artemisinin biosynthesis is supported by both analyses.

Further investigation by fluorescent mRNA in situ hybridization, real time PCR, expression in *Saccharomyces cerevisiae*, and/or creation of transformed *A. annua* plants, is needed to validate these results.

# References


1. Greenwood B, Mutabingwa T: **Malaria in 2002**. *Nature* 2002, **415**(6872):670-672.

2. Olliaro PL, Taylor WR: **Developing artemisinin based drug combinations for the treatment of drug resistant falciparum malaria: A review**. *J Postgrad Med* 2004, **50**(1):40-44.

3. Davis TM, Karunajeewa HA, Ilett KF: **Artemisinin-based combination therapies for uncomplicated malaria**. *Med J Aust* 2005, **182**(4):181-185.

4. Delabays N, Simonnet X, Gaudin M: **The genetics of artemisinin content in Artemisia annua L. and the breeding of high yielding cultivars**. *Curr Med Chem* 2001, **8**(15):1795-1801.

5. Bhakuni RS, Jain DC, Sharma RP, Kumar S: **Secondary metabolites of *Artemisia annua* and their biological activity**. *Current Science* 2001, **80**(1):35-48.

6. Chen D, Ye H, Li G: **Expression of a chimeric farnesyl diphosphate synthase gene in Artemisia annua L. transgenic plants via Agrobacterium tumefaciens-mediated transformation**. *Plant Science* 2000, **155**(2):179-185.

7. Duke M, Paul R, Elsohly H, Sturtz G, Duke S: **Localization of artemisinin and artemisitene in foliar tissues of glanded and glandless biotypes of *Artemisia annua* L.** *International Journal of Plant Sciences* 1994, **155**(3):365-372.

8. Zhang L, Ye HC, Li GF: **Effect of development stage on the artemisinin content and the sequence characterized amplified region (SCAR) marker of high-artemisinin yielding strains of Artemisia annua L.** *Journal of Integrative Plant Biology* 2006, **48**(9):1054-1062.

9. Van Nieuwerburgh FCW, Casteele SRV, Maes L, Goossens A, Inze D, Van Bocxlaer J, Deforce DLD: **Quantitation of artemisinin and its biosynthetic precursors in Artemisia annua L. by high performance liquid chromatography - electrospray quadrupole time-of-flight tandem mass spectrometry**. *Journal of Chromatography A* 2006, **1118**(2):180-187.

10. Gershenzon J, Duffy MA, Karp F, Croteau R: **Mechanized Techniques for the Selective Extraction of Enzymes from Plant Epidermal Glands**. *Analytical Biochemistry* 1987, **163**(1):159-164.

11. Gershenzon J, Mccaskill D, Rajaonarivony JIM, Mihaliak C, Karp F, Croteau R: **Isolation of Secretory-Cells from Plant Glandular Trichomes and Their Use in Biosynthetic-Studies of Monoterpenes and Other Gland Products**. *Analytical Biochemistry* 1992, **200**(1):130-138.

12. McConkey ME, Gershenzon J, Croteau RB: **Developmental regulation of monoterpene biosynthesis in the glandular trichomes of peppermint**. *Plant Physiology* 2000, **122**(1):215-223.

13. Bauw G, Montagu MV: **Two-Dimensional Polyacrylamide Gel Electrophoresis-Based Analysis for the Identification of Proteins and Corresponding Genes**. *Differentially Expressed Genes in Plants: A Bench Manual* 1997:150-118.

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

15. Zukas AA, Breksa AP: **Extraction methods for analysis of Citrus leaf proteins by two-dimensional gel electrophoresis**. *Journal of Chromatography A* 2005, **1078**(1-2):201-205.

16. Tilleman K, Stevens I, Spittaels K, Van den Haute C, Clerens S, Van den Bergh G, Geerts H, Van Leuven F, Vandesande F, Moens L: **Differential expression of brain proteins in glycogen synthase kinase-3 beta transgenic mice: A proteomics point of view**. *Proteomics* 2002, **2**(1):94-104.

17. Tilleman K, Van den Haute C, Geerts H, van Leuven F, Esmans EL, Moens L: **Proteomics analysis of the neurodegeneration in the brain of tau transgenic mice**. *Proteomics* 2002, **2**(6):656-665.

18. Teoh KH, Polichuk DR, Reed DW, Nowak G, Covello PS: **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* 2006, **580**(5):1411-1416.

# 4. Quantitation of artemisinin and its biosynthetic precursors in *Artemisia annua* L. by high performance liquid chromatography – electrospray quadrupole time-of-flight tandem mass spectrometry

JOURNAL OF CHROMATOGRAPHY A, 1118 (2006) 180-187

Filip Van Nieuwerburgh[1], Sofie Vande Casteele[1], Lies Maes[2], Alain Goosens[2], Dirk Inzé[2], Jan Van Bocxlaer[3] and Dieter Deforce[1]

6) Laboratory for Pharmaceutical Biotechnology, Ghent University, Harelbekestraat 72, B-9000 Ghent, Belgium

7) Department of Plant Systems Biology, Flanders Interuniversity Institute for Biotechnology, Ghent University, Technologiepark 927, B-9052 Ghent, Belgium.

8) Laboratory of Medical Biochemistry and Clinical Analysis, Ghent University, Harelbekestraat 72, B-9000 Ghent, Belgium

**ABSTRACT.**

This study reports the development and validation of a rapid, sensitive and selective assay for the quantitation of artemisinin, arteannuin B, artemisitene and artemisinic acid in *Artemisia annua* L. by reversed phase High Performance Liquid Chromatography (HPLC) Electrospray (ESI) Quadrupole – Time Of Flight (Q-TOF) tandem mass spectrometry (MS/MS). A recovery of >97% for all analytes was achieved by immersing one gram of fresh plant material in chloroform for one minute. This result supports the hypothesis that artemisinin and some of its structural analogs present in the leaves *A. annua* L. are localized entirely in the subcuticular space of the glands on the surface of the leaves. We validated the use of this chloroform extract, without additional sample preparation steps, for quantitative Q-TOF MS/MS. No ion suppression (matrix effect) resulting from interference with other compounds was detected. For every concentration within the range of the standard curve (0.1 to 3.00 μg/ml), accuracy was between 85% and 115%. Within- and between-day variations for the analysis of *A. annua* L. samples were <20%.

## 4.1. Introduction

Malaria is one of the world's most important parasitic diseases. There are at least 300 million acute cases of malaria each year globally, resulting in more than a million deaths[1, 2]. Multi-drug resistance of the *Plasmodium* strains to the cheapest and most widely used antimalarials such as chloroquine, mefloquine and sulfadoxine-pyrimethamine is one of the biggest challenges in the fight against malaria[1].

*Artemisia annua* L. (sweet wormwood), a herb of the Asteraceae family has been used for centuries for the treatment of fever and malaria[3]. Artemisinin, an endoperoxide-containing sesquiterpene lactone, is the main component responsible for this therapeutic effect. Based on artemisinin, several semi-synthetic derivatives such as artemether, arteether and artesunate have been produced[3]. The WHO recommends that all countries experiencing resistance to conventional monotherapies should use combination therapies, preferably those containing artemisinin derivatives (ACTs – artemisinin-based combination therapies)[4, 5].

As artemisinin cannot be synthesized chemically in an economically feasible way, *A. annua* is the only practical source of this valuable drug[6]. Unfortunately, *A. annua* contains only very small amounts of artemisinin ranging from 0.001 to 1.54% of dry weight[7]. Several research programs have been set up trying to increase the concentration of artemisinin in *A. annua* by optimizing the growing and harvesting conditions, by selecting high yielding cultivars or by creating transgenic plants[6, 8]. To study the content of artemisinin and its biosynthetic precursors in plants, we developed a very simple extraction method followed by HPLC – ESI MS/MS.

Several other methods have been reported for the extraction, chromatography and detection of artemisinin and its structural analogs in *A. annua*. Liquid solvent extraction of dried plant material is currently the most commonly applied technique. Also more complicated extraction techniques such as Super Critical Fluid Extraction (SFE), (Pressurized Solvent Extraction (PSE) and Microwave-assisted Extraction (MAE) have been used. For the quantitation of artemisinin a large array of techniques have been developed including thin layer chromatography (TLC), high performance liquid chromatography with UV detection (HPLC-UV), HPLC with electrochemical detection (HPLC-ECD), HPLC with evaporative light scattering detector (HPLC-ELSD), gas chromatography with flame ionization detector

(GC-FID), GC coupled to mass spectrometry (GC-MS), GC coupled to tandem mass spectrometry (GC-MS/MS), supercritical fluid chromatography with FID (SFC-FID), ELSD (SFC-ELSD) or MS (SFC-MS) and capillary electrophoresis with UV detection (CE-UV). A review by Christen *et al*.[9] gives an excellent overview of these techniques and discusses some of them in more detail.

Some of these methods such as TLC, EC and UV-detection (artemisinin is UV-transparent therefore derivatisation is required) are time-consuming and not suited for routine analysis. More important is the fact that most of these methods lack specificity (TLC, UV-detection, FID, ECD, ELSD). As an *A. annua* plant extract may contain hundreds of components, some structural analogues of artemisinin, good specificity of the detector is essential.
The high sensitivity and selectivity of MS and certainly MS/MS present a major advantage for the detection of specific components in plant extracts. Several GC-MS[10, 11], HPLC-MS[12-15] and HPLC-MS/MS[16] methods have been developed to analyze artemisinin and its derivatives in blood, plasma or serum. For analysis of *A. annua* extracts a SFC-MS method has been reported[17].

To our knowledge, we report the first MS/MS method developed to analyze artemisinin and its biosynthetic precursors in *A. annua.* The main advantages of our method are not only the excellent specificity but also the extremely short and efficient sample preparation.

## 4.2. Experimental

### 4.2.1. Chemicals

Pure reference standard of artemisinin, 98% was obtained from Sigma-Aldrich (Bornem, Belgium). The other reference standards arteannuin B, artemisitene and artemisinic acid were kindly provided by the Walter Reed Army Institute of Research (Washington, USA). The internal standard (I.S.) β-artemether was a gift from Arenco Pharmaceutica N.V. (Geel, Belgium). LC-MS grade absolute methanol was obtained from Biosolve (Valkenswaard, the Netherlands). Analytical grade chloroform was obtained from Acros (Geel, Belgium). Analytical grade ammonium acetate, analytical grade sodium acetate and acetic acid (99.8%) were obtained from Sigma-Aldrich (Bornem, Belgium). Purified water of 18.2 MΩ/cm was obtained from a Milli-Q system (Millipore, Belgium).

### 4.2.2. *Artemisia annua* L. plants

The plants were grown under controlled conditions (21 °C; 12h day/12h night regime). Seeds were kindly provided by the National Botanic Garden of Belgium (Meise, Belgium).

### 4.2.3. Analytical Standards

Individual stock solutions (1 mg/ml) of artemisinin, arteannuin B, artemisitene, artemisinic acid and internal standard β-artemether were prepared by accurately weighing required amounts into separate volumetric flasks and dissolving in appropriate volumes of methanol. Analytical standards were prepared as a mixture of each analyte (0.1 µg/ml to 3 µg/ml each) and the internal standard (0.4 µg/ml) by serial dilution of stock solutions in methanol – 1 mM ammonium acetate buffer adjusted to pH 5 with acetic acid (50 – 50 v/v).

### 4.2.4. Sample Preparation

Extraction was performed by immersing one gram of plant material in 6 ml chloroform for one minute. An aliquot of 10 µl of this extract was then dissolved in 1 ml methanol – 1 mM ammonium acetate buffer adjusted to pH 5 with acetic acid (50 – 50 v/v) containing 0.4 µg/ml of the I.S. β-artemether. This procedure was carried out on the plants of interest: *Artemisia annua* L. (Asteraceae) and the negative controls *Artemisia Absinthium* L. (Asteraceae), *Mentha spicata* L. (Lamiaceae) and *Mentha piperita* L. (Lamiaceae).

### 4.2.5. Liquid Chromatography

A Waters Alliance 2695 HPLC system was used to deliver the mobile phase [pump A, 1 mM ammonium acetate buffer adjusted to pH 5 with acetic acid; pump B, 100% methanol] for gradient elution at a flow rate of 0.2 mL/min. The initial composition of 50:50 was maintained for 1 minute; next the methanol content was increased linearly to 80% over a period of 6 min and maintained for 18 minutes. Re-equilibration time was 10 min between runs. The sample injection volume was 100 µl for all samples. Chromatographic separations were achieved on an Alltech Ultrasphere $C_{18}$ IP 5µm column (150 x 2.1 mm) protected by a Waters XTerra MS $C_{18}$ 5µm guard column (10 x 2.1). A LC Packings ACUrate ICP-04-20 post-column splitter was used to divert one-fourth of the effluent into the electrospray LC-MS interface.

### 4.2.6. Q-TOF mass spectrometry

Mass spectrometric detection was performed on a Q-TOF Ultima mass spectrometer (Micromass, Manchester, UK) equipped with an electrospray source operating in positive mode (ESP+). The ESI capillary voltage was set at 2.7 kV. The source and desolvation temperatures were optimized at 130 and 300 °C, respectively. Nitrogen was used as desolvation gas with a flow rate of 500 L/h. MS/MS analysis was performed using argon (0.9 bar) as the collision gas. An MS/MS method was used to quantify artemisinin (*m/z* 283 → 219+229+247+265), arteannuin B (*m/z* 249 → 185+189+203+231), artemisitene (*m/z* 281 → 217+227+245+263), artemisinic acid (*m/z* 252 → 189+199+217) and β-artemether (*m/z* 316 → 267+284+307). Cone voltage had an optimum at 40V for all components. The collision energy was optimized at 7 eV for artemisinin, 10 eV for arteannuin B, 7 eV for artemisitene, 11 eV for artemisinic acid and 7 eV for β-artemether. Data acquisition and analysis were carried out using Masslynx version 4.0. software. Analytical standard curves (second-order polynomic regression) were calculated using analyte to I.S. peak area ratios. The concentrations of the respective analytes in test samples are interpolated from the standard curves using the analyte to I.S. peak area ratios from the test samples.

