„Studies on the evolution of complex natural products biosynthetic pathways on basis of Taxol biosynthesis in plants and endophytic fungi"

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Diplom-Chemie-Ingenieur
Uwe Herbert Heinig
aus Freital, Sachsen

Berichter: Universitätsprofessor Dr. rer. nat. Rainer Fischer
Universitätsprofessor Dr. rer. nat. Marc Stadler

Tag der mündlichen Prüfung: 22.06.2012

Eidesstattliche Erklärung:

Hiermit bestätige ich die vorliegende Arbeit selbständig angefertigt zu haben und keine anderen Hilfsmittel und Quellen als die im Text erwähnten verwendet zu haben.

Aachen, Juni 2012

_____

(Uwe Heinig)

Table of contents

I Introduction

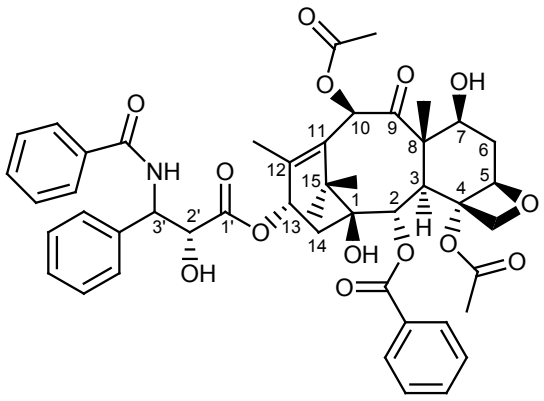
I.1 *Taxol – history, clinical impact and production*

During the 1950s, a joint scientific undertaking between the National Cancer Institute (NCI) and the United States Department of Agriculture (USDA) was initiated to screen natural products for the identification of novel anticancer drugs. This program led to the identification of Taxol and camptothecin, which became two of the world's most widely used anticancer drugs (SUFFNESS and WALL 1995). Taxol (generic name paclitaxel) (Figure I-1) is probably the better known of the two metabolites and its biosynthesis is the focus of this PhD thesis. The compound was first isolated from the bark of the pacific yew tree *Taxus brevifolia*. Its complex, highly oxygenated diterpenoid structure was elucidated in 1971 (WANI *et al.* 1971). Taxol has become one of the most successful treatments for a variety of cancers (including ovary, breast, lung, head and neck carcinomas and the AIDS-related Karposi's carcinoma) despite difficulties in generating a reliable supply (CRAGG *et al.* 1993; GOLDSPIEL 1997). Members of the genus *Taxus* are distributed throughout Asia, North and Central America and Europe, with all species and subspecies producing Taxol-like compounds, also referred to as taxoids (KINGSTON *et al.* 2002). Today, more than 350 structurally distinct taxoid compounds have been isolated (BALOGLU and KINGSTON 1999; ITOKAWA 2003).

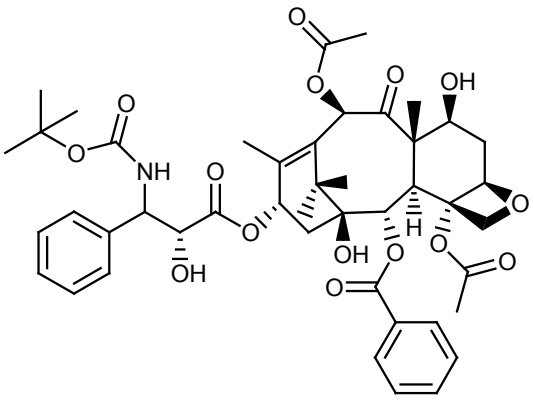
Soon after its isolation, Taxol was shown to have a unique mode of action based on shifting microtubule equilibrium towards assembly, resulting in abnormally stable microtubules that block the cell cycle in G_2/M phase (SCHIFF *et al.* 1979). Even today, 40 years later, only a few compounds with the same mode of action are known, e.g. epothilones (GOODIN 2008). Taxol underwent clinical trials in the 1980s and was approved by the FDA for treatment of refractory ovarian cancer in 1992 (SUFFNESS and WALL 1995). This was followed by approval for treatment of several additional cancers (OBERLIES and KROLL 2004).

Today, Taxol and its chemical derivative Taxotere (Figure I-1), modified at the C13 side chain, are among the most widely used anticancer drugs. Furthermore they are also used for other applications, including coronary heart disease, where the drug reduces the formation of scar tissue following balloon angioplasty (RAJA 2006; TANIMOTO *et al.* 2007). Taxol-

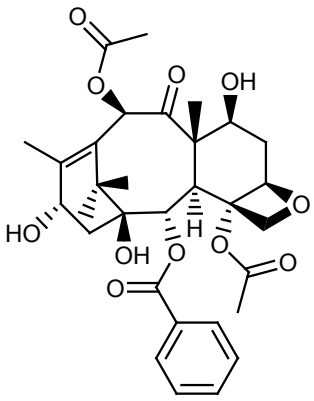
eluting stents (TAXUS® Express2®, Boston Scientific, Natick, MA, USA), which received FDA approval in 2003/2004, have shown to significantly decrease the risk of in-stent restenosis due to neointimal hyperplasia compared to bare-metal intracoronary stents. Due to the slow release of a cytostatic dose of Taxol over an extended period, the drug reduces the neointimal growth after stent deployment. Assuming that a growing number of coronary heart diseases will be treated with cardiovascular stents, the demand of Taxol will further increase (HTAY and LIU 2005; LASALA *et al.* 2006).



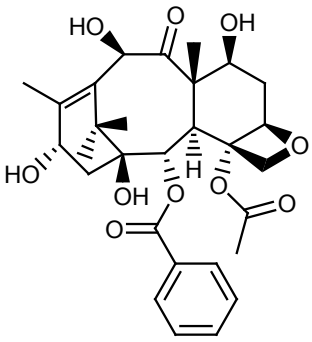
Taxol (Paclitaxel)



Taxotere (Docetaxel)



Baccatin III



10-Deacetylbaccatin III

Figure I-1: Chemical structures of Taxol (with atom numbers), Taxotere and the late Taxol-precursors Baccatin III and 10-Deacetylbaccatin III (HEINIG and JENNEWEIN 2009)

Supply has been a major challenge throughout the clinical development of Taxol. Taxol makes up only a minor proportion of the total taxoid content of *Taxus* trees. As mentioned above the compound was first isolated from yew bark, hence isolation from this natural source leads to the destruction of the tree. Taking into account that *Taxus* is a very slow growing

plant and the amount of Taxol is comparatively low in relation to other Taxoids, natural sources do not represent reliable production systems for Taxol. The commercial isolation of 1 kg of Taxol from *T. brevifolia* requires the bark of 2000-3000 very slow-growing trees (HARTZELL 1991; CROOM 1995; SUFFNESS and WALL 1995). Furthermore the yield of Taxol is highly dependent on the *Taxus* species. Some species such as *Taxus baccata* (the European yew tree) produce hardly any Taxol at all (NADEEM *et al.* 2002). Alternatively taxoids can also be isolated from the needles of *Taxus*. Although the content of Taxol is even lower than in the bark, high concentrations of late precursors can be isolated.

Therefore the major current source of Taxol and Taxotere is semisynthesis (HOLTON *et al.* 1995). The late precursors Baccatin III and 10-deacetylbaccatin III (Figure I-1) can be isolated from yew needles without killing the trees and can be modified with synthesized side chain molecules to obtain the desired products. This production system still relies on yew trees for precursor molecules and therefore depends on epigenetic and environmental factors. An alternative production strategy is the use of *Taxus* cell suspension cultures, obtained from the species *T. brevifolia* (GIBSON *et al.* 1993), *T. baccata* (SRINIVASAN *et al.* 1995) and *T. canadensis* (KETCHUM *et al.* 1999). These cell cultures produce biomass faster than *Taxus* trees and can be grown under reproducible conditions. Under optimized culture conditions and induction of production with methyl jasmonate, up to 23 mg/L/d of taxanes can be generated with a Taxol content of 13-20% (KETCHUM *et al.* 1999). These yields demonstrate the impressive biosynthetic capacity of *Taxus* cell cultures. However, sustaining such high rates of secondary metabolite production in plant cell culture is very difficult (DEUS-NEUMANN and ZENK 1984; HALL and YEOMAN 1987; MORRIS *et al.* 1989; PARR *et al.* 1990; SCHRIPSEMA and VERPOORTE 1992). Since 2002, Bristol-Meyers-Squibb Inc. has switched its sourcing to plant cell culture-derived Taxol (RITTER 2004). Additionally several total synthesis routes have been developed too, however at best providing a maximum yield of 2% of Taxol, hence not representing a useful alternative production platform (HOLTON *et al.* 1994a; HOLTON *et al.* 1994b; NICOLAOU *et al.* 1994; DANISHEFSKY *et al.* 1996; XIAO *et al.* 2003).

All together the production methods used today, although improved a lot over the times still are very difficult and costly regarding production itself but also with concern of purification of either Taxol or late precursors as Baccatin III from complex taxane mixtures. This discrepancy between demand and supply is the biggest challenge in clinical application of

Taxol. It has driven research into new production strategies, such as metabolic engineering of the yeast *Saccharomyces cerevisiae* (JENNEWEIN *et al.* 2005; DEJONG *et al.* 2006; ENGELS *et al.* 2008), *E. coli* (HUANG *et al.* 1998; AJIKUMAR *et al.* 2010) and different plant systems like *Arabidopsis thaliana* (BESUMBES *et al.* 2004) and the moss *Physcomitrella patens* (ANTEROLA *et al.* 2009) However, metabolic engineering of yeast for the total biosynthesis of Taxol or other advanced taxoids is extremely complex and still in its infancy. Today, the total fermentation of taxadiene has been achieved in significant amounts in *Saccharomyces cerevisiae* (ENGELS *et al.* 2008) and *E. coli* (AJIKUMAR *et al.* 2010). Hence, the establishment of recombinant microorganisms, like yeast or bacteria, offers great perspectives not only for the production of Taxol but also for other complex natural products and derivatives thereof (CHANG and KEASLING 2006).

Due to the problems related to supply an extensive search for alternative sources for Taxol and related Taxanes was initiated. This search led to the isolation of endophytic fungi, surprisingly having been shown to contain the identical natural products, in this case Taxol and Baccatin III after cultivation independently from their plant host (STIERLE *et al.* 1993). In general endophytes are bacteria and fungi that live within plants. They are defined as "microbes that colonize living, internal tissues of plants without causing any immediate, overt negative effects" (BACON *et al.* 2000). These species are considered to be a wellspring of novel secondary metabolites with significant potentials for medical use. The number of endophytic species is unknown, but may exceed one million, providing an extremely large pool of biological and hence biochemical diversity (DREYFUSS and CHAPELA 1994). Endophytic fungi have been shown to produce compounds with a range of properties, including antibiotics, antivirals, antioxidants, antidiabetic agents, immunosuppressive compounds, insecticidal products and anticancer agents (STROBEL and DAISY 2003; STROBEL *et al.* 2004; ALY *et al.* 2011). Besides natural products produced by the fungi exclusively, especially interesting are metabolites originally known from plants that are also observed in endophytes. Up to now there are only few examples for the occurrence of identical secondary metabolites in plants as well as in their associated endophytes. Examples are the natural topoisomerase I inhibitor camptothecin originally isolated from *Camptotheca acuminata* which also was found in endophytic fungi as *Nothapodytes foetida* and more recently in cultures of *Fusarium solani* (PURI *et al.* 2005; KUSARI *et al.* 2009a;

KUSARI *et al.* 2009b) or loline alkaloids from grasses (BLANKENSHIP *et al.* 2001). But the most prominent compound in this context is definitely Taxol.

Strobel and coworkers isolated an endophytic fungus from *T. brevifolia*, named *Taxomyces andreanae* in which organic extract after cultivation independently from its host Taxol as well as Baccatin III could be detected. After confirmation of the compound's identity as Taxol by the use of immunological assays with a polyclonal anti-taxane antibody, by thin layer chromatography and by HPLC-mass spectrometry the authors concluded from their data, that the fungus was able to produce the compound (STIERLE *et al.* 1993). Further investigation involving feeding studies with ^{14}C-labeled precursors supported the initial results (STIERLE *et al.* 1993; STROBEL *et al.* 1996). Starting with *Taxomyces andreanae* the collection of endophytic fungi from different yew species all over the world resulted in the isolation and identification of numerous "Taxol-synthesizing" fungi.

Several endophytic fungi have been isolated from *Taxus* species in Asia, including *Fusarium solani* from *T. celebica*, (CHAKRAVARTHI *et al.* 2008) the so-called fungus BT2, an endophyte of the genus *Cliocladium* from Indian yew tree (SREEKANTH *et al.* 2009) and three other unnamed fungi from *T. mairei* (GUO *et al.* 2006; ZHOU *et al.* 2007). In China, *Sporormia minima* and a fungus from the genus *Trichothecium* have been isolated from *T. wallichiana* (WANG *et al.* 2000; SHRESTHA *et al.* 2001) as well as another four endophytes obtained from *T. x media* and *T. yunnanensis* (ZHANG *et al.* 2008; ZHANG *et al.* 2009b). From the European yew tree *T. baccata* collected in northern Italy Caruso and co-workers detected taxanes in extracts of 15 endophytic fungi and 10 endophytic actinomycetes (CARUSO *et al.* 2000b). From other geographic locations for example 16 endophytic ascomycetes were isolated from *Taxus globosa* in Mexico (SOCA-CHAFRE *et al.* 2011) in which extracts taxanes were observed and four fungal isolates were described from *T. hicksii* collected in Canada (DAHIYA 1996).

These observations led to many questions regarding not only the function of taxanes for the fungi but especially about the purpose of this predicted biosynthetic ability for the organisms. Whereas other compounds having a known function, for example alkaloids in defending the plant host from parasites or there is a reasonable theory for the production of identical compounds in both distantly related species, for example for gibberellins in plants as natural growth hormones and in associated fungi as compounds that might be involved in parasitic infection inducing length growth there seems to be no obvious reason why endophytic fungi

should produce taxanes (BÖMKE and TUDZYNSKI 2009; ALY *et al.* 2011). Regarding the yields observed from *Taxomyces andreanae* and other endophytes from yew trees in comparison to production of *Taxus* itself the amounts of fungal taxanes are negligibly small. Hence, there is no obvious reason why the organisms should "waste" resources for the production of an extremely complex metabolite that is anyway present in much higher concentrations in their natural environment. Furthermore the ability of Taxol production is no feature of all isolated *Taxus* endophytes eliminating the general need of taxane production for an unknown reason to survive in the microhabitat inside the *Taxus* tree.

Even more surprising than isolation of "taxane producers" from *Taxus* spp. from various geographic locations was the discovery of taxanes in extracts of fungal endophytes isolated from other plants. *Pestalotiopsis* spp. supposed to synthesize the compounds, such as *Pestalotiopsis microspora* found on *T. wallichiana,* were found not only on yews but also on cypress trees, which do not produce taxoids (STROBEL *et al.* 1996; STROBEL 2002). This unexpected discovery led to an enlargement of the search for potentially "Taxol-producing" microbes beyond *Taxus* species. *Pestalotiopsis guepini* found on *Wollemia nobilis* and *Seimatoantlerium tepuiense* isolated from *Maguireothamnus speciosus* are examples of "Taxol-producing" endophytes from sources outside of the *Taxus* genus (STROBEL *et al.* 1997; STROBEL *et al.* 1999). Other examples in this context were the finding of two "Taxol producing" fungi on *Terminalia arjuna* (GANGADEVI and MUTHUMARY 2009b; a) and the isolation of various *Phyllosticta* species from different plant hosts, as for example *Melochia corchorifolia*, *Citrus medica*, *Cupessus* species and *Hibiscus rosa-sinensis* (KUMARAN *et al.* 2008b; a; c; KUMARAN *et al.* 2009a). A summary of "taxane producing" endophytic fungi, including all literature in peer reviewed journals and patents in English language, is given in Table I-1. This limited summary (all publications in Chinese are excluded) thereby contains >120 different endophytic microorganisms, mainly fungi, from various *Taxus* species and >70 endophytes in which organic extracts taxanes were detected from other plant hosts. Besides the strongly varying origin and multiple fungal genera comparison of the data further reveals that the yields of taxanes detected are very heterogeneous. Although the experimental setup including media composition, cultivation conditions and extraction procedures only differs slightly the amounts of taxanes observed vary from 14 ng/L to nearly 900 µg/L culture medium.

Table I-1: Summary of endophytic fungi in which organic extracts Taxol, Baccatin III or 10-Deacetylbaccatin III or in case of CIEIA taxanes were detected; Table thereby contains only research papers in English journals or patents; first table part shows fungi originated from *Taxus* spp. whereas in the second part all listed endophytes were isolated from different plant species; methods printed bolt were used for Taxane quantification; abbreviations: CIEIA: competitive inhibition enzyme immunoassay; TLC: thin layer chromatography; HPLC: high performance liquid chromatography; MS: mass spectrometry; UV: detection via absorption in the ultraviolet λ-range; IR: infrared spectroscopy; NMR: nuclear magnetic resonance spectroscopy; MIC: minimal inhibitory concentration assay; MTT: cytotoxicity calorimetric assay (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide); * µg/g dry mycelium.

Organism	Origin	Yield [µg/L]	Detection methods	Reference
Taxomyces andreanae	*T. brevifolia*	0.024-0.05	**CIEIA**, TLC, HPLC-UV, MS	(STIERLE *et al.* 1993; STROBEL *et al.* 1994)
10 endophytic fungi	*T. brevifolia*	0.095-2.43	**CIEIA**, TLC, MS	(STIERLE *et al.* 1995)
Pestalotiopsis microspora	*T. wallichiana*	-	TLC, MS, NMR	(STROBEL *et al.* 1996)
Alternaria alternata	*T. hicksii*	332-512	**HPLC-UV**	(DAHIYA 1996)
12 endophytic fungi	*T. cuspidata*	-	**CIEIA**	(KIM *et al.* 1999)
Kitasatospora sp.	*T. baccata*	1.3	**CIEIA**, LC/MS	(CARUSO *et al.* 2000a)
10 actinomycetes/15 fungi	*T. baccata/brevifolia*	0.05-0.15	**CIEIA**	(CARUSO *et al.* 2000b)
Pestalothiopsis microspora	*T. wallichiana*	2.9	**CIEIA**	(METZ *et al.* 2000)
Tubercularia sp. strain TF5	*T. mairei*	-	TLC, HPLC, UV, MS	(WANG *et al.* 2000)
Penicillium spp.	*Taxus species*	-	-	(EL and DIALLO 2000)
3 endophytic fungi	*T. wallichiana*	0.015-0.16	**CIEIA**, HPLC-MS, TLC	(SHRESTHA *et al.* 2001)
21 endophytic fungi	*T. mairei*	-	MTT assay	(HUANG *et al.* 2001)
Multiple endophytes	Various plant species	0.01-165	**CIEIA**, TLC, MS	(STROBEL *et al.* 2001)
Endophytic fungus BT2	*T. chinensis*	4-18	**HPLC-UV**, HPLC-MS	(GUO *et al.* 2006)
3 endophytic fungi	*T. chinensis* var. *mairei*	-	HPLC-MS	(ZHOU *et al.* 2007)
Nodulisporium sylviforme	*T. cuspidata*	450	**HPLC-UV**	(CHI *et al.* 2008)
Fusarium sp.	*T. wallichiana*	-	MIC	(GOGOI *et al.* 2008)
3 endophytic fungi	*T. media/yunnanensis*	112-140 *	TLC, UV, HPLC-MS	(ZHANG *et al.* 2008)
C. cladosporioides	*T. media*	800	TLC, UV, HPLC-MS	(ZHANG *et al.* 2009a)
Aspergillus candidus	*T. media*	112 *	TLC, UV, HPLC-MS, NMR	(ZHANG *et al.* 2009b)
Phomopsis sp.	*T. cuspidata*	418	**HPLC-UV**, MS, NMR, TLC	(KUMARAN and HUR 2009)
Fusarium solani	*T. chinensis*	164	**HPLC-UV**	(DENG *et al.* 2009)
4 endophytic fungi	*T. chinensis*	30	**CIEIA**	(MIAO *et al.* 2009a)
Mucor rouxianus sp.	*T. chinensis*	30	**CIEIA**, HPLC-UV, MS	(MIAO *et al.* 2009a)
Aspergillus niger	*T. cuspidata*	273	**HPLC-UV**, MS	(ZHAO *et al.* 2009)
Gliocladium sp.	*T. baccata*	10	**CIEIA**, HPLC-UV, TLC, MS, NMR	(SREEKANTH *et al.* 2009)

Table I-1 continued

13 endophytic fungi	*T. chinensis*	871	**HPLC-UV**, MS	(LIU *et al.* 2009)
2 *Phomopsis sp.*	*T. cuspidata*	478	**HPLC-UV**, MS, NMR, TLC	(KUMARAN *et al.* 2010a)
endophytic fungi M57	*T. media*	50	**HPLC-UV**	(LI *et al.* 2011)
16 endophytic fungi	*T. globosa*	0.065-0.25	**CIEIA**	(SOCA-CHAFRE *et al.* 2011)
Didymostilbe sp.	*T. chinensis*	8-15	**CIEIA**, HPLC-UV, MS	(WANG and TANG 2011)
9 *Pestalotiopsis microspora*	*Taxodium distichum*	0.014-1.5	**CIEIA**, HPLC-UV, MS	(LI *et al.* 1996)
Pestalotiopsis guepinii	*Wollemia nobilis*	-	**CIEIA**, HPLC-UV, MS, NMR	(STROBEL *et al.* 1997)
Periconia sp.	*Torreya grandifolia*	0.3-0.4	**CIEIA**, TLC, MS, NMR	(LI *et al.* 1998)
hazelnut, *Corylus avellana*	hazelnut, *Corylus avellana*	-	HPLC-MS	(HOFFMAN *et al.* 1998)
4 endophytes	hazelnut, *Corylus avellana*	-	HPLC-MS	(HOFFMAN *et al.* 1998)
18 endophytic fungi	*Ginko biloba*	0.26	**CIEIA**	(KIM *et al.* 1999)
Seimatoantlerium tepuiense	*M. speciosus*	0.25-0.35	**CIEIA**, TLC, UV, MS	(STROBEL *et al.* 1999)
One endophytic fungus	*Cephalataxus fortunei*	-	MTT assay	(HUANG *et al.* 2001)
3 endophytic fungi	*Torreya grandis*	-	MTT assay	(HUANG *et al.* 2001)
10 endophytic fungi	hazelnut, *Corylus avellana*	0.02-16	**HPLC-UV**	(HOFFMAN 2003)
Fusarium solani	*T. celebica*	1.6	TLC, **HPLC-UV**, MS	(CHAKRAVARTHI *et al.* 2008a)
Pestalotiopsis pauciseta	*C. helicacabum*	113	**HPLC-UV**	(GANGADEVI *et al.* 2008)
13 fungal species	-	-	HP-TLC	(GANGADEVI and MUTHUMARY 2008a)
Bartalinia robillardoides	*Aegle marmelos*	188	**HPLC-UV**	(GANGADEVI and MUTHUMARY 2008b)
Phyllosticta melochiae	*Melochia corchorifolia*	478	**HPLC-UV**, MS, NMR	(KUMARAN *et al.* 2008a)
Phyllosticta spinarum	*Cupressus sp.*	235	**HPLC-UV**, MS, NMR, IR, TLC	(KUMARAN *et al.* 2008b)
Phyllosticta citricarpa	*Citrus medica*	265	**HPLC-UV**, MS, NMR, TLC	(KUMARAN *et al.* 2008c)
Chaetomella raphigera	*Terminalia arjuna*	79	**HPLC-UV**, IR, MS, NMR	(GANGADEVI and MUTHUMARY 2009a)
Pestalotiopsis terminaliae	*Terminalia arjuna*	211	**HPLC-UV**	(GANGADEVI and MUTHUMARY 2009b)
Phyllosticta dioscorea	*Hibiscus rosa-sinensis*	298	**HPLC-UV**, MS, NMR, TLC	(KUMARAN *et al.* 2009a)
Phyllosticta tabernaemontanae	*Wrightia tinctoria*	461	**HPLC-UV**, MS, NMR, IR, TLC	(KUMARAN *et al.* 2009b)
Phomopsis sp., *C. langeronii*	*Wollemia nobilis*	-	CIEIA	(STANIEK *et al.* 2010)
Lasiodiplodia theobromae	*Morinda citrifolia*	245	**HPLC-UV**, MS, NMR, IR, TLC, HP-TLC	(PANDI *et al.* 2011)

With regard to the analytical methods used for quantification in case of usage of the anti-taxane immunoassay the yields were lower compared to calculation according to peak areas detected by HPLC-UV, what might indicate the higher specificity of the immunological system.

The yields by UV detection might have been overestimated due to matrix effects meaning other compounds present in the complex raw extracts that also absorb at the used wavelength. Unfortunately no report either compares two methods for quantification or quantifies the compounds via a structure dependent selective analysis technique. Mass spectrometry (MS) and nuclear magnetic resonance spectroscopy (NMR) were only used for verification of the compounds identity. Although some analytical features might have an impact on the results, especially the yields it is also possible that the differences are due to the different fungal species which behave differently under the same screening conditions.

Comparison of all these data so far does not lead to a convincing model regarding the origin of Taxol biosynthetic pathway and reasons for the wide spread of the detection of identical highly complex taxanes through various endophyte genera and in *Taxus*.

As mentioned above the observed production of taxanes by microbes cannot be explained by a defence mechanism, due to the much higher taxane levels produced by *Taxus* spp. This could only be the case for endophytes from not taxane producing plants, where Taxol for example could inhibit the growth of pathogenic fungi, such as *Pythium* spp. and *Phytophthora* spp., which benefits the host (YOUNG *et al.* 1992). Furthermore taxane production in plants is limited to the small genus *Taxus*. If also the plant hosts from which other taxane producing fungi were isolated from could suddenly produce the compound a mechanism of wide spread of the pathway through fungal infection followed by lateral gene transfer would be possible as it was suggested for camptothecin biosynthesis due to the finding of the compound in different plant families (WINK 2008). For gibberellins biosynthesis, the only example up to now where there is phytochemical evidence for the compounds and information about the biosynthetic pathways in both plants and fungi the biosyntheses evolved independently, although in the beginning a connection of evolution was suggested (BÖMKE and TUDZYNSKI 2009). Although possible in theory an independent evolution of a highly complex biosynthesis as towards Taxol, in only distantly related species from different kingdoms seems not probable or at least there is no obvious explanation for such a scenario, whereas in case of gibberellins, having different functions in the different organisms leading to benefits for the respective organisms there is an evolutionary driving

force. Regarding the hypothesis of a connection of evolution of Taxol biosynthesis in plant and endophytic fungi the major questions arising are about the origin of the pathway, the direction of a necessary gene transfer and furthermore the degree of conservation of the genes and enzymes catalysing the multiple reactions towards the end products, found to be identical in plants and fungi. For endophytic fungi nearly nothing is known so far about the pathway. In contrast there is a lot of information on genes and enzymes of the Taxol biosynthetic pathway in *Taxus* species. The information known about the pathway in *Taxus* will be presented in the next chapter prior to further considerations on the possible evolution of the biosynthesis on molecular biological and genetic level.

I.2 *The taxane biosynthetic pathway in Taxus spp.*

All taxanes possess the common taxoid skeleton taxa-4(5),11(12)-diene (HEZARI *et al.* 1995; KOEPP *et al.* 1995). The pathway committing step is the cyclization of the universal diterpene precursor geranylgeranyl diphosphate (GGPP) to this core structure. The tricyclic taxadiene skeleton is then modified by several cytochrome P450-dependent monooxygenases and acyltransferases to yield either Taxol or other taxoid compounds.

As far as only a few of the >350 known taxoid structures have known pharmacological properties (BALOGLU and KINGSTON 1999; ITOKAWA 2003), it is essential to understand the regulation of this biosynthetic pathway to increase flux towards the desired compounds (KETCHUM *et al.* 2003). After initial cyclization of GGPP, there are many branch points that result in the great diversity of taxoid structures, e.g., 14β-hydroxy taxoids and 13-acetyl derivatives. These compounds may play a role in plant defence (DANIEWSKI *et al.* 1998) as antibiotics (YOUNG *et al.* 1992; ELMER *et al.* 1994) or toxins to discourage mammal herbivory (OGDEN 1988).

The biosynthesis of Taxol from geranylgeranyl diphosphate involves at least 19 enzymatic steps and can be divided into several distinct enzymatic reactions (JENNEWEIN *et al.* 2004b). The formation of the taxa-4(5),11(12)-diene backbone is followed by a sequence of eight hydroxylation reactions that require atmospheric oxygen (EISENREICH *et al.* 1998). This indicates that the reactions are catalyzed by cytochrome P450-dependent monooxygenases, which is also typical for many monooxygenation reactions in secondary metabolic pathways (SCHULER 1996). Floss and Mocek (FLOSS and MOCEK 1995)

proposed the order to be C5 and C10, followed by C2 and C9, then C13 and C7 and finally C1 late in the pathway, based on the hydroxylation pattern of known isolated taxoids. Three of these hydroxyl groups are further acylated, including two acetylations and one benzoylation, although the timing of these reactions is not clear (WALKER *et al.* 2002a). Advanced taxoids, such as Baccatin III, require further oxidation of the hydroxyl group at C9 and formation of the oxetane ring at C4,5. The last steps involve attachment of a β-phenylalanoyl side chain at C13 followed by 2'-hydroxylation and *N*-benzoylation. Related reactions lead to the *N*-tigloyl and *N*-hexanoyl derivates Cephalomannine and Taxol C (BALOGLU and KINGSTON 1999). These are the major steps of the pathway.

The taxane core is derived via the plastidial 2-*C*-methyl-D-erythritol phosphate (MEP) pathway (EISENREICH *et al.* 1996), in which isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), the universal precursors of all terpenes, are built from pyruvate and glyceraldehyde-3-phosphate through the intermediate 1-deoxy-D-xylulose-5-phosphate (DXP) (ROHMER 1999; EISENREICH *et al.* 2001; KUZUYAMA and SETO 2003).

Like all diterpenoids, taxanes are based on geranylgeranyl diphosphate, which is derived from one molecule of DMAPP and three molecules of IPP via head to tail condensation. *Taxus* geranylgeranyl diphosphate synthase (GGPPS) was first isolated by the Croteau group from *T. canadensis* cells and later by the Verpoorte group from *T. baccata* cells (HEFNER *et al.* 1998; LASKARIS *et al.* 2000). The protein was characterized as a typical prenyltransferase. In a *T. cuspidata* cDNA library (SCHOENDORF *et al.* 2001), ESTs representing GGPPS were quite abundant, representing 1.7 ‰ of the clones (JENNEWEIN *et al.* 2004b).

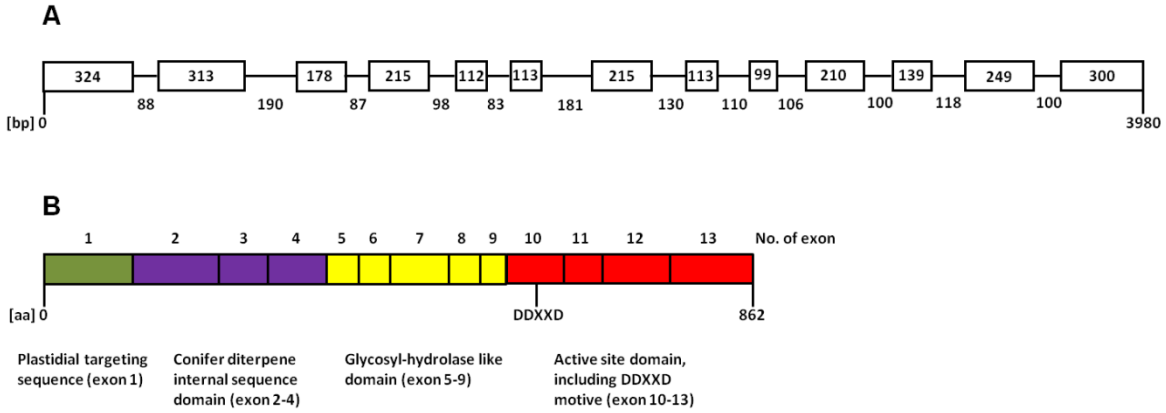


Figure I-2: Genomic structure and schematic enzyme structure of taxadiene synthase A: genomic structure of taxadiene synthase from *Taxus baccata* consisting of 13 exons (boxes) and 12 introns (lines); numbers represent the length of the fragments in bp **B:** schematic enzyme structure including plastidial leader and main characteristic terpene synthase motifs, numbered exons (boxes) and introns (vertical lines).

The committing step in taxoid synthesis is the cyclization of geranylgeranyl diphosphate to taxadiene, catalyzed by taxadiene synthase (TDS) (HEZARI *et al.* 1995).

Attempts to isolate TDS from *Taxus* trees yielded in an enzyme with a molecular mass of ~79 kDa (HEZARI *et al.* 1995). Its properties were similar to other plant terpene synthases, such as a relatively low K_m and Mg^{2+} as the cofactor, but unusually the optimum pH was 8.5. The main product of the enzyme was confirmed as taxa-4(5),11(12)-diene (KOEPP *et al.* 1995), even though chemical analysis predicted preferential formation of the 4(20)-11(12)-isomer (GUERITTE-VOEGELEIN *et al.* 1987).

A 2586-bp cDNA clone encoding a 98-kDa pre-protein was isolated by a homology-based PCR cloning strategy (WILDUNG and CROTEAU 1996). The corresponding enzyme contained an *N*-terminal plastidial targeting sequence, which was cleaved after import into the plastid, although the exact length of this sequence remains unclear. Heterologous expression of pseudomature, *N*-terminally truncated TDS variants in *E. coli* have indicated a sequence length of up to 79 amino acid residues for the plastidial target sequence (Williams et al., 2000b). Further analysis revealed features typical of plant terpene synthases, such as a DDXXD-motif responsible for cofactor binding, a conifer diterpene internal sequence domain and a glycosyl hydrolase-like domain (Figure I-2B) (TRAPP and CROTEAU 2001).