## 4.3. Results and discussion

### 4.3.1. Sample preparation

Most plant extraction methods start with lyophilisation or drying of the plant material, followed by extraction with an organic solvent such as hexane or toluene[9]. As all cells are disrupted by these extraction methods, all soluble components are extracted from the plant. These extracts contain a massive amount of components (e.g. chlorophyll) interfering with HPLC (clogging) and MS (matrix effect). Additional sample preparation has to be performed prior to HPLC – MS/MS. Unfortunately these additional steps (solid phase extraction, filtering, evaporation steps) are not only time-consuming, but are also a possible source of variations in recovery.

In the specific case of the extraction of artemisinin and its bioprecursors from *A. annua*, these problems can be avoided. Duke *et al*.[18] reported that a 5 seconds dip in chloroform extracted 97% of the artemisinin and 100% of artemisitene from *A. annua*. In the report by Duke *et al*., quantitation was performed by HPLC-UV after derivatisation. Light microscopy and Transmission Electron Microscopy revealed that the 5 seconds dip results in collapse of the subcuticular cavity of the glands on the leaf surface but did not disrupt cell membranes. An *A. annua* biotype without glands contained neither artemisinin nor artemisitene[18]. These results indicate that artemisinin and artemisitene present in foliar tissue are localized entirely in the subcuticular space of glands of *A. annua*.

We hypothesized that this chloroform extract can be analyzed on a mass spectrometer without additional sample preparation steps as it contains only a very small quantity of interfering components (e.g. chlorophyll) compared to plant extracts where the plant material is lyophilized, dried or grinded.

### 4.3.2. Extraction time

We decided to prolong the extraction time as long as possible to break open as much glands as possible without introducing interfering compounds. After an extraction time of one minute, chlorophyll starts to be released into the chloroform, indicating that cells with interfering compounds begin to break open. Figure 1 shows a picture of glandular trichomes before and after a 1 minute chloroform extraction. The cuticle is crumpled after chloroform extraction.

The epidermal cells are unaffected by the treatment. The extraction time of one minute was validated during the recovery studies.

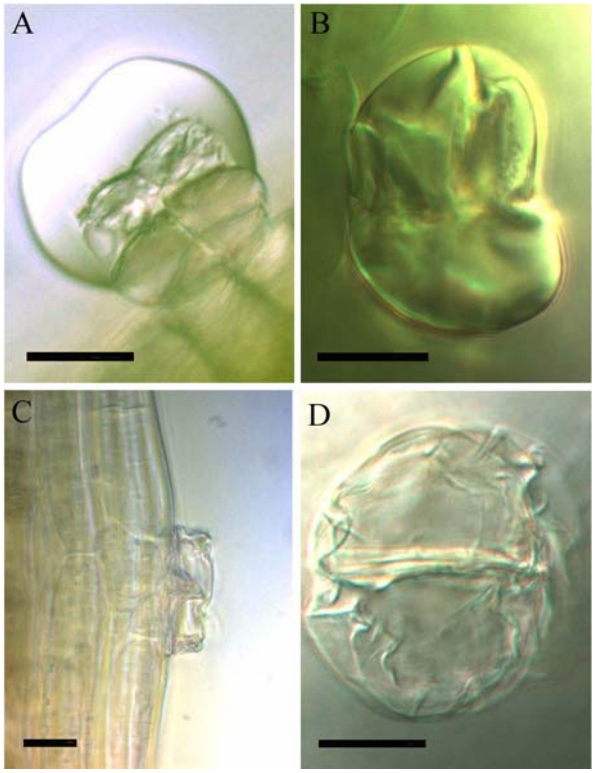


Figure 1: Picture of a glandular trichome on a leaf of *A. annua* L. before (A and B) and after (C and D) chloroform extraction. The cuticle is crumpled after chloroform extraction. The epidermal cells are unaffected by this treatment. Black bar is 10 µm.

### 4.3.3. Recovery

Two different experiments were done to asses the recovery. In a first experiment, 15 equal samples of one gram fresh *A. annua* leaves were prepared, 5 of which were spiked with 60 µl of a 10 mg/ml methanol solution of each analyte. Immediately after evaporation of the methanol, all 15 samples were analyzed. The recoveries of the 5 spiked samples were calculated as the ratio between the measured quantity and the spiked quantity increased with the mean quantity of the analytes in the 10 unspiked samples. Table 1 shows the mean of the recoveries for the different analytes (>97% for each analyte).

| **Spiked quantities** (µg/ml) | **arteannuin B** | **artemisitene** | **artemisinin** | **artemisinic acid** |
|---|---|---|---|---|
| Mean quantity unspiked samples[1] | 0.37 | 0.06 | 0.17 | 0.42 |
| Spiked quantity[2] | 1.00 | 1.00 | 1.00 | 1.00 |
| Total quantity in spiked samples[3] | 1.37 | 1.06 | 1.17 | 1.42 |
| **Recovered quantities** (µg/ml) | **arteannuin B** | **artemisitene** | **artemisinin** | **artemisinic acid** |
| Spiked sample 1[4] | 1.47 (107.39%) | 1.13 (106.30%) | 1.24 (106.12%) | 1.44 (100.87%) |
| Spiked sample 2[4] | 1.27 (93.01%) | 0.97 (91.27%) | 1.06 (90.23%) | 1.26 (88.49%) |
| Spiked sample 3[4] | 1.51 (110.53%) | 1.22 (115.15%) | 1.28 (109.37%) | 1.52 (106.66%) |
| Spiked sample 4[4] | 1.14 (83.47%) | 0.89 (83.70%) | 0.96 (81.59%) | 1.19 (83.48%) |
| Spiked sample 5[4] | 1.31 (95.74%) | 1.01 (95.14%) | 1.15 (98.55%) | 1.56 (109.45%) |
| **Mean spiked samples** (µg/ml)[5] | **1.34 (98.03%)** | **1.04 (98.31%)** | **1.14 (97.17%)** | **1.39 (97.79%)** |
| **Standard deviation** (µg/ml) | **0.15 (11.03%)** | **0.13 (12.46%)** | **0.13 (11.42%)** | **0.16 (11.35%)** |

Table 1: Recovery from spiked samples: Fifteen equal samples of one gram fresh *A. annua* leaves were prepared. Ten samples were not spiked and analyzed[1]. Five samples were spiked with each analyte[2]. The total quantity of the analytes present in the spiked samples was calculated as the sum of the spiked quantity and the mean quantity of the analytes in the 10 unspiked samples[3]. The spiked samples were analyzed and the individual[4] and mean[5] absolute recoveries (% recovery between brackets) were calculated. Quantities are presented as the concentration after sample preparation (multiply by 600 to obtain quantities in µg analyte / g fresh plant material).

This very high recovery (> 97%) of the spiked amounts does not imply a high recovery of the amounts present in the plant. The recovery of the amounts present in the plant with our method, cannot be measured directly. Therefore, we estimated this recovery in a second experiment by comparing the recovery achieved with a previous described extraction method[19] before and after our one-minute chloroform extraction. Six equal samples of one gram fresh leaf material were prepared. Three of them were extracted following a previously described extraction method[19]. Briefly, this method consists of an extraction with 2 x 3 ml toluene after lyophilisation and pulverization of the plant material followed by a normal-phase Silica gel solid-phase extraction (SPE). An aliquot of 1 ml of plant extract was passed through the 500 mg Silica gel column, followed by washing with 2 ml petroleum ether - diethyl ether (9:1) and elution with 2 x 0.5 ml acetonitrile. The eluate was evaporated to dryness under $N_2$ and reconstituted in 1 ml methanol – ammonium acetate buffer (50 – 50 v/v) for further analysis. Note that compared to our method, the analytes are a 100 fold more concentrated by this SPE. The other three samples were subjected to exactly the same extraction protocol, but after they were first extracted by our method (one minute chloroform extraction). Table 2 gives and overview of the results. The amount of artemisinin, arteannuin B and artemisinic acid found in the plant material after chloroform treatment was less than 3% compared to the amount found in the three non-pretreated samples. This experiment shows that >97% of

artemisinin, arteannuin B, and artemisinic acid is extracted by a one minute dip in chloroform. As a one minute dip in chloroform is the only sample preparation step in our method, we conclude that a recovery of >97% can be achieved by our method. For artemisitene, the results are less conclusive as the measured quantity after chloroform extraction falls below the lower limit of quantitation (LLOQ). Nevertheless, the experiment gives a good indication of a high recovery of artemisitene.

| | arteannuin B | artemisitene | artemisinin | Artem. acid |
|---|---|---|---|---|
| Mean quantity WITHOUT preceding chloroform extraction (µg/ml)[1] | 58.64 | 0.28 | 24.17 | 72.07 |
| Mean quantity AFTER preceding chloroform extraction (µg/ml)[2] | 3.60 | 0.02 (< LLOQ) | 1.58 | 5.53 |
| Quantity not extracted by preceding chloroform extraction | 6.14% | 6.80% | 6.55% | 7.68% |
| Residual chloroform in samples after chloroform extraction[3] | > 5% | > 5% | > 5% | > 5% |
| **Recovery** | > 98.86% | > 98.20% | > 98.45% | > 97.32% |

Table 2: Recovery with chloroform extraction of *Artemisia annua* leaves: Six equal samples of one gram fresh leaf material were prepared. Three of them were extracted following a previously described extraction method[19] which uses extraction with toluene after lyophilisation and pulverization of the plant material[1]. The other three samples were extracted in exactly the same way but after they were first extracted for one min with chloroform[2]. The percentage of the chloroform which sticks to the plant material after chloroform extraction (accounting for a part of the not-extracted percentage) was gravimetrically determined[3]. Quantities are presented as the concentration after sample preparation (multiply by 6 to obtain quantities in µg analyte / g fresh plant material).

### 4.3.4. Chromatography

During flow injection analysis on the Q-TOF MS, $[M+Na]^+$ adducts were found to be far more intense than $[M+H]^+$ or $[M+NH4]^+$ adducts. At first we tried to intensify the $[M+Na]^+$ adducts, reducing other adducts by using sodium acetate buffer and performing analysis on the $[M+Na]^+$ adducts. This approach was abandoned due to variation caused by build up of sodium acetate deposits on the ion sampling cone of the mass spectrometer. Finally we decided to use ammonium acetate buffer and to perform MS/MS analysis on the $[M+H]^+$ (artemisinin, artemisitene, arteannuin B) or $[M+NH4]^+$ (artemisinic acid, artemether) adducts.
A total of 3 isocratic and 25 gradient elutions were compared, testing varying methanol - buffer ratios and testing varying gradient speeds. The method with the highest peak resolution was chosen. The reproducibility of the retention times was very dependent on the buffer concentration. Increasing the buffer concentration from 0.1 mM to 1 mM greatly enhanced the reproducibility of the retention times, resulting in a variation of less than 25 seconds. By varying the pH of the ammonium acetate buffer, the retention time of artemisinic acid can be

influenced. Peak resolution was optimal at pH 5. Using final conditions, all analytes were separated from each other with peak resolutions from 1.0 to 2.4 (Fig. 2).

The column and guard column were stable for at least 1000 injections. No signs of column deterioration have been detected yet.

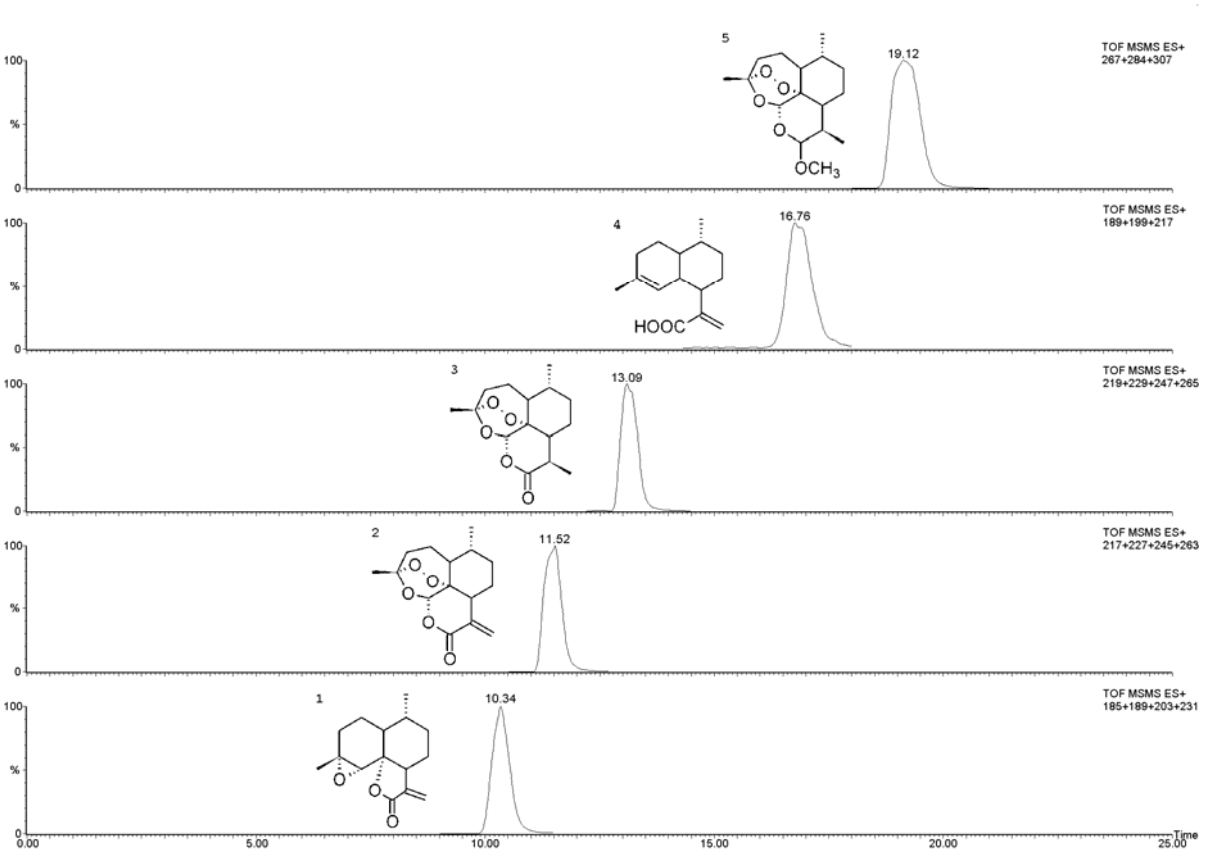


Figure 2: Chromatogram with retention times and chemical structures of (1) arteannuin B, (2) artemisitene, (3) artemisinin, (4) artemisinic acid and (5) the internal standard artemether. This chromatogram is the result of the analysis by electrospray QTOF-MS/MS of an analytical standard containing 1.2 µg/ml of each analyte and 0.4 µg/ml IS.