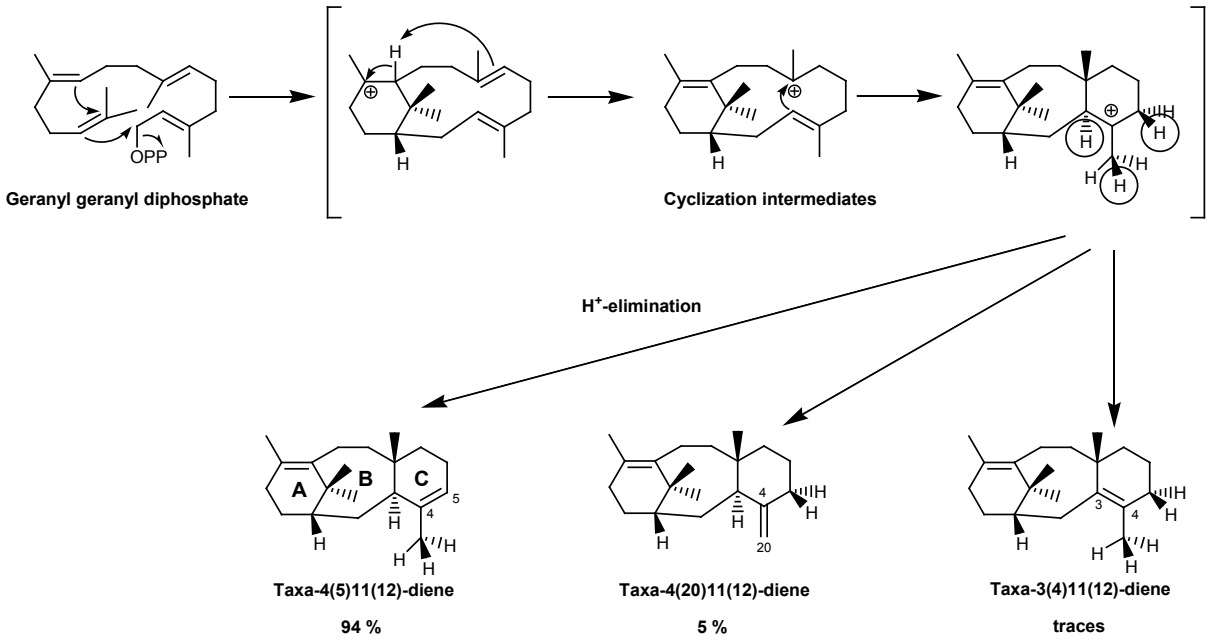


Figure I-3: Reaction mechanism of taxadiene synthase starting from GGPP over the cyclization intermediates to taxenyl cation, which undergoes H^+-elimination, three possibilities lead to three regioisomers with major product taxadiene-4(5),11(12)-diene.

Recently a heterolougously expressed taxadiene synthase truncation variant (-107 amino acids) was purified and crystal structure was elucidated, via co-crystallization with Mg^{2+} and either with 13-aza-13,14-dihydrocopalyl diphosphate or 2-fluorogeranylgeranyl diphosphate (KOKSAL *et al.* 2011). Although this variant was not active anymore, this further truncated protein was the only protein capable of forming crystals.

In feeding experiments with the natural substrate GGPP, the enzyme produced 94% taxa-4(5),11(12)-diene, 5% taxa-4(20),11(12)-diene, 1% verticillene and trace amounts of the tentatively identified 3(4),11(12) isomer (Figure I-3) (WILLIAMS *et al.* 2000b).

The mechanism of the enzyme was discovered during several investigations. A scheme including the proposed cyclization intermediates is shown in Figure I-3 (LIN *et al.* 1996; WILLIAMS *et al.* 2000a; COATES *et al.* 2005; JIN *et al.* 2005).

For the cloning of potential cytochrome P450 genes involved in the Taxol biosynthesis, three complementary cloning approaches were applied. The first approach was differential display for specific cloning of cytochrome P450 genes (SCHOPFER and EBEL 1998) using the methyl jasmonate inducibility of Taxol biosynthesis in plant cell cultures (YUKIMUNE *et al.* 1996). Secondly several candidates could be isolated using PCR with degenerate primers designed on highly conserved P450 motives (PERF motif and the heme-binding motif) (HOLTON and LESTER 1996; JENNEWEIN *et al.* 2004a). In addition to this specific cloning approach of *Taxus* cytochrome P450 genes, random sequencing of a cDNA library from methyl jasmonate-induced *T. cuspidata* cell culture identified additional cytochrome P450 clones (JENNEWEIN *et al.* 2004b).

Using these three approaches, nearly 30 very similar (homology >70% on aa level) candidates for cytochrome P450 genes with potential relevance to taxoid biosynthesis were obtained with most of the cloned genes being identified by all three approaches. The biosynthesis of Taxol involves approximately nine oxygenation reactions, thus implying significant redundancy in hydroxylase functions. Heterologous expression of cloned cytochrome P450 gene candidates in *Spodoptera frugiperda* baculovirus insect cell expression system (JENNEWEIN *et al.* 2001) as well as *Saccharomyces cerevisiae* (backer's yeast) (SCHOENDORF *et al.* 2001; JENNEWEIN *et al.* 2003) and functional testing of the expressed enzymes resulted in the identification of taxoid 2α-, 5α-, 7β-, 10β-, 13α- and 14β-hydroxylases (Figure I-4). All of the isolated sequences were very similar to each other, but more distantly related to other cytochrome P450 genes.

Regarding the order of reactions in the biosynthetic pathway hydroxylation steps of Taxol biosynthesis can be divided into early, intermediate and late reactions. In microsomal fractions of *T. cuspidata,* six hydroxylations to taxadiene-hexaol occur under standard assay conditions (WHEELER *et al.* 2001). The next modifications are probably C13 and C10 hydroxylations. The enzymes were shown to use both 5α-hydroxy-taxadiene and 5α-acetoxy-taxadiene but with opposite substrate selectivities. Although 5α-hydroxy-taxadiene and 5α-acetoxy-taxadiene are substrates for both enzymes, the former is favored for 13-hydroxylation whereas the later is more likely to be modified at C10 (JENNEWEIN *et al.* 2001; SCHOENDORF *et al.* 2001). Results from cell feeding study experiments in *Taxus* cell cultures using the relevant intermediates (KETCHUM and CROTEAU 2006; KETCHUM *et al.* 2007) indicate a bifurcation occuring early in the Taxol biosynthesis pathway, one branch leading to Taxol and the other one to alternative taxoids (or perhaps to Taxol via a different route).

Isolation of a clone encoding taxoid 14β-hydroxylase (JENNEWEIN *et al.* 2003) supports the hypothesis of diversification at this early stage.

C9-hydroxylation is also thought to be an early reaction in the pathway. *In vivo* studies in which yeast were fed 5α-hydroxy-taxadiene showed that one cDNA encoding a P450 candidate might represent a taxoid 9α-hydroxylase, but the product has yet to be confirmed by NMR (CROTEAU *et al.* 2006).

For the analysis of intermediate/late oxygenation steps taxusin (5α,9α,10β,13α-tetraacetoxy-taxa-4(20),11(12)-diene), a compound isolated from yew heartwood that is thought to be a dead-end metabolite rather than an intermediate in Taxol synthesis (KOEPP *et al.* 1995), was used as the test substrate. Thereby, it was possible to identify the taxoid 2α- and 7β-monooxygenases (CHAU and CROTEAU 2004; CHAU *et al.* 2004a).

It was shown that both enzymes can operate sequentially, with 7β-hydroxylation probably followed by 2α-hydroxylation. By incubating microsomes with taxusin, the common hexaol (2α,7β-dihydroxy taxusin) is formed (CHAU *et al.* 2004a).

The other steps leading to hydroxylation at C1 oxygenation at C9 and the oxetane ring formation at C4 and 5 are still unknown. As far as several P450 candidates isolated initially are still without a defined function it is most likely that the reactions are catalyzed by some of these enzymes that are probably P450 mediated. Lack of substrates for testing of the reactions with at least two monooxygenases still unknown and many *Taxus* cytochrome P450 genes

undefined, the order of reactions in Taxol biosynthesis and any corresponding phylogenetic analysis can only be regarded as approximate.

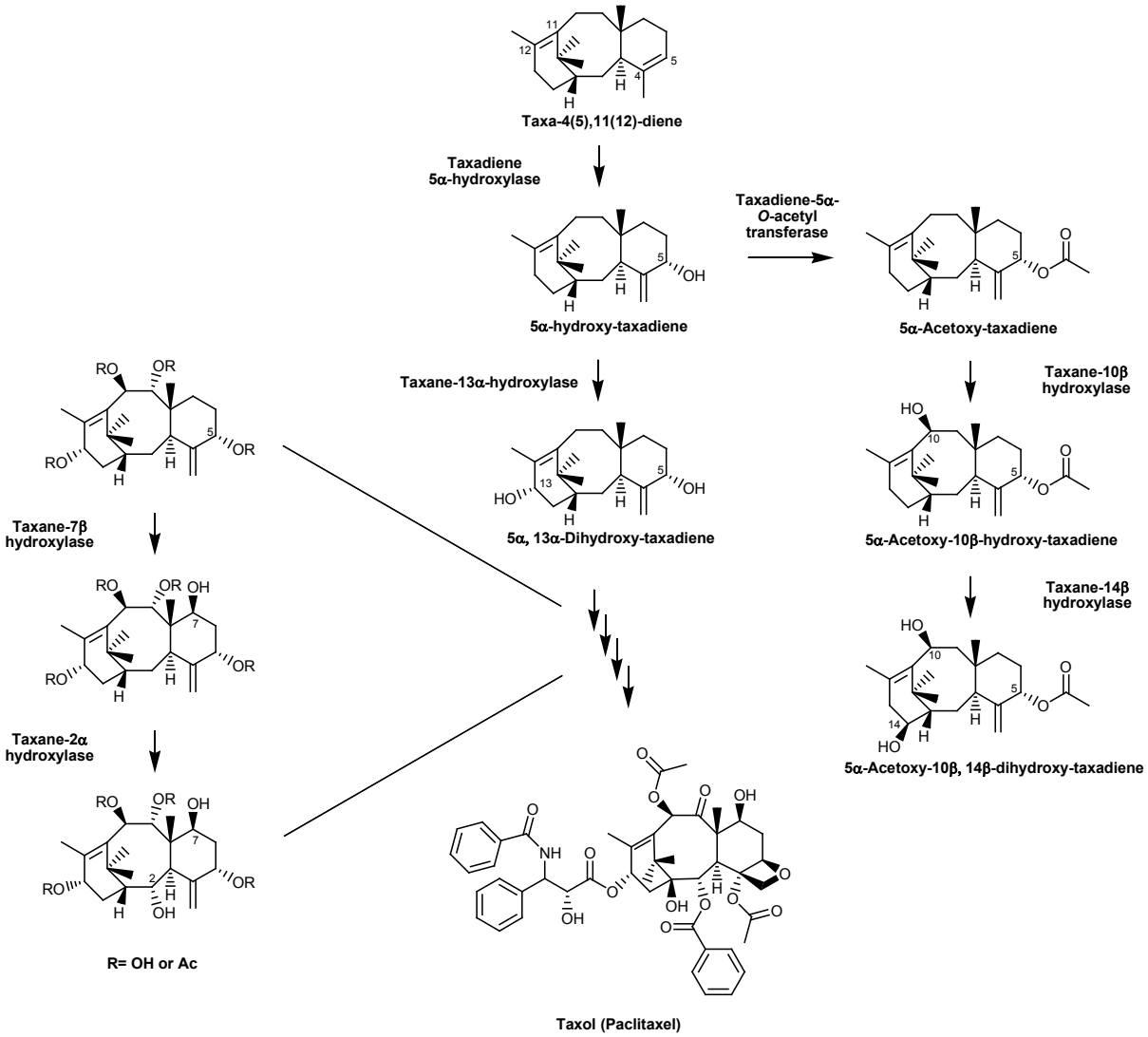


Figure I-4: P450 monooxygenase-mediated hydroxylations of the taxa-4(5),11(12)-diene backbone lead to Taxol. This process includes the early modifications at C5, C10 and C13, the C14 hydroxylation to major side products and the two modifications at C7 and C2, which are thought to be important in the main Taxol biosynthetic pathway (HEINIG and JENNEWEIN 2009).

It is reasonable to assume that the family of taxoid cytochrome P450-dependent monooxygenases evolved through gene duplication and divergence from a common ancestor (PICHERSKY and GANG 2000), as suggested by the >70% similarity among the taxoid hydroxylases and much lower similarity to other plant-derived P450 monooxygenases (JENNEWEIN *et al.* 2004b). This homology-functionality relationship is also observed in steroid biosynthesis, where hydroxylases with the same catalytic capability show up to 68%

similarity, whereas those with different functions also have more diverse sequences (FELDMANN *et al.* 2002; KIM and TSUKAYA 2002). Phylogenetic comparison demonstrates the tight coherence of taxoid cytochrome P450 enzymes and the very distant relationship to other plant cytochrome P450s (<35%). Only catalytically similar enzymes, such as abietadienol/abietadienal oxidase from loblolly pine (RO *et al.* 2005), are grouped together with taxoid P450s (KASPERA and CROTEAU 2006).

For clear evidence on the order of taxoid hydroxylation steps, it will be necessary to identify and characterize the missing monooxygenases. Therefore, it is necessary to get access to more highly functionalized (intermediate) taxoids. Due to the complexity of total chemical synthesis of taxoids and the inaccessibility from natural sources, the current unavailability represents a major obstacle in the functional assignment in the intermediate and late hydroxylation reactions.

Finally a lot of efforts were made to explore the additional modifications of the hydroxylated taxadiene backbone including acetylations, benzoylation and the assembly of the side chain. Taxol contains four ester functional groups at C2 (benzonate), C4 (acetate), C10 (acetate) and C13 (*N*-benzoyl-3-phenylisoserinoyl) on the taxane core. Among the enzymes responsible for these reactions, the first to be investigated was the 5-*O*-acetyltransferase. Acetylation in this position is considered to be the progenitor of the rearrangement reaction leading to the oxetane ring of taxoids (Walker et al., 1999). After demonstrating activity in *T. canadensis* soluble protein extracts, the enzyme was partially purified and shown to be a 50-kDa acetyl-CoA-dependent transferase with a pH optimum of ~9.0, a low μM K_m value and selectivity for less functionalized taxanols (CROTEAU *et al.* 1999). With the objective to isolate more acetyltransferases of the Taxol biosynthesis, degenerate primers were designed according to a consensus protein sequence, obtained from an alignment of transacylase sequences of plant origin, and used to generate PCR probes with which then a *T. cuspidata* cDNA library was screened. This yielded eight full-length cDNAs and seven ESTs, leading to the identification of 15 acyltransferase-type genes (WALKER *et al.* 2000; JENNEWEIN *et al.* 2004b). Functional analysis in *Escherichia coli* soluble protein extracts resulted in the identification of the taxadien-5α-ol-*O*-acetyltransferase, the taxoid-2α-*O*-benzoyl transferase, the 10β-*O*-acetyltransferase and two cDNA clones encoding enzymes involved in transferase reactions at the C13 side chain (WALKER and CROTEAU 2000b; a). Furthermore another transferase was found that also led to 5α-hydroxylated taxadiene which as the other transferases was also

able to functionalize higher hydroxylated taxanes but surprisingly with different regioselectivity of already present hydroxygroups (CHAU *et al.* 2004b).

These results indicated that the acyltransferases are not substrate-specific and only moderately regiospecific. The acylation position appears to depend very much on the substitution pattern of the metabolized precursor (CHAU *et al.* 2004b). Although the precise timing of C5 acetylation is not completely clear, it is still probable that the step occurs early in the pathway and is somehow influenced by division into C13 and C14 taxoid syntheses (KETCHUM *et al.* 2007). Regarding homology of the enzymes, like cytochrome P450-dependent monooxygenases, taxoid acyltransferases are very similar to each other (>65%) and probably have also evolved from a common ancestor by gene duplication and divergence.

Assembly of the side chain at C13 is probably the last modification of the taxane core. The β-phenylalanoyl-type side chains were shown to be formed from α-phenylalanine through the activity of an aminomutase (LEETE and BODEM 1966; PLATT *et al.* 1984). Feeding studies with *Taxus* cells demonstrated that the *N*-benzoyl-3'-phenylisoserinoyl side chain of Taxol also originates from α-phenylalanine metabolism (FLEMING *et al.* 1994). The mutase activity responsible for this committed step in side-chain biosynthesis was first observed in *Taxus* stem extracts by Walker and Floss (WALKER and FLOSS 1998). Further feeding studies with Baccatin III, the supposed substrate, were performed with either β-phenylalanine or phenylisoserine, showing that both molecules were incorporated, although the unbenzylated amino acid was three-times more efficient as a substrate. However, *N*-benzoyl phenylisoserine was not a suitable substrate, indicating the formation of β-phenylalanoyl- or phenylisoserine-baccatin III prior to *N*-benzoylation. These results do not answer the question about the timing of 2'-hydroxylation. The acceptance of phenylisoserine indicates that the reaction occurs before the attachment of the chain, although no amino acid hydroxylase activity has been detected thus far (SILVERMAN 2000). On the other hand, *Taxus* microsomes can catalyze the conversion of β-phenylalanoyl-baccatin III to phenylisoserine-baccatin III (LONG and CROTEAU 2005), indicating the possibility of hydroxylation comparable to oxygenation of the taxane core structure. The undefined cytochrome P450 cDNAs from the *Taxus* library therefore remain possible candidates for a taxoid 2'-hydroxylase.

A taxoid aminomutase cDNA was isolated from the *T. cuspidata* library and expressed functionally in *E. coli* (WALKER *et al.* 2004). Side chain assembly is catalyzed by

C13-propanoyl-CoA transferase (WALKER *et al.* 2002b) and transfer of the benzoyl moiety a *N*-benzoyl transferase (WALKER *et al.* 2002a).

In order to determine whether *N*-debenzoyl-2'-deoxytaxol or *N*-debenzoyltaxol was the natural substrate for taxoid *N*-benzoyl transferase, *N*-debenzoyltaxol was synthesized and the enzyme kinetics for both compounds were determined (LONG *et al.* 2008). The efficiency of benzoyl-CoA transfer to the 2'-hydroxylated substrate was shown to be double that of unsubstituted substrate, indicating that *N*-debenzoyltaxol is the preferred precursor for taxoid *N*-benzoyl transferase. In this context, selectivity for the CoA co-substrate was also tested. Only taxoids varying with respect to 3'-*N*-substitution are observed in *Taxus* cell cultures, e.g., Taxol (*N*-benzoyl-3'-phenylisoserinoyl), cephalomannine (tigloyl) and Taxol C (hexanoyl) (BALOGLU and KINGSTON 1999). The taxoid *N*-benzoyl transferase was found to be highly selective for benzoyl-CoA and did not convert any other substrate (LONG *et al.* 2008). This indicates that some of the still uncharacterized acyltransferase candidates might correspond to enzymes that facilitate the diversification of taxoid composition at this last step of the biosynthetic pathway.

In contrast to all this data about Taxol biosynthesis from *Taxus* there is nearly no information on the pathway responsible for the compound formation in endophytic fungi. Only very few reports show the amplification of short gene fragments of taxadiene synthase, 10-deacetylbaccatin III-10-*O*-acetyl transferase (DBAT) and C-13 phenylpropanoid side chain-CoA acyltransferase (BAPT) from endophytic fungi's genomic DNA (MIAO *et al.* 2009b; STANIEK *et al.* 2009; KUMARAN *et al.* 2010b). However, these results are not very conclusive as far as despite all reports state 98 % identity of plant and fungal gene fragments but the size of the amplification products vary from a size corresponding to the cDNA clone to a size indicating that all introns are present as in *Taxus*.

These data do therefore not provide evidence for the Taxol biosynthesis in fungi. It is also not possible to answer the question about evolution of the pathway, without identification of the genes from fungi and final comparison to the detailed data of the Taxol biosynthesis of *Taxus*.

I.3 *Evolution and gene transfer*

The extremely complex biosynthetic pathway towards Taxol, known from *Taxus* and the predicted occurrence of its biosynthesis in distantly related organisms, plants and fungi, raises

the question if the genes/enzymes in these two species are similar to each other, indicating a connection in evolution or if there is no homology and the biosynthetic pathways evolved completely independent from each other.

As mentioned above the occurrence of "taxane producing microbes" was found not to be limited to one *Taxus* species or to the *Taxus* family and also not to a certain category of plants (i.e. conifers). As for the host plant species no significant pattern is obvious either regarding fungal species in which extracts identical compounds as in plants were detected or a possible geographic occurrence of the phenomenon. Taxol "production" by microbes seems to be rather ubiquitous, although this seems very unlikely for a complex niche natural product synthesized over approximately 19 enzymatic steps, like Taxol.

What kind of evolutionary scenario seems to be reasonable for biosyntheses of complex natural products found in organisms from different kingdoms? Important in this context are considerations about horizontal gene transfer processes, especially between different kingdoms, the genomic organization of secondary metabolite pathways either scattered over the genomes or organized in gene clusters dependent on the kingdom and possible evolutionary driving forces resulting in different evolutionary scenarios.

In general evolution of a biosynthetic pathways observed in different organisms can occur either convergently or divergently. In the context of secondary metabolite pathways convergent evolution means that the benefit of a compound or compound class itself is the driving force for development of a biosynthesis. If the product is essential or beneficial, or identical compounds provide different important benefits to the different organisms a synthesis can develop independently in a variety of species also from different kingdoms. Mechanistically this will most likely happen through gene duplication and diversification of already present genes from primary metabolism of the respective organisms. Of course dependent on the origin these pathways might be fundamentally heterogeneous in gene and protein sequences. Nevertheless catalyzing reactions belonging to pathways leading to formation of the identical products or molecules of a product class, at least the enzymes have to share some conserved characteristics important for the catalytic activity towards the defined structures of intermediates and end products.

In contrast divergent evolution of secondary metabolite pathways is based on a connection of evolution of biosyntheses occurring in different organisms. The biosynthesis, hence the ability to produce a certain compound is transferred from the ancestor organism either vertically by reproduction within the same genus or by horizontal gene transfer to other species living in

the same habitat. Through diversification of either the original genus over time or the biosynthesis genes themselves the biosynthetic pathways diverge during evolution. For example genes of a biosynthesis can duplicate again, followed by mutation of the duplicate that might lead to an enzyme with differing activity. Phenomenons like this are supposed to be a reason for extensive branching of secondary metabolite pathways leading to a certain screening for new biological active compounds (JONES and FIRN 1991; FIRN and JONES 2003). Thus out of an ancestral single pathway in one organism that turned out to produce beneficial compounds and was therefore retained in the genome of the ancestor organism, further wide spread even over the borders of the initial genus or kingdom, many similar biosyntheses evolved. Nevertheless all these syntheses go back to the original ancestral pathway and share certain sequence homology to each other.

Besides these two extreme scenarios also various combinations are possible. For example transfer of a single gene as common ancestor in combination with convergent evolution of further pathway genes/enzymes might lead to identical natural products if the evolutionary force to produce a special bioactive compound is high enough in a certain environment.

In case of Taxol biosynthesis in plants and its associated endophytic fungi many scenarios seem possible. The data presented in I.1 and I.2 through its heterogeneity provides pro and contras for each model.

Assuming a convergent evolution of Taxol biosynthesis all fungi predicted to have the biosynthesis would need an environmental driving force triggering the development of their own Taxol biosynthetic pathway. Of course bioactive taxanes might play a role for the organisms for defence against other microbes in their habitat. But this is only reasonable for predicted producers isolated from plant hosts not able to produce the identical compounds themselves. In case of potential Taxol production by endophytes from *Taxus spp*. this benefit is not present due to the much higher production level of the plant. Hence, even if fungal producers that had been living in another environment before, in which they needed the ability of Taxol biosynthesis, colonize *Taxus* they would most likely stop the production of the anyway very low amounts of the taxoid compounds within yew tree. This is obviously not the case as far as still most isolates in which extracts taxanes were detected originate from *Taxus* (Table I-1).

The other possibility, a divergent evolutionary theory for Taxol biosynthesis that includes a horizontal gene transfer from one species to another does not require necessarily a pressure for the organism to produce the compounds. The association of the majority of predicted

"taxane producing" fungi to the yew tree thereby can be taken as an indication for such a scenario. Nevertheless the low yields of taxanes found in those fungi and the detection of taxanes in extracts of fungi not associated with yew trees contradict this evolutionary assumption.

However with regard to complexity of Taxol biosynthesis a complete convergent scenario seems improbable. Mixed evolutionary theories as for example one trans-kingdom transfer of one or few ancestor genes followed by duplication and divergence in the different species may also be possible. A transfer of fewer genes than the supposed 19 involved in Taxol biosynthesis in yew might be more probable.

Despite all theoretical considerations the existing non consistent data on "Taxol producing" fungi leave the question about the evolutionary development of the biosynthesis completely open.

In general before favouring a theory with or without a horizontal gene transfer and hence divergent or convergent evolution, factors facilitating and restraining transfer processes of DNA from organisms of one kingdom to another have to be taken into account. These considerations include thereby mechanisms for such transfers as well as organisms dependent differences.

Since most evidence for horizontal gene transfer is derived from large scale genome analysis and comparison, mechanisms for eukaryote-eukaryote gene transfer are not completely understood so far (KEELING and PALMER 2008; RICHARDS *et al.* 2009; FITZPATRICK 2011; GOODWIN *et al.* 2011). In case of fungi mechanisms like somatic fusion or interspecies hyphal anastamosis were suggested as possible facilitating processes for genetic exchange (FITZPATRICK 2011).

As major barriers different genetic codes of origin and destination organism have to be mentioned as well as different intron/exon structures what can lead to incorrect splicing (KEELING and PALMER 2008; FITZPATRICK 2011). Furthermore gene promoters can be incompatible or for example in fungi gene silencing mechanisms exist that tend to pseudogenize foreign genes (SHIU *et al.* 2001).

Despite the still undiscovered mechanisms responsible for eukaryote-eukaryote gene transfer and the general barriers lowering the chance of a transfer of a functionally active entire secondary metabolic biosynthesis from one kingdom to another there are also several examples for the contribution of horizontal gene transfer events in natural product pathway evolution.

Although the occurrence of identical natural products from distantly related organisms, like in case of Taxol, is a very rare phenomenon, molecules of the same natural product class have been isolated from organisms of different kingdoms. Pathways studied in most detail in this regard are polyketide biosyntheses. This class of compounds is well known from plants, fungi and bacteria. Polyketides are secondary metabolites that are synthesized via modular multi domain enzyme complexes similar to fatty acids with polyketide synthases catalyzing acyltransfer, reduction, cyclization and different post-cyclization modifications (STAUNTON and WEISSMAN 2001). Due to the similar mechanism of these biosynthetic pathways (PKS types I-III) regardless of species or kingdom they are derived from, a connected evolution seems to be possible. On the other hand a scenario involving gene duplication and diversification in each kingdom is feasible, too, since formation of polyketides is strongly related to fatty acid synthesis, which is present in all organisms. Investigations on that topic during the last years delivered more evidence for a convergent evolution in each kingdom, probably starting from fatty acid biosynthesis genes. Due to the large amount and structural variance of compounds that mainly are antibiotics and therefore provide an obvious advantage for each producer this was not very surprising. Nevertheless there are some examples of lateral gene transfer processes involved as well, although all these horizontal transfers are reported only between different fungi or between actinobacteria and ascomycete fungi (KROKEN *et al.* 2003; KHALDI *et al.* 2008; LAWRENCE *et al.* 2011).

With regard to the topic of this thesis trans-kingdom gene transfers between fungi and plants are of particular interest. As mentioned above there is only very limited evidence for such events regarding secondary metabolite pathways. So far only one report presents a large scale phylogenetic analysis exploring possible gene transfers between plant and fungi. By comparison of six plant and 159 microbial genomes Richards and co-workers were able to identify nine candidate genes that might have been transferred horizontally from one kingdom to another. Although none of these genes encoded a protein from secondary metabolism it has thereby been shown that lateral transfers in both directions occurred during evolution between these two kingdoms. Nevertheless these events are supposed to be ancient and rare (RICHARDS *et al.* 2009).

Despite the observed low probability of a plant-fungus or fungus-plant transfer a specific characteristic of genomic organization, the clustering of whole secondary metabolite pathways in bacteria but also in fungi may facilitate the possibility of a transfer of an entire multi gene secondary metabolite pathway.

Moreover operons are a common feature in bacteria. Occurrence of such genomic units was thought to be rare in eukaryotes. Similar genes next to each other in genomes are often products of gene duplication, followed by divergence (OSBOURN and FIELD 2009).

Through intensive research in the field of natural product pathways it became clear that especially in filamentous fungi most of the secondary metabolite biosyntheses genes are not scattered throughout the genome but are rather organized in gene clusters (KELLER *et al.* 2005). The first clusters identified were bacterial ones, like that for actinorhodin biosynthesis in *Streptomyces* (MALPARTIDA and HOPWOOD 1986). From this starting point many more clusters for polyketide antibiotic biosynthesis were isolated from different *Streptomyces spp*. and other actinobacteria. Also in filamentous fungi more and more secondary metabolites pathway gene clusters could be identified (KELLER *et al.* 2005). Thereby the concept of gene clustering in fungi was found to be not only common for polyketide pathways like aflatoxins from *Aspergillus spp*. (KELLER and HOHN 1997), but also for other natural product classes including peptide antibiotics as penicillin (SMITH *et al.* 1990) and terpenes as gibberellins (TUDZYNSKI and HÖLTER 1998) and trichothecenes (HOHN *et al.* 1993). In contrast to bacteria and fungi the genes are mainly not linked in plants. Nevertheless there are six examples of secondary metabolite gene clusters in plants, too. The first one for the production of 2,4-dihydroxy-7-methoxy-1,4-benzoxazin was discovered in 1997 by Frey and co-workers in maize and further characterized during the following years (FREY *et al.* 1997; FREY *et al.* 2003). With more than 500 kb it is extremely large compared to the gene clusters from microbes described before and furthermore the genes are not continuous. Some genes for the biosynthesis are located more than 15 Mb apart from each other. The other five clusters identified in plants all represent terpene biosynthetic pathways, two clusters for diterpene biosynthesis the momilactone and the phytocassane gene clusters from rice (SHIMURA *et al.* 2007; WU *et al.* 2011) and three clusters for triterpene formation: the avenacin gene cluster from oat (QI *et al.* 2004) and the thalianol and maternal gene clusters from *Arabidopsis thaliana* (FIELD and OSBOURN 2008; FIELD *et al.* 2011). The last two have a size comparable to known fungal clusters (CHU *et al.* 2011).

From fungi four examples are known concerning clustered diterpene biosynthetic pathways. Gibberellin synthesis in *Gibberella fujikuroi* involves a copalyl synthase/kaurene synthase gene, which is clustered with genes encoding a geranylgeranyl diphosphate synthase and three P450 monooxygenases (TUDZYNSKI and HÖLTER 1998). Since then gibberellin biosynthetic gene clusters were isolated from multiple *Fusarium* species as well as from some

other fungi as *Shaceloma manihoticola, Phaeosheria sp. L487* or *Magnaporthe grisea* (BÖMKE *et al.* 2008). Aphidicholin synthesis in *Phoma betae* involves a gene cluster encoding a terpene cyclase, a geranylgeranyl diphosphate synthase and several cytochrome P450 monooxygenases, as well as a transcription factor and a transport factor (TOYOMASU *et al.* 2004). Aflatrem biosynthesis in *Aspergillus flavus* also involves a cluster of genes encoding a terpene cyclase, a geranylgeranyl diphosphate synthase, several cytochrome P450 monooxygenases and transcription and transport factors (ZHANG *et al.* 2004). Finally, two gene clusters for the biosynthesis of diterpenes were isolated from *Phomopsis amygdali*, including geranylgeranyl diphosphate synthases, terpene cyclases and several candidates for P450 monooxygenases (TOYOMASU *et al.* 2008).

These data, giving evidence for diterpene gene clusters in fungi as well as in plants suggest that Taxol biosynthesis genes also could be clustered in fungi and/or in *Taxus.* Thus a horizontal transfer of the entire pathway instead of several coordinated gene transfers from one to the other species might be one possible scenario. Of course the low probability (RICHARDS *et al.* 2009) and the general barriers for a successful transfer leading to a functional biosynthetic pathway in an only distantly related species from a different kingdom (FITZPATRICK 2011) still have to be considered.

However, mechanisms and reasons for clustering of natural product biosynthetic pathways in eukaryotes are still poorly understood as well as the role of lateral gene transfer especially between different kingdoms.

The strongest evidence for the existence of that kind of transfer from bacteria to fungi was found for penicillin and cephalosporin biosynthesis which are present in fungi and are found in bacteria, too (WEIGEL *et al.* 1988). Nevertheless, there are other biosynthetic pathways that definitely evolved without a horizontal gene transfer but show a highly conserved pattern of the same natural products as well like the gibberellins (BÖMKE and TUDZYNSKI 2009).

Other theories focus on the location of secondary metabolite gene cluster within the genome in order to explain the predicted role in evolution. In actinomycetes and in fungi they are found mainly in the telomere regions which are highly dynamic (HOFFMEISTER and KELLER 2007). This would allow relocation and formation of clusters, but could also result in instability of metabolites production. One theory arising from these location considerations is the so-called selfish-cluster hypothesis, which proposes a selective advantage created by clustering itself (WALTON 2000). Genomic flexibility and the known mechanisms of deletion or translocation of genes increase the risk of a loss of a functional pathway. Also

vertical transfer of a complex pathway with its genes scattered over the genome would be more risky than a possible horizontal transfer of all genes in one step as a cluster (WALTON 2000). Therefore it is possible that secondary metabolite gene clusters represent mobile genetic elements that lead to more stability of a specific "none-essential" natural product biosynthesis.

This kind of organization is thought to provide a possibility for either generating more chemical diversity by for example horizontal transfer followed by duplication and diversification or for transfer of entire pathways to other not related species in the same environment giving them an selective advantage through the newly produced compounds. This theory is supported by several reports showing clear evidence for the spread of secondary metabolite gene clusters by horizontal transfer between different fungal species of the same genus but also towards less closely related fungi (KHALDI *et al.* 2008; MEHRABI *et al.* 2011; SLOT and ROKAS 2011).

Altogether it is not possible to construct an evolutionary model explaining all phenomenons observed in secondary metabolic pathways up to now. Most likely many different evolutionary processes are involved and it furthermore seems as if these mechanisms are dependent on the species as well as the respective biosynthesis itself. Nevertheless the organization in gene clusters and horizontal gene transfer both were shown to play a role in secondary metabolite biosynthesis evolution, although only limited information is available on the mechanisms for eukaryote-eukaryote DNA transfer (KEELING and PALMER 2008; FITZPATRICK 2011).

In the field of diterpene biosynthesis only the gibberellin pathway was studied in detail in plants and their endophytic fungi up to now (TUDZYNSKI and HÖLTER 1998; FISCHBACH and CLARDY 2007; BÖMKE and TUDZYNSKI 2009). So far it represents the only example of an endophyte/plant system that shows the same natural products in phytochemical analysis and for which the pathways are known on mechanistic as well as on genetic level. Thus, with regard to the present project this biosynthesis is the only source of data to make assumptions for the possible Taxol biosynthetic pathway evolution in *Taxus* species and its endophytic fungi by comparison to another plant/endophyte relationship.

Gibberellins are found in plants and fungi, but the enzymes catalyzing the initial synthesis step, the formation of *ent*-kaurene, are different.

In plants, the reaction is catalyzed by two enzymes with copalyl diphosphate as an intermediate (OLSZEWSKI *et al.* 2002), whereas in fungi, such as *Gibberella fujikuroi*, there

is a single enzyme that catalyses the reaction directly from geranylgeranyl diphosphate to *ent*-kaurene (TUDZYNSKI and HÖLTER 1998). Furthermore it was found, that not only the initial cyclization but also the following series of hydroxylations towards gibberellic acid GA_3, which is a product in plants as well as in fungi, are different (FISCHBACH and CLARDY 2007; BÖMKE and TUDZYNSKI 2009).

Together with the fact, that all the P450 genes/enzymes in *Gibberella fujikuroi* are much more similar to other P450 genes/enzymes from fungi than to the corresponding genes/enzymes from plants, in this case a convergent evolution of the pathway in both species is probable (TUDZYNSKI and HÖLTER 1998). This is furthermore supported by the genomic organisation, scattered over the genome in plants, for example located on all five chromosomes in *Arabidopsis thaliana*, but clustered in all known fungal producers (TUDZYNSKI and HÖLTER 1998; BÖMKE and TUDZYNSKI 2009).