### 4.3.5. Specificity

In contrast with $[M+Na]^+$ adducts, the $[M+H]^+$ and $[M+NH4]^+$ adducts were easily fragmented with low collision energies. Figure 3 shows the fragmentation spectrum of artemisinin at an optimal collision energy of only 7eV. Between 3 and 4 fragments were chosen to be monitored for each analyte. Using the sum of several fragments for MS/MS quantitation, has the advantage of increased signal strength and enhanced signal stability, but the disadvantage of lower specificity. As fragments with higher m / z values tend to be more specific, fragments with the highest m / z values were selected.

To check the specificity of the method, chloroform extracts of *Mentha piperita*, *Mentha spicata* and *Artemisia absinthium* were analyzed. These three plant species also have epidermal glands on their leaves, but are not reported to produce artemisinin. In these control extracts, no MS/MS signal could be detected for the components of interest.

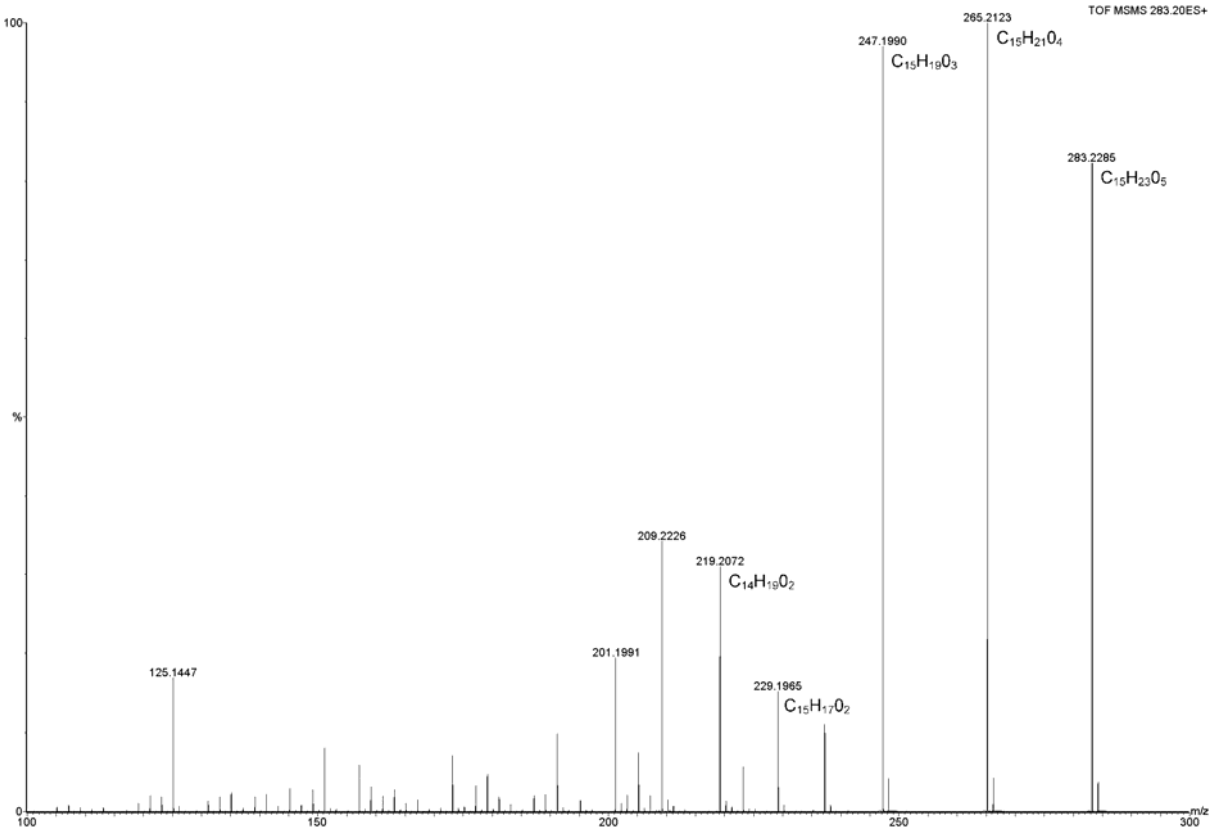


Figure 3: This spectrum shows the fragmentation of artemisinin (m/z 283) with the optimal collision energy of 7 eV. The MS/MS signal is calculated as the sum of the fragments with m/z 219, 229, 247 and 265. The molecular formulas show that the fragments are mainly formed by dehydration of their parent ion.

### 4.3.6. Ion suppression (matrix effect)

To check for interferences from other compounds by ion suppression (matrix effect) a standard curve obtained from standards prepared in mobile phase was compared to a standard curve obtained from standards made in mobile phase spiked with matrix (10 µl chloroform extract of *Mentha piperita* /ml). HPLC-MS/MS analysis (3 measurements for each sample) of these standards, resulted in almost identical measurements for the spiked and the non-spiked standards. The statistical method of Bland *et al.*[20] was used for assessing the agreement between the two methods. The p-values for the t-test with the nulhypothesis that the mean of the differences between both methods is equal to zero, were 0.53, 0.26, 0.67 and 0.74 for arteannuin B, artemisinin, artemisitene and artemisinic acid respectively. No significant

difference could be found for any of the analytes, meaning that no ion suppression could be detected. For this reason, in the final method, standards were not spiked with chloroform extract to include matrix.

### 4.3.7. Accuracy, precision, limit of detection (LOD) and lower limit of quantitation (LLOQ)

The definitions for accuracy, precision, LOD and LLOQ were adopted from the FDA guidelines for bioanalytical method validation[21]. The LOD was defined as the lowest observable peak response for an analyte above the background noise, 3 times the system noise in the matrix. The LLOQ was defined as the lowest concentration for an analyte with a response signal 5 times the system noise in the matrix, a precision of 20% and an accuracy of 80-120%. Within-day accuracy and precision were calculated with three determinations on one day. Between-day accuracy and precision were calculated from 7 determinations on 3 days spanning a two week period. Accuracy and precision were calculated for each of 7 spiked concentrations (0.1; 0.2; 0.4; 0.8; 1.2; 2.0 and 3.0 µg/ml) within the range of the standard curve. Within the range of the standard curve, coefficient of variation (CV%) was <15% and accuracy was between 85% and 115% for all analytes and all 7 spiked concentrations (Table 3).

Within- and between-day variation was also calculated for unspiked *A. annua* samples. Twenty equal samples of one gram fresh leaf material were prepared and kept between 4°C to 8°C until extraction. Ten of these samples were extracted with chloroform on day 1, five on day 2 and again five on day 3. The extracts were stored at -20°C until HPLC-MS/MS analysis. The first 10 extracts (extraction on day one) were analyzed in one day allowing calculation of within-day variation. The other 10 extracts were analyzed on two different days spanning a two week period. All 20 independent samples were used to calculate the between-day variation. The within- and between-day variation of the complete procedure (extraction and quantitation by MS/MS) is <20 % (Table 3) for all analytes except for artemisitene for which the amount present in the unspiked samples was below the LLOQ. The variation for the unspiked samples is higher than for the spiked samples. A possible reason is variation in the release of the analytes out of the glandular trichomes. Another possibility is an actual variation in the 20 samples as the leaves for these samples were collected from three different plants.

| | arteannuin B | | artemisitene | | artemisinin | | artemisinic acid | |
|---|---|---|---|---|---|---|---|---|
| | Within | Between | Within | Between | Within | Between | Within | Between |
| LOD (µg/ml) | 0.001 (S/N: 3) | | 0.0005 (S/N: 5) | | 0.0001 (S/N: 9) | | 0.04 (S/N: 4) | |
| LLOQ (µg/ml) | 0.1 (S/N: 86) | | 0.1 (S/N: 52) | | 0.1 (S/N: 22) | | 0.1 (S/N: 5) | |
| Accuracy at LLOQ | 108±9% | 111±8% | 107±13% | 111±8% | 114±10% | 112±6% | 104±7% | 108±8% |
| CV% at LLOQ | 8.5% | 7.5% | 12.6% | 6.7% | 8.4% | 5.1% | 6.6% | 7.1% |
| Accuracy (0.1 – 3 µg/ml spiked) | 94 - 108% | 94 - 111% | 89 - 107% | 85 - 110% | 90 - 113% | 85 - 111% | 97 - 104% | 96 - 107% |
| CV% (0.1 – 3 µg/ml spiked) | 1.0 - 8.5% | 0.9 - 7.5% | 2.0 - | 1.3 - 8.5% | 1.1 - | 1.3 - | 0.8 - 7.1% | 0.9 - 7.1% |
| Mean unspiked extracts (µg/ml) | 0.36 | 0.33 | 0.07 | 0.08 | 0.16 | 0.15 | 0.39 | 0.36 |
| St. dev. unspiked extracts (µg/ml) | 0.05 | 0.06 | 0.02 | 0.02 | 0.03 | 0.03 | 0.06 | 0.07 |
| CV% unspiked extracts | 14% | 19% | 24% | 22% | 15% | 19% | 17% | 20% |

**Table 3. Accuracy, precision, LOD and LLOQ**

LOD and LLOQ are presented with peak-to-peak signal-to-noise ratio. Within- and between-day accuracy and precision are presented at LLOQ and for 7 spiked concentrations (0.1; 0.2; 0.4; 0.8; 1.2; 2.0 and 3.0 µg/ml) within the range of the standard curve. Within- and between-day variation was also calculated for 20 unspiked *A. annua* samples. Quantities are presented as the concentration after sample preparation.

### 4.3.8. Dynamic range and polynomial regression

Based on the LLOQ and dynamic range of the MS/MS signal, standard curves were established from 0.1 to 3.00 µg/ml for artemisinin, arteannuin B, artemisitene and artemisinic acid. Several regression models were evaluated to establish these curves. For the range of 0.1 – 3 µg/ml, a best-fitted second-order polynomial regression $(y = Ax^2+Bx+C)$ described the measurements of the analytical standards at best (typically $R^2 > 0.99$). Limiting the range to 0.1 – 0.8 µg/ml, a linear regression $(y = Ax+B)$ would also be acceptable with $R^2 > 0.99$, but still a second-order polynomial regression describes this range better with $R^2 > 0.999$. In practice, the use of a second-order polynomial regression not only extended the useful dynamic range, but also reduced the between-day variation.

The range of the standard curves may not extend high enough to analyze high yielding plants[7]. Dilutions can be made from the extracts of these plants. To check if these dilutions do not present any ill effects, a sample of one gram fresh leaf material from *A. annua* was spiked with 16 mg of artemisinin in a methanol solution. After evaporation of the methanol, the artemisinin was extracted and prepared with the standard sample preparation procedure. Immediately before HPLC – MS/MS analysis, the sample was diluted 16 fold with methanol – ammonium acetate buffer (50 – 50 v/v) containing 0.4 µg/ml of the internal standard. The diluted sample was measured 3 times; recovery was 100.0% ± 8.4%.

## 4.4. Conclusions

This study reports the development and validation of a rapid, sensitive and selective assay for the quantitation of artemisinin, arteannuin B, artemisitene and artemisinic acid in *A. annua* L. by reversed phase HPLC ESI Q-TOF MS/MS. An absolute recovery of >97% was achieved by immersing one gram of plant material in chloroform for one minute. This result supports the hypothesis that artemisinin and some of its structural analogs present in the leaves of *A. annua* L. are localized entirely in the subcuticular space of the glands on the surface of the leaves. We validated the use of this chloroform extract for quantitative MS/MS without additional sample preparation steps. No ion suppression (matrix effect) resulting from interference with other compounds was detected. To check the specificity of the method, chloroform extracts of *Mentha piperita*, *Mentha spicata* and *Artemisia absinthium* were analyzed. These three plants also have epidermal glands on the leaves, but do not synthesize artemisinin. No signal for the components of interest was detected in these control extracts. With a LOD of at least 0.04 µg/ml, a LLOQ of 0.10 µg/ml and a dynamic range from 0.10 to 3.00 µg/ml for each analyte, the method has enough sensitivity and flexibility to measure low and high yielding cultivars. For every concentration within the range of the standard curve (0.1 to 3.00 µg/ml), accuracy was between 85% and 115%. Within- and between-day variations for the analysis of unspiked *A. annua* L. samples were <20%.

# References


1. Greenwood B, Mutabingwa T: **Malaria in 2002**. *Nature* 2002, **415**(6872):670-672.

2. Snow RW, Craig M, Deichmann U, Marsh K: **Estimating mortality, morbidity and disability due to malaria among Africa's non-pregnant population**. *Bull World Health Organ* 1999, **77**(8):624-640.

3. Klayman DL: **Qinghaosu (artemisinin): an antimalarial drug from China**. *Science* 1985, **228**(4703):1049-1055.

4. Olliaro PL, Taylor WR: **Developing artemisinin based drug combinations for the treatment of drug resistant falciparum malaria: A review**. *J Postgrad Med* 2004, **50**(1):40-44.

5. Davis TM, Karunajeewa HA, Ilett KF: **Artemisinin-based combination therapies for uncomplicated malaria**. *Med J Aust* 2005, **182**(4):181-185.

6. Delabays N, Simonnet X, Gaudin M: **The genetics of artemisinin content in Artemisia annua L. and the breeding of high yielding cultivars**. *Curr Med Chem* 2001, **8**(15):1795-1801.

7. Bhakuni RS, Jain DC, Sharma RP, Kumar S: **Secondary metabolites of *Artemisia annua* and their biological activity**. *Current Science* 2001, **80**(1):35-48.

8. Chen D, Ye H, Li G: **Expression of a chimeric farnesyl diphosphate synthase gene in Artemisia annua L. transgenic plants via Agrobacterium tumefaciens-mediated transformation**. *Plant Science* 2000, **155**(2):179-185.

9. Christen P, Veuthey JL: **New trends in extraction, identification and quantification of artemisinin and its derivatives**. *Curr Med Chem* 2001, **8**(15):1827-1839.

10. Mohamed SS, Khalid SA, Ward SA, Wan TS, Tang HP, Zheng M, Haynes RK, Edwards G: **Simultaneous determination of artemether and its major metabolite dihydroartemisinin in plasma by gas chromatography-mass spectrometry-selected ion monitoring**. *J Chromatogr B Biomed Sci Appl* 1999, **731**(2):251-260.

11. Theoharides AD, Smyth MH, Ashmore RW, Halverson JM, Zhou ZM, Ridder WE, Lin AJ: **Determination of dihydroqinghaosu in blood by pyrolysis gas chromatography/mass spectrometry**. *Anal Chem* 1988, **60**(2):115-120.

12. Naik H, Murry DJ, Kirsch LE, Fleckenstein L: **Development and validation of a high-performance liquid chromatography-mass spectroscopy assay for determination of artesunate and dihydroartemisinin in human plasma**. *J Chromatogr B Analyt Technol Biomed Life Sci* 2005, **816**(1-2):233-242.

13. Rajanikanth M, Madhusudanan KP, Gupta RC: **An HPLC-MS method for simultaneous estimation of alpha,beta-arteether and its metabolite dihydroartemisinin, in rat plasma for application to pharmacokinetic study**. *Biomed Chromatogr* 2003, **17**(7):440-446.

14. Rajanikanth M, Madhusudanan KP, Gupta RC: **Liquid chromatographic-mass spectrometric method for the determination of alpha-,beta-arteether in rat serum**. *J Chromatogr B Analyt Technol Biomed Life Sci* 2003, **783**(2):391-399.

15. Souppart C, Gauducheau N, Sandrenan N, Richard F: **Development and validation of a high-performance liquid chromatography-mass spectrometry assay for the determination of artemether and its metabolite dihydroartemisinin in human plasma**. *J Chromatogr B Analyt Technol Biomed Life Sci* 2002, **774**(2):195-203.

16. Sabarinath S, Rajanikanth M, Madhusudanan KP, Gupta RC: **A sensitive and selective liquid chromatographic/electrospray ionization tandem mass spectrometric assay for the simultaneous quantification of alpha-,beta-arteether and its metabolite dihydroartemisinin in plasma, useful for pharmacokinetic studies**. *J Mass Spectrom* 2003, **38**(7):732-742.

17. Dost K, Davidson G: **Analysis of artemisinin by a packed-column supercritical fluid chromatography-atmospheric pressure chemical ionisation mass spectrometry technique**. *Analyst* 2003, **128**(8):1037-1042.

18. Duke M, Paul R, Elsohly H, Sturtz G, Duke S: **Localization of artemisinin and artemisitene in foliar tissues of glanded and glandless biotypes of *Artemisia annua* L.** *International Journal of Plant Sciences* 1994, **155**(3):365-372.

19. Vandenberghe DR, Vergauwe AN, Van Montagu M, Van Den Eeckhout EG: **Simultaneous determination of artemisinin and its bioprecursors in Artemisia annua**. *J Nat Prod* 1995, **58**(5):798-803.

20. Bland JM, Altman DG: **Statistical methods for assessing agreement between two methods of clinical measurement**. *Lancet* 1986, **1**(8476):307-310.

21. FDA: **Guidance for Industry: Bioanalytical Method Validation**. 2001.

# SUMMARY

## PART I: Polymorphism association with Obsessive-Compulsive Disorder

Currently OCD is viewed mainly as a neuropsychiatric disease. The current definition opposes obsessions (anxiety inducing) to compulsions (anxiety reducing), which patients recognise as excessive or unreasonable. Recent prevalence studies report prevalence rates around 3 %. OCD is treated only in a minority of the cases. The main reason for this "hidden epidemic" is that the patients want to keep their disease secret and don't want to disclose their "crazy" symptoms.

Currently the most effective treatment for OCD is a combination of behaviour therapy and pharmacotherapy. This combined approach improves the condition of most patients substantially, and occasionally leads to a complete recovery. All Serotonin Reuptake Inhibitors (SRIs) have shown to be effective in OCD. Their exclusive efficacy has given grounds to the "serotonin hypothesis" that serotonin plays an important role in the pathogenesis of OCD. However, this serotonin hypothesis is too limited. It is clear that the dopaminergic system is also involved in the pathogenesis of OCD. In case of non-response or partial response to SRIs, evidence has accumulated that the addition of antipsychotic drugs to SRIs might lead to symptom improvement.

Twin studies showed a concordance rate of approximately 50–60% in monozygotic twins, compared to 10% in dizygotic twins. Family studies show a significant aggregation of illness within families compared to population prevalence. This evidence suggests a genetic basis for OCD, but also shows that obsessive and compulsive behaviours are influenced by non-genetic and environmental factors.

During the last 2 decades, a large number of association studies have been dedicated to disentangle the genetic components that may be involved in OCD. The preliminary and frequently inconsistent nature of the data represented in the majority of OCD genetic-association studies may seem discouraging. Failure to confirm previously identified susceptibility loci could result from a number of reasons, including the potential for population admixture, the clinical heterogeneity of OCD, small sample sizes (and subsequent lack of statistical power) or epistasis (genetic interaction).

In light of the putative role of the serotonergic, dopaminergic and possibly (nor)adrenergic system in OCD, following polymorphisms were analysed in a sample of >100 OCD patients and a control sample of >100 ethnically matched Caucasian subjects by means of a case-control study:

- Taq IA polymorphism in the non-coding region flanking the 3' end of the dopamine D2 receptor (DRD2) gene
- Catechol-O-Methyl Transferase (COMT) NlaIII high/low activity polymorphism
- 1438 A/G polymorphism within the promoter region of the postsynaptic 5-$HT_{2A}$ receptor
- 5-$HT_{1D\beta}$ G861C polymorphism: silent G-to-C substitution at nucleotide 864 of the coding region of the 5-$HT_{1D\beta}$ autoreceptor gene
- 5-HTTLPR: serotonin transporter gene (5-HTT) promoter 44-bp deletion/insertion

We found indication towards an association of the 5-HTTLPR S-allele with female OCD patients, and toward an association of the 5-$HT_{2A}$ G-allele and GG genotype with patients with a positive family history of OCD and an early onset of disease. There was a significant predominance of the DRD2 A2A2 genotype ($p = 0.049$), a higher frequency of the DRD2 A2 allele ($p = 0.020$) and a higher frequency of the low-activity COMT allele ($p = 0.035$) in male OCD patients compared to male controls. In addition, we observed an association of the DRD2 A2A2 genotype in patients with an early onset of OCS (age $\leq$ 15 years) ($p= 0.033$).

Research by which the treatment of patient is tuned on the basis of their genotype, is called pharmacogenetics. Strong evidence suggests that genetic variation plays an important role in inter-individual differences in medication response and toxicity. We studied whether polymorphisms of the 5-HTT, 5-$HT_{1D\beta}$ and 5-$HT_{2A}$ genes affect the efficacy of venlafaxine and paroxetine treatment in OCD. The results of this study strongly suggest that response in paroxetine treated OCD patients is associated with the GG genotype of the 5-$HT_{2A}$ polymorphism ($\chi^2 = 8.66$, df=2, p=0.013). In venlafaxine treated OCD patients, response is associated with the SL genotype of the 5-HTTLPR polymorphism ($\chi^2 = 9.71$, df=2, p=0.008). This study demonstrates that pharmacogenetics could solve the current trial and error approach with different SRIs for the treatment of OCD.

## PART II: proteomics on *Artemisia annua*

Three strategies were followed to discover genes of the plant *Artemisia annua* L. that are involved in the production of the antimalarial artemisinin: a proteome analysis, a quantitative cDNA amplified fragment length polymorphism analysis (cDNA AFLP) and the construction of full length Expressed Sequence Tag (EST) cDNA libraries. The results of these 3 strategies can be compared and complemented with each other.

The proteome of *A. annua* was investigated to discover candidate genes related to artemisinin production. We identified proteins that were differentially expressed between trichomes and whole leaf tissue. We also identified differentially expressed proteins between lower and upper leafs of *A. annua* plants. In a third proteome analysis, we explored the proteins present in a one-minute chloroform extract of *A. annua* plants. Trichomes are considered to be the artemisinin factories of *A. annua*. Upper leafs produce more artemisinin than the lower leafs. Proteins that are upregulated in these samples might thus be involved in artemisinin production. With the chloroform extraction, the content of the capitate glands of the trichomes is extracted. These glands possibly contain proteins involved in artemisinin production.

Proteins of interest were identified by tandem mass spectrometry. The fragmentation spectra were searched against the public Mass Spectrometry protein sequence DataBase (MSDB) and against three Expressed Sequence Tag (EST) libraries (one from flower buds of *A. annua*, one from the trichomes on the flower buds and one being the result of a cDNA subtraction of both plant tissues). To narrow down the results to the most valid gene candidates, the proteome data was compared with a cDNA AFLP analysis that investigated samples of *A. annua* leafs, taken at different time points during a 72h time period after exposure to jasmonic acid. We were able to compile a list of EST candidates, which could be useful for further investigation.

To be able to find which plant hormones trigger *A. annua* to produce more artemisinin, to check the correlation between artemisinin content of *A. annua* leafs and the number of trichomes on these leafs, and to be able to check transformed plants or shoot cultures for enhanced artemisinin production, a quantitation method for artemisinin and its bioprecursors had to be developed. A recovery of >97% for all analytes was achieved by immersing one gram of fresh plant material in chloroform for one minute. This result supports the hypothesis that artemisinin and some of its structural analogs are localized entirely in the subcuticular space of the glands on the surface of the leaves. We validated the use of this chloroform extract, without additional sample preparation steps, for quantitative HPLC – ESI – QTOF MS/MS. No ion suppression (matrix effect) resulting from interference with other compounds was detected. For every concentration within the range of the standard curve (0.1 to 3.00 µg/ml), accuracy was between 85% and 115%. Within- and between-day variations for the analysis of *A. annua* samples were <20%. We can conclude that this one minute lasting extraction procedure is a good alternative for other time consuming extraction procedures.

# SAMENVATTING

## DEEL I: Associatie van polymorfismen met obsessieve-compulsieve stoornis

De obsessieve-compulsieve stoornis (OCS) wordt tegenwoordig beschouwd als een neuropsychiatrische ziekte. OCS wordt gekarakteriseerd door zowel obsessies (angst inducerend) als compulsies (angst verminderend). De patiënten voelen deze obsessies en compulsies zelf aan als excessief of redeloos. De incidentie van OCS in de algemene bevolking bedraagt 3%. Slechts een minderheid van de gevallen wordt behandeld. Een belangrijke reden voor deze "verborgen epidemie" is het feit dat patiënten hun ziekte geheim houden en beschaamd zijn over hun symptomen.

De meest effectieve behandelingsmethode voor OCS combineert gedragstherapie met farmacotherapie. De toestand van de meeste patiënten verbetert opmerkelijk door deze combinatietherapie en occasioneel treedt zelfs volledige genezing op. De effectiviteit van Serotonine Reuptake Inhibitoren (SRIs) voor de behandeling OCS heeft de basis gelegd voor de "serotonine hypothese", namelijk dat serotonine een belangrijke rol speelt in de pathogenese van OCS. Deze serotonine hypothese is echter te beperkt, aangzien het duidelijk is dat het dopaminerge systeem ook betrokken is in de pathogenese van OCS. Er is toenemende evidentie dat de combinatie van SRIs met een antipsychotisch geneesmiddel de symptomen verbetert in geval van ontoereikende respons op SRIs.

Tweeling-studies tonen een concordantie van ongeveer 50–60% in monozygotische tweelingen, vergeleken met 10% in dizygotische tweelingen. Familie-studies tonen een significante aggregatie van OCS binnen éénzelfde familie. Deze feiten suggereren een genetische basis voor OCS, maar tonen ook dat obsessief en compulsief gedrag beïnvloed wordt door omgevingsfactoren.

Gedurende de laatste 2 decennia, werden een groot aantal associatiestudies uitgevoerd om de genetische componenten te ontrafelen die betrokken zijn in OCS. Een groot deel van de uit deze associatiestudies afkomstige resultaten lijken tegenstrijdig. Geïdentificeerde susceptibiliteitsloci worden meestal niet geconfirmeerd. Dit kan het gevolg zijn van een aantal factoren, zoals mogelijke vermenging van de OCS-populatie met andere populaties, de klinische heterogeniteit van OCS, analyse van een te kleine groep patiënten (en een daaruit volgend gebrek aan statistische power) of epistase (genetische interactie).