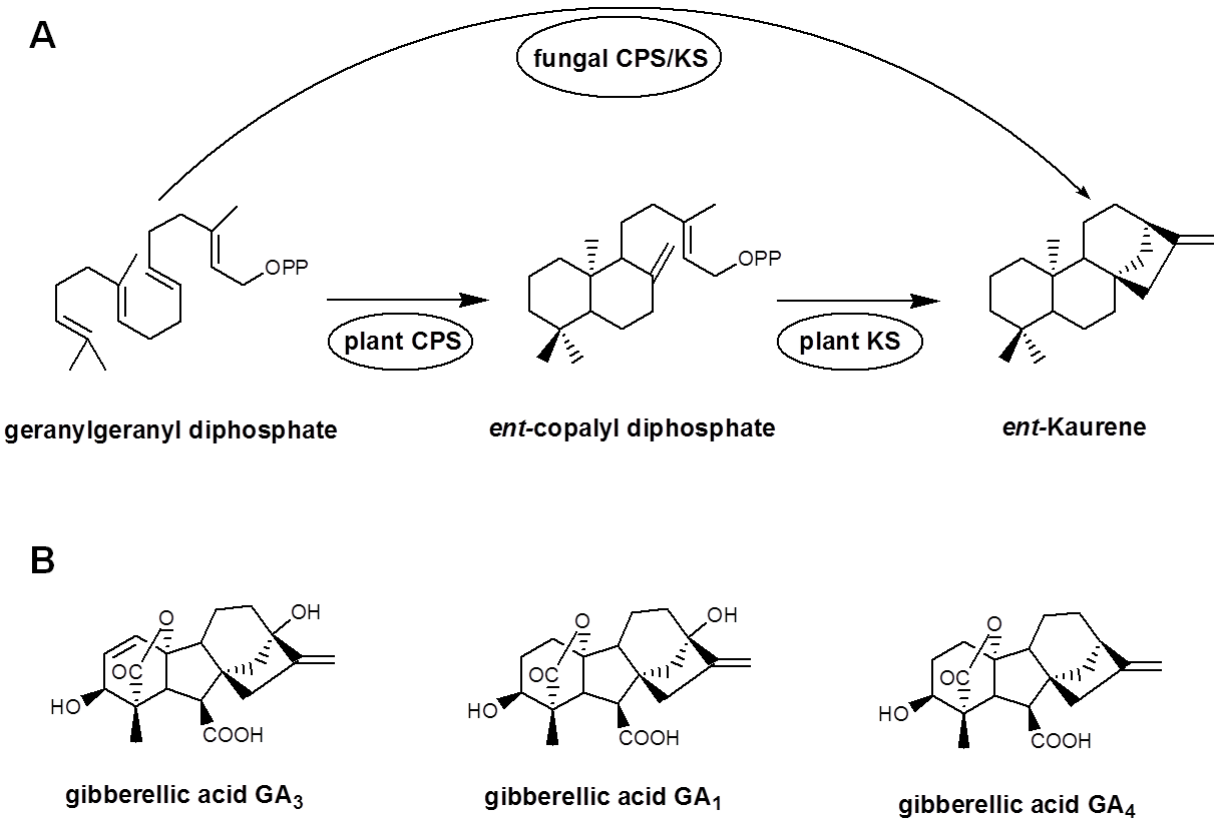


Figure I-5: Gibberellins biosynthesis; A: Comparison of enzymatic reactions towards *ent*-kaurene; two terpene synthases involved in plants with *ent*-copalyl diphosphate as real intermediate; direct conversion of geranylgeranyl diphosphate to *ent*-kaurene in fungi; **B:** three examples of bioactive gibberellic acids GA_3, GA_1 and GA_4 produced by plants and fungi; GA: gibberellic acid, CPS: *ent*-copalyl diphosphate synthase, KS: *ent*-kaurene synthase, CPS/KS: bifunctional enzymes from fungi.

The explanation for this evolution here most likely was due to the biological activity and its role for the respective organisms.
For both fungi and plants the compounds have a defined function, regulating length growth or are important for infection. For the pathogenic fungus it is feasible to mimic compounds of the plant to influence plant processes, but avoiding possible inhibition of its own production by developing a different biosynthetic pathway.
The physiological functions of taxoids in plants and endophytes are not known. There is no indication that endophytic fungi derived taxanes influence the host or that the fungi are somehow affected by plant derived taxoids.
All evolutionary mechanisms discussed, like clustering of pathways, horizontal gene transfer of single genes or entire gene clusters between different organisms as well as evidence for the opposite scenario, a convergent evolution of two biosyntheses with similar product distribution as described last for gibberellins lead to even more open questions regarding the origin of Taxol biosynthesis in plants and fungi.
No evolutionary scenario can be excluded or favoured. Only by clarification of the predicted fungal pathway and comparison to plant Taxol biosynthesis it might be possible to conciliate the very heterogeneous data on the in many respects interesting and so far mysterious phenomenon.

I.4 *Objective of the project*

The aim of this PhD project was the clarification of the Taxol biosynthetic pathway in *Taxus* associated endophytic fungi. As described in detail in the introduction chapters, up to the start of my work in 2007 there was no information on genes/enzymes that are involved in fungal Taxol biosynthesis. All reports concerning the Taxol biosynthetic pathway were obtained through examination of *Taxus* species. Thus, the choice of the approach for the project was done according to the existing phytochemical data of endophytic fungi on the one hand and because of theoretical considerations about possible scenarios of evolution of such a complex secondary metabolite pathway in two distantly related species.
Complex biosynthetic pathways leading to identical natural products seem to be unlikely to develop by convergent evolution, although the only present example the gibberellin biosynthesis evolved in accordance to this theory. In nature not only the occurrence of this

phenomenon of identical natural product production by not related organisms but also evidence for processes including trans-kingdom gene transfers are extremely rare.

However, the existence of such pathways for structurally extremely elaborate natural products, like Taxol, in distantly related organisms is rather surprising, even more though when obvious beneficial factors for the organisms are lacking. The occurrence of the Taxol biosynthesis in the evolutionary ancient species *Taxus* and in with *Taxus* associated endophytic fungi is rather extraordinary. Lateral gene transfer seemed to be one plausible explanation. However, the Taxol biosynthesis in *Taxus* tree is most likely not clustered and the acquisition of the Taxol biosynthesis by the tree in form of a complete biosynthetic cluster from the fungus seems also obscure as far as at this time there was no example of an event like this for any biosynthetic gene or gene cluster.

Through the identification of several of the Taxol biosynthetic genes it now became possible to study this assumed lateral gene transfer of a highly complex biosynthetic pathway between two only very distantly related species. Through to the availability of genetic data on the Taxol biosynthetic pathway in yew the *Taxus*/Taxus Endophyte system offers an excellent and unique model to study this lateral gene transfer and the evolution of a biosynthetic pathway for a structurally complex natural product.

All reports on Taxol producing endophytes up to that time focused on the detection of natural products via different analytical techniques, like immunological assays (CARUSO *et al.* 2000b), thin layer chromatography and mass spectrometry (STIERLE *et al.* 1993; HEINIG and JENNEWEIN 2009; ZHOU *et al.* 2010) (Table I-1), only. Also preliminary studies examining the biosynthesis during my own diploma thesis, for the first time on molecular biological level did not led to more information about the evolutionary origin of Taxol biosynthesis (HEINIG 2006). As discussed already the yields of taxanes in endophytes (ng-µg/L range) in comparison to *Taxus* (µg-mg/L range) are very low. Because of this reason, an approach trying to isolate the genes responsible for production by fungi from expression libraries, as done in case of *Taxus* genes seemed not promising, due to the assumption that the pathway is not well expressed.

Therefore the hypothesis made that the two pathways evolved not independently but gene transfer processes must have been involved together with the molecular biological assumptions led to the plan of experimental work.

The experimental work was structured as follows. The analytical part involved first the examination of the analysis of fungi already described to produce Taxol or taxanes from

different origin, meaning including an endophyte not originated from *Taxus* and including especially *Taxomyces andreanae*, the first fungus in which organic extracts taxanes were detected (STIERLE *et al.* 1993; HOFFMAN 2003). Nevertheless, to exclude the not proven but possible scenario of a loss of capability for the synthesis of the secondary metabolite due to loss of one or all genes responsible for the production through cultivation under laboratory environments also new *Taxus* endophytes were isolated from their natural environment.

The molecular biological part of the project involved the homology based hybridization screening of taxane producing endophyte's genomes in order to isolate and characterize the genes involved in the fungal pathway towards Taxol. This was done via Southern Blotting, through genomic lambda phage library screening and sequencing of isolated clones' inserts.

By these approaches the aim was not only to find evidence for the Taxol biosynthesis in fungi on genetic level, but also to discover the evolutionary origin of this secondary metabolite pathway, which, as one of very few examples, occurs in very distantly related species, plants and its endophytic fungi.

II Material and methods

II.1 *Material*

II.1.1 Chemicals and consumables

All chemicals used had at least *pro analysis* quality. Solvents for extraction were *for synthesis* grade, whereas solvents for chromatography were *gradient grade* or *LC/MS grade.* Chemicals, enzymes and consumables were purchased from the following companies.
Agilent Technologies (Waldbronn), Applied biosystems (Darmstadt), BioRad (München), BioTrend (Köln), Hartmann Analytik (Braunschweig), Eppendorf (Hamburg), GE Healthcare (Freiburg), GeneScript, Greiner (Solingen), Invitrogen (Leek, Netherlands), Kodak (Stuttgart), Millipore (Eschborn), New England Biolabs (Frankfurt am Main), Qiagen (Hilden), Roth (Karlsruhe), Sarstedt (Nümbrecht), Sigma (Deisenhofen), Stratagene (Amsterdam), VWR (Darmstadt), Whatman (Bender & Hobein, Bruchsal).

II.1.2 Kit systems

Macherey und Nagel, Düren	Nucleo Spin® Plasmid Kit
	Nucleo Spin® Extract II Kit
	NucleoFast 96 PCR plates
Qiagen, Hilden	QIAquick® Oligo Nucleotide Removal Kit
	QIAquick® Gel Extraction Kit
	Qiagen Lambda Midi/Maxi Kit
	Qiagen Oligitex mRNA Mini Kit
Inivitrogen, Leek, Netherlands	Zero Blunt® TOPO® PCR Cloning Kit
	pTrcHis2-TOPO® TA Expression Kit
Stratagene, Amsterdam	Lambda Dash® II / Gigapack® III XL Genomic Library Kit
	Lambda ZAP II/ Gigapack III Gold Kit

Thermo Scientific	Pierce® ECL Western Blotting Substrate
Roche, Grenzach-Wyhlen	Expand High Fidelity PCR System
Cardax Pharmaceuticals; Hawaii	Anti-Taxan Immunoassay
Fermentas, St. Leon-Rot	HexaLabel™ DNA Labeling Kit
Clonetech, Mountain View; USA	Marathon cDNA Amplification Kit

II.1.3 Enzymes and vectors

For DNA restriction the following enzymes were used: *Eco*RI, *Bam*HI, TSP509I, *Kpn*I, *Sac*I, *Eco*RV, *Hind*III, *Not*I, *Dpn*I. All enzymatic restrictions were performed according to manual provided by New England Biolabs (NEB).

Table II-1 shows the vectors used for cloning, expression and phage library construction.

Table II-1: **Vectors used in this thesis**

Vector	Genotype	Reference
pBlueskript SK-	f1(-), pUC, lacZ', *lac*, *T3*, *T7*, amp^R	Stratagene
pTrcHis2-TOPO	pBR322, *trc*, ampR, HIS-6-tag	Invitrogen
pCR Blunt	pUC, *lac*, *T7*, lacZα-ccdB, kan^R, zeo^R	Invitrogen
Lambda Dash II	*red*, *gam*, *T3*, *T7*	Stratagene
Lambda Zap II		Stratagene

II.2 *Microorganisms*

II.2.1 *Escherichia coli*

For amplification of plasmid DNA or for expression of proteins *Escherichia coli* (*E. coli*) strains *E. coli* DH5α™ (chemical competent) und TOP10 (electro competent) were used (Invitrogen; Leek, Netherlands).

Genotype DH5α™:
F⁻ 80*lac*ZΔM15Δ(*lac*ZYA-*arg*F), U169, *rec*A1, *end*A1, *hsd*R17 (r_k-, m_k+), *pho*A, *sup*E44λ⁻ *thi*⁻1, *gyr*A96, *rel*A1

Genotype *One Shot*® TOP10:
F⁻ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) 80*lac*ZΔM15ΔlacX74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L *end*A1 *nup*G λ-

For phage infection with either WT lambda DNA or recombinant lambda DNA and *in vivo excision* the following *E. coli* strain were used (Stratagene, Amsterdam)

Genotype XL1-Blue MRA (control strain for WT λ DNA)
Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 gyrA96 relA1 lac*

Genotype XL1-Blue MRA (P2)
XL1-Blue MRA (P2 lysogen)

Genotype SOLR
e14⁻(McrA⁻) Δ(*mcrCB-hsdSMR-mrr*)*171 sbcC recB recJ uvrC umuC::*Tn*5* (Kan^r) *lac gyrA96 relA1 thi-1 endA1* λR [F′ *proAB lac*IqZΔ*M15*] Su⁻ (nonsuppressing)

Genotype XL1-Blue MRF′
Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* [F′ *proAB lac*IqZΔ*M15* Tn*10* (Tet^r)]

II.2.2 Endophytic fungi from culture collection

Beside newly isolated and characterized endophytic fungi from *Taxus* three species described in literature as taxane producers were purchased from culture collections and examined phytochemically.

Taxomyces andreanae	CBS 279.92 described by (STROBEL *et al.* 1994)
H10BA2	NRRL 21209 described by (STIERLE *et al.* 2000)
UPH-12	NRRL 30405 described by (HOFFMAN 2003)

II.3 *Media and antibiotics*

All media listed were prepared with double distilled water and sterilized for 20 min at 121 °C prior to usage. For production of solid media 1.5-2 % agar were added before sterilization. Dependent on the experiment antibiotics (Table II-2) were added after sterilization and cooling to <60 °C or room temperature.

S7 medium

1	g/L	glucose
3	g/L	fructose
6	g/L	saccharose
1	g/L	sodium acetate
1	g/L	soytone
1	mg/L	thiamine
1	mg/L	biotin
1	mg/L	pyridoxale
1	mg/L	calcium pantothenate
3.6	mg/L	magnesium sulfate
6.5	mg/L	calcium nitrate
1	mg/L	potassium nitrate
2.5	mg/L	$ZnSO_4$
5	mg/L	MnCl
2	mg/L	$FeCl_3$
5	mg/L	phenylalanine
100	mg/L	sodium benzoate
1	mL	potassium dihydrogensulfate (1M,pH 6.8)

Potato dextrose medium

4	g/L	potato infus
20	g/L	glucose
pH 5.6		

YM – 6.3 medium

4	g/L	glucose
4	g/L	yeast extract
20	g/L	glucose
pH 6.3		

M1D medium

20	g/L	sucrose
360	mg/L	magnesium sulfate [7 H_2O]
200	mg/L	calcium nitrate [4 H_2O]
80	mg/L	potassium nitrate
200	mg/L	sodium sulfate
0.1	mg/L	thiamine
0.1	mg/L	pyridoxale
65	mg/L	potassium chloride
16.5	mg/L	sodium dihydrogenphosphate [H_2O]
4.5	mg/L	manganese sulfate [4 H_2O]
1.5	mg/L	zinc sulfate [7 H_2O]
1.5	mg/L	boric acid
2	mg/L	ferric citrate
0.75	mg/L	potassium iodide
0.5	mg/L	2,4-Dichlorophenoxyacetic acid
0.5	mg/L	nicotinic acid

***Luria-Bertani* (LB) medium**

10	g/L	NaCl
10	g/L	tryptone
5	g/L	yeast extract
pH 7.0		

2 YT medium

16	g/L	tryptone
10	g/L	yeast extract
5	g/L	NaCl
pH 7.0		

SOC medium

20	g/L	tryptone
5	g/L	yeast extract
0.5	g/L	NaCl
2.5	mL	KCl (1 M solution)
20	mL	glucose (1 M solution) after sterilization
5	mL	$MgCl_2$ (2 M solution) after sterilization

pH 7.0

NZY medium

5	g/L	NaCl
2	g/L	$MgSO_4 * 7H_2O$
5	g/L	yeast extract
10	g/L	NZ – amine

pH 7.5

NZY-Top medium

5	g/L	NaCl
2	g/L	$MgSO_4 * 7H_2O$
5	g/L	yeast extract
10	g/L	NZ-amine

pH 7.5

0.7 %(w/v) agarose

In Table II-2 the antibiotics are listed with concentrations of stock solutions as well as final working concentrations.

Table II-2: **Antibiotics used, concentrations of stock solutions and final concentration in the media**

Antibiotic	c (stock solution)	Working concentration
ampicillin	100 mg/mL	100 µg/mL
kanamycin	50 mg/mL	50 µg/mL
streptomycin	100 mg/mL	25 µg/mL

II.4 *Isolation and cultivation of microorganisms*

II.4.1 Isolation of endophytic fungi from *Taxus* bark

To isolate endophytic fungi from *Taxus* bark 0.5x0.5 cm pieces of the bark were cut out with a sterile scalpel and were surface sterilized for 5 min in 70 % ethanol. Then the inner bark was separated from the outer one and put on PDA agar supplemented with 25 mg/L streptomycin. The plates were incubated at room temperature until fungal growth was visible. The mycelium was then transferred to fresh plates by hyphal tip method (GUO *et al.* 2006).

II.4.2 Cultivation of endophytic fungi

The isolated endophytic fungi were cultivated on solid media, either on PDA+streptomycin or on YM-6.3 agar. Every week the fungi were transferred to fresh plates, by cutting out a piece of overgrown agar from an old one.

In liquid culture the fungi were grown in YM-6.3 medium until no more glucose is detectable. Furthermore the fungi were cultivated in S7 and M1D medium already described for taxane producing endophytes for 3 weeks (STIERLE *et al.* 1993). After cultivation the mycelium was harvested via filtration through two layers of miracloth and stored at -80 °C. Medium was extracted for taxane analytics.

II.4.3 Cultivation of *E. coli*

Cultivation of *E. coli* was done in either test tubes (3 mL medium) or in Erlenmeyer flasks (up to 500 mL medium). The small cultures were inoculated with a single colony from agar plate. For flask cultures 0.5-2 mL of an overnight culture pre-culture were inoculated.

II.4.3.1 Cultivation of *E. coli* for plasmid isolation

E. coli were cultivated overnight at 37 °C in 2YT medium supplemented with the appropriate antibiotic on a rotary shaker with 160 rpm. The cells (cultures up to 2 mL) were harvested by centrifugation for 1 min in a table centrifuge (5415D, Eppendorf). The obtained biomass was than either stored at -20 °C or used for plasmid isolation.

II.4.3.2 Cultivation of *E. coli* for protein expression

To express recombinant proteins in *E. coli* cultures (2YT, antibiotic) were inoculated with a defined volume of an over-night pre-culture (OD_{600nm}=0.05-0.1). These cells were grown at 37 °C, 160 rpm until optical density of 0.5. At this point expression was induced by addition

of 1 mM IPTG. Induced cultures were than incubated over night at 17-28 °C. Biomass was harvested by centrifugation at 4500 rpm, 4 °C (3S-R multifuge, Heraeus) and either stored at -20 °C or directly used for protein extraction. Protein expression in deep well plates was done using 24-square-well plates (Greiner). Cultures with a total volume of 2.5 mL (2YT, antibiotic, 1 mM IPTG) were incubated at 28 °C, 750 rpm using microtron shaker (Infors HT).

II.4.3.3 Preparation of *E. coli* for infection with λ-phages & *in vivo* excision

Prior to infection with λ-phage *E. coli* XL1Blue MRA (P2) or for control of packaging efficiency *E. coli* XL1Blue MRA were cultivated up to an OD_{600nm} of 0.5 to 1 in LB medium supplemented with 0.2 %(w/v) maltose and 10 mM $MgSO_4$. For phage liquid cultures *E. coli* XL1Blue (P2) was cultivated to saturation over night. All cultivations were done at 37 °C with 160 rpm.

For infection with Lambda Zap II containing phages *E. coli* MRF' strain was cultivated to an OD_{600nm} of 0.5-1 (LB supplemented with 0.2 %(w/v) maltose and 10 mM $MgSO_4$) at 37 °C with 160 rpm. *E. coli* MRF' and SOLR strains for *in vivo* excision protocol were grown over night in LB with supplements at 30 °C and 160 rpm.

II.4.4 Cultivation of λ-phage

For preparation of λ-phage liquid cultures 100 µL of a saturated *E. coli* XL1Blue MRA (P2)-culture, 100 µL of 10 mM $MgCl_2$/$CaCl_2$-solution and 100 µL of phage suspension were incubated at 37 °C for 15 min. The infected bacteria were transferred into 50 mL NZY medium and cultivated at 37 °C until the solutions were clear, meaning all bacteria were lysed by the phages (8-10 h). After incubation for another 15 min with 1 mL of $CHCl_3$ water/medium phase was removed and cell debris were separated from phage suspension via centrifugation. The lysate was used directly for isolation of λ-phage DNA.

II.4.5 Determination of cell density

The cell density of bacterial and yeast cultures was determined by measuring the optical density at 600 nm with Biophotometer (Eppendorf). 0.5-1 mL fermentation broths were analyzed in a plastic cuvette. As standard for calibration of the instrument the pure medium of the cultures was used.

II.5 *Molecular biological methods*

II.5.1 Isolation and purification of nucleic acids

II.5.1.1 Isolation of genomic DNA from plants and fungi

One gram of the biomaterial (mycelium of endophytic fungi or *Taxus* needles) was homogenized using liquid nitrogen, mortar and pistil. After addition of 10 volumes of CTAB-buffer (100 mM Tris, pH8, 20 mM EDTA, 1.4 M NaCl, 2 % (v/v) β-mercaptoethanol, 2 % (w/v) CTAB) and incubation for 1 h at 65 °C, cell debris was removed by centrifugation (15 min, 2000*xg*). The supernatant was extracted twice with equal volume of $CHCl_3$ / Isoamylalcohol 24:1. DNA was precipitated from the resulting solution by addition of 0.66 volumes cold isopropanol and pelleted by centrifugation (15 min, 4000xg) or in case of *Taxus* by spooling on a Pasteur pipette. After washing with 76 % EtOH, 10 mM NaOAc DNA was dissolved in 10mM Tris pH8.5, 100 ng/mL RNAseA at 4 °C over night. Concentration was determined by Nanodrop ND-1000. Purity was checked via agarose gel electrophoresis. All samples were stored at 4 °C.

II.5.1.2 Isolation of plasmid DNA and λ-phage DNA

Isolation of plasmid DNA from *E. coli* was done with the Macherey & Nagel Nucleospin plasmid kit according to the manufacturer's guidelines.

Phage DNA was isolated from lysates with Qiagen Lambda Midi/Maxi kits. This was also done according to the protocol provided by the manufacturer.

II.5.1.3 Precipitation of DNA

For increasing DNA concentration and removal of short DNA fragments and salts, respectively DNA solutions were precipitated by addition of ethanol (99.8 %) to a final concentration of 70 % (v/v). DNA was pelleted by centrifugation, washed with 70 % ethanol, air dried and solved in water.

II.5.1.4 Agarose gel electrophoresis

For analysis, separation or purification DNA fragments were supplemented with 1/5 of DNA loading buffer (0.1 % (w/v) bromophenol blue, 0.1 % (w/v) xylencyanol, 50 % (v/v) glycerol, to 50 mL with 1 x TBE buffer pH 8.3). After loading to agarose gel pockets (gels 0.7 % - 1.2 % (w/v) in TBE buffer) electrophoreses was carried out at 40 – 120 V for 30 min to 4 h.

For visualization either ethidium bromide was added to the gel (25 µg/L). In case of Southern Blotting 1/10 volume of SBGR solution was added to the samples. DNA signals were visualized using gel imager at a wavelength of 302 nm. The size of DNA fragments could be determined by comparison with DNA markers (1 kb, NEB, 100 bp, NEB, DNA Size Standard λ DNA *Hind* III + ΦX174 DNA *Hae* III digests, Finnzymes).

II.5.1.5 Purification of DNA fragments from agarose gels

DNA fragments produced by PCR or via restriction digest were purified over agarose gel electrophoresis. The desired fragment out of the mixture was cut out with a scalpel. The DNA from the gel slice was purified with NucleoSpin Extract II Kit from Macherey & Nagel according to manufacturer's protocol.

II.5.1.6 Determination of DNA concentration

DNA concentration was determined by measurement of absorption at 260 nm, using Nanodrop ND-1000. Therefore the instrument is calibrated with the buffer/water in which nucleotide sample is solved, followed by recording the absorption spectrum of 1 µL of the sample. Concentration was directly calculated by the software in ng/µL.

II.5.1.7 Isolation of total RNA & construction of cDNA library

Total RNA from endophytes was isolated by the so called Borax-method. All solutions and buffers were prepared with diethylpyrocarbonate (DEPC) treated water. Four grams of the biomaterial (mycelium of endophytic fungi) was homogenized using liquid nitrogen, mortar and pistil. After transfer to a SS-34 centrifugation tube (Beckmann-Coulter) and addition of 15 mL borax buffer (0.2 M sodium tetraborate, 30 mM EGTA, 1 %(w/v) SDS, 1 %(w/v) deoxycholate, 1 %(v/v) Nonidet NP-40, 2 %(w/v) polyvinylpyrolidone, 10 mM DDT, pH9.0) the biomass was further disrupted by treatment with ultraturrax (15.000 rpm) for 5 min. The resulting suspension was incubated for 1 h at 42 °C, followed by adding 1.2 mL of 2 M KCl solution and incubation for 1 h on ice, leading to precipitation of SDS. After clearing of the solution by centrifugation (12.000*xg,* 20 min, 4 °C) and filtration through sterile Miracloth 5 mL of a 8 M LiCl solution were added for selective precipitation of RNA over night at -20 °C. RNA precipitate was obtained by centrifugation (12.000*xg*, 30 min, 4°C) and washed three times with cold 2 M LiCl solution prior to re-suspension in 2.8 mL TES buffer (50 mM Tri/HCl pH 5.7, 5 mM EDTA, 50 mM NaCl) supplemented with 1 M CsCl. This suspension was used to overlay 1.2 mL of TES buffer supplemented with 5.7 M CsCl in an ultra

centrifuge tube. Density gradient centrifugation was carried out at room temperature with 100.000*xg* for 16 h. Pure RNA occurs as a pellet, whereas other components, for example DNA, are in the supernatant according to their lower density. RNA was dissolved in 500 µL TE buffer (10 mM Tris/HCl pH8.0, 1 mM EDTA) and used for isolation of messenger RNA via the batch protocol of Qiagen Oligotex mRNA Mini Kit. With the mRNA a cDNA-RACE library was constructed using Marathon cDNA Amplification Kit (Clonetech) according to the manufacturer's instructions.

II.5.2 *In vitro* amplification of DNA by polymerase chain reaction

Polymerase chain reaction (PCR) is a method for selective amplification of DNA fragments from plasmid or genomic DNA (templates). Two oligo nucleotides complementary to the 5'-prime-end and inverse complementary to 3'-prime-end of the desired DNA fragments are used as primers (see tables in the following chapters). After annealing of the primers at their respective binding temperature to the template they are elongated by the added thermostable polymerases. All primers used in this thesis were ordered from Invitrogen or MWG Biolabs. As polymerases Herculase II (Stratagene) or Expand high fidelity system (Roche) were used. PCR reactions were pipetted according to manufacturer's protocols. PCR was performed with GeneAmp PCR System (Applied Biosystems).

Table II-3: General setup of PCR programs; T_m is the melting temperature of the primers, it is calculated with the formula: T_m=(A+T)*2+(G+C)*4-6.

Cycle No.	Reaction step	Time	Temperature
1	denaturation of template DNA	3-5 min	95 °C
	denaturation	30 sec	95
25-35	annealing of primers	30 sec	T_m
	elongation	30 sec-1 min /kb	68 °C or 72 °C
1	final elongation	7 min	68 °C or 72 °C

II.5.2.1 Primers for amplification of terpene synthase 0021_TS_1762

For amplification of terpene synthase 0021_TS_1762 from EF0021 the primers listed in Table II-4 were used. 0021_TS_1762_cosyn synthesized by the company GeneScript was amplified out of pUC57 construct delivered. All other PCRs were made with pTrcHis2-0021_TS_1762_cosyn construct as a template, amplifying the whole vector (plasmid backbone and gene) leading to variance according to primer sequences. For rapid amplification of cDNA ends (RACE) from EF0021 cDNA-library, primers were used in PCR according to protocol of Marathon cDNA amplification kit (Clonetech) against adapter primer1 (AP1: 5'–CCATCCTAATACGACTCACTATAGGGC–3') delivered by the manufacturer.

Table II-4: Primer used for amplification of 0021_TS_1762_cosyn, 0021_TS_1762_del, intron1 variants of 0021_TS_1762_cosyn and amplification of 0021_TS_1762 cDNA and the two housekeeping genes citrate synthase (Cit-primer) and pyruvate kinase (PKi-primer) via RACE-PCR.

Primer name	Primer sequence
1762_i1_forw_1	5'GCACGTCTGACCTTCCCGCAGGAATCTGCTGTTGGTCAGTTCTCTTGG-3'
1762_i1_forw_2	5'-CGTCTGACCTTCCCGCAGGAATCTGCTGTTGGTCAGTTCTCTTGG-3'
1762_i1_forw_3	5'-CTGACCTTCCCGCAGGAATCTGCTGTTGGTCAGTTCTCTTGG-3'
1762_i1_forw_4	5'-ACCTTCCCGCAGGAATCTGCTGTTGGTCAGTTCTCTTGG-3'
1762_i1_forw_5	5'-TTCCCGCAGGAATCTGCTGTTGGTCAGTTCTCTTGG-3'
1762_i1_forw_6	5'-CCGCAGGAATCTGCTGTTGGTCAGTTCTCTTGG-3'
1762_i1_forw_7	5'-CAGGAATCTGCTGTTGGTCAGTTCTCTTGG-3'
1762_i1_forw_8	5'-GAATCTGCTGTTGGTCAGTTCTCTTGG-3'
1762_i1_rev_1_P	5'-phosphate-GTACAGACGTTCGATAACGTCGGTGTAAGACTGAGACGG-3'
1762_i1_rev_2_P	5'-phosphate-CAGACGTTCGATAACGTCGGTGTAAGACTGAGACGG-3'
1762_i1_rev_3_P	5'-phosphate-ACGTTCGATAACGTCGGTGTAAGACTGAGACGG-3'
1762_i1_rev_4_P	5'-phosphate-TTCGATAACGTCGGTGTAAGACTGAGACGG-3'
1762_i1_rev_5_P	5'-phosphate-GATAACGTCGGTGTAAGACTGAGACGG-3'
1762_i1_rev_6_P	5'-phosphate-AACGTCGGTGTAAGACTGAGACGG-3'
1762_i1_rev_7_P	5'-phosphate-GTCGGTGTAAGACTGAGACGG-3'
1762_del_forw_P	5'-phosphate-GACATGAAAGAGGACTCTCTG-3'
1762_del_rev	5'-GTCCTGGATGAAGATGAATTCG-3'
1762_cosyn_forw	5'-ATGCCGCCAGCAGGTATCAGCTTCCG-3'
1762_cosyn_rev	5'-TTATACACGCAGTTTACCCAGG-3'
CycC1762 Race for1	5'-GATGCCTCCCGCTGGGATCTCCTTTCGATCTC-3'

Table II-4 continued

CycC1762 Race rev1	5'-GATATCACCAATATCACCAAGCTTCTTCTCAG-3'
CycC1762 Race for2	5'-CGAACCGCGTCAGGCCTGCCGAACACTTGC-3'
CycC1762 Race rev2	5'-GTACGTACTGCTTCATAGCGTATGTCATCAC-3'
Cit_forw	5'-GCTCCTCACCGGGCAAGTTCCATCCACAAGC-3'
Cit_rev	5'-CGATGATGCATAGTCCATCAAAGCCTCAAATCG-3'
PKi_forw	5'-CCAAGGTCATCCAACCAGGACGCATCATCTACG-3'
PKi_rev	5'-CATGACACAATCGGCTCCATCCGTAACAGCG-3'

II.5.2.2 Amplification of Internal Transcribed Spacer (ITS) regions

ITS regions of isolated *Taxus* endophytes were amplified via PCR using the universal primers ITS1 (5'-TCCGTAGGTGAACCTGCGG-3') and ITS4 (5'-TCCTCCGCTTATTGATATGC-3') (SIM *et al.* 2010). For amplification the 2x PCR-Master mix Solution (i-Max II, INtRON Biotechnology), 1 µL of each Primer (50 µM) and 20 ng genomic DNA (water to 25 µL) were used. PCR products were purified with NucleoFast 96 PCR plates (Machery-Nagel) according to manufacturer's protocol and sequenced directly using ITS1 and ITS4 as sequencing primers.

II.5.3 Enzymatic modifications of DNA and transfer to *E. coli*

II.5.3.1 Restriction with endonucleases

Enzymatic hydrolysis with endonucleases was used for control of ligations, modification of DNA ends for further ligation reactions, linearization of plasmids, partial digest of genomic DNA for phage library construction, insert size determination of phage vector inserts and complete restriction for Southern Blotting analysis. In case of plasmid controls and insert size determination 300 ng-1 µg DNA were used and reactions performed for 3 h. For partial restriction for library construction 3-5 µg of DNA and 1.25 U of restriction enzyme were used in reaction. Restriction was stopped after 30 min by cooling on ice and immediate precipitation.

For Southern Blotting analysis 110 µg DNA were cut by 100 U of restriction enzyme over night.

II.5.3.2 Ligation

Linking DNA fragments was performed with T4 DNA ligase (NEB). Reaction was done using a molar ratio vector:insert of 1:3 or 1:5 together with buffer, ligase and water (total volume 5-10 µL) by incubation for 3-24 h at 16 °C or 4 °C. In case of 1762 intron1 variants the linear PCR products were re-circulated.

II.5.3.3 Transformation of DNA into competent *E. coli*

All transformations were performed with commercially available competent cells (II.2.1). For electro-transformation cell aliquots were thawed on ice while DNA (3-5 µL of ligation reaction or 50-200 ng of plasmid) was placed in a pre-cooled sterile electro cuvette. After adding cells for electroporation with 2.5 kV for 5 msec 500-750 µL SOC medium were added. The suspension was incubated at 37 °C for up to 1 h for regeneration of the cells prior to plating on selective agar plates.

Transformation into chemical competent *E. coli* was done according to manufacturer's protocol.

II.5.4 Construction of genomic λ-phage libraries

II.5.4.1 Ligation into Lambda Dash II

The DNA fragments obtained by non complete restriction digest were cloned into phage vector Lambda Dash II (predigested with *Bam*HI). Therefore 1 µL vector (1 µg), 0.5 µL 10x T4 Ligase buffer, 0.5 µL T4-DNA Ligase and 0.4 µg (1-3 µL) of insert DNA were combined in a plastic vial. After filling reaction to a total volume of 5 µL, ligation was incubated over night at 4 °C.