In het kader van de zoektocht naar de mogelijke rol van het serotonerge, dopaminerge en (nor)adrenerge systeem in OCS, werden volgende polymorfismen geanalyseerd in een steekproef van >100 OCS-patiënten en >100 etnische aangepaste controle stalen:

- Taq IA polymorfisme in de niet-coderende regio die het 3' uiteinde van het dopamine D2 receptor (DRD2) gen flankeert
- Catechol-O-Methyl Transferase (COMT) NlaIII hoge/lage activiteit polymorfisme
- 1438 A/G polymorfisme in de promotor regio van de postsynaptische 5-$HT_{2A}$ receptor
- 5-$HT_{1D\beta}$ G861C polymorfisme: stille G-naar-C substitutie op nucleotide 864 van de coderende regio van het 5-$HT_{1D\beta}$ autoreceptor gen
- 5-HTTLPR: 44-bp deletie/insertie in de promotor regio van het serotonine transporteur gen (5-HTT)

Onze studie leverde aanwijzingen voor een associatie van het 5-HTTLPR S-allele met vrouwelijke OCS-patiënten, alsook voor een associatie van het 5-$HT_{2A}$ G-allel en GG genotype met patiënten met een familiale OCS historiek en een vroeg begin van de ziekte. Er was een significant overwicht van het DRD2 A2A2 genotype ($p = 0.049$), een hogere frequentie van het DRD2 A2 allel ($p = 0.020$) en een hogere frequentie van het lage-activiteit COMT allel ($p = 0.035$) in mannelijke OCS-patiënten vergeleken met mannelijk controle-patiënten. Daarenboven observeerden we een associatie van het DRD2 A2A2 genotype in patiënten met een vroeg begin van de ziekte (leeftijd $\leq$ 15 jaar) ($p= 0.033$).

Farmacogenetica is een onderzoeksdomein waardoor de behandeling van de patiënt kan afgesteld worden op basis van zijn genotype. Genetische variatie speelt een belangrijke rol in inter-individuele verschillen in reactie op, en toxiciteit van medicatie. Vanuit deze invalshoek bestudeerden we of polymorfismen van de 5-HTT, 5-$HT_{1D\beta}$ en de 5-$HT_{2A}$ genen een effect hebben op de werkzaamheid van venlafaxine en paroxetine in de behandeling van OCS. Een positieve respons van OCS-patiënten op paroxetine behandeling blijkt sterk geassocieerd met het GG genotype van het 5-$HT_{2A}$ polymorfisme ($\chi^2 = 8.66$, df=2, p=0.013), terwijl een positieve respons op venlafaxine geassocieerd blijkt met het SL genotype van het 5-HTTLPR polymorfisme ($\chi^2 = 9.71$, df=2, p=0.008). Deze studie promoot farmacogenetica als een veelbelovend alternatief voor de huidige proefondervindelijke behandeling van OCS met verschillende SRIs.

## DEEL II: proteomics op *Artemisia annua*

Drie strategieën werden gevolgd met als doel de identificatie van genen betrokken in de productie van het antimalaria middel artemisinine door de plant *Artemisia annua*: een proteoom analyse, een kwantitatieve "cDNA geAmplificeerde Fragment Lengte Polymorfisme analyse" (cDNA AFLP) en de aanmaak van EST- (Expressed Sequence Tag) banken. De resultaten van deze 3 strategieën zijn complementair en kunnen met elkaar vergeleken worden.

Het proteoom van *A. annua* werd onderzocht om kandidaat genen te identificeren die in verband staan met artemisinine productie. We identificeerden proteïnen die differentieel tot expressie komen in klierharen t.o.v. volledig blad-materiaal. We identificeerden ook proteïnen die differentieel tot expressie komen in hoog t.o.v. laag gelokaliseerde bladeren van *A. annua*. In een derde proteoom analyse, deden we onderzoek naar de proteïnen die aanwezig zijn in een chloroform extract van *A. annua*. De klierharen worden beschouwd als de "artemisinine fabrieken" van *A. annua*. Hoge bladeren produceren meer artemisinine dan lage bladeren. Proteïnen die meer tot expressie komen in deze stalen zijn dus mogelijks betrokken in de productie van artemisinine. Door middel van chloroform extractie, wordt de inhoud van de blaasjes op de klierharen geëxtraheerd. Deze blaasjes bevatten mogelijks proteïnen die betrokken zijn in de productie van artemisinine.

Mogelijks interessante proteïnen werden geïdentificeerd door middel van nano-vloeistofchromatografie gekoppeld aan tandem massaspectrometrie. De fragmentatie spectra werden vergeleken met de publieke MassaSpectrometrie proteïne sequentie DataBank (MSDB) en met 3 EST-banken (één van de bloemknoppen van *A. annua*, één van de klierharen op de bloemknoppen en één die resulteert na cDNA subtractie van de 2 voorgenoemde plantweefsels). Om een selectie te maken van de meest verdedigbare resultaten, werden de proteoom data vergeleken met een cDNA AFLP analyse van stalen van *A. annua* bladeren, genomen op verschillende tijdspunten gedurende een periode van 72 uur na blootstelling aan jasmonaat. Deze vergelijking heeft geleid tot een lijst van EST-kandidaten die interessant zijn voor verder onderzoek.

We ontwikkelden een kwantificatiemethode voor artemisinine en zijn bioprecursors om te kunnen nagaan welke planthormonen een verhoogde artemisinine productie teweeg brengen, om te kunnen nagaan of er een correlatie is tussen artemisinine inhoud van *A. annua* bladeren en het aantal klierharen op deze bladeren, en om te kunnen nagaan of getransformeerde planten of kiemculturen een verhoogde artemisinineproductie vertonen. Een recuperatie van >97% voor alle analyten werd bekomen door één gram vers plantmateriaal gedurende 1 minuut onder te dompelen in chloroform. Dit resultaat ondersteunt de hypothese dat artemisinine en sommige van zijn structurele analogen volledig gelokaliseerd zijn in de subcuticulaire ruimte van de blaasjes op de oppervlakte van de bladeren. We valideerden het gebruik van dit chloroform extract, zonder bijkomende staalvoorbereidende stappen, voor kwantitatieve HPLC – ESI – QTOF MS/MS. Er werd geen ionisatie onderdrukking (matrix effect) door interferentie met andere componenten gedetecteerd. Voor elke concentratie binnen het bereik van de standaard curve (0.1 tot 3.00 µg/ml), lag de accuraatheid tussen 85% and 115%. De dag-tot-dag variatie voor de analyse van *A. annua* stalen was <20%. We kunnen concluderen dat deze zeer korte extractieprocedure een goed alternatief is voor andere arbeidsintensieve procedures.

# APPENDICES

# Appendix 1

Identified proteins in the trichomes-*versus*-whole leafs proteome study

*Note:*
*Protein spots that identified as histones, chaperonins, proteasomes, DNA binding proteins, porins, Ribulose-1,5-bisphosphate carboxylase and ATPases, were omitted from the list.*