II.5.4.2 Packaging reaction

The packaging reaction represents the critical step of library construction since it affects the number of phage clones obtained in the end. 2-3 µL of each ligation reaction were transferred to a vial of Gigapack III XL Packaging extract. The suspension was mixed by stirring with a pipette tip and incubated for 1 hour and 45 min at 22 °C. Due to their high temperature sensitivity the packaging extracts have to be thawed quickly. DNA has to be added to the extract when it just began to thaw. Furthermore packaging time should not be enlarged because of decreasing packaging efficiency in case of longer incubation. The reaction was stopped by addition of 500 µL of SM buffer (50 mM Tris/HCl pH7.5, 100 mM NaCl, 8 mM

$MgSO_4$*7 H_2O, 0.01 % gelatin). After extraction with 20 µL of $CHCl_3$ for protein removal the final library (water phase) was transferred to a fresh 1.5 mL vial and stored at 4 °C.

II.5.4.3 Determination of library titer

To determine the number of phage clones in the library *E. coli* XL1Blue(P2) cells were incubated to an optical density (OD_{600nm}) between 0.5 and 1. After harvesting the cells they were resuspended in 10 mM $MgSO_4$ and diluted to an OD_{600nm} of 0.5. One µL of the library was added to 200 µL cell suspension and the resulting mixture was incubated for infection for 15 min at 37 °C. Afterwards cells were pipetted into 3 mL of pre-warmed (50 °C) NZY-Top agar, mixed and spilled onto pre-warmed NZY agar plates. The liquid agar was dispensed over the plate by gently shaking. Plates were stored at room temperature until the Top-agar was solid. Plaque development was carried out by incubation at 37 °C over night. The size of the library could be determined by multiplication of plaque number with the total library volume. In general all titers were specified in "plaque forming unit per milliliter" (pfu/mL).

II.5.4.4 Determination of packaging efficiency

The packaging efficiency of Gigapack III XL extracts could be determined with WT-λ-DNS (λcI857 *Sam7*), delivered with the construction kit from Stratagene. Therefore 0.2 µg were packed into phages as described above. The reaction was stopped by adding 1 mL SM buffer (50 mM Tris/HCl, pH7.5, 100 mM NaCl, 8 mM $MgSO_4$*7 H_2O, 0.01 % gelatin) and diluted 1:10000. 200 µL *E. coli* XL1Blue MRA cell suspension in 10 mM $MgSO_4$ with an OD_{600nm}=0.5 were infected with 10 µL of the phage dilution and plated as described before. For a good efficiency around 400 plaques should grow under these conditions.

II.5.4.5 Plating of the library

Volume of phage suspension used for infection was chosen according to the titer of the library. Aim was to get as many clearly separated plaques per plate as possible. Analog to the titer determination procedure libraries were plated on NZY agar plates.

II.5.4.6 Plaque lifts

To screen library plates via hybridization it was necessary to transfer the plaques or at least part of the phages from each plaque onto positively charged Nylon membranes (Hybond N+, diameter 82 mm or 136 mm, GE Healthcare). These membranes were put onto the plates for 5 min, avoiding bubbles between membrane and agar. During this period the signs made on

the membranes before were copied on the plates to allow later identification of plaques responsible for radioactive signals. After removal from the plates membranes were incubated 2 min in denaturation buffer (1.5 M NaCl, 0.5 M NaOH) followed by neutralization for 5 min (1.5 M NaCl, 0.5 M Tris/HCl, pH 7.5), short washing in 0.2 M Tris pH 7.5 and 2xSSC (0.3 M NaCl, 34 mM sodium citrate, pH 7.0) and 30 min of air drying on a piece of whatman-paper. The dry filters were heated to 120 °C for 30 min in order to irreversibly fix the DNA on the membranes.

II.5.5 Construction of Lambda Zap II libraries from Lambda Dash II clones

II.5.5.1 Ligating the inserts

The DNA fragments of λ-Dash II positive clones obtained by restriction digest with TSP509I (NEB) were cloned into phage vector Lambda Zap II (predigested with *Eco*RI). Therefore 1 µL of vector (1 µg), 0.5 µL of 10x T4 Ligase buffer, 0.5 µL of T4-DNA Ligase and 0.4 µg of insert DNA were combined in a plastic vial. After filling up the reaction to a total volume of 5 µL, ligation was incubated over night at 4 °C.

II.5.5.2 Packaging and plating of Lambda Zap II libraries

Packaging and plating of libraries was done as described for λ-Dash II libraries. In contrast to the size selective packaging extracts Gigapack III XL here non size selective packaging extracts Gigapack III Gold were used (insert size 0-10 kb). For infection bacterial strain *E. coli* MRF' was used. Positive clones were cored from the agar plate and inserts were prepared for sequencing by *in vivo* excision into pBlueskript SK-, instead of isolation of Lambda phage DNA from liquid phage cultures.

II.5.5.3 *In vivo* excision of Lambda Zap II inserts

Excision was done according to the manufacturer's protocol (instruction manual, Lambda Zap II Predigested *Eco*RI/CIAP-Treated Vector Kit, Stratagene). *E. coli* clones were transferred to liquid culture for plasmid isolation.

II.5.6 Southern Blotting

Southern blotting experiments were performed according to a protocol for neutral transfer (SAMBROOK and RUSSELL 2001). Complete digest of up to 110 µg genomic DNA was separated on an 0.7 %(w/v) agarose gel for 4-5 h at 40 V. Preparation of the gel for blotting

involved one 30 min denaturing step in denaturation buffer (1.5 M NaCl, 0.5 M NaOH) and two neutralization steps of 30 min and 15 min, respectively in neutralization buffer (1.5 M NaCl, 0.5 M Tris/HCl pH7.5). Afterwards the blot was built as shown in Figure II-1. Capillary transfer of the DNA (transfer buffer 10x SSC, 1.5 M NaCl, 170 mM sodium citrate) from gel onto positively charged Nylon membranes (Hybond N+, size of the respective gel, GE Healthcare) took 12-24 h. After heat fixation (120 °C, 30 min) hybridization was performed as described below (II.5.8).

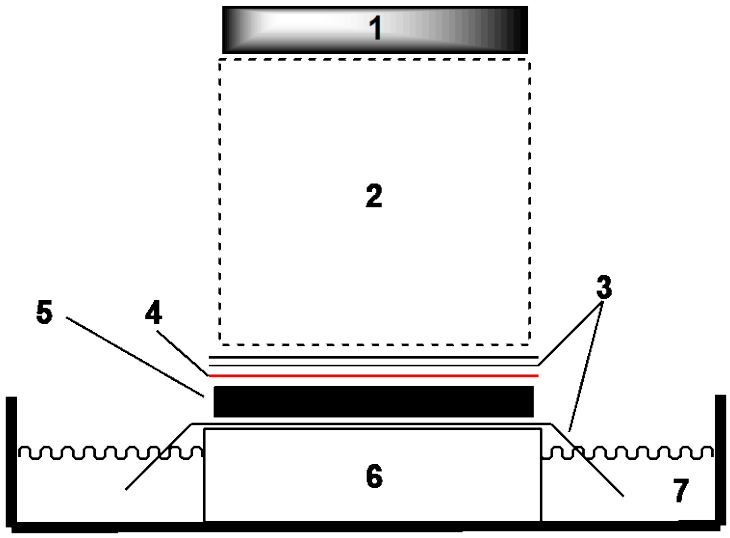


Figure II-1: Schematic construction of a capillary driven Southern blot: 1 weight, 2 pack of paper towels (5-10 cm) for creation of capillary force, 3 layers of whatman paper, 4 positively charged nylon membrane, 5 agarose gel, 6 inverted gel tray as support, 7 glass basin filled with transfer buffer.

II.5.7 Construction of probes for hybridization screening

II.5.7.1 Construction of probe templates

For hybridization templates for labeling have to be constructed first. This can either be done by PCR followed by labeling of the resulting fragments with Klenow-polymerase or by synthesis of oligo-nucleotides labeled afterwards with polynucleotide-kinase. Oligo-nucletides are listed in Table II-5. PCR was performed as described before with primers listed in Table II-6.

Table II-5: Oligo-nucleotides used as templates for labeling for taxadiene-5α-hydroxylase (T5H).

Oligo name	Oligo sequence
T5H-1	5'-GGCATCCCACAGTAGTACTCTGCGGCCCTGCGGGAAACCGGCTT ATTCTGTCCAACGAGGAGAAGCTGGTGCAGATGTCG-3'
T5H-2	5'-CCACCACTTCGCCAATGGCTTTGATTTTCAAGCTCTTGTCTTCCAA TCCAGAATGCTATCAAAAAGTAGTTCAAGAGC-3'

Table II-6: Primers for amplification of Taxol biosynthesis gene fragments for construction of hybridization probes; TDS: taxadiene synthase; T13H: taxane-13α-hydroxylase, T5H: taxadiene-5α-hydroxylase.

Primer name	Primer sequence
TDS1-forw	5'-GCAGCGCTGAAGATGAATGC-3'
TDS1-rev	5'-CGATTCGATACCCCACGATCC-3'
TDS2-forw	5'-GCCCTCGGCCTCCGAACCC-3'
TDS2-rev	5'-GCCATGCCGGATTCTTTCCACC-3'
TDS3-forw	5'-GGTGGAAGGAATCCGGCATGGCAG-3'
TDS3-rev	5'-GTCGCCAGCTCAAGGATACAAGCTC-3'
T13H-forw	5'-ATGGATGCCCTTAAGCAATTGGAAGTTTCCCC-3'
T13H-rev	5'-GCTCCTGCAGGTGCTCC-3'
T5H-forw	5'-CCAACGAGGAGAAGCTGGTGC-3'
T5H-rev	5'-GGTGGTGTCATAGGAGGCATGGAGC-3'

II.5.7.2 Labeling of double-stranded DNA with [α-^{32}P]dATP

Labeling with radioactive nucleotides was done with HexaLabel™ DNA labeling kit (Fermentas) according to manufacturer's protocol. All reactions were done using 100 ng of template DNA and 1.85 MBq [α-^{32}P]dATP (Hartmann Analytic). In case of TDS 33 ng of each fragment were used. Purification of the probes (removal of not incorporated nucleotides) was done using Sephadex G-25 gelfiltration columns (GE Healthcare) according to manufacturer's protocol. Before adding the probes to hybridization reaction they had to be denatured at 99 °C for 10 min.

II.5.7.3 Labeling of oligo-nucleotides with [γ-^{32}P]dATP

Labeling of oligo-nucleotides was done via transfer of one radioactive phosphate to each oligo using polynucleotide-kinase (PNK). Therefore 20 pmol of each oligo (Table II-5), 5 µL 10x PNK buffer (NEB), 1.85 MBq [γ-^{32}P]dATP (Hartmann Analytic), 20 U polynucleotide-kinase (NEB, 10 U/µL) and water to 50 µL were put together , mixed and incubated at 37 °C for 30 min. After incubation the reaction mixture was purified with QIAquick Nucleotide Removal Kit (Qiagen) according manufacturer's protocol for cleanup of radioactive samples. The eluted probe was directly used for hybridization.

II.5.8 Lambda phage library hybridization screening

II.5.8.1 Hybridization with long probes (II.5.7.2)

Heat fixed membranes (Nylon N+, GE Healthcare) were transferred in hybridization rolls (GFL) and supplemented with 20 mL Roti-Hybri-Quick (Ready to use hybridization buffer, Carl Roth GmbH) + 100 µg/mL of salmon sperm DNA (Sigma) in. Prehybridization was carried out at 55 °C for 3h. Probes were added and hybridization was performed over night at 55 °C. Washing steps were processed according to the manufacturer's protocol (Roti Hybri Quick Manual, Carl Roth GmbH, Karlsruhe) for 30 min 1:2 dilution of the buffer, for 30 min 1:5 dilution of the buffer and 15 min 1:10 dilution of the buffer at 55°C (IPS around 20-100 for phage membranes). Southern blots were washed till the detectable radioactivity was reduced to max. 100 IPS. In case of re-used probes either the complete hybridization solution was transferred directly to a new hybridization roll or the stored solution (-20 °C) was thawed, denatured for 10 min at 99 °C and then added to the new roll.

Visualization of signals occured by autoradiography on preflashed (Preflash unit, GE Healthcare) X-Ray films (Hyperfilm MP, GE Healthcare) after 4d incubation at -80°C in case of library screening experiments. Southern Blots were incubated up to 21 days, depending on stringency of the washing steps. For development, developing and fixing solutions from Sigma-Aldrich were used.

II.5.8.2 Hybridization with oligo-nucleotides (II.5.7.3)

Hybridization with oligo-nucleotides was carried out at 42 °C instead of 55 °C because of the reduced length of the probes. Re-usage of probes did not require denaturation, since single stranded probes were used.

II.5.8.3 Identification and isolation of plaques

Identification of potential candidate plaques detected by signals obtained from hybridization experiments was done by comparing marks on the membranes/films to the films to those on the plates. Due to the unique possible orientation plaques fitting to signals could be clearly identified.

The “positive” plaques were cut out including the agar under and around it with the bigger end of a Pasteur pipette. The phages were eluted from these agar/plaque pieces by incubation in 300 µL of SM buffer for two hours at 4 °C. Phage solutions could be stored at 4 °C for months.

II.5.8.4 Re-screening of plaques isolated in first screening round

Due to the high plaque density on the initial library plates all isolated phages were re-screened. Therefore the complete procedure including plating, plaque lifting, hybridization and detection was repeated with phage solutions obtained according to II.5.8.3. In order to be sure to be able to identify and isolate single plaques phages were plated less dense. Clones from plates, where a strong enrichment of positive signals could be detected, were regarded as positive. Plaques were isolated and eluted in 200-300 µL of SM buffer. These suspensions were used for infection of liquid phage cultures.

II.5.9 Sequencing

II.5.9.1 Sanger sequencing

Sequencing of plasmids and PCR products was performed *in house* of the IME by Raphael Soeur using 3730 DNA analyzer (Applied Biosystems) by Raphael Soeur at Fraunhofer IME. For this purpose 50 ng-1 µg of DNA were supplemented with 100 pmol of sequencing primer in a total volume of 30 µL.

II.5.9.2 Next generation sequencing

Phage DNA of PC4 and PC9 was shotgun sequenced by MWG Biotech GmbH. Fungal genomes were sequenced by Seq-It GmbH with 454 technology (EF0021) and by Source BioScience imaGenes GmbH by paired end sequencing (HiSEQ, *Taxomyces andreanae*). The assembled sequence data was delivered by the companies in fasta format.

II.5.9.3 Programs for sequence analysis

Analysis of sequencing results was done using CLC workbench 6.3 for local blast x analysis of fungal genomes and phage clone DNA, DNASTAR Lasergene package for assembly and alignment and construction of phylogenetic trees, "FGENESH"-software (http://linux1.softberry.com/) for intron/exon prediction and Clone Manager Suite 8 for restriction analysis. Annotation of genes/enzymes was done using NCBI blast x/n/p search.

II.5.9.4 Proteins used for local blast analysis

Following tables list the protein sequences used for local blastx analysis with CLC workbench in order to identify terpene synthases and P450 oxygenases that might be involved in secondary metabolism of the fungal species and are possible candidates for enzymes involved in fungal Taxol biosynthetic pathway (Table II-9, Table II-8). Furthermore the sequence list of already known proteins of Taxol biosynthesis from *Taxus* also used for blast x analysis is shown in Table II-7.

Table II-7: Taxol biosynthesis protein sequences from Genebank used for local blastx analysis.

Accession No.	Description
Q41594	taxa-4(5),11(12)-diene synthase [*Taxus brevifolia*]
AAS89065	taxoid 2-α-hydroxylase [*Taxus canadensis*]
AAU93341	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*]
Q8W4T9	taxoid 13-α hydroxylase [*Taxus cuspidata*]
Q9AXM6	5-α-taxadienol-10-β-hydroxylase [*Taxus cuspidata*]
AAX59902	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*]
AAO66199	taxane 14β-hydroxylase [*Taxus cuspidata*]
Q8LL69	3'-*N*-debenzoyl-2'-deoxytaxol *N*-benzoyltransferase [*Taxus canadensis*]
AAL92459	phenylpropanoyltransferase [*Taxus cuspidata*]
AAY16196	10-deacetyl baccatin III acetyltransferase [*Taxus wallichiana var. mairei*]
Q9M6F0	taxadien-5-α-ol O-acetyltransferase [*Taxus cuspidata*]
AAT79354	taxane 2-α-O-benzoyltransferase [*Taxus x media*]
AAQ75553	taxoid 7-β-hydroxylase [*Taxus cuspidata*]

Table II-8: P450 oxygenase protein sequences from Genebank used for local blastx analysis.

Accession No.	Description
AAS89065	taxoid 2-α-hydroxylase [*Taxus canaden*sis]
NP_180997	CYP710A1; C-22 sterol desaturase/ oxygen binding [*Arabidopsis thaliana*]
NP_196416	flavonoid 3'-monooxygenase/ oxygen binding [*Arabidopsis thaliana*]
O13317	isotrichodermin C-15 hydroxylase [*Fusarium sporotrichioides*]
CAA75566	cytochrome P450 monooxygenase [*Gibberella fujikuroi*]
AAX59902	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*]
XP_001881086	cytochrome P450 monooxygenase [*Laccaria bicolor* S238N-H82]
AAU93341	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*]
CAE09055	cytochrome P450 oxidoreductase [*Gibberella fujikuroi*]
Q8W4T9	taxoid 13-α hydroxylase [*Taxus cuspidata*]
CAA75567	cytochrome P450 monooxygenase [*Gibberella fujikuroi*]
CAA75565	cytochrome P450 monooxygenase [*Gibberella fujikuroi*]
BAA33717	cytochrome P450 [*Coprinopsis cinerea*]
XP_001835122	cytochrome P450 [*Coprinopsis cinerea* okayama7#130]
NP_197962	GA3; *ent*-kaurene oxidase/ oxygen binding [*Arabidopsis thaliana*]
AAK11564	ent-kaurenoic acid hydroxylase [*Arabidopsis thaliana*]
NP_850337	CYP98A3; p-coumarate 3-hydroxylase [*Arabidopsis thaliana*]
ABQ22962	cytochrome P450 [*Bacillus subtilis* subsp. subtilis str. 168]
CAH64679	cytochrome P450 monoxygenase [*Botryotinia fuckeliana*]
AAO64248	trichothecene C-8 hydroxylase [*Fusarium sporotrichioides*]
XP_001886909	cytochrome P450 monooxygenase CYP63 [*Laccaria bicolor* S238N-H82]
BAI52803	fusicoccadiene 8-ol C-15 hydroxylase [*Alternaria brassicicola*]
CAE76652	cytochrome P450 monooxygenase [*Botryotinia fuckeliana*]
BAI52800	fusicoccadiene C-8 hydroxylase [*Alternaria brassicicola*]
BAI52801	cytochrome P450 [*Alternaria brassicicola*]
XP_001830548	cytochrome-450 hydroxylase [*Coprinopsis cinerea* okayama7#130]
Q9AXM6	5-α-taxadienol-10-β-hydroxylase [*Taxus cuspidata*]
CAP58781	cytochrome P450 monooxygenase [*Botryotinia fuckeliana*]
XP_747185	cytochrome P450 [*Aspergillus fumigatus* Af2□3]
EDP47672	cytochrome P450 alkane hydroxylase [*Aspergillus fumigatus* A1163]
EDP55514	cytochrome P450 alkane hydroxylase [*Aspergillus fumigatus* A1163]
XP_002910681	cytochrome P450 oxidoreductase [*Coprinopsis cinerea* okayama7#130]
NP_851105	BR6OX1 brassinosteroid-6-oxidase [*Arabidopsis thaliana*]
NP_566462	CYP90D1; oxidoreductase, putative cytochrome P450 [*Arabidopsis thaliana*]

Table II-9: Terpene synthase protein sequences from Genebank used for local blastx analysis.

Accession No.	Description
BAD29971	aphidicolan-16β-ol synthase [*Phoma betae*]
Q6WP50	Presilphiperfolan-8-β-ol synthase [*Botryotinia fuckeliana*]
ABC4641	*ent*-kaurene synthase [*Gibberella intermedia*]
XP_001832573	CND15p [*Coprinopsis cinerea* okayama7#130], Cop1
XP_001836556	terpene synthase [*Coprinopsis cinerea* okayama7#130], Cop2
ABC4641	*ent*-kaurene synthase [*Gibberella intermedia*]
XP_001832925	hypothetical protein CC1G_12294 [*Coprinopsis cinerea* okayama7#130], Cop3
XP_001836356	hypothetical protein CC1G_06441 [*Coprinopsis cinerea* okayama7#130], Cop4
XP_001834007	hypothetical protein CC1G_09421 [*Coprinopsis cinerea* okayama7#130], Cop5
XP_001832549	hypothetical protein CC1G_03563 [*Coprinopsis cinerea* okayama7#130], Cop6
Q38710	(-)-abieta-7(8),13(14)-diene synthase [*Abies grandis*]
BAB39207	diterpene cyclase-2 [*Kitasatospora griseola*]
BAG16278	diterpene cyclase [*Nocardia brasiliensis*]
BAG30962	copalyl diphosphate synthase [*Phomopsis amygdali*]
NP_001053841	*ent*-kaur-16-ene synthase [*Oryza sativa* Japonica Group]
BAB62102	aphidicolan-16β-ol synthase [*Phoma betae*]
O04408	*ent*-copalyl diphosphate synthase [*Pisum sativum*]
Q947C4	levopimaradiene synthase [*Ginkgo biloba*]
Q675L5	isopimaradiene synthase [*Picea abies*]
AAT65717	geranylgeranyl diphosphate synthase [*Aspergillus flavus*]
Q8NJA1	trichodiene synthase [*Fusarium mesoamericanum*]
BAB39206	diterpene cyclase-1 [*Kitasatospora griseola*]
XP_002396668	hypothetical protein MPER_03050 [*Moniliophthora perniciosa* FA553]
O13284	*ent*-kaur-16-ene synthase [*Phaeosphaeria sp*. L487]
Q03471	aristolochene synthase [*Penicillium roqueforti*]
P13513	trichodiene synthase [*Fusarium sporotrichioides*]
Q675L4	levopimaradiene [*Picea abies*]
BAF45924	fusicoccadiene synthase [*Phomopsis amygdali*]
Q50EK2	levopimaradiene synthase [*Pinus taeda*]
XP_001265719	terpene synthase metal binding domain protein [*Neosartorya fischeri* NRRL 181]
BAG30961	phyllocladan-16α-ol synthase [*Phomopsis amygd*ali]
Q41594	taxa-4(5),11(12)-diene synthase [*Taxus brevifolia*]
XP_003013365	terpene synthase family protein [*Arthroderma benhamiae* CBS 112371]
Q00G37	*ent*-cassa-12,15-diene synthase [*Oryza sativa* Japonica Group]
XP_001887869	mycorrhiza-upregulated terpene synthase mbd protein [*Laccaria bicolor* S238N-H82]
Q9UVY5	*ent*-kaur-16-ene synthase [*Gibberella fujikuroi*]
XP_002384087	lanosterol synthase, putative [*Aspergillus flavus* NRRL3357]

II.6 *Biochemical Methods*

II.6.1 Protein extraction

For protein extraction harvested biomass was re-suspended in 3 mL/g cells of buffer T (100 mM Tris/HCl pH7.4, 10 mM $MgCl_2$, 5 mM β-mercaptoethanol). Cells were disrupted mechanically with glass beads (0.2–0.5 mm, Roth) and raw extract obtained by centrifugation (4500 rpm, Heraeus 3S-R, or in case of multi – well plates 3280*xg* Eppendorf, 5810R). All centrifugation steps were performed at 4 °C and samples were always stored on ice. Raw extracts were either used directly for analysis and test reactions or stored on ice for a maximum of one week. Protein concentration was determined in triplicates via Bradford assay (Roth, 1:5 dilution).

II.6.2 Terpene synthase activity assay

For *in vitro* testing crude extract was used (1 volume protein extract; 3 volumes 100 mM Tris/HCl pH7.4, 10 mM $MgCl_2$, 5 mM β-mercaptoethanol; 50 µM substrate ^{3}H-GGPP or ^{14}C-IPP, Biotrend/DMAPP, Sigma, total volume 500 µL). Reactions were incubated at 30 °C over night. After addition of 500 µL of brine reactions were extracted twice with the same volume of ethyl acetate. Extracts were concentrated in air-stream and analyzed via radio-TLC (silica plates, Merck, cyclohexane/ethyl acetate 9:1). Detection was done with Radio-TLC Scanner RITA Star (Raytest, Straubenhardt, Germany). Test reactions for 1762 intron1 variants were performed in carbon coated 96 well plates in order to adsorb the none polar products generated during reaction over night at 30 °C (HEINIG *et al.* 2010). Wells were emptied, dried and extracted with 500 µL of cyclohexane each for analysis. Extracts were concentrated in air-stream and analyzed via radio-TLC (silica plates, Merck, cyclohexane/ethyl acetate 9:1). Detection was done with Radio-TLC Scanner RITA Star (Raytest, Straubenhardt, Germany).

II.6.3 SDS polyacrylamid gel electrophoresis

Besides functional testing of proteins raw extracts were analyzed via sodium dodecyl sulfate polyacrylamid gel electrophoresis (SDS-PAGE). Therefore samples (12.5 µL) were first supplemented with 2.5 µL loading buffer (10x: 62.5 mM Tris/HCl pH 6.8, 30 % (v/v) glycerol, 10 % (v/v) β-mercaptoethanol, 4 % (w/v) SDS, 0.05 % (w/v) bromophenol blue) and denatured at 99 °C for 10 min. Separation was done using a discontinuous PAA gel consisting of a 4 % collection gel (625 µL of 1 M Tris/HCl pH6.8, 830 µL of 30 % (v/v) polyacrylamide, 50 µL of 10 % (w/v) SDS, 15 µL of 20 % (w/v) APS, 5 µL TEMED, 3625 µL ddH_2O) and a 10 % separation gel (3750 µL of 1 M Tris/HCl pH8.8, 3300 µL of 30 % (v/v) polyacrylamide, 100 µL of 10 % (w/v) SDS, 30 µL of 20 % (w/v) APS, 10 µL TEMED, 2785 µL ddH_2O). Electrophoresis was performed in a Mini-protean chamber (Biorad) for 45 min at 180 V in SDS – running buffer (125 mM Tris/HCl pH8.3, 960 mM glycerol, 0.5 % (w/v) SDS). As size standard 5 µL of a protein ladder (Fermentas) were used.
After separation of protein extracts the collection gel was cut off and the separation gel was either stained with Coomassie according to (WONG *et al.* 2000) or used for Western Blot analysis.

II.6.4 Western Blot analysis

Transfer of proteins to nitrocellulose membranes (0.45 µm, Millipore) was carried out in a Biorad "Tank Blot" apparatus (transfer buffer: 1.44 % (w/v) glycine, 0.3 % (w/v) Tris, 20 % (v/v) methanol) by applying an electric field vertical to the gel (120 V, 90 min).
The immunological detection procedure is summarized in the following table. It is based on the binding of an anti – HIS antibody to the proteins on the membrane that carry an artificial HIS-6 tag. A secondary antibody, which is labeled with alkaline phosphatase binds to the first antibody. By final incubation with alkaline phosphatase substrate NBT/BCIP (nitro-blue tetrazolium chloride/5-bromo-4-chloro-3'-indolyphosphate; Roth, 2:1 mixture with 20 % (v/v) formamide) proteins carrying HIS-tag appear as violet bands.
Buffers used were 10x PBS (1.37 M NaCl, 27 mM KCl, 81 mM Na_2HPO_4, 15 mM KH_2PO_4), PBS-T (1x PBS, 0.05 % (v/v) Tween20) and alkaline phosphatase buffer (AP: 100 mM Tris-HCl pH9.6, 100 mM NaCl, 100 mM $MgCl_2$)

Table II-10: Western blotting procedure, including all working steps, reagents and incubation times.

Step	Reagent	Time
blocking of membrane	5 % (w/v) milk powder in PBS-T	90 min
binding of first antibody	rabbit anti HIS antibody	60 min
washing	PBS-T	3x 15 min
binding of secondary antibody	goat anti-rabbit, AP labeled	60 min
washing	PBS-T	3x 15 min
equilibration in AP buffer	AP buffer	20 min
detection reaction	NBT/BCIP 1:100 in AP buffer	until signal development
stopping reaction	water	5 min

II.7 *Phytochemical methods*

II.7.1 Organic solvent extraction of taxanes

The media of endophytic fungi's cultures were extracted twice with equal volume of chloroform. The resulting organic solution was dried over magnesium sulfate and evaporated to dryness. The raw product was resolved in methanol.

Plant material, *Taxus* needles and Tobacco leafs, was lyophilized and afterwards extracted with dichloromethane/methanol 1:1 by soxhlet extraction. The resulting organic solution was evaporated to dryness, and resolved in dichloromethane. After two times extraction with water the organic layer was dried over magnesium sulfate, the solvent evaporated and the raw product dissolved in methanol.

II.7.2 Anti-Taxane Competitive Inhibition Enzyme Immuno Assay (CIEIA)

The immunoassay was carried out according to the protocol of manufacturer (Immunoassay for the quantitative detection of taxanes in biological matrices, Cardax Pharmaceuticals, Hawaii). The test represents a solid phase competitive immunoassay. Taxane concentration is measured by the amount of inhibition of the reaction of anti-taxane-antibody with the

immobilized Taxol protein conjugate (antigen). First 100 µL of a 1:100 dilution of taxol-protein coating antigen in PBS-buffer (50 mM Na_3PO_4, 0.15 M NaCl, pH 7.0) in the required number of wells of an 96 well microtiter plate. The covered plate was incubated at room temperature for one hour. After washing six times with TBS-T buffer (50 mM Tris, 0.15 M NaCl, 0.05 %(v/v) Tween-20, pH 7.0) wells were blocked by incubation for 1 hour at room temperature with 200 µL of 1 % BSA in PBS, followed by another 4 washing steps with TBS-T. Next the extracts or taxol standard solutions were applied in a total volume of 50 µL in BPT-M buffer (0.25 %(w/v) BSA, 0.05 %(v/v) Tween-20, 20 %(v/v) methanol in PBS). As background control at least 5 wells were incubated with BPT-M buffer. Detection was performed by first addition and 1 hour incubation of 1:100 dilution of anti-taxane antibody in BPT buffer (0.25 %(w/v) BSA, 0.05 %(v/v) Tween-20 in PBS), four times washing with TBS-T and secondly 1 hour incubation with alkaline-phosphatase labelled goat anti-rabbit IgG conjugate diluted in BPT. For colour development alkaline phosphatase substrate (1 mg/mL in alkaline phosphatase buffer, pH 9.5) was added. The covered plates were incubated for another hour and afterwards analysed with a dual wavelength ELISA reader (reference wavelength = 650 nm, sample wavelength = 405 nm).

For determination of standard curve taxol in concentrations of 111, 37, 12.33, 4.11, 1.37, 0.46, 1.15 ng/mL was used (Table II-11, Figure II-2). The samples were tested in different dilutions. Values lying in linear range of the standard curve were used for calculation of the concentration.

Table II-11: standard curve for anti-taxane immunoassay, background absorption was 0.109, B_0 is the absorbance without inhibition.

Sample	Absorbance [405 nm]	B/B_0	Corrected absorbance [405 nm]	Concentration [ng/mL]
Taxol 1	0.139	0.05	0.03	111
Taxol 2	0.136	0.04	0.028	37
Taxol 3	0.143	0.05	0.034	12.33
Taxol 4	0.158	0.08	0.049	4.11
Taxol 5	0.2	0.14	0.092	1.37
Taxol 6	0.276	0.26	0.167	0.46
Taxol 7	0.401	0.46	0.293	0.15
B_0	0.639	0.83	0.53	0

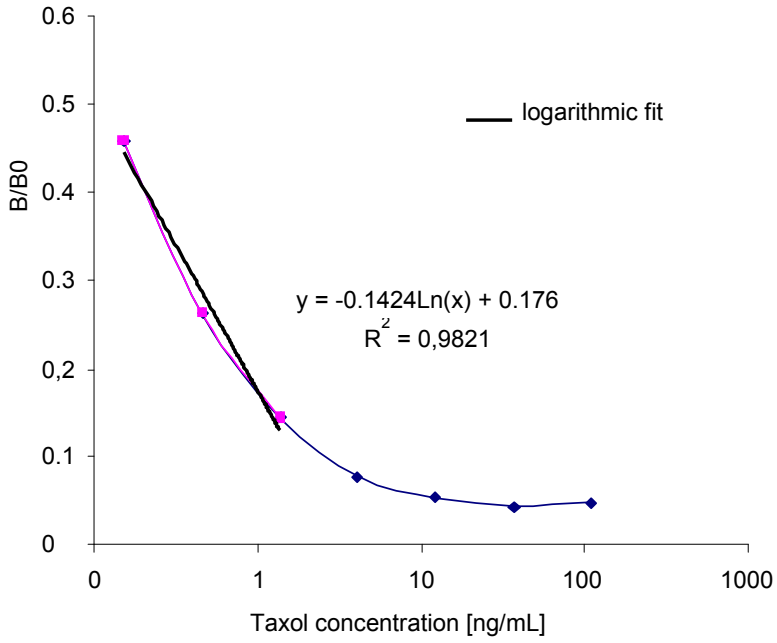


Figure II-2: Standard curve for determination of amount of taxane production via CIEIA

II.7.3 LC/MS/MS analysis of fungal and plant taxoid extracts

LC/MS/MS was performed using "multiple reaction monitoring" scan mode with the QTrap3200 system from Applied Biosystems. The three most intensive mass transitions for three standard substances Taxol, Baccatin III and 10-Deacetyl-Baccatin III (Sigma-Aldrich) were used for detection (Table II-12) in ESI negative ionisation mode (conditions: curtain gas 25 psi, CAD gas medium, ionspray voltage -4500 V, temperature 450°C, gas1 50 psi, gas2 65 psi). The MS method was combined with separation of the substances via HPLC (column: Curosil PFP 150x3 mm, 3 µm, Phenomenex, oven: 25 °C, flow rate: 300 µL/min, solvent A: 98 % water 2 % ACN + 10 mM NH_4OAc, solvent B: 2 % water 98 % ACN + 10 mM NH_4OAc, gradient: 0 min 70 % A, 0.5 min 70 % A, 15 min 0 % A, 20 min 0 % A, 21 min 70 % A, 23 min 70 % A).

Table II-12: MS parameters of MRM method for detection of Taxol, 10-Deacetyl baccatin III and Baccatin III; DP: declustering potential, EP: entrance potential, CE: collision energy, CXP: collision cell exit potential.

Compound	Mother ion [m/z]	Fragments [m/z]	DP [V]	EP [V]	CE [V]	CXP [V]
10-Deacetyl baccatin III	543.2	120.9	-55	-10	-30	-2
		391.2	-55	-10	-20	-4
		76.8	-55	-10	-76	0
Baccatin III	645.2	120.8	-35	-6.5	-28	0
		543.2	-35	-6.5	-24	-6
		76.9	-35	-6.5	-88	0
Taxol	852.3	525.2	-35	-9	-22	-6
		120.9	-35	-9	-44	-2
		319.2	-35	-9	-46	-4

II.7.4 Semi-preparative HPLC

Semi preparative HPLC was used for control of authenticity of detected 10-Deacetyl-baccatin III from EF0021 organic extract. Chromatography was performed with Shimadzu LC-20AD system using column Gemini C18 (250x10 mm, 5 µm, Phenomenex) and Diode array detector (254 nm) for detection and fraction collection (oven: 25 °C, flow rate: 4 mL/min, solvent A: 98 % water 2 % ACN + 10 mM formic acid, solvent B: 2 % water 98 % ACN + 10 mM formic acid, gradient: 0 min 90 % A, 3 min 70 % A, 50 min 50 % A, 55 min 0 % A, 60 min 0 % A, 63 min 90 % A, 65 min, fractions collected: 2.3 mL).

Retention time for 10-DAB was determined to 31.7 min (fractions 51, 52) analyzing pure substance (Sigma-Aldrich). In the analysis of EF0021 extract (injection volume: 100 µL) these fractions were collected. Volume was reduced to ~200 µL and samples were analyzed via LC/MS/MS (II.7.3).

III Results

III.1 *Analytical methods for the detection of taxanes*

As described in detail in the introduction the detection of taxanes from endophytes' organic extracts was done with several different methods up to now, including immunological detection, thin layer chromatography (TLC), liquid chromatography with different detection methods as UV or mass spectrometry, infrared or nuclear magnetic resonance spectroscopy (NMR) (ZHOU *et al.* 2010).

Of course all these methods have advantages as well as disadvantages and the choice of the detection system has to be well considered. The major challenge in the present case was the analysis of crude organic extracts consisting of a complex mixture of different organic molecules and besides the searched taxanes. Furthermore according to literature the molecules of interest were supposed to be present only in very low concentrations. Thus the detection method had to be highly sensitive and additionally also selective for taxanes to avoid elaborate purification steps. Several of the methods mentioned do not fulfill all of these requirements. Some of the methods seem to be not sensitive enough and require pure substances as for example NMR spectroscopy although by using it the most detailed information about the analyzed molecule could be obtained. Others, like TLC or simple UV detection lack the desired selectivity; the doubtless identification of certain compounds from crude extracts is difficult if not impossible at all.

On this background two fundamentally different methods were used in this study, one immunological assay that probably represents the most sensitive detection method and triple quadrupol LC/MS/MS for identification of taxanes from complex crude extracts according to structural features meaning in characteristic molecule fragmentation.

Pre-screening was done using the commercially available immunological system, based on competitive inhibition (CIEIA). An anti-taxane antibody is used, which is able to capture Taxol and taxoid compounds from the organic extract as well as to bind to taxanes coated on an immobilized antigen meaning there is a competition between these two possible binding partners.

For detection a secondary antibody labelled with alkaline phosphatase is used. Hence, if taxanes are present in the sample less antibody binds to the taxane coated antigen and the colour development catalyzed by alkaline phosphatase is decreased in comparison to samples

not containing taxanes. Concentration of taxoids is proportional to this reaction's inhibition (Cardax Pharmaceuticals, Hawaii). This method was for example already used for screening of taxane producing endophytes from *Taxus baccata* (CARUSO *et al.* 2000b). This competitive immuno assay is extraordinary sensitive, but does not deliver proof for taxoid structures. Due to this lack of structure related information cross reactivities to other compounds present in complex organic raw extracts are possible.

Therefore as a second detection method LC/MS/MS in "**m**ultiple **r**eaction **m**onitoring" scan mode (MRM) was used. By using MRM it is possible to select for a specific ion in Quadrupole 1 (Q1), for example the molecular ion of a compound. Quadrupole 2 (Q2) is used as collision cell. The ions isolated in Q1 are fragmented into molecule specific ions, prior to selection of these fragments in Quadrupole 3 (Q3) and detection (Figure III-1).

The major advantage of this analysis method is the high selectivity for the respective compound by detection of characteristic fragments obtained after selection for the molecule ion. Although the total sensitivity might be lower in comparison to scanning for the mother ion, due to loss of all ions additionally obtained through fragmentation, the signal to noise ratio increases because of the strict selection. Hence, this method is highly specific on structural features of the analyte molecules. In combination with the characteristic retention times on HPLC for the compounds this approach allows the doubtless detection and identification of the desired molecules from complex natural product extracts obtained from plants and fungi and is therefore the best available detection method for the present analytical problem.

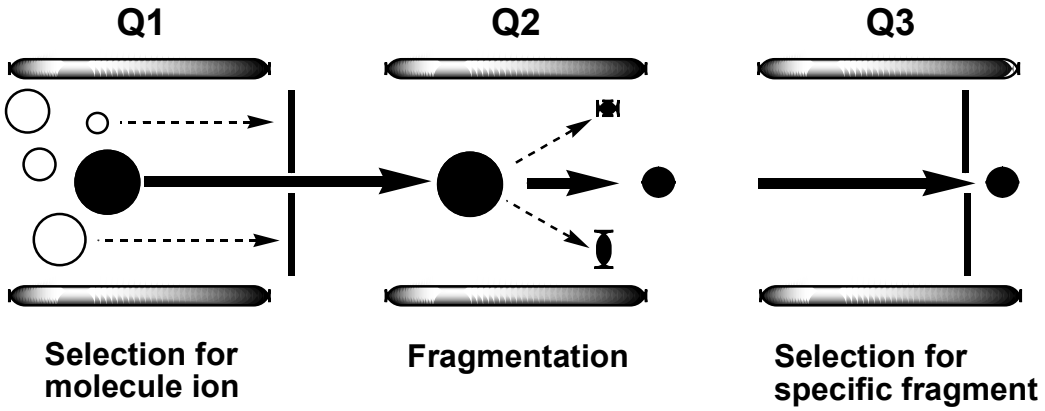


Figure III-1: Scheme of "multiple reaction monitoring" (MRM) triple quadrupol mass spectrometry scan mode. Q1, Q2, Q3: quadrupol 1, 2, 3.

For the analysis of endophytic fungi's organic extracts a method for LC/MS/MS analysis was created, using three commercially available reference substances, 10-Deacetylbaccatin III, Baccatin III and Taxol. For each compound the optimal mass spectrometry parameters were determined, including declustering potential, entrance potential, collision cell entrance and exit potential. Furthermore a collision energy ramp was used to find the three most intense mass transitions with corresponding collision energies for the molecules.

These resulting individual MRM methods were merged and a mixture of all three standards was injected multiple times for optimization of electron spray ionization (ESI) conditions, including voltage, temperature and gas parameters (II.7.3). For all three compounds the negative ionization mode was found to be favorable (ESI-).

For separation of the compounds via HPLC prior to mass detection a column specially developed for taxane analysis was used (II.7.3). The transitions and retention times of this method, used for all later analyses are summarized in Table III-1.

Table III-1: **MRMs and retention times characteristic for Baccatin III, 10-Deacetyl-baccatin III and Taxol.**

Compound	Mother ion [m/z]	Frag. 1 [m/z]	Frag. 2 [m/z]	Frag 3 [m/z]	Retention time [min]
Baccatin III	645.2 (BacIII+acetate)	543.2	120.8	76.9	7.02
10-Deacetylbaccatin III	543.2 (10-DAB-H)	391.2	120.8	76.8	4.72
Taxol	852.3 (Taxol-H)	525.2	319.2	120.9	9.90

A chromatogram of the reference compound mixture (1 µg/mL, injection volume 10 µL) is shown in Figure III-2. Furthermore the mass transitions of each compound are shown. By use of this method the detection limits were determined for Taxol, Baccatin III and 10-Deacetyl baccatin III to 35 fmol, 23 fmol and 28 fmol, respectively.

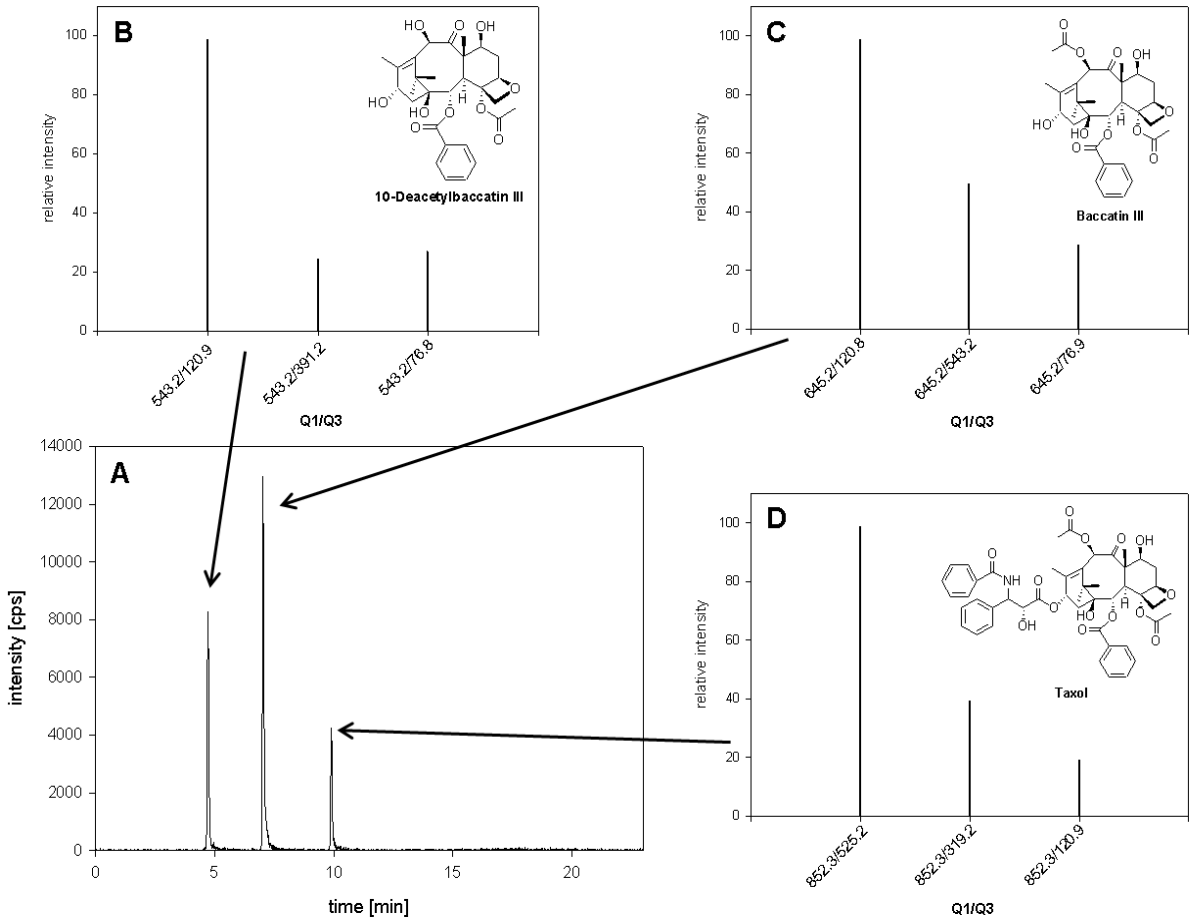


Figure III-2: **LC/MS/MS analysis of taxane references; A:** Chromatogram of LC/MS/MS analysis of a mixture of 10-Deacetylbaccatin III, Baccatin III and Taxol (1 µg/mL each, 10 µL injected); **B:** characteristic mass transitions of 10-Deacetylbaccatin III; **C:** characteristic mass transitions of Baccatin III; **D:** characteristic mass transitions of Taxol.

III.2 *Analysis of endophytes described as taxane producers*

The project was started with the phytochemical examination of three fungal species that were described as taxane producers in literature. These three strains obtained from culture collection were *Taxomyces andreanae* (CBS 279.92) (STROBEL *et al.* 1994), UPH-12 (NRRL 30405) (HOFFMAN 2003) and H10BA2 (NRRL 21209) (STIERLE *et al.* 2000). Thereby *Taxomyces andreanae* was the first fungus described to produce Taxol, as mentioned already. H10BA2 is another endophyte from *Taxus* also described by the group of Gary Strobel. UPH-12 was selected to include also a fungus that was not found on *Taxus* spp. It was isolated from hazelnut, an angiosperm plant not able to produce taxanes itself. The

organisms were transferred to agar plates first followed by cultivation in suspension cultures and organic solvent extraction as described in chapters II.4.2 and II.7.1

Neither by immunoassay nor by LC/MS/MS analysis of culture extracts any taxane could be detected (chapters II.7.2, II.7.3). These fungal species were described before as taxane producers, for example *Taxomyces andreanae* produced Taxol and Baccatin III (STIERLE *et al.* 1993). Cultivation and extraction conditions were chosen according to the analyses of these species. Hence it should have been possible to detect taxanes in these extracts.

Although the molecular biological working plan is independent from transcription of the genes and thus independent from taxanes production, detection of the products is an essential hint that the predicted pathway might be present. Since no other information is available this phytochemical examinations are the only possibility to at least assume a biosynthesis in fungi.

The reasons for these unexpected results are not absolutely clear. Two different scenarios might be feasible. The fungi analyzed were deposited to culture collections years ago. For strain conservation they of course were not cultivated in the complex screening media mimicking the environment within their natural habitat. These different conditions might have led to a silencing of the synthesis. So far there are not many reports about loss of the ability of secondary metabolite production because of handling under laboratory conditions. For some secondary metabolite pathways it was shown that some species of a genus have the biosynthesis whereas some other species lost genes or parts of a cluster and are not able to produce the compounds any more (BÖMKE *et al.* 2008; PROCTOR *et al.* 2009). But these are of course natural evolutionary processes. The frequency of occurrence of such mechanisms is not known. Furthermore it was shown, at least for the model organism *Aspergillus* that transposable elements play a role in secondary metabolism gene cluster regulation (SHAABAN *et al.* 2010). In another study serial transfers were used in order to create nonaflatoxinogenic *A. flavus* strains from previously aflatoxinogenic ones, in that case to study expression profile differences of genes involved in the aflatoxin biosynthesis (CHANG *et al.* 2007). The silencing of the aflatoxin production was aimed in this study. A possible Taxol biosynthesis in endophytic fungi might have been silenced accidently through serial transfer over years in combination with adaption of the fungal species to standard cultivation media which differ from their natural environment and is supposed to be a major factor influencing secondary metabolite production.

Nevertheless all these explanations are highly speculative. Because of this uncertainness of the proposed reasons for no detection of taxanes from the three culture collection strains an

alternative additional strategy had to be created. When looking into literature it seems quite easy to isolate endophytic fungi from plant tissue, in this case *Taxus* species. As mentioned, detection of the natural product is proposed to be limited to rather fresh cultures, due to silencing through serial transfers or through adaption to the laboratory environment. Thus isolation of new endophytes from *Taxus* might lead to the detection of taxanes which is thought to be at least an indication for the presence of a fungal Taxol biosynthetic pathway.

III.3 *Identification of taxane producing endophytic fungi*

III.3.1 Isolation & characterization of endophytic fungi from *Taxus* spp.

The possible reasons for a loss of the ability to produce certain secondary metabolites under laboratory conditions led to the approach to isolate new endophytic fungi from various *Taxus* species.

Several studies demonstrated already the isolation of endophytic fungi from *Taxus* spp. in general and furthermore of Taxol or Taxoid producing ones from different locations all over the world.

Therefore *Taxus* woody material was collected in different locations mainly in Germany but also around the world (Table III-2).

For the isolation of Taxol or Taxoid producing fungi we used the previously described method by Guo et al. (GUO *et al.* 2006). Similarly we transferred surface sterilized inner bark of yew trees onto PDA agar plates supplemented with streptomycin to avoid bacterial contamination and incubated the plates for up to four weeks, until fungal growth occurred (Figure III-3 A).

In order to finally obtain pure cultures hyphal tips were transferred to new plates. Cultures that obviously still contained more than one fungus were separated again on fresh plates resulting for example in EF0001 and EF0001B (Table III-2). The isolates showed very heterogenous morphological characteristics and growth behavior on plates (examples Figure III-3 B-D) as well as in liquid cultures. Growth over a complete plate took between 5 days to 4 weeks. In liquid culture the fungi grew either as suspensions or formed aggregates. Coloring of medium or mycelia as well as viscosity of the cultures, probably due to formation of polysaccharides could be shown to be species dependent.

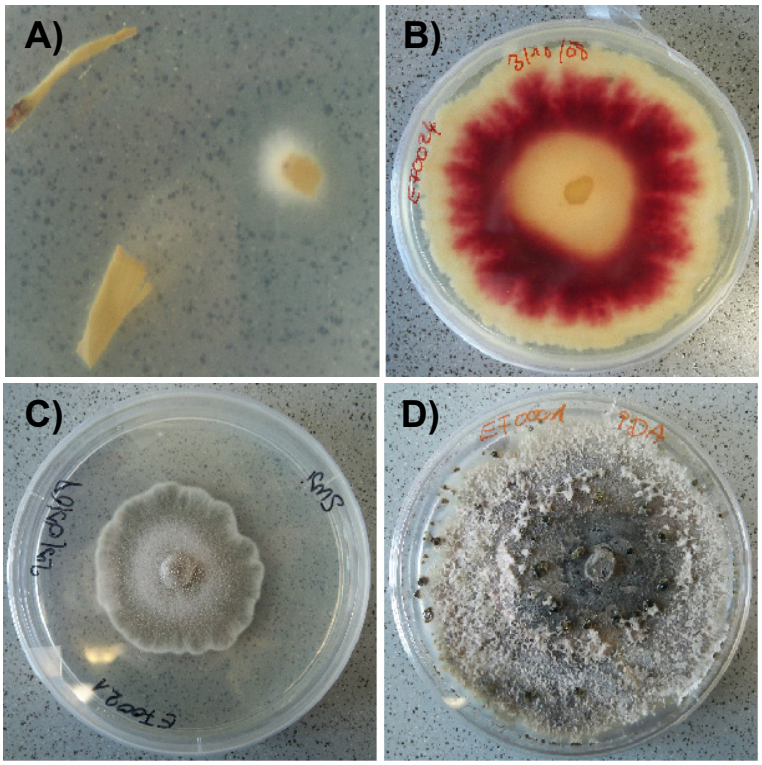


Figure III-3: Isolation of endophytic fungi from *Taxus* bark; A) appearance of fungal growth from inner bark of yew tree; **B)-D)** examples for different phenotype of endophyte isolates; **B)** EF0024; **C)** EF0021; **D)** EF0001.

Furthermore the phenotype was influenced by the medium composition (Figure III-4). This was reported before and was expected as far as it is known, that fungal natural product synthesis, e.g. phenolic compounds responsible for culture colour, is very much dependent on the culture conditions, e.g. carbon source, trace elements.

Figure III-4: Influence of culture medium on phenotype of newly isolated endophytic fungus EF0018 (left: M1D agar; right: PDA agar)

Using this procedure 34 different new endophytic fungi from *Taxus* species were isolated. For the identification of taxane producing endophytes we used the two different, complementary methods as described for the analysis of the fungi obtained from culture collections (III.1).
Prior to inoculation of liquid cultures for phytochemical analysis fungal strains were transferred onto plates with respective screening medium (M1D, S7, YM). Table III-2 summarizes the isolated endophytic fungi determined via the conserved internal transcribed spacer (ITS) regions by BLAST search as well as the *Taxus* host species.

Table III-2: Endophytic fungi isolated from *Taxus* species; including isolate name, original host species and proposed fungal species according to ITS BLAST search with coverage/identity and accession number of best ITS BLAST result.

Isolate	*Taxus* species	Best BLAST result, species (from ITS BLAST)	coverage, identity	Acc. No.
EF0001	*Taxus baccata*	*Phomopsis sp.* NY7255c	100%, 99%	HM999947.1
EF0001B	*Taxus baccata*	*Diaporthe perniciosa* strain ATCC 38578	100%, 100%	HQ908492.1
EF0002	*Taxus baccata*	Phomopsis sp. I414b	100%, 100%	GU584957.1
EF0002A	*Taxus baccata*	*Phomopsis sp.* I414b	100%, 100%	GU584957.1
EF0003	*Taxus media*	Fungal endophyte sp. ECD-2008 isolate 70 *Xylaria sp.* MT0810	100%, 100% 99%, 99%	EU686007.1 HQ414612.1
EF0004	*Taxus baccata*	Fungal endophyte sp. ECD-2008 isolate 178 *Hypoxylon serpens*	100%, 100% 100%, 99%	EU686082.1 HM036598.1
EF0005	*Taxus baccata*	*Diaporthe perniciosa* strain ATCC 38578	100%, 100%	HQ908492.1
EF0006	*Taxus media*	*Hypoxylon cohaerens* var.microsporum	100%, 99%	AJ390399.1
EF0007	*Taxus baccata*	*Coniothyrium fuckelii*	100%, 100%	FR667993.1
EF0008	*Taxus baccata*	*Microsphaeropsis olivacea* strain CBS 442.83	100%, 99%	GU237865.1
EF0009	*Taxus baccata*	*Discostroma fuscellum*	100%, 100%	GU244511.1
EF0011	*Taxus baccata*	*Beauveria bassiana* isolate s044	100%, 100%	HQ649861.1
EF0012	*Taxus sp.*	nd	nd	nd

Table III-2 continued

EF0013	*Taxus baccata*	Ascomycota sp. PIMO_429	100%, 99%	JF705947.1
		Fungal endophyte sp. AP458	100%, 99%	FM200718.1
		Pleosporales sp. agrAR078	100%, 99%	FN435680.1
EF0014	*Taxus baccata*	*Fusarium tricinctum* culture-collection WAC:12337	100%, 100%	JF776665.1
EF0014A	*Taxus baccata*	Fungal endophyte sp. ECD-2008 isolate 185	99%, 100%	EU686089.1
		Biscogniauxia nummularia isolate BI21	99%, 99%	EF155488.1
EF0015	*Taxus baccata*	*Coniothyrium fuckelii*	100%, 100%	FR667993.1
EF0016	*Taxus sp.*	*Leotiomycetes sp.* 4922	100%, 100%	FR668003.1
EF0017	*Taxus baccata*	*Paraconiothyrium sporulosum* strain M43	100%, 100%	JF340257.1
EF0018	*Taxus baccata*	Phomopsis sp. I414b	100%, 100%	GU584957.1
EF0019	*Taxus baccata*	*Phomopsis sp.* I414b	100%, 100%	GU584957.1
EF0020	*Taxus baccata*	*Phomopsis sp.* I414b	100%, 100%	GU584957.1
EF0021	*Taxus baccata*	*Phialocephala fortinii* strain 92-109	95%, 93%	AY078135.1
EF0022	*Taxus baccata*	*Fusarium lateritium*	100%, 100%	AJ269850.1
EF0023	*Taxus baccata*	*Fusarium lateritium*	100%, 100%	AJ269850.1
EF0024	*Taxus baccata*	*Fusarium tricinctum* strain gf10	100%, 100%	HQ262502.1
EF0025	*Taxus baccata*	nd	nd	nd
EF0026	*Taxus baccata*	*Phomopsis sp.* I414b	100%, 100%	GU584957.1
EF0027	*Taxus baccata*	nd	nd	nd
EF0028	*Taxus baccata*	nd	nd	nd

Table III-2 continued

EF0029	*Taxus cephallus*	nd	nd	nd
EF0030	*Taxus cephallus*	nd	nd	nd
EF0031	*Taxus sp.*	nd	nd	nd
EF0032	*Taxus sp.*	nd	nd	nd

nd: not determined

The internal transcribed spacer (ITS) region is a non functional RNA sequence located in the structural ribosomal RNA. Sequence analysis and comparison of this genetic element is widely used in taxonomy and molecular phylogeny (GARDES and BRUNS 1993; GANLEY *et al.* 2004; SMITH *et al.* 2007). Figure III-5 shows the ITS amplification products used for taxonomic determination of the newly isolated endophytic fungi.

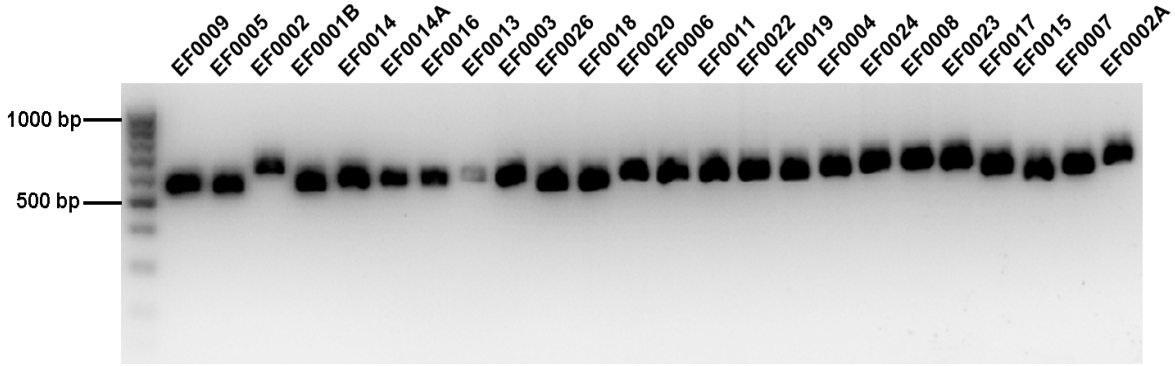


Figure III-5: Electrophoretic separation of PCR products of fungal ITS regions; Lane 1: 100 bp marker (NEB), **Lane 2-24:** amplification products.

The DNA samples were purified and sequenced using PCR amplification primers ITS1 and ITS4. The resulting sequences were assembled and subjected to BLAST search (Supplementary information, CD). This analysis yielded in high homology matches (Table III-2) and respective species in the GenBank database of NCBI. The major species observed belonged to the genera *Phomopsis*, *Fusarium* and *Diaporthe* already described as endophytes on *Taxus* and other plants.

According to the aim of the project being the identification of Taxol producers and fungal Taxol biosynthetic pathway, further characterization on morphological level was not performed.

III.3.2 Immunological assay for taxane detection from fungal extracts

Immunoassay was used as first screening method for the identification of Taxol/Taxane containing endophytes. For determining the concentration of total taxanes in the medium-extracts a standard curve was set up (II.7.2). As a positive control an organic extract of *Taxus* needles was used.

The antibody assay yielded in two organic extracts of endophytic fungi EF0001 and EF0016 in which taxanes could be detected. The concentrations were calculated to 7.8 ng/L and 2.5 ng/L culture medium taxoids, respectively. These amounts of taxoid compounds, determined from the Taxol standard curve (II.7.2) in comparison to the values obtained for the positive control *Taxus* were thereby 50,000 fold lower (Table III-3). This was not surprising in the first place, taking into account that secondary metabolite production is very much dependent on the culture medium, culture conditions and the time of cultivation. Furthermore taxanes from *Taxus* are higher in concentration due to drying of the needles prior to extraction. The values are given in ng/g dry weight.

In case of the screening performed in this study, all fungi were cultivated in the same screening media which were previously described in the literature for Taxol endophyte cultivation for 2 weeks followed by harvesting the mycelia for molecular biological work and extraction of the medium with chloroform. YM-6.3 was not used for cultivation of taxane producing endophytes so far but is a standard full medium for the cultivation of fungi. It was for example used for fermentation of basidiomycetes for the production of anti bacterial sesquiterpenoid natural products, the melleolides (ENGELS *et al.* 2011). Nevertheless these low values together with the observation, that no extract of the three fungi from culture collection previously analyzed showed taxane production left doubts about the reliability of this data and hence the specifity of the assay.

Table III-3: Results of immunoassay using polyclonal anti-taxane-antibody kit (Cardax Pharmaceuticals).

Organism	c [ng/mL], according to standard curve	c [ng/L], per liter culture medium	c [ng/g], per gram biomaterial
Taxus baccata	10401.7	/	$173.3*10^3$
EF0001	3.1	7.8	/
EF0016	1.5	2.5	/
N. tabacum	52.8	/	17.6

The Caradax kit used for the detection of Taxol/Taxanes a polyclonal antibody, therefore it was necessary to include a negative control for examining the specificity of the assay. Therefore an organic extract from *Nicotiana tabacum cv. Petit Havana* SR1 (green house, IME) was used as a negative control. Surprisingly this extract also gave a response in the assay, which was with 17.6 ng/g dried leaf material up to 7 times higher than the signals from the endophytes' extracts. This led to the conclusion that this method was not reliable for determination of taxane production in very low concentration ranges, as observed for the endophytes, whereas the assay works specifically for yew tree, due to the over 10,000 fold higher signal observed here.

As a consequence of the unreliable results, the immunological assay was not used further, because the results are not reliable. This observation was thereby the first time that a background activity was described for the assay; however previous studies did not include neg. controls. The only described negative control was a pure medium control, where of course no natural products are present that might lead to false signals. Analysis of the remaining fungi EF0020-EF0032 was therefore performed via LC/MS/MS.

III.3.3 LC/MS/MS analysis of fungal extracts

All fungal extracts were subjected to LC/MS/MS analysis. Thereby in two fungal extracts out of the total 34 endophytes' samples taxanes were found. In extract of EF0001 Baccatin III was detectable as shown in Figure III-6. The clear identification of the compound was possible due to the definite mass transitions compared to the pure substance at the respective retention time (Figure III-6).

The second fungal sample in which a taxoid compound could be detected was the extract of the fungus named EF0021. In this extract 10-Deacetylbaccatin III was detected (Figure III-7). Again the mass transitions and the retention time were identical to the standard.
In contrast to immunological assay in EF0016 extract previously showed a response indicating the presence of taxoids no taxanes were detected in LC/MS/MS analysis. Thus the signal obtained before was either due to unspecific binding of additional compounds from the extract or by other taxanes that are not detected by the highly selective MRM method, which is as described in detail before limited to three taxoid natural products. This result again showed the probable unreliability of the immunological assay alone. It is absolutely necessary to control the results via an assay based on molecule structure and physical properties like LC/MS/MS.
In all other fungal extracts none of the three reference substances could be detected. To examine the stability of taxane production in the newly isolated endophytes the whole procedure, cultivation, extraction and analysis was repeated for EF0021. Thereby it could be shown that the signal intensity decreased strongly. 10-Deacetybaccatin III was detected but only in trace amounts. This indicates an influence of the treatment of the fungi under laboratory conditions in contrast to the natural environment in relation with the plant host for the natural product formation. This result further supports the hypothesis of a silencing of secondary metabolite production by multiple transfers under laboratory conditions. According the aim and strategy of the project this phenomenon does not influence the further molecular biological experiments. The screening for the Taxol biosynthetic pathway on genetic level is not dependent on the existence of transcript or proteins of the biosynthesis and hence also not dependent on the natural product titer.

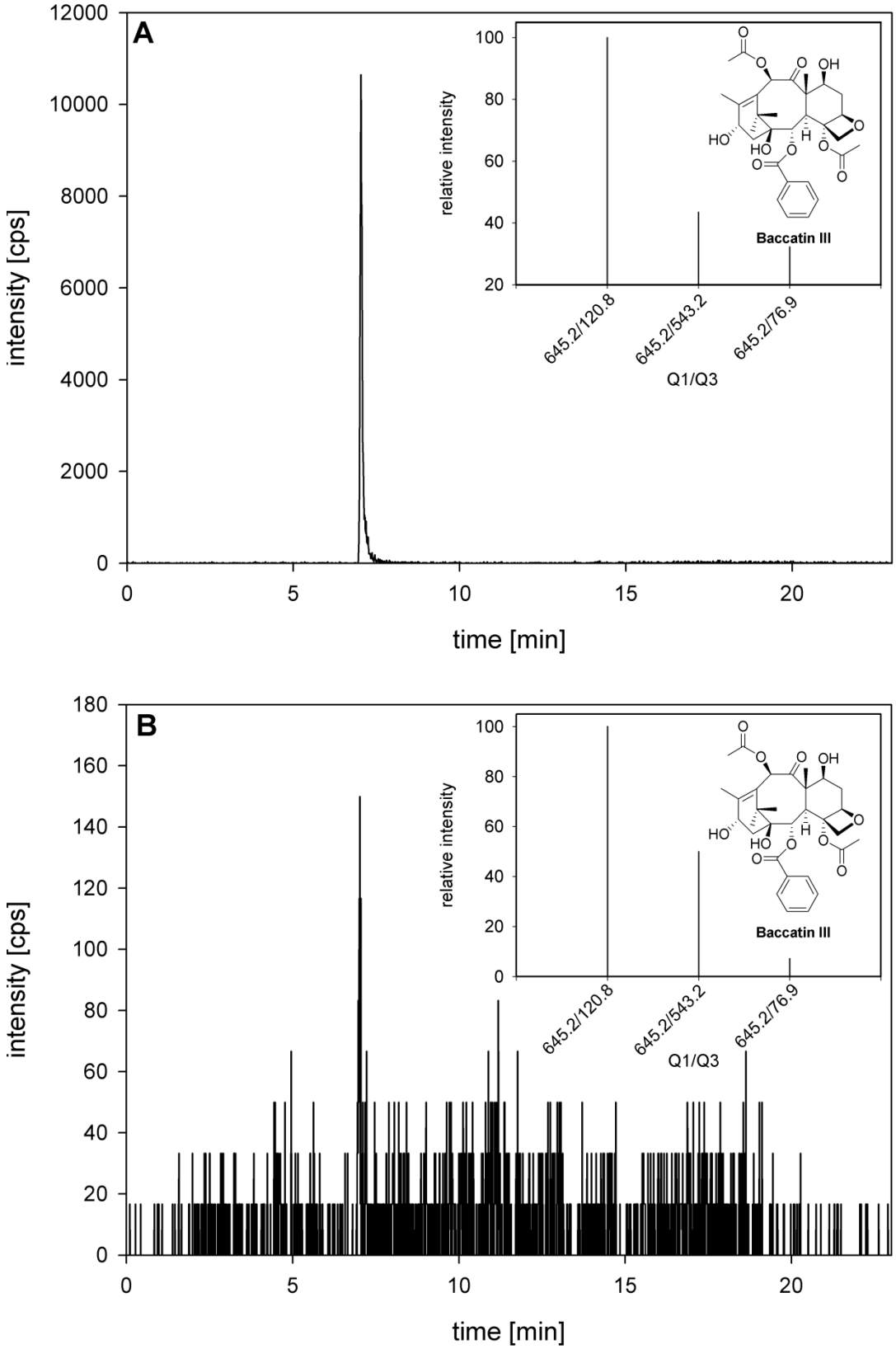


Figure III-6: LC/MS/MS analysis of EF0001 organic extract; A: chromatogram of Baccatin III pure substance and mass transition spectrum; **B:** chromatogram of EF0001 analysis and mass transition spectrum for signal at retention time 7.02 min.