| Code spot | D | I | Mascot Score | Proteïne | Code Proteïne | Interpro | GSTSUB | AAGST | AAFB |
|---|---|---|---|---|---|---|---|---|---|
| 3409 | 194 | 184 | 0 | | | | | | 005-F08: metallothionein 1 [Aster tripolium] |
| | | | | | | | | | *009-D09: At1g05600 [Arabidopsis thaliana]* |
| 4908 | 151 | 161 | 101 | Methionine synthase.- Glycine max (Soybea | Q71EW8_SOYBN | Methionine synthase | | | |
| | 151 | 161 | 55 | Hypothetical protein.- Medicago truncatula ( | Q1SW13_MEDTR | Hypothetical | | | |
| 5826 | 238 | 170 | 103 | Cytosolic phosphoglucomutase.- Populus to | Q6S3D6_9ROSI | Phosphoglucomutase and phosphomannon | 055 C09 | 051-C10 | |
| 7109 | 92 | 96 | 97 | Putative 33kDa oxygen evolving protein of p | Q943W1_ORYSA | Photosystem II manganese-stabilizing prote | | 021-A02: pathogenesis-related protein [Zinr | 007-G01: Oxygen-evolving enhancer protei |
| | | | | | | | | | 010-B08: pathogenesis-related protein [Zinr |
| 8419 | 168 | 160 | 36 | PSA17329 NID: - Pisum sativum | CAA76741 | Calreticulin/calnexin | 015 E02 | | 007-G01: Oxygen-evolving enhancer protei |
| | | | | | | | 047 D03: heat stress transcription factor [Ly | | |
| 8909 | 29 | 231 | 75 | Protein disulfide-isomerase (EC 5.3.4.1) pre | ISAASS | Disulphide isomerase / Thioredoxin | 073 E11: putative gag-pol polyprotein [Solar | 050-D08: human RAN binding protein 16-lik | |
| | | | | | | | 045 D10: retrotransposon protein, putative, | 032-D10: 3-beta-hydroxy-delta5-steroid deh | |
| 3316 | 250 | 219 | 0 | | | | 009 G05: amorpha-4,11-diene synthase [Ar | | |
| 5822 | 193 | 189 | 198 | AP004338 NID: - Oryza sativa (japonica cul | BAC83515 | Fumarate reductase/succinate dehydrogena | | | 014-D08: WAV2; catalytic [Arabidopsis thali |
| | 193 | 189 | 50 | Putative CCT chaperonin gamma subunit.- ( | Q5Z6U5_ORYSA | Chaperonin | | | |
| 8209 | 118 | 233 | 66 | Pathogenesis-related protein.- Zinnia elegar | Q8LNX9_ZINEL | Bet v I allergen | 061 H06: ribosomal protein L14 [Helianthus | 021-A02: pathogenesis-related protein [Zinr | 014-H08: pathogenesis-related protein [Zinr |
| | | | | | | | | | 014-A10: pathogenesis-related protein [Zinr |
| | | | | | | | | | 010-B08: pathogenesis-related protein [Zinr |
| 9505 | 33 | 245 | 89 | Heat shock protein 80 (Fragment).- Solanur | Q8LRU7_SOLTU | Heat shock protein | 007 F05: Ycf3 [Ranunculus macranthus] | 029-G08: cytochrome c oxidase subunit 6b- | 003-D02 |
| | 33 | 245 | 87 | Sequence 65 from Patent WO0168863 (Fra | CAC93759 | Heat shock protein | | 050-B05: RPL15 [Arabidopsis thaliana] dbj\| | 001-F10: unknown protein [Arabidopsis thal |
| 1204 | 241 | 192 | 0 | | | | 056 F02: Integrase, catalytic region [Medica | 006-F04 | |
| | | | | | | | | 047-G07 | |
| 2506 | 190 | 202 | 118 | G protein beta subunit-like protein.- Solanur | Q38JI7_SOLTU | WD40 / Quinonprotein alcohol dehydrogena | | | 012-E01: NADPH quinone oxidoreductase-l |
| | 190 | 202 | 110 | GTP-binding regulatory protein beta chain h | T02300 | WD40 | | | |
| | 190 | 202 | 91 | Guanine nucleotide regulatory protein - rape | S48839 | | | | |
| 3818 | 196 | 180 | 71 | 43.6K hypothetical protein F10O5.4 - Arabid | E86480 | Bacterial surface antigen | | 047-D07 | |
| | 196 | 180 | 71 | Outer envelope membrane protein homolog | T12975 | Bacterial surface antigen / Chloroplast enve | | | |
| 5703 | 22 | 218 | 40 | Cytochrome P450.- Vigna radiata | Q9LKH7_9FABA | Cytochrome_P450 | | 032-A11: monodehydroascorbate reductase | |
| 5824 | 229 | 188 | 98 | Succinate dehydrogenase (EC 1.3.99.1) flav | T51815 | Fumarate reductase/succinate dehydrogena | | | |
| 6816 | 138 | 122 | 173 | Malate oxidoreductase (malic enzyme) [imp | E84508 | Malic enzyme | 034 F04 | 028-A07: 2,3-bisphosphoglycerate-indepen | 001-G05: cyclin-dependent protein kinase [ |
| | 138 | 122 | 50 | Phosphoglycerate mutase (EC 5.4.2.1), 2,3- | S44373 | Phosphoglucomutase / Metalloenzyme | | 021-F02: cofactor-independent phosphogly | |
| | 138 | 122 | 47 | Hypothetical protein OSJNBa0077M12.113. | Q8H334_ORYSA | Kinase | | 018-B10: caffeic acid O-methyltransferase I | |
| 2711 | 54 | 240 | 145 | Glyceraldehyde-3-phosphate dehydrogenas | Q8VWP4_CAPAN | Glyceraldehyde 3-phosphate dehydrogenas | | 034-H10: phi-1 [Nicotiana tabacum] | 022-D02: glyceraldehyde-3-phosphate dehy |
| | 54 | 240 | 101 | Putative RNA-binding protein.- Arabidopsis | Q8LAU2_ARATH | NAD-dependent epimerase/dehydratase | | 020-B11 | 007-E08: glyceraldehyde-3-phosphate dehy |
| | 54 | 240 | 84 | RNA binding protein, putative, expressed.- ( | Q2QSR7_ORYSA | RNA | | 030-G08: putative cinnamyl alcohol dehydro | 017-H08: hypothetical protein MtrDRAFT_A |
| 2805 | 3 | 178 | 274 | Catalase 3 (EC 1.11.1.6).- Helianthus annu | Q9M503_HELAN | Catalase | 072 E07: CPRD12 protein [Vigna unguicula | 029-C10: catalase 2 [Helianthus annuus] | 009-A04: OSJNBb0024F06.20 [Oryza sativa |
| | 3 | 178 | 91 | AlaT1.- Vitis vinifera (Grape). | Q45RS3_VITVI | 1-aminocyclopropane-1-carboxylate syntha | 043 H03: Integrase, catalytic region [Medica | | |
| | | | | | | | 001 H08: glutathione peroxidase [Helianthu | | |
| 3508 | 3 | 178 | 31 | Hypothetical protein OSJNBa0077M12.113. | Q8H334_ORYSA | Kinase | | | |
| 4714 | 156 | 133 | 133 | DnaK-type molecular chaperone LIM18 - tru | JC2215 | Heat shock protein | | | |
| | 156 | 133 | 57 | GDHB glutamate dehydrogenase (EC 1.4.1 | Q1HDV6_VITVI | Glu/Leu/Phe/Val dehydrogenase | | | |
| 6107 | 119 | 246 | 53 | Emb\|CAA71173.1 (Hypothetical protein At5 | Q9LTE1_ARATH | Heavy metal transport/detoxification protein | | | 014-A11: ribulose-1,5-bisphosphate carbox |
| 4702 | 64 | 234 | 327 | Phosphoglycerate kinase.- Medicago trunca | Q1SNT1_MEDTR | Phosphoglycerate kinase | 072 E07: CPRD12 protein [Vigna unguicula | 010-E10: cytosolic phosphoglycerate kinase | 015-A12: protein binding / serine-type endo |
| | | | | | | | 026 A03: GCPE protein [Catharanthus rose | 015-H05 | 004-B03: expressed protein [Oryza sativa (j |
| | | | | | | | | 025-E08 | 008-E01 |
| | | | | | | | | 001-A01: ribosomal protein L37 [Glycine ma | |
| | | | | | | | | 055-G01: fructokinase-like [Solanum tubero | |
| 6708 | 246 | 186 | 65 | Phosphoglycerate kinase.- Medicago trunca | Q1SNT1_MEDTR | Phosphoglycerate kinase | | 032-A11: monodehydroascorbate reductase | |
| 8809 | 89 | 130 | 397 | Protein At4g24190.- Arabidopsis thaliana (M | Q1NZ06_ARATH | Cytokine / ATPase / Endoplasmic reticulum | | | 011-C10 |
| | 89 | 130 | 53 | Lipoxygenase (EC 1.13.11.12).- Adelostem | Q4FCM5_9GENT | Lipoxygenase | | | |
| 2509 | 174 | 172 | 0 | | | | | | 022-F12 |
| 7113 | 145 | 132 | 51 | Putative polynucleotide phosphorylase.- Ory | Q69LE7_ORYSA | RNA | 062 G09: amorpha-4,11-diene synthase [Ar | 032-F11: MADS-box transcription factor CD | |
| | | | | | | | 009 C06 | | |
| | | | | | | | 026 C12 | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | 063 C09 | | |
| 8905 | 47 | 211 | 535 | Chaperonin 60 alpha subunit.- Canavalia lin | Q9ZTV1_CANLI | Chaperonin | 073 F09 | | 006-E09: early tobacco anther 1 [Medicago |
| | 47 | 211 | 69 | Glycosyl transferases-like protein (AT3g159 | Q9LSB5_ARATH | Cytochrome b5 / Glycosyl transferase | | | |
| 9105 | 63 | 168 | 0 | | | | | 021-A02: pathogenesis-related protein [Zinr | |
| | | | | | | | | 030-G06: 1-deoxyxylulose 5-phosphate syn | |
| | | | | | | | | 030-B02: peroxiredoxin [Hyacinthus orienta | |
| 1622 | 56 | 225 | 57 | OSJNBa0053K19.11 protein.- Oryza sativa | Q7XPR2_ORYSA | Glycine cleavage T-proteins | | | |
| 3703 | 105 | 206 | 59 | Aspartate aminotransferase-like.- Solanum | Q2XTE6_SOLTU | Aspartate aminotransferase | | 022-E11: hypothetical protein MtrDRAFT_A | |
| | 105 | 206 | 73 | NADH-ubiquinone oxidoreductase 49 kDa s | NUCM_ARATH | NADH dehydrogenase / NADH-ubiquinone | | | |
| 5817 | 232 | 212 | 126 | NADH2 dehydrogenase (ubiquinone) (EC 1 | S52737 | NADH dehydrogenase / NADH-ubiquinone | 042 B11: RuBisCo large subunit [Cucumis s | | |
| | 232 | 212 | 115 | Transketolase (EC 2.2.1.1) precursor, chlor | T09015 | Transketolase (DXP synthase) | 047 A04: glucose acyltransferase [Solanum | | |
| 9316 | 71 | 119 | 179 | Putative 14-3-3 protein.- Zea mays (Maize). | Q5XPI0_MAIZE | 14-3-3 (protein binding) | | 032-B11: 70 kDa heat shock cognate protei | 014-F05 |
| | 71 | 119 | 73 | Plastid-lipid-associated protein pap - comm | T03635 | Plastid lipid-associated proteins (PAPs) / fil | | 007-B04: 14-3-3-like protein [Helianthus an | |
| | 71 | 119 | 52 | Heat shock protein 70 (Hsc70-5) - Arabidop | T49939 | Chaperonin | | 014-H02: 3-oxoacyl-[acyl-carrier protein] re | |
| 5823 | 245 | 228 | 159 | AP004338 NID: - Oryza sativa (japonica cul | BAC83515 | Fumarate reductase/succinate dehydrogena | | | 014-D08: WAV2; catalytic [Arabidopsis thali |
| | 245 | 228 | 44 | Putative CCT chaperonin gamma subunit.- | Q5Z6U5_ORYSA | Chaperonin | | | |
| 2706 | 59 | 204 | 204 | NADH-glutamate dehydrogenase (EC 1.4.1 | Q8W1X4_LYCES | Glu/Leu/Phe/Val dehydrogenase | | 008-G03: chloroplast photosystem I reactio | |
| 3417 | 179 | 213 | 69 | Putative peroxidase (EC 1.11.1.7) (Fragmer | Q5QQS6_ZINEL | Peroxidase | | | |
| 4907 | 172 | 173 | 58 | NADH2 dehydrogenase (ubiquinone) (EC 1 | S52737 | NADH dehydrogenase / NADH-ubiquinone | | | |
| 5706 | 46 | 238 | 461 | Phosphoglycerate kinase-like.- Solanum tub | Q2V9B3_SOLTU | Phosphoglycerate kinase | 063 D12 | 029-C10: catalase 2 [Helianthus annuus] | 004-B03: expressed protein [Oryza sativa (j |
| | 46 | 238 | 68 | DP000009 NID: - Oryza sativa (japonica cul | ABF93789 | Glyceraldehyde 3-phosphate dehydrogenas | 001 C08 | 027-D09 | 022-D12: unknown protein [Arabidopsis tha |
| | 46 | 238 | 60 | NB-ARC domain containing protein, expres | Q2QPV5_ORYSA | Disease resistance | 003 F04: polyprotein [Cynara scolymus] | 027-C08 | |
| | | | | | | | | 010-E10: cytosolic phosphoglycerate kinase | |
| | | | | | | | | 035-B09 | |
| | | | | | | | | 026-D11 | |
| 5917 | 19 | 208 | 139 | Transketolase (EC 2.2.1.1) precursor - pota | S58083 | Transketolase (DXP synthase) | | 013-H10: thioredoxin/transketolase fusion p | |
| | | | | | | | | 051-B03: cytosolic ascorbate peroxidase [N | |
| 6601 | 62 | 239 | 105 | Fructose-bisphosphate aldolase (EC 4.1.2.1 | ADSPAP | Fructose-bisphosphate aldolase | | 005-H04 | 017-H03: plastid fructose bisphosphate aldo |
| | 62 | 239 | 97 | Putative fructose bisphosphate aldolase.- T | Q2PER4_TRIPR | Fructose-bisphosphate aldolase | | 021-E09: fructose-bisphosphate aldolase [A | 013-C07: plastidic aldolase NPALDP1 [Nico |
| | | | | | | | | | 017-B05: unknown protein [Arabidopsis tha |
| 1312 | 140 | 162 | 0 | | | | 035 A10: Lipolytic enzyme, G-D-S-L [Medic | | |
| 1609 | 40 | 187 | 234 | Putative NADH-ubiquinone oxireductase (H | Q9SK66_ARATH | NAD-dependent epimerase/dehydratase | 011 C12 | 005-C12: malate dehydrogenase [Vitis vinife | 004-F11: putative mitochondrial NAD-deper |
| | 40 | 187 | 122 | Malate dehydrogenase (EC 1.1.1.37) precu | T51311 | L-lactate/malate dehydrogenase | 042 B11: RuBisCo large subunit [Cucumis s | 025-F12: succinyl-CoA ligase alpha 1 subu | 011-B10: rbcL [Oryza sativa (indica cultivar- |
| 1807 | | | 285 | Glycine hydroxymethyltransferase (EC 2.1.2 | S40212 | Glycine hydroxymethyltransferase | | 016-F12: At4g28450/F20O9_130 [Arabidop | 003-E12: glycine hydroxymethyltransferase |
| | | | 56 | Catalase 4 (EC 1.11.1.6).- Helianthus annu | Q9M502_HELAN | Catalase | | | |
| 3612 | 230 | 223 | 213 | Malate dehydrogenase (EC 1.1.1.37).- Lupi | Q8GZN3_LUPAL | L-lactate/malate dehydrogenase | | 030-G08: putative cinnamyl alcohol dehydr | 007-B03: malate dehydrogenase [Plantago |
| | 230 | 223 | 102 | Glyceraldehyde-3-phosphate dehydrogenas | Q8VXM5_CHAVU | Glyceraldehyde 3-phosphate dehydrogenas | | 008-G09: glyceraldehyde 3-phosphate dehy | |
| | | | | | | | | 029-E12: malate dehydrogenase [Plantago | |
| | | | | | | | | 030-G04 | |
| 4822 | 201 | 207 | 112 | AP004338 NID: - Oryza sativa (japonica cul | BAC83515 | Fumarate reductase/succinate dehydrogena | | | 015-C11: BiP-isoform D [Glycine max] |
| 8302 | 13 | 242 | 107 | Chlorophyll a/b-binding protein Lhcb2 [impo | T52322 | Chlorophyll A-B binding protein | | | |
| | 13 | 242 | 74 | S46215 NID: - Lactuca sativa | AAB23371 | Triosephosphate isomerase | | | |
| | 13 | 242 | 53 | NADH:ubiquinone oxidoreductase-like.- Sol | Q38M79_SOLTU | NADH dehydrogenase / NADH-ubiquinone | | | |
| 2815 | 219 | 158 | 226 | Formate-tetrahydrofolate ligase, FTHFS.- M | Q1SCS3_MEDTR | Formate-tetrahydrofolate ligase | | 048-A12: putative yeast pheromone recepto | |
| 4717 | 236 | 229 | 53 | Acyl-CoA dehydrogenase.- Medicago trunca | Q1RWA6_MEDTR | Acyl-CoA dehydrogenase | | | |
| | 236 | 229 | 44 | Glyceraldehyde-3-phosphate dehydrogenas | DEPMNB | Glyceraldehyde 3-phosphate dehydrogenas | | | |
| 5825 | 239 | 175 | 100 | Succinate dehydrogenase (EC 1.3.99.1) fla | T51815 | Fumarate reductase/succinate dehydrogena | 042 B11: RuBisCo large subunit [Cucumis s | | |
| | 239 | 175 | 54 | Cytosolic phosphoglucomutase.- Populus to | Q6S3D6_9ROSI | Phosphoglucomutase and phosphomannon | | | |
| 1718 | 162 | 128 | 120 | Aspartate transaminase (EC 2.6.1.1), cytos | T14311 | Aspartate aminotransferase | | | 016-B11: putative acid phosphatase [Oryza |
| | | | | | | | | | 023-G12: polygalacturonase inhibitor [Actin |
| 5210 | | | 0 | | | | | | 007-H11: cytosolic ascorbate peroxidase [F |
| 6108 | 110 | 221 | 0 | | | | 055 C09 | | |
| | | | | | | | 060 H09 | | |
| 6819 | | | 0 | | | | | 028-A07: 2,3-bisphosphoglycerate-indepen | 007-D07 |
| 3311 | 170 | 197 | 104 | Carbonic anhydrase (Fragment).- Striga asi | Q1PG92_9LAMI | Carbonic anhydrase | | | |
| 3724 | 127 | 149 | 0 | | | | | 034-B07: unknown protein [Arabidopsis tha | |

# Appendix 2

Identified proteins in the upper-*versus*-lower leafs proteome study

*Note:*
*Protein spots that identified as histones, chaperonins, proteasomes, DNA binding proteins, porins, Ribulose-1,5-bisphosphate carboxylase and ATPases, were omitted from the list.*

| Code spot | D | I | Mascot Score | Proteïne | Code Proteïne | Interpro | GSTSUB | AAGST | AAFB |
|---|---|---|---|---|---|---|---|---|---|
| 2505 | 22 | 20 | 247 | Malate dehydrogenase (EC 1.1.1.37) precu | T51311 | L-lactate/malate dehydrogenase | 042 B11: RuBisCo large subunit [Cucumis | 005-C12: malate dehydrogenase [Vitis vir | 004-F11: putative mitochondrial NAD-dep |
| 5207 | 16 | 22 | 0 | | | | | 012-G03: peroxiredoxin [Ipomoea batatas<br>055-F02: putative peptidylprolyl isomeras<br>048-F10 | 007-G01: Oxygen-evolving enhancer prot<br>007-H09: calcium ion binding [Arabidopsis |
| 701 | 31 | 34 | 342 | Glycolate oxidase.- Brassica napus (Rape). | Q3L1H0_BRANA | Alpha-hydroxy acid dehydrogenase | | | 015-D05: glycolate oxidase/ oxidoreducta<br>007-D02: BURP [Medicago truncatula] |
| 4207 | 20 | 19 | 0 | Putative Rieske Fe-S protein (Fragment).- C | Q6E4A8_CYNDA | Rieske iron-sulfur protein (cytochrome C lin | | | 015-H07: chloroplast Rieske FeS protein |
| 6 4205 | 18 | 37 | 141 | Photosystem II oxygen-evolving complex pr | JS0771 | Photosystem II oxygen evolving complex pr | | | |