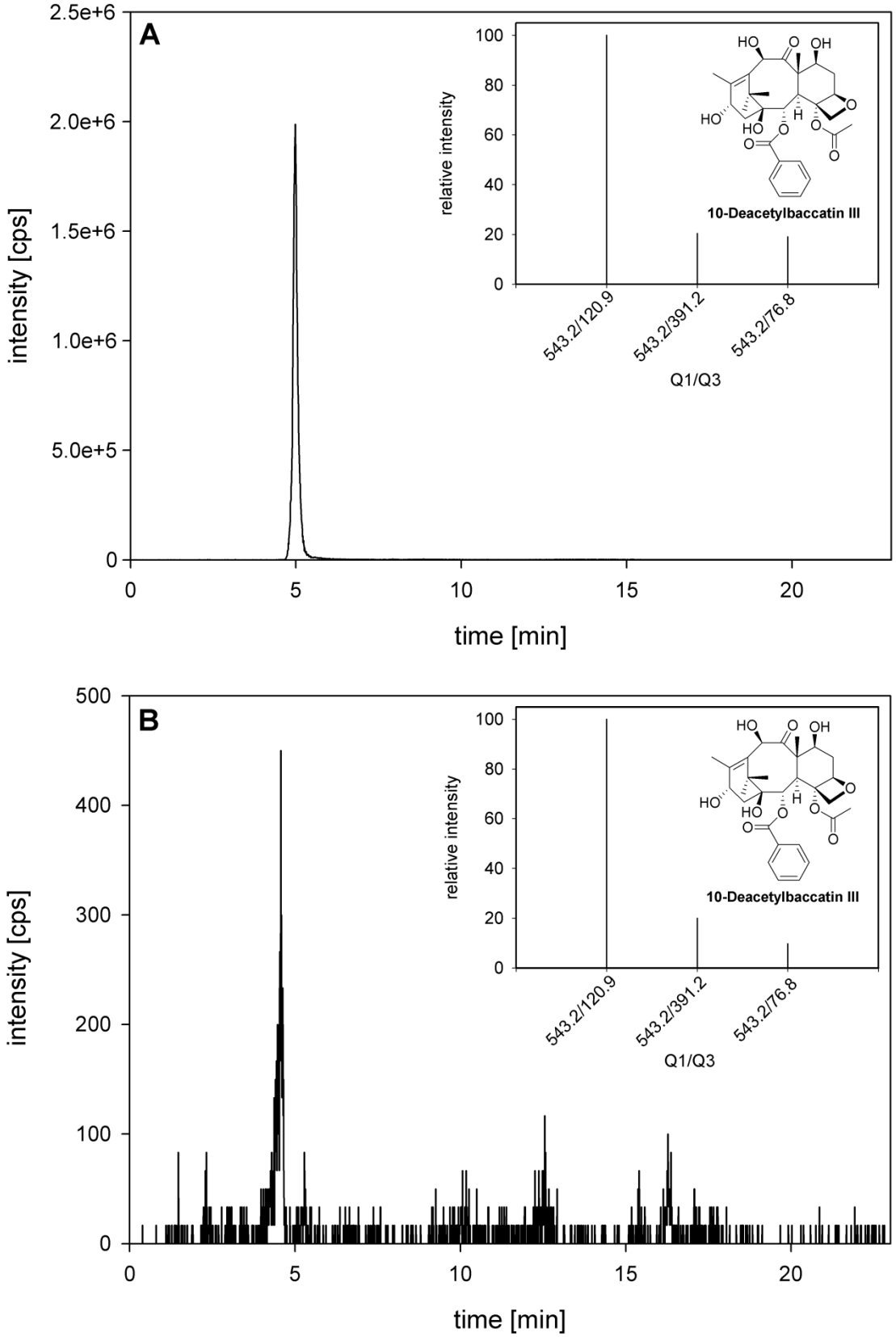


Figure III-7: LC/MS/MS analysis of EF0021 organic extract; A: chromatogram of 10-Deacetylbaccatin III pure substance and mass transition spectrum; **B:** chromatogram of EF0021 analysis and mass transition spectrum for signal at retention time 4.72 min.

A further control experiment was performed with EF0021 extract. The aim was to ensure the identity of the compound 10-Deacetybaccatin III detected from raw extract. Therefore first an endophytic extract containing no taxanes was spiked with pure 10-Deacetylbaccatin III. This sample was subjected to preparative HPLC using a semi-preparative C18 column with fraction collection (see chapter II.7.4). The fractions obtained were analyzed afterwards with the taxane LC/MS/MS method and the compound was recovered in fractions 51 and 52. If the original signal in EF0021 organic extract was not 10-Deacetylbaccatin III, but coming from another substances giving the correct mass transitions and retention time accidently the usage of this different chromatographic conditions would lead to a loss of the signal in theory.

EF0021 extract was treated as described and fractions 51 and 52 were tested. In the samples 10-Deacetylbaccatin III was detected. Hence, the signal detected from EF0021 raw extract was 10-Deacetybaccatin III without any doubts.

EF0001 and EF0021 were used for the further molecular biological examination, including Southern Blot analysis and genomic library screening.

III.4 *Molecular biological examination of endophytic fungi*

III.4.1 Southern Blotting as molecular biological "screening method"

Besides the phytochemical analysis of organic fungal extracts genomic DNA of all the species was extracted in order to perform Southern blotting analysis. According to the hypothesis that the genes of both plant and fungal Taxol biosynthesis are similar, hybridization with probes designed on the sequences of plant genes should lead to signals from *Taxus* DNA as well as from endophytic fungi's DNA.

Initial experiments were performed with *Taxus* DNA to optimize the blotting and hybridization conditions. Therefore *Taxus* DNA was digested with two different restriction endonucleases, *Eco*RV and *Hind* III. After separation via agarose gel (Figure III-8A) DNA was blotted onto nylon membrane according to protocol for neutral transfer described by Sambrook and Russel (SAMBROOK and RUSSELL 2001). For taxadiene synthase probe preparation the gene (*tds*) from *Taxus chinensis* was amplified from pGem-T easy-tds plasmid (Figure III-8B, Stefan Jennewein). The purified DNA fragment was used as template for amplification of three 500 bp fragments that were labeled with ^{32}P and used for

hybridization experiments. The resulting X-ray film is shown in Figure III-8A. In both lanes 3 clear signals were detected at different sizes. Hence, blotting and hybridization conditions were regarded as useful for Southern blotting analysis of endophytic fungi to detect similar genes in fungi's genomes on the one hand and getting indication on the degree of homology by comparison of signal intensity on the other hand.

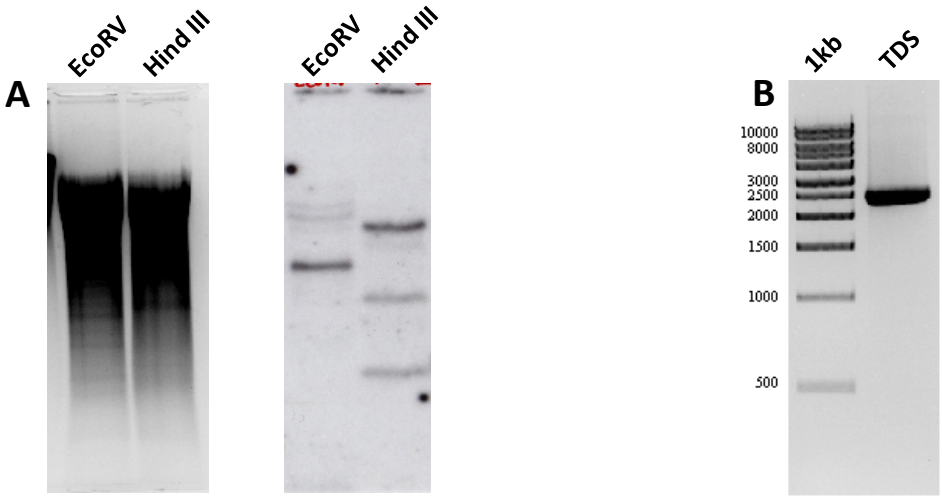


Figure III-8: Southern Blotting analysis of *Taxus baccata* genomic DNA; A) agarose gel separation of complete digest lane 1 *Eco*RV, lane 2 *Hind* III of *Taxus bacatta* genomic DNA; X-Ray film of Southern blotting membrane, after hybridization with ^{32}P-radiolabeled taxadiene synthase probe (55°C, 16h) and exposure time of 21d at -80°C with intensifier screen; **B)** agarose gel analysis of amplification of taxadiene synthase (cDNA) from pGem-T easy-tds plasmid used as a probe template for labelling.

For "screening" of endophytes, genomic DNA samples were digested with *Hind* III and hybridization was performed with probes for taxadiene synthase (TDS) and taxane-13α-hydroxylase (T13H) or taxane-5α-hydroxylase (T5H) under the conditions determined previously for *Taxus* gDNA hybridization. As negative control genomic DNA of *Nicotiana tabacum cv. Petit Havana* SR1 (green house, IME) was used.

In addition to the newly isolated fungi the three endophytes received from culture collections were also analyzed.

In theory only one or two signals should be visible in case of taxadiene synthase, assuming that the genomes did not contain multiple copies. In case of experiments using hydroxylase probes multiple signals were probable, due to the similarity of P450 hydroxylases known from *Taxus* in combination with the medium stringency of the hybridization conditions.

The results for EF0001 and EF0021 are shown in Figure III-9, where also the results for *Taxus* and the negative control are presented.

For both fungal DNA samples EF0001 and EF0021 one weak but nevertheless clear band was visible for hybridization with taxadiene synthase probe. In comparison to the positive control the signals had a very low intensity indicating only low sequence homology between plant and fungal terpene synthase. In the negative control no specific signal was observed.

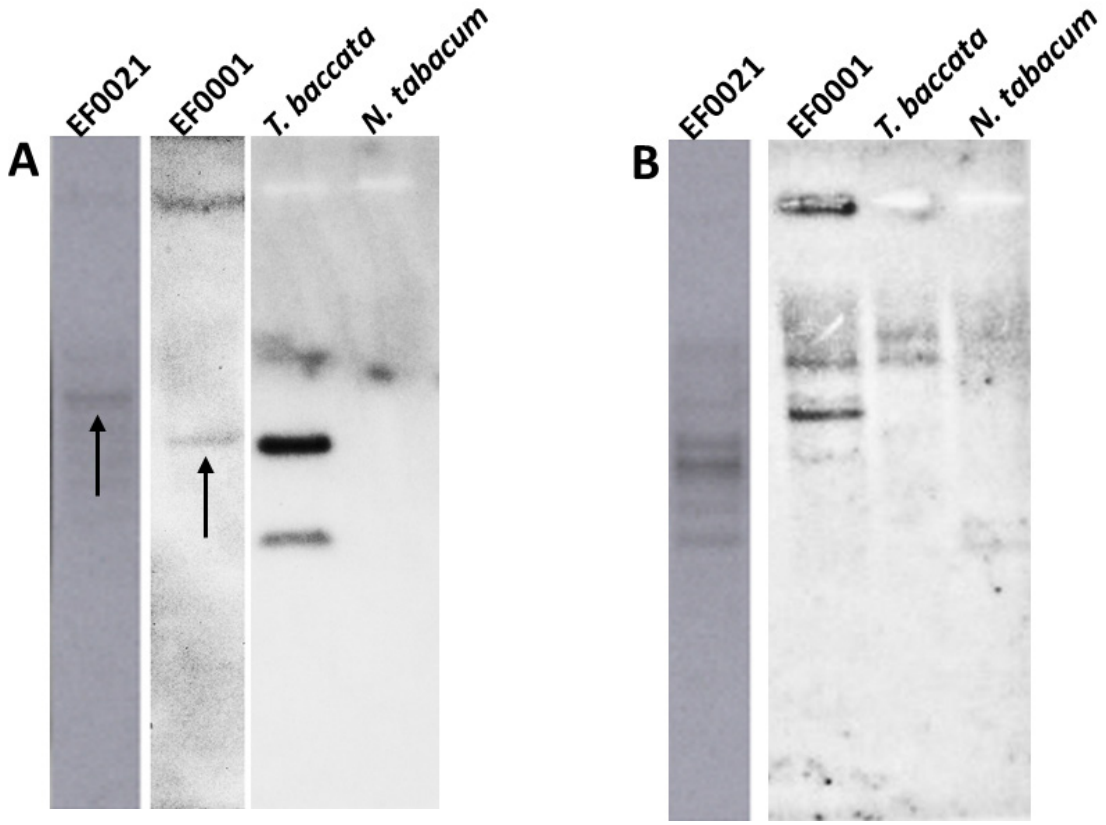


Figure III-9: Southern blotting experiments with EF0001, EF0021, *Taxus baccata* and *N. tabacum* as a negative control with different probes; DNA samples digested with *Hind* III; A) X-Ray films of Southern blotting membranes, after hybridization with ^{32}P-radiolabeled taxadiene synthase probe (55°C, 16h) and exposure time of 4d at -80°C with intensifier screen; lane 1: EF0021 DNA, lane 2: EF0001 DNA, lane 3: *T. baccata* DNA, lane 4: *N. tabacum* DNA, arrows mark signals of fungal taxadiene synthase. **B)** X-Ray films of Southern blotting membranes, after hybridization with ^{32}P-radiolabeled taxane P450 hydroxylase gene probes (55°C, 16h) and exposure time of 4d at -80°C with intensifier screen; lane 1: EF0021 DNA, hybridization with T5H, lane 2: EF0001 DNA, lane 3: *T. baccata* DNA, lane 4: *N. tabacum* DNA, lane 2-4 hybridization with T13H.

The hybridization experiments with P450 hydroxylase probes the results were according to the initial expectations. As in the positive control *Taxus baccata* the probes bound multiple times in both fungal genomes. This observation gave additional hint for Taxol biosynthetic pathway in endophytic fungi. But the more reliable indication was the result for taxadiene synthase gene since P450 genes genes are more similar than terpene synthase genes in general.

All other fungal species, newly isolated and obtained from culture collection, were analyzed the same way. For none of these fungi a signal was observed in Southern blot with taxadiene synthase or P450 hydroxylase probes.
In summary for two newly isolated endophytic fungi there are indications for a possible fungal Taxol biosynthetic pathway. Taxanes could be detected in fungis organic extracts and furthermore Southern Blotting results gave hints for genes at least somehow similar to genes of Taxol biosynthesis from *Taxus* species.
Hence, these two endophytes were chosen for genomic phage library hybridization screening.

III.4.2 Genomic phage library construction and screening

As initially discussed the yields of taxanes that were detected up to now were found to be in ng or µg amount. Hence it can be assumed that a possible biosynthetic pathway is only poorly expressed. Since there is also no method available to induce transcription/production a similar approach to the isolation of *Taxus* Taxol biosynthesis using an EST library constructed from a with methyljasmonate induced *Taxus* cell culture seemed to be not promising
For this reason the search for Taxol biosynthesis in fungi has to be done on genomic level. Since no definite gene sequences are known so called “Genome Walking” based on amplification using primers of known or conserved gene parts is not possible. The method of choice was genomic phage library construction and screening.
Thereby the whole genome is cut into pieces and the resulting DNA fragments are transferred into λ-phage. By plating the phages followed by immobilization of their DNA on membranes the library can be screened via hybridization with radiolabelled probes.
The success of the approach is dependent on two major factors, the genome size and the amount of genetic information in every phage. These variables determine the size of a representative library as well as the probability of finding a single copy gene.
The genome sizes of the endophytes are unknown. Taking already sequenced fungi as a basis, genome sizes are probably between 20 and 70 Mb.
In this thesis lambda replacement vector λ Dash II (Stratagene) was used. This type of vectors is created by deleting the part for the lytic cycle from the phage WT genome via restriction. This so called “stuffer fragment” can be replaced by the desired DNA fragments. Due to the small size of the vectors without the “stuffer fragment” it is impossible to pack this DNA into phage. Only with DNA inserts lager than 15 kb inserts packaging becomes possible again.

Hence, these types of phage systems are selective for large genomic DNA inserts. With the Stratagene system used here a 20 kb insert containing vector is 95 % more effectively implemented than a vector carrying a 14 kb DNA fragment. Furthermore λ Dash II takes advantage spi ("sensitive to P2 interference") selection. Lambda phages containing *red* and *gam* genes are unable to grow on *E. coli* strain having a P2 phage lysogen. By replacement of the "stuffer fragment" by the inserts phages become gam^- and are in contrast to WT lambda phage able to form plaques on a P2 containing *E. coli* host strain.
Besides these selection mechanisms avoiding WT vectors to be packed as well as all vectors containing only small inserts, the number of recombinant clones is determining the quality of a constructed library. A library is representative when it is assured that it contains all genetic information of the organism to be analyzed. The number of clones that have to be screened for have a high probability to find a single clone in a library can be calculated by the following formulas.

$$P = 1 - (1 - F/G)^N ; N = \ln (1 - P) / \ln (1 - F/G)$$

P: probability for the presence of a gene in a genomic library with N clones
F: average size of cloned DNA inserts [kb]
G: genome size of the organism [kb]
N: number of independent clones in the genomic library

For example the genomes of *Saccharomyces cerevisiae* and *Arabidopsis thaliana* have a size of 12.8 Mb (NC 001133-48) and around 100 Mb (NC 003070, NC 003071, NC 3074-76), respectively. Assuming an average insert size of 20 kb/clone (F) it would be necessary to screen at least 2950 clones for *Saccharomyces cerevisiae* and over 23000 clones for *Arabidopsis thaliana* to find a single clone with 99 % probability (P). Thus, to find a gene in a genome with more than 99 % probability, it is necessary to screen a number of clones representing around five genome equivalents. Hence, for an estimated average genome size of 50 Mb for fungi a useful library has to contain at least 11500 individual clones.

III.4.2.1 Construction and screening of EF0001 & EF0021 genomic libraries

Three genomic phage libraries from EF0001 DNA were constructed for homology based hybridization screening (Lib0001-080108, Lib0001-150108 and Lib0001-030708). Mycelium for genomic DNA isolation was obtained from initial endophyte's liquid cultures, where medium was extracted for immunoassay and LC/MS/MS analysis. After optimization of incomplete digest with *Bam*HI and precipitation of the resulting DNA fragments, EF0001 DNA was ligated into λ Dash II phage replacement vector and packed into Gigapack XL III phage packaging extracts. The titers of the resulting libraries were $2.5*10^5$, $2.2*10^5$ and $1.8*10^5$ pfu/mL, respectively. Although fungal genome size can be estimated to around 50 Mb the real size of EF0001 genome was unknown. To be sure that the number of clones screened was sufficient both libraries were plated completely. An example plate obtained after plaque formation is shown in Figure III-10A.

The plaques of 20 (Lib0001-080108), 19 (Lib0001-150108) and 20 (Lib0001-030708) plates, respectively were lifted onto membranes, to screen the libraries with on the one hand taxane-13α-hydroxylase (T13H) probe and taxadiene synthase (TDS) probe and on the other hand with taxadiene-5α-hydroxylase (T5H) probe.

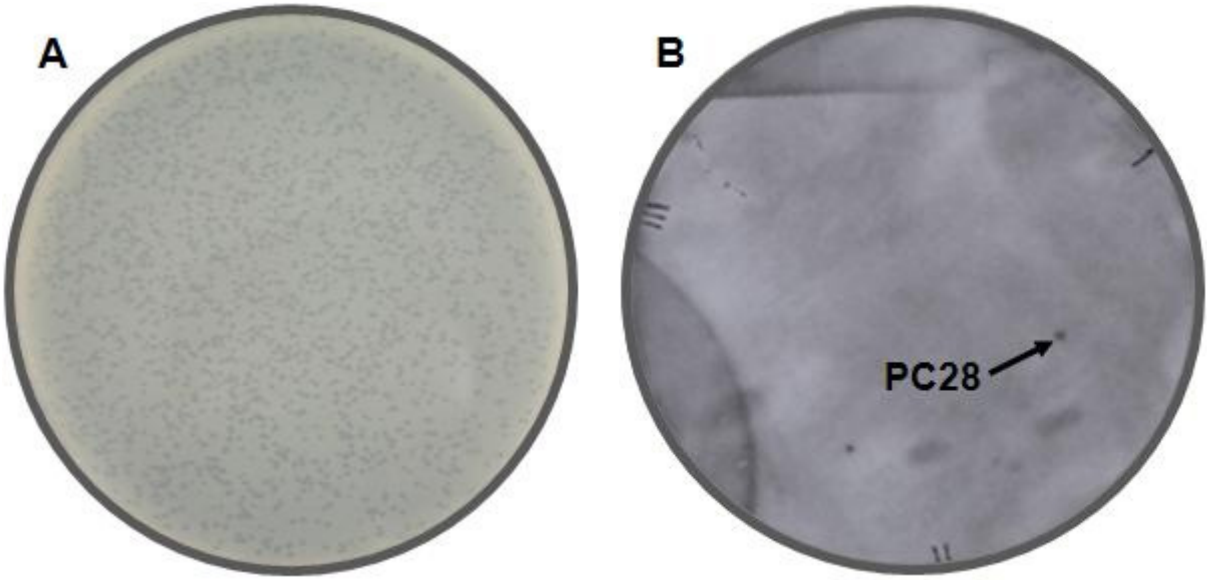


Figure III-10: Plate and X-ray film of membrane of first screening round; A: Example for a 145 mm agar plate of first screening round, plaques appear as clear spots in bacterial surface; **B:** X-ray film of membrane, hybridization with ^{32}P-radiolabeled taxadiene-5α-hydroxylase (55°C, 16h) and exposure time of 4d at -80°C on preflashed film with intensifier screen, marked dot is signal from positive clone (in this case PC28).

In case of EF0021 one library with a titer of $5*10^5$ pfu/mL was used. Construction was done as described for EF0001. For screening with taxadiene synthase (TDS) probe and oligo nucleotide probes homologous to taxadiene-5α-hydroxylase 80,000 plaques (20 plates) were lifted onto membranes (40,000 clones per probe).

According to hybridization procedure developed and validated via Southern Blotting all membranes were screened. For EF0001 two different probes TDS and T13H were used as a mixture for initial screening. T5H and in case of EF0021 screening TDS and T5H – oligo nucleotides were used as pure probes. For verification of results and isolation of pure individual plaques all identified candidates from this "first screening round" (example see Figure III-10B) were eluted and plated again less dense than in the first step. An enrichment of signals through repetition of hybridization ("second screening round") was the criteria for positive clones (Figure III-11). Table III-4 summarizes the obtained results. Screening of genomic phage libraries of EF0001 with radio labeled probes homologue to genes of Taxol biosynthesis from yew tree resulted in three positive clones named PC1, PC2 and PC4 for taxane-13α-hydroxylase, two positive clones PC9 and PC10 for taxadiene synthase and five clones PC23, PC24, PC25, PC26 and PC28 for taxadiene-5α-hydroxylase by screening around 300.000 genomic phage clones.

With a minimum average insert size of 20,000 kb this resulted in screening of at least 6,000 Mb what corresponds to a 120 fold genome coverage assuming an average genome size of 50 Mb for fungal genomes.

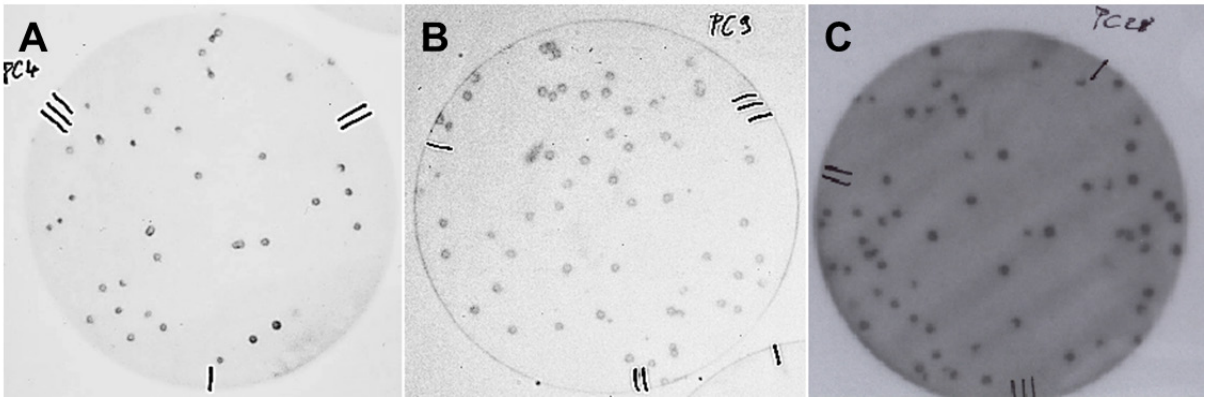


Figure III-11: X-ray films of colony blot membranes of second screening round; hybridization with ^{32}P-radiolabeled probes (55°C, 16h) and exposure time of 4d at -80°C on preflashed film with intensifier screen; Plaque solutions for infection from clones positive in first screening round with a mixture of taxoid-13α-hydroxylase and taxadiene synthase probe of Lib0001-150108 (A, B) and taxadiene-5α-hydroxylase probe from Lib0001-030708 (C); **A:** PC4, positive for taxoid-13α-hydroxylase probe; **B:** PC9, positive for taxadiene synthase probe; **C:** PC28, positive for taxadiene-5α-hydroxylase probe.

In case of EF0021 phage library screening of 40.000 clones (920 Mb; 18 fold coverage) resulted in 5 potential positive phage clones for taxadiene-5α-hydroxylase (oligo nucleotide) probe, whereas no clone for taxadiene synthase probe could be identified from hybridization screening.

From all described phage clones liquid lysates were prepared for isolation of DNA. The insert sizes could be determined by restriction digest using *Bam*HI, which cuts out the insert that initially was ligated into the unique *Bam*HI site. Because of the big inserts (in theorie >20 kb) created by none complete digest prior to library construction all DNA fragment sizes, except the vector band in summary could be used for estimation of the insert size of the clones. It was approximately between 16 kb and >35 kb.

The DNA was used as sequencing template either by Primer walking starting with the T3 and T7 promoter sequences flanking the inserts or in case of PC4 and PC9 was send for complete shotgun sequencing (MWG Biotech GmbH).

Some of the clone's DNA was furthermore used for the construction of small insert Lambda Zap II (insert size 0-10 kb) libraries in order to create more starting points for primer walking and also to identify the definite sequence regions that bound to the probes.

Table III-4: Summary of genomic phage library screening; columns show library name, number of screened plaques, probe and positive clone in the primary and secondary hybridization round. (*T13H and TDS used as a mixture in primary hybridization), T5Hol: oligo nucleotide probe.

Library	No. of plaques	Probe	Positive clones 1st Screening	Positive clones 2nd screening
Lib0001-080108	125,000	T13H*	3	2
		TDS*	3	/
Lib0001-150108	110,000	T13H*	3	1
		TDS*	3	2
Lib0001-030708	94,000	T5H	7	5
Lib0021	40,000	T5Hol	6	5

III.4.2.2 Screening of λ Zap II sub-libraries from EF0001 & EF0021 phage clones

The isolated phage clones from genomic library screening using Lambda Dash II replacement vector had the major advantage of large inserts, due to the size selective packaging. Thereby it is possible to screen a large amount of genetic information with a limited number of phage clones. The inserts obtained in the previous screening had sizes between 16 to 35 kb. Hence, sequencing of the recombinant DNA by primer walking is slow and time consuming.

To overcome this problem the isolated phage DNA was used for construction of sub-libraries based on the Lambda Zap II system (Stratagene).

In contrast to Lambda Dash II the Lambda Zap II insertion vector has a size of 41 kb and was created by deletion of none essential genes and insertion of a multiple cloning site. It has 6 unique restriction sites and DNA fragments between 0 and 10 kb can be inserted. Since there is no selection through vector size, also empty, re-ligated vectors can be present in the library. Therefore the multiple cloning site contains *lac*Z' gene encoding for the α-fragment of β-galactosidase. If a DNA fragment is inserted into the MCS LAcZ the protein can't be expressed. There through the library quality can be controlled via blue-white screening. The second part of the β-galactosidase (*ΔM15 lacZ* domain) is expressed by the F' episome of the host strain. Plaques containing an insert appear white whereas all empty vector carrying plaques are blue.

The other special feature of Lambda Zap II is the possibility to recover the recombinant DNA via *in vivo* excision. The vector contains the complete pBlueskript vector sequence around the MCS. By co-infection of host strain XL1-Blue MRF' with Lambda phages and ExAssist helper phage proteins of f1 phage recognize the f1 origin of replication present in pBlueskript and creates ssDNA from the double stranded lambda DNA. The resulting supernatant contains this pBluescript phagemid besides the lambda phages and the helper phage. SOLR strain (Su^-) cannot be infected with lambda phages and ExAssist helper phage. Only the newly created phagemid is inserted and the pBluescript ssDNA is converted to double stranded plasmid DNA containing the recombinant insert. *E. coli* clones can be selected on ampicillin plates.

To create suitable insert fragments for Lambda Zap II libraries the Lambda Dash II phage clone DNA was digested with TSP509I, resulting in compatible ends for ligation into *Eco*RI pretreated vector. The phage clone DNA was first partially digested to determine conditions leading to fragments of around 4,000 bp or smaller, that can be sequenced later directly from pBluescript (Figure III-12). Hybridization screening was done as described for Lambda Dash

II libraries. Due to the limited genetic information used as template for library construction only 3,000 – 4,000 individual clones were screened.

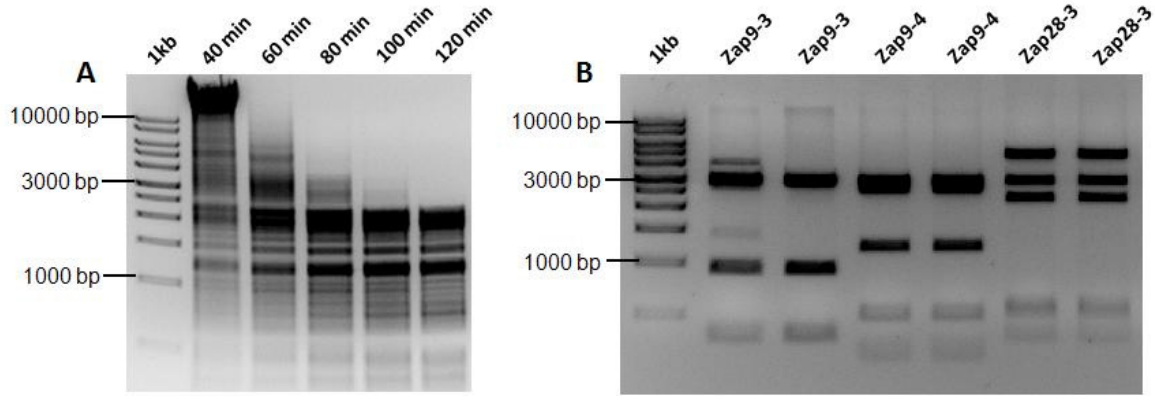


Figure III-12: Lambda Zap II library construction; A: time dependent restriction digest of PC28 phage DNA with TSP509I, 80 min sample was chosen for library construction; **B:** restriction digest with *Kpn*I and *Sac*I of three pBluescript clones after *in vivo* excision of positive clones from hybridization screening, band at 3 kb represents the vector backbone, other bands are derived from recombinant inserts, insert size Zap9-3: ~1200 bp, Zap9-4: ~1900 bp, Zap28-3: ~8000 bp.

Plaques found to give a positive hybridization result were eluted and the resulting phage suspension was used for *in vivo* excision. Vector constructs were isolated afterwards by standard plasmid isolation and tested for inserts by restriction with *Kpn*I and *Sac*I sites flanking the MCS of pBluescript (λ-Zap II).

Through this procedure, three pBluescript constructs from PC9 (Zap9-2, Zap9-3, Zap9-4) and three constructs from PC28 (Zap28-1, Zap28-2, Zap28-3) from EF0001 could be isolated. For PC4 and the clones from EF0021, PCT5H3a and PCT5H4a screening did not lead to isolation of λ–Zap II clones. The number of plaques examined might have been to low in these cases because the quality of the libraries was not sufficient, means to many empty vectors. Nevertheless the screening was not continued because PC4 was sequenced completely and EF0021 was chosen for whole genome sequencing.

III.4.2.3 Sequencing of clones identified via hybridization screening

Phage clones PC1, PC4 and PC9 isolated from Lambda Dash II library screening were sequenced first by primer walking starting from the T3 and T7 promoter sequences on the vector backbone. This effort resulted in two consensus sequences, one for each direction, for the three clones. PC28 positive in hybridization with T5H probe was directly used for construction of Lambda Zap II library. Clones Zap28-1-3 were sequenced using pBlueskript

plasmid preparations from *in vivo* excision. In case of PC9 also three Lambda Zap II clones were sequenced. Two EF0021 phage clones were selected for sequencing.
The results are summarized in Table III-5. Furthermore the best BLAST result is shown and the maximum homologous sequence length to the probe used for hybridization.

Table III-5: Summary of sequencing and analysis of "positive" phage clones by primer walking using T3 and T7 promoter primers on λ Dash II and pBluescript.

Clone	Length [bp]	Max. homology to probe [bp]	NCBI blastx result
PC4-T3	2071	/	hypothetical protein SNOG_00438 [*Phaeosphaeria nodorum* SN15]; XP_0017 1125
PC4-T7	2597	/	hypothetical protein VDAG_05916 [*Verticillium dahliae* VdLs.17]; EGY14752
PC9-T3	799	11	Polygalacturonase inhibitor protein [*Medicago truncatula*]; AES78029
PC9-T7	725	12	no significant blast match
PC1-T3	2334	11	hypothetical protein LEMA_P003780.1 [*Leptosphaeria maculans* JN3]; CBY01591
PC1-T7	1531	/	gamma-glutamyltranspeptidase periplasmic precursor [*Metarhizium anisopliae* ARSEF 23]; EFY94054
Zap9-2	1926 (T3), 1847 (T7)	11	capsid component [Enterobacteria phage lambda]; NP_040583
Zap9-3	1214	17	ferrienterobactin receptor domain protein [*Escherichia coli*]; EGW96796
Zap9-4	3220 (T3), 1832 (T7)	17	hypothetical protein MYCTH_2315847 [*Myceliophthora thermophila* ATCC 42464]; AEO59743
Zap28-1	2780	12	hypothetical protein, partial [*Podospora anserina* S mat+]; XP_001908716
Zap28-2	2550 (T3), 2435 (T7)	11	hypothetical protein SMAC_00380 [*Sordaria macrospora* k-hell]; XP_003351833
Zap28-3	1969 (T3), 2248 (T7)	13	phosphopyruvate hydratase [*Glomerella graminicola* M1.001]; EFQ29251
0021-T5H3a-T3	2043	12	predicted protein [*Ajellomyces capsulatus* NAm1]; XP_001537182
0021-T5H3a-T7	4433	11	hypothetical protein SS1G_13060 [*Sclerotinia sclerotiorum* 1980 UF-70]; XP_001585968
0021-T5H4a-T3	3722	11	hypothetical protein [*Botryotinia fuckeliana*]; CCD44329
0021-T5H4a-T7	4074	20	hypothetical protein SS1G_01906 [*Sclerotinia sclerotiorum* 1980 UF-70]; XP_001597710

None of the sequences contained a gene of Taxol biosynthetic pathway or showed significant homology to probe templates. The maximum sequence length homologous to a probe template was 20 bp for 0021-T5H4a-T7. For some sequences there was even no homology at

all. NCBI blastx analysis did not lead to the identification of possible Taxol biosynthetic pathway sequences or possible candidates for secondary metabolism enzymes either. Regarding Zap clones Zap9-2 and Zap9-3 the inserts were most closely related to λ-phage or *E. coli* sequences. That result was unexpected taking into account, that if the hybridization probes bind unspecific to the vector backbone all screened phage clones should have been positive.