# Appendix 3

Identified proteins in the one-minute chloroform extract proteome study

*Note:*
*Protein spots that identified as histones, chaperonins, proteasomes, DNA binding proteins, porins, Ribulose-1,5-bisphosphate carboxylase and ATPases, were omitted from the list.*

| Code spot | Mascot Score | Proteïne | Code Proteïne | Interpro | GSTSUB | AAGST | AAFB |
|---|---|---|---|---|---|---|---|
| G1 | 0 | | | | 025 G09: IVD (ISOVALERYL-COA-DEHYD | 028-G03 | |
| G15 | 48 | Beta-galactosidase.- Sandersonia aurantiac | Q6EM02_SANAU | Glycoside hydrolase | 044 H09 | 026-D02: glucose-6-phosphate isomerase [ | 021-F11 |
| G15 | 0 | Gcyc (Fragment).- Goyazia rupicola. | Q6UQK6_9LAMI | transcription factor | | | 023-D06 |
| G18 | | | | | | 010-C02: caffeic acid O-methyltransferase [ | |
| G19 | 70 | Polygalacturonase-like protein.- Fragaria an | Q84LI7_FRAAN | Glycoside hydrolase | | 010-C02: caffeic acid O-methyltransferase [ | 013-B01 |
| G22 | 0 | | | | | 031-F06: lipid transfer protein precursor [Gc | |
| G26 | 51 | LEXYL1 protein.- Lycopersicon esculentum | Q76MS5_LYCES | Glyceraldehyde 3-phosphate dehydrogenas | | | |
| G27 | 75 | LEXYL1 protein.- Lycopersicon esculentum | Q76MS5_LYCES | Glyceraldehyde 3-phosphate dehydrogenas | | | |
| G28 | 127 | LEXYL1 protein.- Lycopersicon esculentum | Q76MS5_LYCES | Glyceraldehyde 3-phosphate dehydrogenas | | 055-G02 | |
| G28 | 67 | beta-xylosidase-like protein - Arabidopsis th | T49925 | Glycoside hydrolase | | | |
| G29 | 132 | LEXYL1 protein.- Lycopersicon esculentum | Q76MS5_LYCES | Glyceraldehyde 3-phosphate dehydrogenas | | | |
| G29 | 69 | beta-xylosidase-like protein - Arabidopsis th | T49925 | Glycoside hydrolase | | | |
| G31 | 0 | | | | | | 023-G12: polygalacturonase inhibitor [Actin |
| G32 | 0 | | | | | 005-H03 | |
| G33 | 0 | | | | | 005-H03 | |
| G37 | 0 | | | | 026 C12 | | 009-D09: At1g05600 [Arabidopsis thaliana] |
| G40 | | | | | | 020-B07: putative hypersensitive-induced re | 004-H06 |
| G48 | 0 | | | | 026 C12 | 001-B10: beta-ketoacyl-ACP synthase I [Pe | 007-B06 |
| G48 | | | | | 005 E12: At1g19360 [Arabidopsis thaliana] | 006-F12: isoflavone-7-o-methytransferase 9 | 021-A09 |
| G48 | | | | | 052-G06: gag-pol polyprotein-related [Medi | 034-E02: nodule-enhanced malate dehydro | 013-C09 |
| G49 | 0 | | | | | 001-B10: beta-ketoacyl-ACP synthase I [Pe | |
| G56 | 0 | | | | | 034-D11: poly(A)-binding protein [Nicotiana | |
| G56 | | | | | | 031-B04: O-methyltransferase/ S-adenosylr | |
| G56 | | | | | | 010-C02: caffeic acid O-methyltransferase [ | |
| G57 | | | | | | 055-G02 | |
| G66 | 56 | Polygalacturonase-like protein.- Fragaria an | Q84LI7_FRAAN | Glycoside hydrolase | | | |
| G68 | 0 | | | | 047 H03 | 037-B12: violaxanthin deepoxidase [Chrysa | |
| G68 | | | | | 053 C08: cyclopropane fatty acid synthase | 014-C03: transaldolase [Solanum tuberosur | |
| G68 | | | | | 026 C12 | | |
| G79 | 0 | | | | | 010-C05: hypothetical protein MtrDRAFT_A | 017-D09: chloroplast hypothetical protein [Z |
| G81 | 64 | Galactan:galactan galactosyltransferase 1.- | Q6TW99_AJURE | Glycoside hydrolase | | | |
| G91 | 0 | | | | | | 013-B09: transposon protein, putative, Pon |
| O1 | 0 | | | | 009 C06 | 002-G03: clp-like energy-dependent proteas | 022-E08: unknown [Hyacinthus orientalis] |
| O1 | | | | | 033-B03 | 001-B10: beta-ketoacyl-ACP synthase I [Pe | |
| O1 | | | | | 047-D11 | 032-F11: MADS-box transcription factor CD | |
| O10 | 0 | | | | 054 A02 | | 006-D08 |
| O10 | | | | | | | 010-A06: thiosulfate sulfurtransferase [Datis |
| O13 | 0 | | | | 039-B12 | 008-H10: plastid ribosomal protein S9 precu | |
| O14 | | | | | 039-B12 | | 020-F06: putative thioredoxin peroxidase 1 |
| O8 | 0 | | | | | 015-G03: synptobrevin-related protein [Pyru | 006-D08 |
| R141 | 0 | | | | | | *009-D09: At1g05600 [Arabidopsis thaliana]* |
| R16 | 51 | Polygalacturonase-like protein.- Fragaria an | Q84LI7_FRAAN | Glycoside hydrolase | | | |
| R42 | 50 | MTD2.- Medicago truncatula (Barrel medic). | Q9LLM2_MEDTR | Zinc finger | | | |
| R46 | 0 | | | | 020 D11 | | |
| R55 | 0 | | | | 058-A04 | 012-C06: fiber protein Fb19, putative, expre | |
| R71 | 47 | Hypothetical protein OSJNBa0077M12.113. | Q8H334_ORYSA | transferase | 061 C02: polyprotein [Cynara scolymus] | | 013-A04 |
| R71 | 0 | | | | 043 G08: (3R)-linalool synthase [Artemisia a | | 014-H05: Chlorophyll a-b binding protein, cl |
| R78 | | | | | | | 001-A01 |
| R79 | | | | | | 010-C02: caffeic acid O-methyltransferase [ | |
| R125 | | | | | | 010-C02: caffeic acid O-methyltransferase [ | |
| R126 | | | | | | | 001-A01 |
| R129 | | | | | | 010-C02: caffeic acid O-methyltransferase [ | |
| R143 | | | | | 053-G02 | | |

# Appendix 4

BLAST results for the trichomes-*versus*-whole leafs proteome study

*Note:*
*BLAST results with proteins that identified as histones, chaperonins, proteasomes, DNA binding proteins, porins, Ribulose-1,5-bisphosphate carboxylase and ATPases, were omitted from the list.*

**BLASTN 2.2.15 [Oct-15-2006]**

Database: TrichomesVsLeafs.txt
Posted date: Oct 23, 2006 4:22 PM
Number of letters in database: 399,230
Number of sequences in database: 432

---

**Query=** AA186 (430 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| 8302 Chlorophyll a/b binding protein – A. thaliana | 107 | 7e-025 |

**Query=** AA202 (495 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| 8302 Chlorophyll a/b binding protein – A. thaliana | 100 | 2e-022 |

**Query=** AA693 (215 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| 8302 Chlorophyll a/b binding protein – A. thaliana | 72 | 2e-014 |

**Query=** AA320 (553 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| 8302 Chlorophyll a/b binding protein – A. thaliana | 64 | 1e-011 |

**Query=** AA321 (535 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| 8302 Chlorophyll a/b binding protein – A. thaliana | 56 | 3e-009 |

**Query=** AA785 (323 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| 8302 Chlorophyll a/b binding protein – A. thaliana | 58 | 4e-010 |

- Matching for spot 8302:
  - **MASCOTscore107: chlorophyll a/b-binding protein Lhcb2 [imported] - Arabidopsis thaliana**
  - MASCOTscore74: S46215 NID: Lactuca sativa (Triosephosphate isomerase)
  - MASCOTscore53: NADH:ubiquinone oxidoreductase-like - Solanum tuberosum (Potato)
  - No match with GSTSUB, AAGST and AAFP
- This data supports the selection of AAFB 014_H05: Chlorophyl a/b binding protein (See Blast analysis for the chloroform extract proteome experiment)

**Query=** AA201 (113 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| GSTSUB_UP_047_A04_26OCT2005_032.ab1 | 46 | 5e-007 |

**Query=** AA244 (464 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAGST_UP_013_H10_25AUG2004_066.ab1 | 684 | 0.0 |
| 5817 Transketolase precursor, chloroplast – spinach | 90 | 2e-019 |

**Query=** AA323 (330 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAGST013_H10 | 98 | 5e-022 |
| 5817 Transketolase precursor, chloroplast – spinach | 68 | 5e-013 |

**Query=** AA332 (369 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAGST013_H10 | 155 | 3e-039 |
| 5817 Transketolase precursor, chloroplast – spinach | 141 | 4e-035 |

- Matching for spot GR5917 en R5917:
  - MASCOTscore139: Transketolase precursor - potato (DXP synthase)
  - AAGST_013_H10: Thioredoxin/transketolase fusion protein [synthetic construct]
  - AAGST_051_B03: Cytosolic ascorbate peroxidase [Nicotiana tabacum]
- Matching for spot GR5817 en R5817:
  - MASCOTscore126: NADH2 dehydrogenase (ubiquinone) 76K chain precursor - potato
  - **MASCOTscore115: Transketolase precursor, chloroplast – spinach (DXP synthase)**
  - MASCOTscore85: AT3g60750/T4C21_160 - Arabidopsis thaliana (DXP synthase)
  - GSTSUB_042_B11: RuBisCo large subunit [Cucumis sativus]
  - GSTSUB_047_A04: Glucose acyltransferase [Solanum berthaultii]

**Query=** AA243 (563 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| 6601 Fructose-bisphosphate aldolase precursor | 52 | 5e-008 |

**Query=** AA551 (144 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAFB013_C07 | 107 | 2e-025 |
| 6601 Fructose-bisphosphate aldolase precursor | 36 | 7e-004 |

**Query=** AA564 (217 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAFB_UP_017_H03_06DEC2005_017.ab1 | 272 | 1e-074 |
| AAFB013_C07 | 172 | 7e-045 |
| 6601 Fructose-bisphosphate aldolase precursor | 170 | 3e-044 |
| AAFB017_H03 | 153 | 6e-039 |
| AAFB003_E06 | 153 | 6e-039 |

**Query=** AA763 (136 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAFB013_C07 | 92 | 1e-020 |

**Query=** AA764 (163 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAFB013_C07 | 143 | 5e-036 |
| 6601 Fructose-bisphosphate aldolase precursor | 48 | 2e-007 |

- Matching for spot 6601:
  - **MASCOTscore105: Fructose-bisphosphate aldolase precursor, chloroplast - spinach**
  - AAGST_005_H04
  - AAGST_021_E09: Fructose-bisphosphate aldolase [Arabidopsis thaliana]
  - AAFB_017_H03: Plastid fructose bisphosphate aldolase [Stevia rebaudiana]
  - AAFB_013_C07: Plastidic aldolase NPALDP1 [Nicotiana paniculata]
  - AAFB_017_B05: unknown protein [Arabidopsis thaliana] gb|AAM47942.1| unknown protein [Arabidopsis thaliana] gb|AAL62377.1| unknown protein [Arabidopsis thaliana] dbj|BAB03030.1| unnamed protein product [Arabidopsis thaliana]

**Query=** AA282 (371 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAGST_UP_018_B10_26AUG2004_078.ab1 | 339 | 9e-095 |
| AAGST018_B10 | 50 | 1e-007 |

**Query=** AA610 (128 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAGST018_B10 | 44 | 2e-006 |

**Query=** AA633 (370 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAGST_UP_018_B10_26AUG2004_078.ab1 | 385 | e-108 |
| AAGST018_B10 | 62 | 3e-011 |

- Matching for spot 6816:
  - MASCOTscore173: malate oxidoreductase (malic enzyme) [imported] - Arabidopsis thaliana
  - MASCOTscore50: phosphoglycerate mutase (EC 5.4.2.1), 2,3-bisphosphoglycerate-independent - common tobacco Phosphoglucomutase / Metalloenzyme
  - MASCOTscore47: Hypothetical protein OSJNBa0077M12.113.- Oryza sativa (japonica cultivar-group). kinase
  - GSTSUB_ 034 F04
  - AAGST_028_A07: 2,3-bisphosphoglycerate-independent phosphoglycerate mutase (Phosphoglyceromutase) (BPG-independent PGAM) (PGAM-I) gb|AAA33499.1| 2,3-bisphosphoglycerate-independent phosphoglycerate mutase
  - AAGST_018_B10: caffeic acid O-methyltransferase II [Nicotiana tabacum]
  - AAGST_021_F02: cofactor-independent phosphoglyceromutase [Apium graveolens]
  - AAFB_001_G05: cyclin-dependent protein kinase [Arabidopsis thaliana] gb|AAC27476.1| putative PREG1-like negative regulator [Arabidopsis thaliana] gb|AAY57312.1| At2g44740 [Arabidopsis thaliana] gb|AAY17415.1| At2g44740 [Arabidopsis thaliana]

**Query=** AA340 (460 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| 8809 Protein At4g24190.- A. thaliana | 141 | 5e-035 |

**Query=** AA327 (225 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| 9505 Sequence 65 from Patent WO0168863 | 141 | 3e-035 |
| 8809 Protein At4g24190 - A. thaliana | 52 | 2e-008 |
| 9505 Heat shock protein 80 | 44 | 4e-006 |

- Matching for spot 9505:
  - **MASCOTscore89: Heat shock protein 80 (Fragment).- Solanum tuberosum (Potato).**
  - **MASCOTscore87: Sequence 65 from Patent WO0168863 (Fragment).- Lycopersicon esculentum (Tomato). (Heat shock protein like: reacts with geminivirus products)**
  - GSTSUB_007_F05: Ycf3 [Ranunculus macranthus] (large substrates, multiple functions)
  - AAGST_029_G08: cytochrome c oxidase subunit 6b-1 [Oryza sativa]
  - AAGST_050_B05: 50S ribosomal protein L15, chloroplast precursor [Arabidopsis thaliana]
  - AAFB_003_D02
  - AAFB_001_F10: unknown protein [Arabidopsis thaliana] gb|AAP88362.1| At5g02050 [Arabidopsis thaliana] emb|CAB82978.1| putative protein [Arabidopsis thaliana]
- Matching for spot 8809
  - **MASCOTscore397: Protein At4g24190.- Arabidopsis thaliana (cytokine / ATPase / Endoplasmic reticulum targeting sequence / Heat shock protein)**
  - MASCOTscore53: Lipoxygenase (EC 1.13.11.12) - Adelostemma gracillimum
  - AAFB_011_C10