Since the primer walking approach was very time consuming and the inserts of the phage vector were supposed to be very large, PC4 and PC9 DNA additionally was send to MWG Biotech GmbH for complete shotgun sequencing. Comparison of the borders of the inserts of PC4 and PC9 obtained by primer walking (Table III-5) the complete inserts could be identified. They had a size of 19,995 bp and 18,800 bp, respectively. A summary of MWG sequencing results is shown in Table III-6. Inserts thereby were part of the large contigs for both clones, in case of PC4 on c3 1-11,050 bp and c2 1-7,999 bp and PC9 on c2 1-4,131, c1 24,532-35,906 and c1 1-4,533 (supplementary information, CD).

Table III-6: Summary of MWG Biotech GmbH shotgun sequencing of λ-phage clones PC4 and PC9; including number of total reads, total bases and contigs divided into large contigs and all contigs (large contigs: > 500 bp).

phage clone	large run results		sum of large contigs		sum of all contigs	
PC9	Total Reads	*5,038*	Total Reads	*4,809*	Total Reads	*4,930*
	Total Bases	*1,112,115*	Large Contigs	*3*	All Contigs	*37*
			Total Bases	*48,335*	Total Bases	*52,507*
PC4	Total Reads	*7,781*	Total Reads	*7,454*	Total Reads	*7,711*
	Total Bases	*1,764,493*	Large Contigs	*17*	All Contigs	*105*
			Total Bases	*61,925*	Total Bases	*84,324*

Analysis of the sequences did not lead to the identification of any gene or sequence homologue to genes of Taxol biosynthetic pathway again. Even more surprising was the fact that there were no homologies to the probes used for hybridization either, although

hybridization results were highly reproducible. Alignment of probe sequences to the inserts delivered a maximum homologous part of 17 nucleotides for T13H and PC4 and 16 nucleotides for TDS and PC9. The only explanation for the positive signals in hybridization was the binding to these short sequences. Further analysis via NCBI BLAST confirmed these results. There was no indication for any gene of interest or genes encoding for enzymes probably involved in fungal Taxol biosynthesis.

The only explanation for this observation was that all clones isolated were false positives and hybridization results occurred through binding of probes to very short homologous parts of the recombinant phages under the conditions chosen according to initial Southern Blotting hybridization results.

In summary homology based screening revealed that fungal and plant pathways did not evolve by gene transfer from one species to the other because no homologous gene sequence could be isolated.

III.4.3 Whole genome sequencing of EF0021 & *Taxomyces andreanae*

Due to the unexpected results of genomic library screening approach not finding any gene homologous to genes from Taxol biosynthesis from *Taxus* species the possibility of two unrelated biosyntheses towards Taxol in plants and fungi had to be examined. Even if the genes are not similar it can be assumed that at least the proteins involved in the natural product formation might share some characteristics. As described in the introduction several enzymes of Taxol biosynthesis from yew tree are highly substrate and product specific, especially the key enzyme taxadiene synthase and the cytochrome P450 dependent monooxygenases catalyzing early hydroxylation reactions in the pathway, for example the taxadiene-5α-hydroxylase. The detection of identical taxoid products in fungal extracts as produced by *Taxus* and the enzyme characteristics make it seem very improbable that there are really no similarities also on protein level. By sequencing of the complete genomes of endophytic fungi it should be possible to identify any candidate gene/enzyme that might to be involved in terpene pathways. By prediction of the proteins on the one hand or direct comparison of the genome data to protein data base on the other hand similarities on enzyme level, assumed to be due to mechanistic features of the enzymes catalyzing complex reactions using structural elaborate substrates should become obvious.

To find evidence for a unrelated fungal Taxol biosynthetic pathway it is necessary to sequence at least two endophytes genomes. By comparison of both sequence data sets it should be possible to proof or rule out the hypothesis of an independent evolution of Taxol biosynthesis in plant and fungi by identification of terpene pathway genes/enzymes similar in both fungal species. This of course implies that there are similarities of the biosynthetic pathways in the different endophytic fungi.

Genomic DNA of EF0021 was therefore sent to the company Seq-It GmbH in Kaiserslautern for 454 sequencing. Furthermore *Taxomyces andreanae* genome was sequenced (Source BioScience imaGenes GmbH, Berlin, paired end sequencing). Assembly of the raw sequence data was done by the respective companies.

Sequencing of EF0021 resulted in 2,234,101 total sequence reads with a total number of bases of 871,644,690. This gives an average read length of 390 bases per read.

Alignment of the raw data led to 98.55 % aligned reads and 99.02 % aligned bases, respectively, resulting in 2,623 contigs. Of these total contigs 1,205 large contigs were identified. They have an average contig size of 36,591 bases, with the largest contig of 1,138,940 bases. Hence, the large contigs represent 99.85 % of the complete sequence information. All contigs together cover 44.45 Mbases and the estimated genome size was calculated to 45.9 Mb.

Genome project of *Taxomyces andreanae* delivered 235,442,880 reads of around 100 bp out of a 500 bp paired end library. Assembly led to 16,279 contigs with 2,274 large contigs (greater than 1000 bp) with an average contig size of 18,083 bp, covering 93.5 % of the assembled sequence information. All 16,279 contigs (minimum contig size 140 bp) cover a cumulative sequence length of 45.08 Mbases. The average coverage (reads/contigs) is 555.55. Out of this data genome size of *Taxomyces andreanae* can be approximated to around 45 Mb.

The complete contigs were submitted to NCBI gene bank. Furthermore the multi fasta data of both genomes is added as supplementary information (CD).

III.4.4 Sequence analysis strategy

For analysis of the two genomes with regard to the identification of Taxol biosynthetic pathway key enzymes/genes, total sequences were analyzed by local blastx with CLC workbench. For targeted comparison three different data sets were created.

The first set contained 38 terpene synthase protein sequences from GeneBank (II.5.9.4), including plant and fungal proteins for sesqui- and diterpene synthesis as well as sequences from sterol biosynthesis and geranylgeranyl diphosphate synthases.

The second comparison was performed using 34 Cytochrome P450 oxygenase protein sequences of proteins proven or at least supposed to be involved in secondary metabolite pathways from plants and fungi (II.5.9.4). In contrast to terpene synthases which are nearly all functionally characterized hydroxylases are mainly annotated over homology or in case of fungal enzymes found in gene clusters with for example a terpene synthase what is a strong indication for their role in the respective pathway.

The third data set consisted of all known proteins known to be involved in Taxol biosynthetic pathway in *Taxus* species (14 sequences, II.5.9.4). As a threshold for significance an E-value of E^{-5} was used.

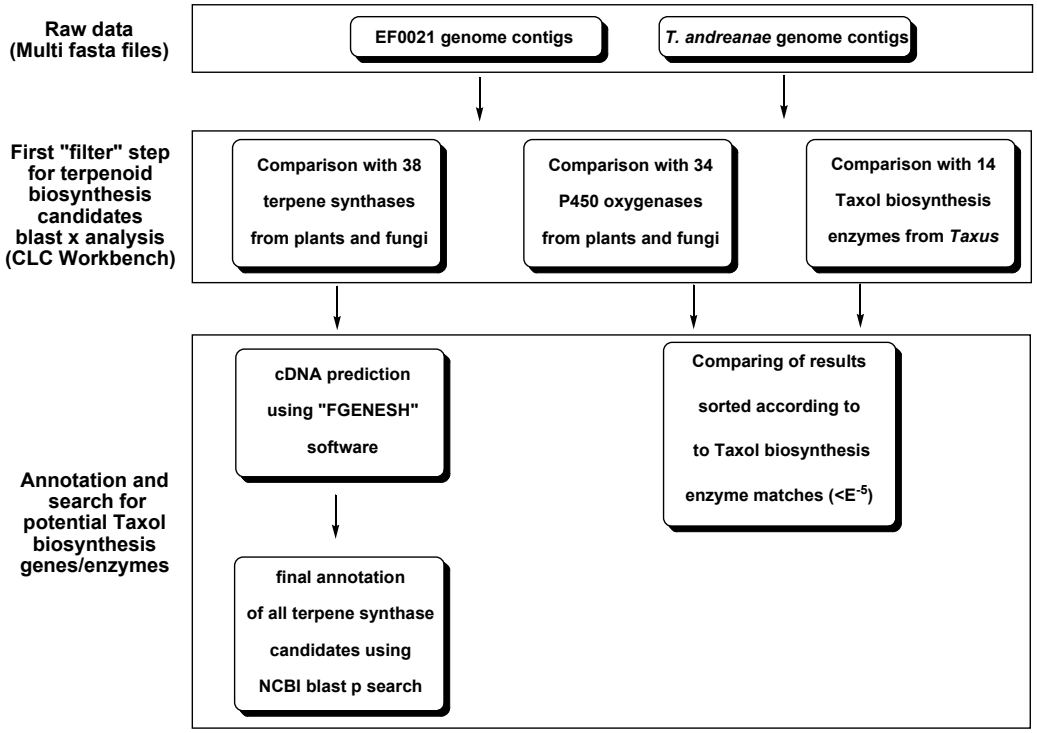


Figure III-13: Schematic overview of "targeted genome analysis approach" for identification of possible fungal Taxol biosynthesis gene/enzyme candidates.

After this first "filter" step the major focus lied on analysis of the terpene synthases, since they are the key enzymes of these biosynthetic pathways. Furthermore this enzyme class shows features characteristic not only for the mechanism catalyzed but also for the organism derived from. For example intron/exon structure, size of proteins and certain motives and the

position of these motives differ significantly in plant enzymes compared to fungal terpene synthases.

The sizes of the open reading frames and translated proteins were obtained by calculation with the "FGENESH"-software (http://linux1.softberry.com/) for prediction of the intron/exon structure. This structure can be rather different in respect to the organism. Therfore the calculation was performed with control genes first. The best algorithm was then used for the prediction of the cDNA of the identified terpene synthase genes. These predicted proteins were than analyzed again via blastp search for final annotation and identification of the closest ortholog.

In case of P450 oxygenases many more search matches were expected, due to motives shared not only by oxygenases involved in secondary metabolism. Therefore the results observed in this analysis were compared with all hits that were significant to Taxol biosynthetic enzymes. In theory by this comparison it should be obvious if the identified P450 candidates are more closely related to fungal or plant (*Taxus*) enzymes.

Additional to this two part targeted analysis approach whole genomes of EF0021 and *Taxomyces andreanae* were aligned with the sequences of an EST library of an induced *Taxus* cell culture, in order to identify genes in the genomes that show homology to *Taxus* genes transcribed in general. Besides the identification of fungal Taxol biosynthesis genes the aim of this search was to figure out if there are any genes homologous between the plant and its endophytes supporting the hypothesis of a gene transfer.

III.4.5 Terpene synthases from EF0021 & *Taxomyces andreanae*

Local blastx analysis with terpene synthase data set (II.5.9.4) and the two fungal genomes led to the identification of 6 genes/proteins in EF0021 genome and 20 genes/proteins from *Taxomyces andreanae*, respectively.

Regarding EF0021 terpene synthases the candidates can be divided into two proteins highly homologous to lanosterol synthase from *Aspergillus nidulans*, one protein similar to geranyl geranyl diphosphate synthase, three sesquiterpene synthases and one diterpene synthase. Due to their very high similarity to lanosterol synthase and the high conservation of steroid biosynthesis in general the first two proteins are most likely to be involved in sterol biosynthesis. More interesting according to the aim of identification of a possible taxadiene synthase from endophytes are the 4 proteins 0021_TS_1580, 0021_TS_1762, 0021_TS_320

and 0021_TS_2010, that are probably involved in natural product synthesis of EF0021 (Table III-7). The genes and calculated proteins show the highest homology to different fungal terpene synthases, but in accordance to the genomic library screening results not with genes of plant origin in particular with the taxadiene synthase gene from *Taxus* species.

The most interesting gene/enzyme is the only diterpene synthase found. It is an unusual chimeric diterpene synthase so far only isolated by Toyomasu and coworkers (TOYOMASU *et al.* 2007) from the plant-pathogenic fungus *Phomopsis amygdali*. This enzyme was shown to have two domains one for the cyclization of geranylgeranyldiphosphate (GGPP) to fusicoccadiene and one prenyltransferase domain for the synthesis of GGPP from the universal terpene precursors isopentenyldiphosphate and dimethylallyldiphosphate. Such multifunctional enzymes are not known from plants where the synthesis of GGPP and the terpene backbone is catalyzed always by two enzymes.

Analysis of *Taxomyces andreanae* genome led to the identification of 20 terpene synthase candidates (19 terpene synthases, one prenyltransferase, Table III-8). Surprisingly all candidates are most probably sesquiterpene synthases, regarding homology in local blastx analysis as well as after calculation of predicted protein sequences.

There was no indication for an enzyme or gene for diterpene synthesis. All identified open reading frames had a size of around 1000 bp resulting in proteins of 305 to 378 amino acids. The only exception is terpene synthase c7630_ts_8 having a cDNA size of 2172 bp resulting in a 713 aa protein. A closer view on the sequence revealed that the predicted protein consists of two parts, both homologue to sesquiterpene synthase Cop6 (XP_001832549). Therefore it is most likely that the prediction was wrong and it is actually two sesquiterpene synthases, although there are predicted proteins in Genebank having the same characteristics.

Table III-7: Predicted terpene synthases and prenyltransferases identified in EF0021 genome

Contig	Name	Contig size [bp]	Best match (local blastx)	E-value	orf size [bp]	Protein size [aa]	Best NCBI blastp (function), organism; accession No.	E-value
1580	0021_TS_1580	412494	Protoilludene synthase; *Armillaria gallica*	$1*E^{-70}$	1023	341	Hypothetical protein CHGG_03509 (Isoprenoid_Biosyn_C1) *Chaetomium globosum* CBS 148.51; XP_001230025.1	$1*E^{-97}$
1762	0021_TS_1762	167955	Fusicoccadiene synthase; *Phomopsis amygdali*	$2*E^{-45}$	2169	723	Polyprenyl synthetase (Isoprenoid_Biosyn_C1) *Glomerella graminicola* M1.001; EFQ35158.1	$5*E^{-83}$
320	0021_TS_320	14785	Cop6; *Coprinopsis cinereus*	$9*E^{-19}$	1608	536	Hypothetical protein SNOG_03562 (Isoprenoid_Biosyn_C1) *Phaeosphaeria nodorum* SN15; XP_001794120.1	0
2010	0021_TS_2010	281509	Cop6; *Coprinopsis cinereus*	$2*E^{-11}$	1257	419	Hypothetical protein ANI_1_396154 (Isoprenoid_Biosyn_C1) *Aspergillus niger* CBS 513.88; XP_001398274.1	$3*E^{-35}$
1411		717764	Lanosterol synthase; *Aspergillus nidulans*	0	2505	835	Lanosterol synthase *Neurospora crassa* OR74A; XP_961026.1	0
1058		100152	Lanosterol synthase; *Aspergillus nidulans*	$6*E^{-147}$	2304	768	Squalene-hopene-cyclase *Aspergillus fumigatus* Af293; XP_751356.1	0
1708		1031903	GGPPS; *Aspergillus nidulans*	$1*E^{-92}$	1047	349	Trans_Isoprenyl Diphosphate Synthases *Sclerotinia sclerotiorum* 1980 UF-70; XP_001588566.1	$3*E^{-158}$

Table III-8: Predicted terpene synthases and prenyltransferases identified in *Taxomyces andreanae* genome

Contig	Name	Contig size [bp]	Best match (local blastx)	E-value	orf size [bp]	Protein size[aa]	Best NCBI blastp (function), organism; accession No.	E-value
7448	c7448_ts	39656	Cop1; *Coprinopsis cinereus*	$4.8*E^{-124}$	1107	368	Isoprenoid_Biosyn_Enz_C1, *Laccaria bicolor* S238N-H82; XP001881043.1	$9*E^{-134}$
7630	c7630_ts_1	118498	Cop4; *Coprinopsis cinereus*	$3.6*E^{-105}$	1050	349	Isoprenoid_Biosyn_Enz_C1, *Postia placenta* Mad-698-R; XP002473599.1	$1*E^{-144}$
	c7630_ts_2		Cop4; *Coprinopsis cinereus*	$5*E^{-64}$	1038	345	Isoprenoid_Biosyn_Enz_C1, *Postia placenta* Mad-698-R; XP002473599.1	$4*E^{-78}$
	c7630_ts_3		Cop4; *Coprinopsis cinereus*	$1*E^{-60}$	1029	342	Isoprenoid_Biosyn_Enz_C1, *Postia placenta* Mad-698-R; XP002473599.1	$5*E^{-89}$
	c7630_ts_4		Cop4; *Coprinopsis cinereus*	$5*E^{-54}$	1002	333	Isoprenoid_Biosyn_Enz_C1, *Postia placenta* Mad-698-R; XP002473599.1	$1*E^{-97}$
	c7630_ts_5		Cop4; *Coprinopsis cinereus*	$4*E^{-50}$	1014	337	Isoprenoid_Biosyn_Enz_C1, *Postia placenta* Mad-698-R; XP002473599.1	$9*E^{-100}$
	c7630_ts_6		Cop4; *Coprinopsis cinereus*	$2*E^{-44}$	1026	341	Isoprenoid_Biosyn_Enz_C1, *Postia placenta* Mad-698-R; XP002473599.1	$3*E^{-95}$
	c7630_ts_7		Cop6; *Coprinopsis cinereus*	$2*E^{-12}$	978	325	predicted protein, *Postia placenta* Mad-698-R; XP002472767.1	$8*E^{-41}$
	c7630_ts_8		Cop6; *Coprinopsis cinereus*	$9*E^{-9}$	2142	713	predicted protein, *Postia placenta* Mad-698-R; XP002473977.1	$4*E^{-69}$
	c7630_ts_9		Cop6; *Coprinopsis cinereus*	$1*E^{-6}$	948	315	predicted protein, *Postia placenta* Mad-698-R; XP002472767.1	$2*E^{-46}$
7735	c7735_ts_1	15912	Pro1; *Armillaria gallica*	$2*E^{-49}$	585	194 (partial)	Isoprenoid_Biosyn_Enz_C1, *Postia placenta* Mad-698-R; XP002472842.1	$5*E^{-38}$
	c7735_ts_2		Pro1; *Armillaria gallica*	$3*E^{-25}$	594	198 (partial)	Isoprenoid_Biosyn_Enz_C1, *Postia placenta* Mad-698-R; XP002475233.1	$3*E^{-21}$
7797	c7797_ptr	75014	GGPPS; *Aspergillus nidulans*	$3.7*E^{-33}$	918	305	Trans_Isoprenyl Diphosphate Synthase; *Postia placenta* Mad-698-R; XP002472170.1	$4*E^{-113}$
	c7797_ts		Cop6; *Coprinopsis cinereus*	$1*E^{-15}$	954	317	Isoprenoid_Biosyn_Enz_C1, *Postia placenta* Mad-698-R; XP002475902.1	$4*E^{-43}$

Table III-8 continued

7466	c7466_ts	76539	Cop6; *Coprinopsis cinereus*	$2*E^{-24}$	1137	378	Isoprenoid_Biosyn_Enz_C1, *Laccaria bicolor* S238N-H82; XP001885710.1	$8*E^{-22}$
8026	c8026_ts	11399	Cop6; *Coprinopsis cinereus*	$4*E^{-16}$	855	284	Isoprenoid_Biosyn_Enz_C1, *Postia placenta* Mad-698-R; XP002475902.1	$9*E^{-23}$
7741	c7741_ts	93200	Cop6; *Coprinopsis cinereus*	$5*E^{-14}$	954	317	Isoprenoid_Biosyn_Enz_C1, *Postia placenta* Mad-698-R; XP002475902.1	$6*E^{-37}$
8371	c8371_ts	13851	Cop4; *Coprinopsis cinereus*	$5*E^{-9}$	1029	342	Isoprenoid_Biosyn_Enz_C1, *Postia placenta* Mad-698-R; XP002475451.1	$7*E^{-27}$
5849	c5849_ts	13348	Cop6; *Coprinopsis cinereus*	$8*E^{-9}$	966	321	Isoprenoid_Biosyn_Enz_C1, *Postia placenta* Mad-698-R; XP002475902.1	$7*E^{-42}$
7493	c7493_ts	142612	Cop1; *Coprinopsis cinereus*	$9*E^{-7}$	987	328	Isoprenoid_Biosyn_Enz_C1, *Schizophyllum comune* H4-8; XP003033206.1	$2*E^{-23}$

Comparison of terpene synthases from EF0021 with enzymes predicted from *Taxomyces andreanae* was done via phylogenetic analysis (Figure III-14).

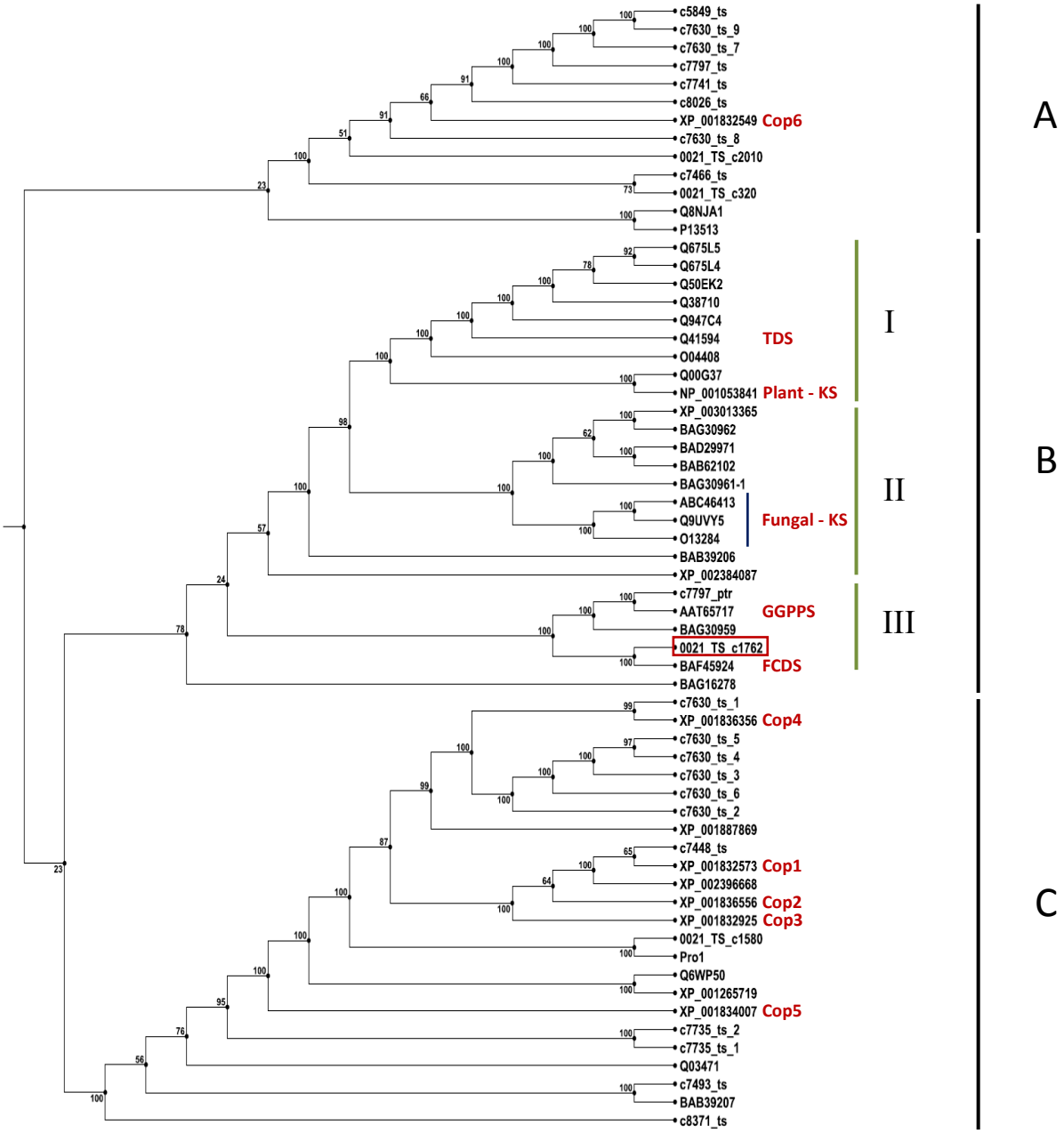


Figure III-14: Phylogenetic tree constructed with UPGMA (unweighted pair group method with arithmetic means) from all predicted terpene synthase protein sequences (Table III-7, Table III-8) and the reference proteins used for local blastx analysis of the two fungal genomes (II.5.9.4); separation into three major clades (A-C); **A:** terpene synthases of Cop6 (XP_001832549) and trichodiene synthase type (P13513); **B:** terpene synthases of plant origin (B.I), involved in fungal gibberellin synthesis (B.II); fungal prenyltransferases and fusicoccadiene synthase (BAF45924), including the only diterpene synthase from both fungal species 0021_TS_c1762, annotated as fusicoccadiene synthase (B.III); **C:** terpene synthases of Cop1-5 type; in red: functionally characterized enzymes.
TDS: taxadiene synthase, KS: *ent*-kauren synthase, GGPPS: geranyl geranyl diphosphate synthase, FCDS: fusicoccadine synthase, Cop1-Cop6: sesquiterpene synthases from *Coprinopsis cinereus* **(AGGER *et al.* 2009)**.

All protein sequences predicted from the two genomes were aligned with the protein sequences initially used for targeted search of terpene synthases (II.5.9.4). The aligned data set was used to construct a phylogenetic tree using UPGMA (**u**nweighted **p**air **g**roup **m**ethod with **a**rithmetic means) with bootstrapping (100 replicates, bootstrap values shown at the nodes, Figure III-14).

The analyzed protein sequences clustered into three major clades (Figure III-14, A-C). Clade B could be divided into three sub-clades, regarding plant or fungal protein origin and proposed or known protein function, for example their role in gibberellin biosynthesis. The abbreviations in red are marking functionally characterized enzymes.

Clade A consisted of proteins annotated as sesquiterpene synthases most similar to Cop6 from *Coprinopsis cinereus*, including 0021_TS_2010 and 0021_TS_320 from EF0021 and 8 proteins from *Taxomyces andreanae*. Furthermore Cop6 (XP_001832549) itself and the two trichodiene synthases from the terpene synthase reference data set belonged to this clade, although the last mentioned sequences were clearly separated from the predicted proteins and Cop6.

All other predicted sesquiterpene synthases clustered together with Cop1-Cop5 in clade C. Within this clade the sequences could be divided into homologues to Cop4, homologues to Cop1-3 and proteins similar to these types of synthases. Six proteins from *Taxomyces andreanae* formed a group with Cop4 (upper part of C, Figure III-14). Of these c7630_ts_1 was most similar to Cop4, whereas the other five sequences were more closely related to each other than to the reference sequence. Looking back to the genome, all the genes encoding for the predicted proteins c7630_ts_1-6 were found not only on the same contig but within a 25 kb DNA fragment. It seems most likely that the genes evolved through duplication events, what explains the high similarity of the predicted amino acid sequences. c7448_ts seemed to be most similar to Cop1-Cop3 in accordance to blastx analysis (Table III-8), whereas c7493_ts initially also most similar to Cop1 now clustered with BAB39207 (predicted di-terpene synthase, *Kitasatospora griseola*). The three additional proteins from *Taxomyces andreanae* (c7735_ts_1, 2 and c8371_ts) in C represented probably new sesquiterpene synthases sharing some features with Cop1-5 type enzymes (grouped in C) but were not as highly related to the reference sequences as the other proteins in the clade. The only enzyme from EF0021 in C is most closely related to protoilludene synthase from *Armillaria gallica*. In contrast to the sequences clustering with sesquiterpene synthase

reference sequences in clades A and C, only two predicted proteins from EF0021 and *Taxomyces andreanae* genomes were part of clade B. Furthermore B contained all reference proteins from plants, including taxadiene synthase, all enzymes involved in gibberellins biosynthesis from plants as well as from fungi and the few known di-terpene synthases known from fungi aphidicolan-16β-ol synthases from *Phoma betae* (BAD29971, BAB62102) and the fusicoccadiene synthase from *Phomopsis amygdali* (BAF45924).

Both predicted endophytic enzymes belonged to sub-clade B.III. They clustered with prenyltransferases from fungi like the GGPPS from *Aspergillus nidulans* and *Phomopsis amygdali* (AAT65717, BAG30959) on the one hand and with fusicoccadiene synthase from *Phomopsis amygdali* (BAF45924) on the other hand. Fusicoccadiene synthase and 0021_TS_1762 are enzymes containing both a prenyltransferase part and a terpene synthase part. Nevertheless, the higher conservation of prenyltransferases in comparison to terpene cyclases made the position of the enzymes in a sub-group together with prenyltransferases look reasonable.

Clade B.III was clearly separated from the other ones (B.I, B.II). Sub-clade B.I contained all sequences derived from plants, including taxadiene synthase and the enzymes involved in plant gibberellins biosynthesis, like *ent*-kauren synthase from rice (NP_001053841), but no sequence derived from fungi. Fungal diterpene synthases as aphidicolan-16β-ol synthase from *Phoma betae* (BAD_29971) and other predicted terpene synthases like diterpene cyclase 1 from *Klitasatospora griseola* (BAB39206) were found in sub-clade II. Furthermore the sequences of fungal copalyl diphosphate synthases and kauren synthases (BAG30962, ABC46413, Q9UVY5) were included here, clearly showing the difference of enzymes catalyzing reactions towards gibberellins between plants and fungi.

In summary, only two enzymes identified in the endophytic genomes were placed in the same major clade as known plant terpene synthases by phylogenetic analysis. Nevertheless, both proteins were more closely related to fungal enzymes, especially to prenyltransferases. By including enzymes from gibberellins pathway it became clear, that the origin of the proteins played the major role for sub-grouping in clade B. Although involved in biosyntheses leading to the same natural end products single enzymes are significantly different and their phylogenetic placement is dominated by the origin. There was further no indication for a terpene synthase that shared features with plant derived cyclases that would support the hypothesis of a gene transfer of Taxol biosynthesis from plant to fungus.

Analysis of clades A and C supports the role of the predicted cyclases from the endophytes grouped here in sesquiterpene biosynthetic pathways in accordance to the initial annotation. The division in A, B and C seemed to be due to mechanistic features resulting in characteristic differences of protein sequences of sesquiterpene and diterpene forming enzymes in fungi. Likewise plant sesquiterpene synthases were clustered dependent on origin derived features, leading to their position in B.I together with the other plant sequences rather than to sesquiterpene biosyntheses enzymes.

Taking mechanistic considerations into account and the general difference of plant and fungal synthases, clades A and C only contained sesquiterpene synthases from fungi, while all predicted and functionally characterized diterpene synthases, all enzymes involved in the synthesis of the diterpenoid gibberellins and all plant derived enzymes regardless of their function belonged to clade B. The only reasonable candidate for a fungal taxadiene synthase evolved independently from plant pathway would be 0021_TS_1762, although no homologue for an enzyme like this was found in *Taxomyces andreanae.*

III.4.6 Cytochrome P450 oxygenases from EF0021 & *Taxomyces andreanae*

As initially discussed in the introduction, P450 oxygenases as well as acyltransferases play major roles in Taxol biosynthesis besides the key enzyme taxadiene synthase. In contrast to terpene synthases many P450 enzymes are not functionally characterized but only annotated by homology in genome projects or are supposed to be involved in biosynthetic pathways due to their occurrence in gene clusters physically linked to for example terpene synthase genes. Cytochrome P450 enzymes are involved in a multitude of biosyntheses and are specific for a respective biosynthetic pathway, for example in the formation of steroids. Due to these considerations the analysis for identification of P450 genes/enzymes from the genomes was performed differently than for terpene synthases. The two genomes were subjected for blastx analysis using the known and characterized protein sequences of Taxol biosynthesis in *Taxus* on the one hand (II.5.9.4) and a data set of 34 protein sequences annotated as P450 oxygenases that might be involved in secondary metabolite pathways on the other hand (II.5.9.4). The two obtained data sets were than sorted by E-value with a threshold of E^{-5} and compared to each other. Results are listed in Table III-9 and Table III-10.

Table III-9: Local blastx analysis of EF0021 genome against Taxol biosynthesis protein data set in comparison to results from local blastx analysis of the genome against P450 oxygenase reference data set.