**Query=** AA254 (616 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| 2805 1-aminocyclopropane-1-carboxylate synthase | 264 | 7e-072 |

**Query=** AA328 (107 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAGST029_C10 | 82 | 9e-018 |

- Matching for spot 2805:
  - MASCOTscore277: Catalase 3 (EC 1.11.1.6) - Helianthus annuus
  - **MASCOTscore91: AlaT1 - Vitis vinifera 1-aminocyclopropane-1-carboxylate synthase**
  - GSTSUB_ 043_H03: Integrase, catalytic region [Medicago truncatula]
  - GSTSUB_ 072_E07: CPRD12 protein [Vigna unguiculata] (oxidoreductase)
  - GSTSUB_ 001_H08: glutathione peroxidase [Helianthus annuus]
  - AAGST_029_C10: catalase 2 [Helianthus annuus]
  - AAFB_009_A04: OSJNBb0024F06.20 [Oryza sativa (japonica cultivar-group)]

- Matching for spot 5706:
  - MASCOTscore461: Phosphoglycerate kinase-like - Solanum tuberosum
  - MASCOTscore68: DP000009 NID: Oryza sativa (GAPDH)
  - MASCOTscore60: NB-ARC domain containing protein - Oryza sativa (Disease resistance)
  - GSTSUB_063_D12
  - GSTSUB_001_C08
  - GSTSUB_003_F04: polyprotein [Cynara scolymus]
  - AAGST_029_C10: catalase 2 [Helianthus annuus]
  - AAGST_027_D09
  - AAGST_010_E10: cytosolic phosphoglycerate kinase 1 [Populus nigra]
  - AAGST_027_C08
  - AAGST_035_B09
  - AAGST_026_D11
  - AAFB_004_B03: gb|AAT81723.1| striated muscle activator-like protein [Oryza sativa]
  - AAFB_022_D12: gb|ABE65868.1| auxin-responsive family protein [Arabidopsis thaliana]

**Query=** AC8 (136 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAGST_UP_032_A11_27AUG2004_095.ab1 | 151 | 2e-038 |

- Matching for spot 5703:
  - MASCOTscore40: Cytochrome P450.- Vigna radiata
  - AAGST_ 032_A11: monodehydroascorbate reductase [Mesembryanthemum crystallinum]
- Matching for spot 6708:
  - AAGST_ 032_A11: monodehydroascorbate reductase [Mesembryanthemum crystallinum]

**Query=** AC9 (867 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAGST_UP_008_G09_07MAY2004_067.ab1 | 52 | 7e-008 |

- Matching voor spot GR3612 en R3612:
  - MASCOTscore217: Malate dehydrogenase (EC 1.1.1.37).- Lupinus albus
  - MASCOTscore112: Putative GAPDH (Fragment).- Orobanche minor
  - AAGST_029_E12: malate dehydrogenase [Plantago major]
  - AAGST_008_G09: glyceraldehyde 3-phosphate dehydrogenase [Linum usitatissimum]
  - AAGST_030_G04
  - AAGST_030_G08: putative cinnamyl alcohol dehydrogenase [Zinnia elegans]
  - AAFB_014_007_B03: malate dehydrogenase [Plantago major]

# Appendix 5

BLAST results for the upper-*versus*-lower leafs proteome study

*Note:*
*BLAST results with proteins that identified as histones, chaperonins, proteasomes, DNA binding proteins, porins, Ribulose-1,5-bisphosphate carboxylase and ATPases, were omitted from the list.*

**BLASTN 2.2.15 [Oct-15-2006]**

Database: UpperVsLowerCDS.txt
    Posted date:  Oct 23, 2006  4:23 PM
  Number of letters in database: 44,506
  Number of sequences in database:  54

---

**Query=** AA296 (230 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAFB_UP_015_H07_06DEC2005_049.ab1 | 34 | 5e-004 |

**Query=** AA682 (261 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAFB_UP_015_H07_06DEC2005_049.ab1 | 34 | 6e-004 |

- Matching for spot 4207:
  - MASCOTscore 59: Putative Rieske Fe-S protein (Fragment).- Cynodon dactylon (Bermuda grass).
  - AAFB_015_H07: chloroplast Rieske FeS protein [Pisum sativum] sp|P26291|UCRIA_PEA Cytochrome b6-f complex iron-sulfur subunit, chloroplast precursor (Rieske iron-sulfur protein) (Plastohydroquinone:plastocyanin oxidoreductase iron-sulfur protein) (ISP) (RISP)

**Query=** AA786 (152 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAFB015_D05 | 96 | 1e-022 |
| 0701 Glycolate oxidase.- Brassica napus (Rape) | 96 | 1e-022 |

**Query=** AC11 (377 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| 0701 Glycolate oxidase.- Brassica napus (Rape) | 214 | 4e-058 |
| AAFB015_D05 | 198 | 2e-053 |

- Matching for spot 0701:
  - MASCOTscore 342: Glycolate oxidase.- Brassica napus (Rape).
  - AAFB_015_D05: glycolate oxidase/ oxidoreductase [Arabidopsis thaliana] ref|NP_188060.1| glycolate oxidase/ oxidoreductase [Arabidopsis thaliana] gb|AAL69528.1| AT3g14420/MOA2_2 [Arabidopsis thaliana] gb|AAL16164.1| AT3g14420/MOA2_2 [Arabidopsis thaliana] gb|AAK96642.1| AT3g14420/MOA2_2 [Arabidopsis thaliana] dbj|BAB01334.1| glycolate oxidase [Arabidopsis thaliana] sp|Q9LRR9|GOX2_ARATH Probable peroxisomal (S)-2-hydroxy-acid oxidase 2 (Glycolate oxidase 2) (GOX 2) (Short chain alpha-hydroxy acid oxidase 2)
  - AAFB_007_D02: BURP [Medicago truncatula]

# Appendix 6

BLAST results for the one-minute chloroform extract proteome study

*Note:*
*BLAST results with proteins that identified as histones, chaperonins, proteasomes, DNA binding proteins, porins, Ribulose-1,5-bisphosphate carboxylase and ATPases, were omitted from the list.*

**BLASTN 2.2.15 [Oct-15-2006]**

Database: ChloroformExtractCDS.txt
    Posted date:  Oct 23, 2006  4:22 PM
  Number of letters in database: 47,895
  Number of sequences in database:  61

---

**Query=** AA86 (309 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAGST_UP_020_B07_26AUG2004_061.ab1 | 531 | e-153 |
| AAGST020_B07 | 147 | 1e-037 |

- Matching for spot G40:
  - AAGST_020_B07: putative hypersensitive-induced response protein [Oryza sativa]
  - AAFB_004_H06

**Query=** AA104M (70 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| GSTSUB_UP_009_C06_20AUG2004_044.ab1 | 32 | 6e-004 |

**Query=** AA699 (265 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAGST032_F11 | 36 | 2e-004 |

- Matching for spot O1:
  - GSTSUB_009_C06
  - AAGST_002_G03: clp-like energy-dependent protease [Fritillaria agrestis]
  - AAGST_001_B10: beta-ketoacyl-ACP synthase I [Perilla frutescens]
  - AAGST_032_F11: MADS-box transcription factor CDM41 [Chrysanthemum x morifolium]
  - AAFB_022_E08: unknown [Hyacinthus orientalis]

- Matching for spot 7113
  - MASCOTscore51: Putative polynucleotide phosphorylase.- Oryza sativa (japonica cultivar-group).
  - GSTSUB_062_G09: amorpha-4,11-diene synthase [Artemisia annua]
  - GSTSUB_009_C06
  - GSTSUB_026_C12
  - GSTSUB_063_C09
  - AAGST_032_F11: MADS-box transcription factor CDM41 [Chrysanthemum x morifolium]

**Query=** AA282 (371 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAGST010_C02 | 36 | 3e-004 |

- Matching for spot G18, R79, R125 en R129:
  - AAGST_010-C02: caffeic acid O-methyltransferase [Rosa chinensis var. spontanea]

- Matching for spot G19:
  - MASCOTscore70: Polygalacturonase-like protein [Fragaria ananassaca]
  - AAGST_010-C02: caffeic acid O-methyltransferase [Rosa chinensis var. spontanea]
  - AAFB_013-B01

- Matching for spot G56:
  - AAGST_034-D11: poly(A)-binding protein [Nicotiana tabacum]
  - AAGST_031-B04: putative caffeoyl-CoA O-methyltransferase [Arabidopsis thaliana]
  - AAGST_010-C02: caffeic acid O-methyltransferase [Rosa chinensis var. spontanea]

Remark:
- E-value is higher than 0.0001
- 6 spots are linked to AAGST010_C02
- 3 AFLP sequences in the link to “caffeic acid O-methyltransferase” in the Blast analysis for the trichomes-*versus*-whole leaf experiment

**Query=** AA294 (105 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAGST_UP_010_C05_07MAY2004_043.ab1 | 48 | 2e-008 |
| AAGST010_C05 | 40 | 4e-006 |

**Query=** AA322 (69 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAGST_UP_010_C05_07MAY2004_043.ab1 | 98 | 1e-023 |
| AAGST010_C05 | 58 | 1e-011 |
| AAFB017_D09 | 46 | 4e-008 |

- Matching for spot G79:
  - AAGST_010-C05: hypothetical protein MtrDRAFT_AC151668g11v1 [Medicago truncatula]
  - AAFB_017-D09: chloroplast hypothetical protein [Zea mays] ref|YP_588293.1| chloroplast hypothetical protein [Zea mays subsp. mays]

**Query=** AA312M (177 letters)

| Sequences producing significant alignments: | Score (bits) | E Value |
|---|---|---|
| AAFB_UP_014_H05_06DEC2005_033.ab1 | 212 | 8e-058 |
| AAFB014_H05 | 90 | 8e-021 |

- Matching for spot R71:
  - MASCOTscore47: Hypothetical protein OSJNBa0077M12.113.- Oryza sativa (japonica cultivar-group).
  - GSTSUB_061_C02: polyprotein [Cynara scolymus]
  - GSTSUB_043_ G08: (3R)-linalool synthase [Artemisia annua]
  - AAFB_013_A04
  - AAFB_014_H05: Chlorophyll a-b binding protein, chloroplast precursor (LHCI type II CAB) gb|AAA33711.1| chlorophyll binding protein precursor prf||1503272A chlorophyll binding protein

# Dankwoord

Ik wil iedereen bedanken die op welke manier dan ook een bijdrage heeft geleverd aan dit onderzoek en aan mijn proefschrift. Zonder de inzet, bereidwilligheid, steun en hulp van veel mensen was dit onderzoek niet mogelijk geweest.

Mijn promotor, Prof. Dieter Deforce gaf mij de kans om dit onderzoek uit te voeren. Zijn grote enthousiasme voor wetenschappelijk onderzoek heeft ertoe geleid dat ik met uiteenlopende onderzoeken en technieken in contact kwam. De overvloed aan strategieën en ideeën om biotechnologische uitdagingen aan te pakken, prikkelden de wetenschapper in mij. Ik wens hem dan ook oprecht te bedanken voor de wetenschappelijke en persoonlijke groei die ik gedurende de laatste jaren mocht meemaken.

Prof. Damiaan Denys is een autoriteit binnen de wereld van angststoornissen en obsessieve-compulsieve stoornis. Hij concipieerde het plan om DNA polymorfismen te onderzoeken bij OCS patiënten. Hierdoor kwam ik in aanraking met de medische genetica, een onderzoeksdomein waarin ik graag verder onderzoek zou willen uitvoeren.

Alain Goossens, Lies Maes, Nancy Terryn, Prof. Dirk Inzé en Prof. Marc Van Montagu van het Vlaams Interuniversitair Instituut voor Biotechnologie en Prof. Pat Covello van het Plant Biotechnology Institute, Saskatoon, Canada wens ik te bedanken voor de samenwerking betreffende *Artemisia annua*. Zonder hen zou dit wetenschappelijk onderzoek nooit van de grond zijn gekomen. Alain, bedankt voor de wetenschappelijke discussies en het opzetten van de samenwerking met Prof. Pat Covello. Lies, speciaal bedankt voor de trichoomstalen en de prachtige microscopische opnames.

Mijn collega's wens ik bijzonder te bedanken voor een aangename periode in mijn loopbaan. Zij zorgden voor de nodige afleiding en gemeende belangstelling die ik nodig had tijdens het werk en de onvermijdelijke moeilijkheden. Ieder van hen heeft door technische, wetenschappelijke of administratieve ondersteuning op zijn/haar manier bijgedragen tot dit doctoraat. In het kader van dit proefschrift wens ik het hele DNA-fingerprintingteam, Kelly, Stijn, Maarten, Ben, Ivan en Kin-Jip speciaal the bedanken voor de hulp bij de genetische- en proteoom-analyses.

Ook voor de wetenschappelijke samenwerkingen buiten dit proefschrift wens ik een aantal mensen te bedanken: Ann-Sophie, Aline, Mado, Katleen, de mensen van het labo microbiologie (FFW), de reumatologie- en de infertiliteits-kliniek (UZ Gent), thesisstudenten en alle bloedgevers voor de zoektocht naar foetale cellen.

Sofie Vande Casteele was vooral betrokken bij de arbeidsintensieve praktijk van het proteïne identificatie werk en het analytische onderzoek. Haar gedrevenheid om de metingen op een wetenschappelijk verantwoorde manier uit te voeren en om de probleemgevoelige massaspectrometers, HPLC- en nanoLC toestellen draaiende te houden, was van onschatbare waarde.

Tot slot wens ik mijn familie te bedanken voor de interesse in mij bezigheden. Bieke bedank ik voor de haar onvoorwaardelijke steun. Zij heeft mij geholpen bij de keuze om vanuit de farmaceutische industrie terug te keren naar de ‘onderzoekswereld’; een keuze waar ik geen spijt van heb.