Query	local blastx, Taxol biosynthesis proteins	E-value	local blastx, P450 reference proteins	E-value
contig03148	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$3.57*E^{-90}$	cytochrome P450 oxidoreductase [*Gibberella fujikuroi*], CAE09055	0
contig02727	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$1.47*E^{-54}$	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$5.75*E^{-54}$
contig00613	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$5.86*E^{-43}$	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$2.23*E^{-42}$
contig01448	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$1.81*E^{-29}$	fusicoccadiene C-8 hydroxylase [*Alternaria brassicicola*], BAI52800	$3.70*E^{-41}$
contig02012	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$6.87*E^{-24}$	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$2.62*E^{-23}$
contig00496	taxoid 13-α hydroxylase [*Taxus cuspidata*], Q8W4T	$1.89*E^{-19}$	ent-kaurenoic acid hydroxylase [*Arabidopsis thaliana*], AAK11564	$7.71*E^{-21}$
contig00421	taxoid 13-α hydroxylase [*Taxus cuspidata*], Q8W4T9	$6.56*E^{-19}$	cytochrome-P450 hydroxylase [*Coprinopsis cinerea* okayama7#130], XP_001830548	$1.72*E^{-43}$
contig00194	taxoid 2-α-hydroxylase [*Taxus canadensis*], AAS89065	$3.94*E^{-18}$	cytochrome P450 monoxygenase [*Botryotinia fuckeliana*], CAH64679	$5.74*E^{-47}$
contig02051	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	$2.71*E^{-16}$	cytochrome P450 monooxygenase [*Botryotinia fuckeliana*], CAE76652	$7.06*E^{-65}$
contig00377	taxoid 14-β-hydroxylase *[Taxus cuspidata*], AAO66199	$3.33*E^{-16}$	cytochrome P450 alkane hydroxylase [*Aspergillus fumigatus* A1163], EDP55514	$7.98*E^{-119}$
contig00459	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	$2.56*E^{-15}$	fusicoccadiene 8-ol C-15 hydroxylase [*Alternaria brassicicola*], BAI52803	$1.65*E^{-54}$
contig00064	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$9.26*E^{-15}$	cytochrome P450 monooxygenase [*Laccaria bicolor* S238N-H82], XP_001881086	$5.08*E^{-21}$
contig01709	taxoid 2-α-hydroxylase [*Taxus canadensis*], AAS89065	$1.06*E^{-14}$	fusicoccadiene 8-ol C-15 hydroxylase [*Alternaria brassicicola*], BAI52803	$1.86*E^{-43}$

Table III-9 continued

contig02000	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	$1.17*E^{-13}$	cytochrome P450 alkane hydroxylase [*Aspergillus fumigatus* A1163], EDP55514	$7.95*E^{-79}$
contig01580	taxoid 13-α hydroxylase [*Taxus cuspidata*], Q8W4T9	$2.32*E^{-13}$	fusicoccadiene 8-ol C-15 hydroxylase [*Alternaria brassicicola*], BAI52803	$1.10*E^{-71}$
contig00418	taxoid 13-α hydroxylase [*Taxus cuspidata*], Q8W4T9	$9.87*E^{-13}$	cytochrome P450-2 [*Coprinopsis cinerea* okayama7#130], XP_001835122	$1.33*E^{-17}$
contig00435	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	$1.06*E^{-12}$	cytochrome P450 monooxygenase [*Laccaria bicolor* S238N-H82], XP_001881086	$3.90*E^{-31}$
contig01280	taxoid 13-α hydroxylase [*Taxus cuspidata*], Q8W4T9	$1.88*E^{-12}$	cytochrome P450 alkane hydroxylase [*Aspergillus fumigatus* A1163], EDP55514	$1.91*E^{-108}$
contig00453	taxoid 2-α-hydroxylase [*Taxus canadensis*], AAS89065	$2.06*E^{-12}$	cytochrome P450 alkane hydroxylase [*Aspergillus fumigatus* A1163], EDP55514	$6.76*E^{-72}$
contig02056	5-α-taxadienol-10-β-hydroxylase [*Taxus cuspidata*], Q9AXM6	$2.12*E^{-12}$	cytochrome P450-2 [*Coprinopsis cinerea* okayama7#130], XP_001835122	$5.71*E^{-58}$
contig00380	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	$2.50*E^{-12}$	cytochrome P450 monooxygenase [*Laccaria bicolor* S238N-H82], XP_001881086	$5.34*E^{-47}$
contig01900	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	$3.94*E^{-12}$	cytochrome P450 monooxygenase [*Laccaria bicolor* S238N-H82], XP_001881086	$1.12*E^{-132}$
contig02343	5-α-taxadienol-10-β-hydroxylase [*Taxus cuspidata*], Q9AXM6	$1.70*E^{-11}$	cytochrome P450 monooxygenase [*Botryotinia fuckeliana*], CAP58781	$2.74*E^{-86}$
contig00414	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	$1.73*E^{-11}$	cytochrome P450 alkane hydroxylase [*Aspergillus fumigatus* A1163]□EDP55514	$6.28*E^{-67}$
contig00510	5-α-taxadienol-10-β-hydroxylase [*Taxus cuspidata*], Q9AXM6	$2.10*E^{-11}$	cytochrome P450 monooxygenase [*Laccaria bicolor* S238N-H82], XP_001881086	$1.44E^{-19}$
contig02617	taxoid 13-α hydroxylase [*Taxus cuspidata*], Q8W4T9	$2.33*E^{-11}$	cytochrome P450 monooxygenase [*Botryotinia fuckeliana*], CAE76652	$1.28*E^{-25}$
contig00684	taxoid 7-β-hydroxylase [*Taxus cuspidata*], AAQ75553	$3.76*E^{-11}$	fusicoccadiene C-8 hydroxylase [*Alternaria brassicicola*], BAI52800	$2.79*E^{-67}$
contig00394	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	$7.54*E^{-11}$	cytochrome-P450 hydroxylase [*Coprinopsis cinerea*□okayama7#130], XP_001830548	$1.18*E^{-48}$

Table III-9 continued

contig00431	taxoid 13-α hydroxylase [*Taxus cuspidata*], Q8W4T9	1.73*E^{-10}	cytochrome P450 alkane hydroxylase [*Aspergillus fumigatus* A1163], EDP55514	4.21*E^{-80}
contig01343	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	1.74*E^{-10}	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	6.64*E^{-10}
contig01970	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	7.17*E^{-10}	cytochrome P450 alkane hydroxylase [*Aspergillus fumigatus* A1163], EDP55514	5.50*E^{-69}
contig00432	5-α-taxadienol-10-β-hydroxylase [*Taxus cuspidata*], Q9AXM6	1.15*E^{-9}	isotrichodermin C-15 hydroxylase [*Fusarium sporotrichioides*], O13317	1.64*E^{-70}
contig00539	taxoid 13-α hydroxylase [*Taxus cuspidata*], Q8W4T9	2.05*E^{-9}	cytochrome P450 alkane hydroxylase [*Aspergillus fumigatus* A1163], EDP55514	5.31*E^{-82}
contig02421	taxoid 13-α hydroxylase [*Taxus cuspidata*], Q8W4T9	2.45*E^{-9}	cytochrome P450 alkane hydroxylase [*Aspergillus fumigatus* A1163], EDP47672	9.98*E^{-19}
contig00234	5-α-taxadienol-10-β-hydroxylase [*Taxus cuspidata*], Q9AXM6	2.45*E^{-9}	cytochrome P450 monooxygenase [*Laccaria bicolor* S238N-H82], XP_001881086	5.29*E^{-20}
contig01990	taxoid 2-α-hydroxylase [*Taxus canadensis*], AAS89065	2.72*E^{-9}	trichothecene C-8 hydroxylase [*Fusarium sporotrichioides*], AAO64248	7.22*E^{-29}
contig02038	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	3.42*E^{-9}	cytochrome P450 monoxygenase [*Botryotinia fuckeliana*], CAH64679	1.32*E^{-37}
contig01951	5-α-taxadienol-10-β-hydroxylase [*Taxus cuspidata*], Q9AXM6	3.49*E^{-9}	cytochrome P450-2 [*Coprinopsis cinerea* okayama7#130], XP_001835122	2.23*E^{-64}
contig0101	taxoid 13-α hydroxylase [*Taxus cuspidata*], Q8W4T	5.51*E^{-9}	cytochrome P450 alkane hydroxylase [*Aspergillus fumigatus* A1163], EDP55514	1.05*E^{-84}
contig00509	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	1.15*E^{-8}	CYP710A1; C-22 sterol desaturase/ oxygen binding [*Arabidopsis thaliana*], NP_180997	3.67*E^{-39}
contig00493	5-α-taxadienol-10-β-hydroxylase [*Taxus cuspidata*], Q9AXM6	3.94*E^{-8}	isotrichodermin C-15 hydroxylase [*Fusarium sporotrichioides*], O13317	2.80*E^{-46}
contig00672	5-α-taxadienol-10-β-hydroxylase [*Taxus cuspidata*], Q9AXM6	4.94*E^{-8}	cytochrome P450-1 [*Coprinopsis cinerea*], BAA33717	2.89*E^{-32}
contig01498	taxoid 7-β-hydroxylase [*Taxus cuspidata*], AAQ75553	4.99*E^{-8}	cytochrome P450-1 [*Coprinopsis cinerea*], BAA33717	1.53*E^{-57}

Table III-9 continued

contig00376	taxoid 2-α-hydroxylase [*Taxus canadensis*], AAS89065	$7.95*E^{-8}$	fusicoccadiene C-8 hydroxylase [*Alternaria brassicicola*], BAI52800	$4.09*E^{-24}$
contig00452	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	$1.09*E^{-7}$	cytochrome P450 monooxygenase [*Laccaria bicolor* S238N-H82], XP_001881086	$3.19*E^{-116}$
contig02405	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	$2.13*E^{-7}$	cytochrome P450 alkane hydroxylase [*Aspergillus fumigatus* A1163], EDP55514	$6.58*E^{-82}$
contig00222	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	$2.68*E^{-7}$	cytochrome P450-2 [*Coprinopsis cinerea* okayama7#130], XP_001835122	$5.64*E^{-64}$
contig00339	taxoid 14-β-hydroxylase *[Taxus cuspidata*], AAO66199	$2.78*E^{-7}$	cytochrome P450 alkane hydroxylase [*Aspergillus fumigatus* A1163], EDP55514	$2.06*E^{-26}$
contig00605	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	$3.73*E^{-7}$	fusicoccadiene C-8 hydroxylase [*Alternaria brassicicola*], BAI52800	$1.08*E^{-38}$
contig01501	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$6.28*E^{-7}$	cytochrome P450 monooxygenase [*Botryotinia fuckeliana*], CAE76652	$1.10*E^{-67}$
contig01472	taxoid 2-α-hydroxylase [*Taxus canadensis*], AAS89065	$1.99*E^{-6}$	cytochrome P450 monooxygenase [*Gibberella fujikuroi*], CAA75566	$8.52*E^{-36}$
contig00320	taxoid 14-β-hydroxylase *[Taxus cuspidata*], AAO66199	$5.63*E^{-6}$	cytochrome-P450 hydroxylase [*Coprinopsis cinerea* okayama7#130], XP_001830548	$1.39*E^{-20}$
contig01879	taxoid 13-α hydroxylase [*Taxus cuspidata*], Q8W4T9	$6.25*E^{-6}$	cytochrome P450 monooxygenase [*Gibberella fujikuroi*], CAA75565	$5.46*E^{-82}$
contig00473	taxoid 13-α hydroxylase [*Taxus cuspidata*], Q8W4T9	$8.03*E^{-6}$	cytochrome P450 monooxygenase [*Laccaria bicolor* S238N-H82], XP_001881086	$1.03*E^{-46}$
contig00402	taxoid 7-β-hydroxylase [*Taxus cuspidata*], AAQ75553	$8.46*E^{-6}$	cytochrome P450 monoxygenase [*Botryotinia fuckeliana*], CAH64679	$5.17*E^{-18}$
contig00407	taxoid 7-β-hydroxylase [*Taxus cuspidata*], AAQ75553	$9.84*E^{-6}$	cytochrome P450-2 [*Coprinopsis cinerea* okayama7#130], XP_001835122	$6.57*E^{-26}$
contig00396	taxoid 7-β-hydroxylase [*Taxus cuspidata*], AAQ75553	$2.26*E^{-5}$	fusicoccadiene C-8 hydroxylase [*Alternaria brassicicola*], BAI52800	$6.56*E^{-37}$
contig01673	taxoid 2-α-hydroxylase [*Taxus canadensis*], AAS89065	$6.88*E^{-5}$	cytochrome P450 alkane hydroxylase [*Aspergillus fumigatus* A1163], EDP55514	$4.27*E^{-65}$

Table III-10: Local blastx analysis of *Taxomyces andreanae* genome against Taxol biosynthesis protein data set in comparison to results from local blastx analysis of the genome against P450 oxygenase reference data set.

Query	Local blastx, Taxol biosynthesis proteins	E-value	Local blastx, P450 reference proteins	E-value
contig_7943	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$3.09*E^{-79}$	cytochrome P450 oxidoreductase [*Coprinopsis cinerea* okayama7#130], XP_002910681	0
contig_7798	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$8.6*E^{-34}$	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$3.28*E^{-33}$
contig_8600	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$2.06*E^{-28}$	cytochrome P450 oxidoreductase [*Gibberella fujikuroi*], CAE09055	$3.51*E^{-28}$
contig_8399	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$3.97*E^{-26}$	cytochrome P450 oxidoreductase [*Gibberella fujikuroi*], CAE09055	$8.90*E^{-26}$
contig_7768	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$8.18*E^{-26}$	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$3.12*E^{-25}$
contig_7609	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$2.17*E^{-25}$	cytochrome P450 oxidoreductase [*Gibberella fujikuroi*], CAE09055	$1.16*E-^{26}$
contig_7848	taxoid 13-α hydroxylase [*Taxus cuspidata*], Q8W4T9	$1.10*E^{-15}$	CYP710A1; C-22 sterol desaturase/ oxygen binding [*Arabidopsis thaliana*], NP_180997	$5.76*E^{-42}$
contig_7962	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$6.50*E^{-14}$	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	$2.48*E^{-13}$
contig_7383	5-α-taxadienol-10-β-hydroxylase [*Taxus cuspidata*], Q9AXM6	$2.96*E^{-12}$	cytochrome P450-1 [*Coprinopsis cinerea*], BAA33717	$1.04*E^{-32}$
contig_7354	taxoid 2-α-hydroxylase [*Taxus canadensis*], AAS89065	$3.73*E^{-11}$	cytochrome P450-1 [*Coprinopsis cinerea*], BAA3371	$3.92*E^{-85}$
contig_8569	5-α-taxadienol-10-β-hydroxylase [*Taxus cuspidata*], Q9AXM6	$1.41*E^{-10}$	cytochrome P450-2 [*Coprinopsis cinerea* okayama7#130□, XP_001835122	$2.04*E^{-58}$
contig_7647	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	$2.21*E^{-10}$	cytochrome P450 monooxygenase [*Laccaria bicolor* S238N-H82], XP_001881086	$3.60*E^{-178}$
contig_7767	5-α-taxadienol-10-β-hydroxylase [*Taxus cuspidata*], Q9AXM6	$4.12*E^{-10}$	cytochrome P450-1 [*Coprinopsis cinerea*], BAA33717	$2.16*E^{-39}$

Table III-10 continued

contig_7849	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	1.64*E^{-9}	cytochrome P450 oxidoreductase [*Coprinopsis cinerea* okayama7#130], XP_002910681	2.15*E^{-9}
contig_7744	taxoid 14-β-hydroxylase [*Taxus cuspidata*], AAO66199	6.23*E^{-9}	cytochrome P450 [*Aspergillus fumigatus* Af293], XP_747185	1.68*E^{-22}
contig_7668	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	1.15*E^{-8}	cytochrome P450 monooxygenase [*Laccaria bicolor* S238N-H82], XP_001881086	3.39*E^{-116}
contig_7815	5-α-taxadienol-10-β-hydroxylase [*Taxus cuspidata*], Q9AXM6	2.91*E^{-7}	cytochrome P450-1 [*Coprinopsis cinerea*], BAA33717	1.85*E^{-29}
contig_8308	taxoid 2-α-hydroxylase [*Taxus canadensis*], AAS89065	3.46*E^{-7}	cytochrome P450-1 [*Coprinopsis cinerea*], BAA33717	1.31*E^{-33}
contig_7370	cytochrome P450 reductase [*Taxus wallichiana var. chinensis*], AAX59902	4.55*E^{-7}	cytochrome P450 oxidoreductase [*Gibberella fujikuroi*], CAE09055	1.02*E^{-6}
contig_7542	5-α-taxadienol-10-β-hydroxylase [*Taxus cuspidata*], Q9AXM6	4.89*E^{-7}	cytochrome P450 monooxygenase CYP63 *[Laccaria bicolor* S238N-H82], XP_001886909	2.94*E^{-47}
contig_8062	taxoid 2-α-hydroxylase [*Taxus canadensis*], AAS89065	5.05*E^{-7}	ent-kaurenoic acid hydroxylase [*Arabidopsis thaliana*], AAK11564	2.25*E^{-9}
contig_7461	taxoid 2-α-hydroxylase [*Taxus canadensis*], AAS89065	6.44*E^{-7}	cytochrome P450-2 [*Coprinopsis cinerea* okayama7#130], XP_001835122	9.77*E^{-34}
contig_7479	taxoid 13-α hydroxylase [*Taxus cuspidata*], Q8W4T9	8.28*E^{-7}	cytochrome-450 hydroxylase [*Coprinopsis cinerea* okayama7#130], XP_001830548	1.22*E^{-66}
contig_7394	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	9.20*E^{-7}	cytochrome P450 [*Aspergillus fumigatus* A□293], XP_747185	1.09*E^{-18}
contig_7636	5-α-taxadienol-10-β-hydroxylase [*Taxus cuspidata*], Q9AXM6	1.78*E^{-6}	cytochrome P450-2 [*Coprinopsis cinerea* okayama7#130], XP_001835122	4.07*E^{-35}
contig_7570	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	1.79*E^{-6}	cytochrome P450 monooxygenase [*Laccaria bicolor* S238N-H82], XP_001881086	3.28*E^{-8}
contig_8187	taxoid 2-α-hydroxylase [*Taxus canadensis*], AAS89065	1.92*E^{-6}	cytochrome P450-2 [*Coprinopsis cinerea* okayama7#130], XP_001835122	6.17*E^{-48}
contig_780	taxoid 7-β-hydroxylase [*Taxus cuspidata*], AAQ75553	2.11*E^{-6}	ent-kaurenoic acid hydroxylase [*Arabidopsis thaliana*], AAK11564	2.50*E^{-7}

Table III-10 continued

contig_8380	5-α-taxadienol-10-β-hydroxylase [*Taxus cuspidata*], Q9AXM6	$2.75*E^{-6}$	cytochrome P450-1 [*Coprinopsis cinerea*], BAA33717	$1.61*E^{-22}$
contig_7889	taxoid 2-α-hydroxylase [*Taxus canadensis*], AAS89065	$4.00*E^{-6}$	cytochrome P450-1 [*Coprinopsis cinerea*], BAA33717	$2.56*E^{-22}$
contig_7777	taxadiene 5-α hydroxylase [*Taxus wallichiana var. chinensis*], AAU93341	$4.61*E^{-6}$	cytochrome-450 hydroxylase [*Coprinopsis cinerea* okayama7#130], XP_001830548	$2.14*E^{-27}$
contig_7607	taxoid 2-α-hydroxylase [*Taxus canadensis*], AAS89065	$6.89*E^{-6}$	cytochrome P450 alkane hydroxylase [*Aspergillus fumigatus* A1163], EDP55514	$8.97*E^{-11}$
contig_8191	taxoid 2-α-hydroxylase [*Taxus canadensis*], AAS89065	$7.87*E^{-6}$	cytochrome P450-1 [*Coprinopsis cinerea*], BAA33717	$5.96*E^{-84}$

The analysis resulted in 58 contigs in case of EF0021 and 33 contigs for *Taxomyces andreanae* showing similarity to known Taxol biosynthesis proteins from yew tree. In case of cytochrome P450 dependent monooxygenase data set the total number of enzyme candidates of this enzyme class was determined to around 90 candidates for EF0021 and more than 200 candidates for *Taxomyces andreanae*, respectively.

The contigs showing similarity to Taxol biosynthesis enzymes were taken as a query for manual comparison with the best results obtained with P450 reference proteins. This was done in order to figure out if the matches for Taxol biosynthesis enzymes were the most significant ones or if the matches identified were more related to other, especially fungal enzymes. Regarding Taxol biosynthesis protein data set all known enzymes were taken into account, including cytochrome P450 dependent monooxygenases (P450ox), acyltransferases, enzymes involved in side chain assembly and a cytochrome P450 reductase required for reduction equivalent regeneration were used. For both fungi this reductase delivered the most significant matches. With E-values of $3.57*E^{-90}$ for EF0021's contig03148 and $3.09*E^{-79}$ for *Taxomyces andreanae*'s contig_7943 the proteins identified were most likely oxido-reductases. This hypothesis was supported by the comparative analysis to the local blastx results of the other data set. Best matches for the mentioned contigs showed an E-value of 0 for cytochrome P450 oxidoreductase from *Gibberella fujikuroi* (contig03148) and cytochrome P450 oxidoreductase from *Coprinopsis cinerea* (contig_7943), respectively. In both genomes multiple matches sharing these characteristics could be identified. However, the high conservation of oxidoreductases in general and requirement of reduction equivalents, hence, reductases, in many biological processes did not allowed to take this as a hint for fungal Taxol biosynthetic pathway. Regarding matches to other sequences of plant Taxol biosynthesis best hits were found for taxane-13α-hydroxylase on contig00496 in EF0021 and on contig_7848 in *Taxomyces* genomes. With E-values of $\sim E^{-19}$ and $\sim E^{-15}$ the similarity was significant. These best P450ox matches were most similar to plant derived sequences from the cytochrome P450ox data set, to *ent*-kaurenoic acid hydroxylase and C-22 sterol desaturase from *Arabidopsis thaliana*. However, significance for these reference proteins was better than for Taxol biosynthesis and it seems quite probable that the two fungi might be capable of sterol or gibberellins biosynthesis. Besides these two hits, one in each genome, all other sequences initially identified by comparison to Taxol biosynthetic enzymes showed much lower E-values, meaning similarity to cytochrome P450 dependent hydroxylases from fungi,

like *Coprinopsis cinereus* for example contigs_7383, 7354 and 8569 from *Taxomyces,* or *Botryotinia fuckeliana* and *Aspergillus* for example contigs_00194, 0251 and 00377 from EF0021. These findings led to the conclusion that there was no clear indication for P450ox from Taxol biosynthetic pathways in the genomes. Nearly all proteins predicted were more closely related to fungal sequences in accordance to terpene synthase results. The only interesting contigs were the described ones showing similarity to plant derived proteins. Nevertheless, for both fungi only one hydroxylase was found not clearly predicted as fungal P450ox. Taxol biosynthetic pathway involves several hydroxylation steps and the respective enzymes were found to be very similar to each other (JENNEWEIN and CROTEAU 2001) more than to hydroxylases from other pathways. Although for example the best six hits for taxane-hydroxylases in EF0021 genome shown in Table III-9 had similar E-values (E^{-19}-E^{-15}) and their best matches for different cytochrome P450 enzymes from Taxol biosynthesis, comparative analysis clearly showed not only a higher similarity to fungal kingdom, but also that these best matches were derived from different species. If for example all the sequences would have been most closely related to one fungal sequence or at least to different sequences from the same host, this might have been a hint for membership of the predicted enzymes of the same pathway.

Furthermore there was no protein predicted showing significant similarity to any of the other enzymes from Taxol biosynthesis in yew.

Taking all this together the data suggested the absence of Taxol biosynthesis in the endophytes. Again the results were in accordance to negative hybridization screening observations. Thus, the possibility of a gene transfer of Taxol biosynthetic pathway can be ruled out. Nevertheless, this analysis did not completely proof that none of the here predicted cytochrome P450 proteins have the capability to hydroxylate Taxoids.

III.4.7 Additional analyses

Additionally to targeted analysis performed in order to identify Taxol biosynthetic pathway genes or candidates for a not homologous biosynthesis in fungi, both genomes were compared to each other and to >3500 sequences from an induced EST library of *Taxus cuspidata*. Neither by direct comparison nor by using BLAST analysis against *Taxus* ESTs a significant homology was found for genes that might be involved in secondary metabolism. Only highly

conserved genes from primary metabolism were identified having similarity in EF0021, *Taxomyces andreanae* and *Taxus* for example genes encoding for proteins of ubiquitin biosynthesis.

III.4.8 Examination of diterpene synthase 0021_TS_1762

Summarizing all the results from genome analysis together the only probable candidate for an enzyme involved in an independently evolved fungal Taxol biosynthesis is 0021_TS_1762. It is the only annotated diterpene synthase and phylogenically seen at least more similar to fungal diterpene synthases and terpene synthases from plants than all the other predicted proteins. To further characterize the enzyme first a cDNA-RACE library was constructed from RNA from EF0021 mycelia. Using gene specific primers according to the isolated genomic clone and the library as template it proved impossible to isolate cDNAs of the gene, indicating that the gene might not have been expressed at the time point of harvesting of the mycelia under the applied cultivation conditions. As far as the conditions were chosen for screening for taxane production and not for optimal expression of certain genes this was not unexpected.

Therefore a synthetic open reading frame for the putative diterpene synthase was designed (Table III-7) and codon-optimized for expression in *E. coli*.

This synthetic gene was cloned into pTrcHis2 vector and expressed in *E. coli.* Functional testing for terpene synthase activity *in vitro* using crude *E. coli* protein extract and ^{3}H labeled geranylgeranyl diphosphate as a substrate under terpene synthase assay conditions varying the pH from 7.2 to 8.0. Unfortunately this did not lead to a conversion of the diterpene precursor GGPP to a cyclic diterpene.

The terpene synthase activity assay was controlled using protoilludene synthase and could be excluded as a probable reason for the obtained negative results (ENGELS *et al.* 2011). The probable explanation for the failure had to be searched in the predicted enzyme sequence. For calculation of the open reading frame from the genome sequence FGENESH software was used. The determining factor for the quality of a prediction like this is the correct choice of organism used for distinguishing introns from exons.

In case of 0021_TS_1762 the best fitting organism available was *Aspergillus nidulans*. As far as no cDNA for 0021_TS_1762 could be amplified via RACE-PCR two genes from primary

metabolism were used to control the correct intron/exon structure. The pyruvate kinase and the citrate synthase were therefore predicted in the same manner, cloned from the EF0021 RACE-library and sequenced. Analysis showed that for these genes the prediction was 100 % right. Hence, the prediction using *Aspergillus nidulans* as basis in general was useful for EF0021 genes.

Taking a closer look at the specific putative sequence of 0021_TS_1762 using blastp analysis showed that the possible problem leading to a non functional protein was most likely located in the beginning of the sequence. Comparison with fusicoccadiene synthase from *Phomopsis amygdali* revealed that in 0021_TS_1762 the terpene synthase domain is lacking one of two motifs relevant for catalytic activity. The probable position of the motif thereby lied in the region of intron1 predicted initially. BLAST search of the sequence fragment around intron1 of predicted 0021_TS_1762 led to the identification of a 165 bp fragment including the 93 bp original intron1 that had no homology to fusicoccadiene synthase. Deleting this sequence from the open-reading frame as a modified intron1 leaded to a new putative terpene synthase 0021_TS_1762_del now having a DDXXE motif (Figure III-15).

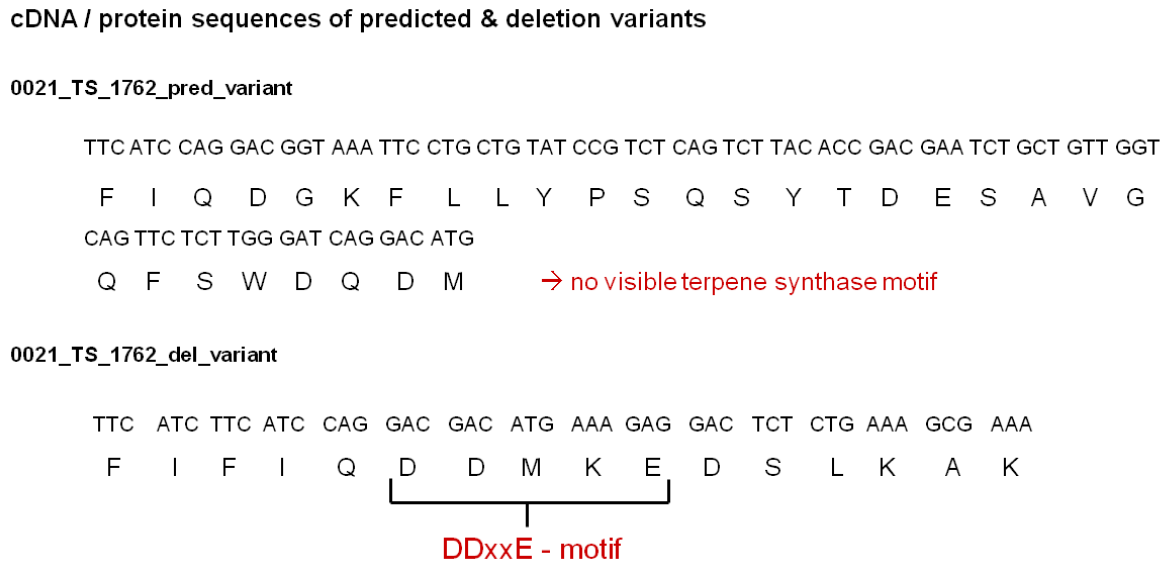


Figure III-15: cDNA and protein sequence of predicted 0021_TS_1762 and the deletion variant, in which intron1 was enhanced from 93 bp to 165 bp due to a gap in sequence homology to fusicoccadiene synthase from *Phomopsis amygdali.* The region corresponding to gene sequence around intron1 is lacking a terpene synthase motif. By deletion of the manually determined intron1 in the resulting protein a DDxxE motif occurs.

An expression clone was created by whole plasmid PCR using pTrcHIS2-0021_TS_1762 construct and relegation of the PCR product forming pTrcHIS2-0021_TS_1762_del. The new protein was expressed and tested as described before.

Although the protein was expressed as shown in Figure III-16, the new enzyme was not functional either. Both recombinant enzymes were furthermore tested using ^{14}C labeled isopentenyl diphosphate and dimethylallyl diphosphate as substrates instead of ^{3}H-GGPP, suggesting a role of prenyltransferase domain for reaction towards terpenoid backbone through a mechanism of metabolic channeling. But also by these experiments it was not possible to show terpene synthase activity.

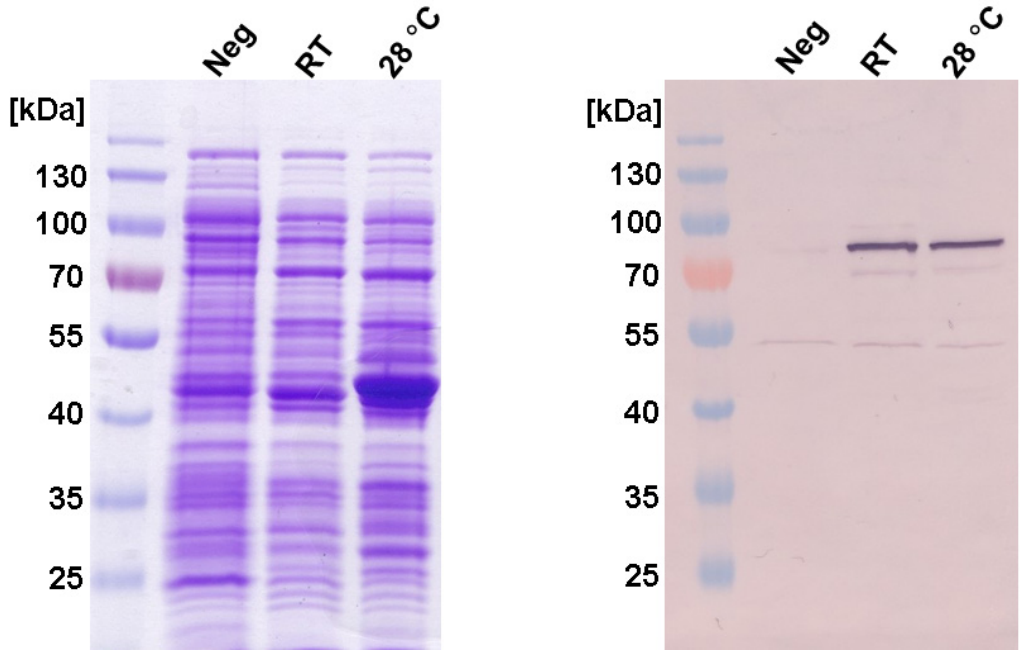


Figure III-16: SDS PAGE stained with Coumassie and western blot analysis of expression of 0021_TS_1762_del in *E. coli*; Neg: protein sample of *E. coli* not containing expression vector; **RT:** protein sample of *E. coli* expression culture of 0021_TS_1762_del performed at room temperature over night, induction at OD_{600nm}=0.5with 1 mM IPTG; **28 °C:** protein sample of *E. coli* expression culture of 0021_TS_1762_del performed at 28 °C over night, induction at OD_{600nm}=0.5with 1 mM IPTG.

Another feature of intron1 was its divisibility by three, indicating the possibility of more theoretical variants containing additional amino acids compared to 0021_TS_1762. Analysis of 0021_TS_1762 sequence showed a maximum of 56 variants (13 aa). Although all these enzymes are theoretical possible without destroying the open-reading frame none of them contained a DDXXE or DDXXD motif as already observed for original 0021_TS_1762.

It was possible to clone 35 of these variants. Protein expression was done here in 24 well plates. Test reactions as well as extractions were performed in plasma carbon coated 96 well plates (HEINIG *et al.* 2010). In accordance to the results before none of the recombinant proteins showed a terpene synthase activity neither with ^{3}H-GGPP nor with ^{14}C-IPP and DMAPP as substrates.

In conclusion, all results indicate that diterpene synthase 0021_TS_1762 represents a non functional enzyme, although it cannot completely ruled out that one of the variants not cloned is the active one or that incorrect folding due to expression in *E. coli* takes place.

Nevertheless, although it was not possible to show activity towards taxadiene or another terpenoid backbone for whatever reason taking all the results together and furthermore comparing 0021_TS_1762_del with taxadiene synthase from yew tree regarding general features the enzyme is probably not a fungal taxadiene synthase.

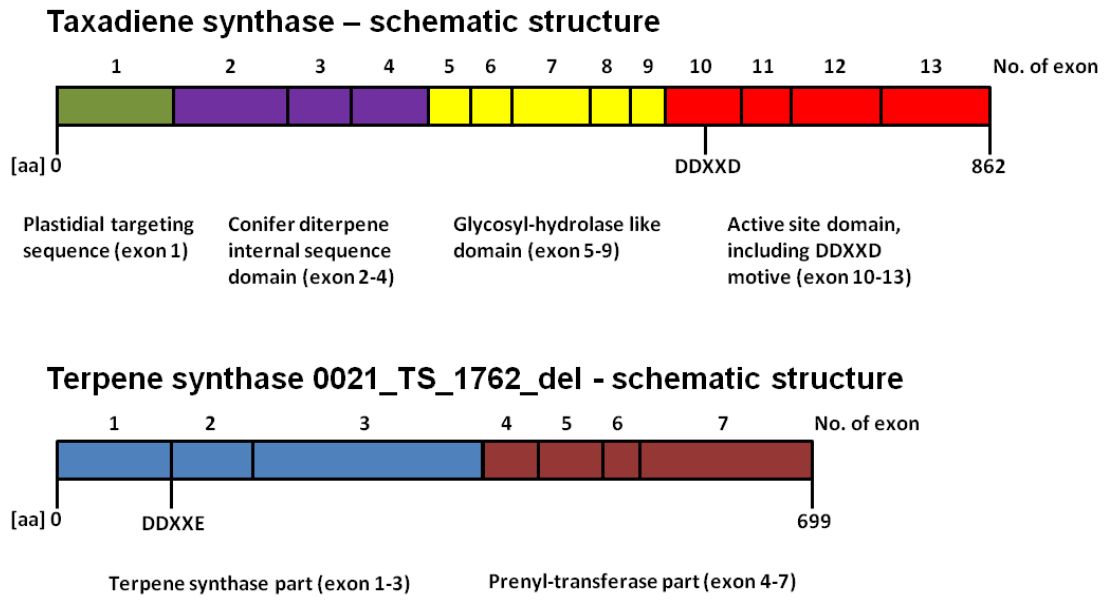


Figure III-17: Schematic protein structures of taxadiene synthase and 0021_TS_1762_del, including domains and position of catalytic DDXXD/E motif, exons as boxes, introns as vertical lines, length in amino acids.

All features of 0021_TS_1762 are typical for a fungal terpene synthase, including intron/exon structure, number of introns, size of the protein and most important the position of DDXXD/E motif within the protein. These characteristics are fundamentally different to the structure of terpene synthases from plants and in particular to taxadiene synthase from *Taxus*. Figure

III-17 shows a comparison between taxadiene synthase from *Taxus baccata* and 0021_TS_1762_del, with respect to protein architecture, pointing out the major differences between the two enzymes. Whereas the plant derived taxadiene synthase has characteristic domains known from all plant terpene synthases like targeting sequence for plastids or glycosyl-hydrolase like domain and specific sequences for its origin in this case conifers diterpene synthase 0021_TS_1762_del is lacking all these typical parts and consists only of around 300 aa containing the features relevant for synthase activity.

Furthermore there is no example for plant terpene synthases having both a terpene synthase domain and a prenyltransferase domain as 0021_TS_1762 has. These multifunctional enzymes seem to occur exclusively in fungal terpenoid biosyntheses and even here they are unusual.

Another strong indication that the protein is not a taxadiene synthase was its absence in *Taxomyces andreanae*. Even if the theory of an independent evolution of the pathway in the species is right, at least similarity between the fungal enzymes catalyzing the highly complex reaction towards taxadiene should be there.

IV Discussion

The objective of the project was the elucidation of the evolutionary origin of complex secondary metabolisms pathways. Up to today very little knowledge exists on the evolutionary origin of natural products (secondary metabolites). The diterpenoid natural product Taxol, paclitaxel, offers a particular interesting case. Taxol and related taxanes occur in plants only in the small genus *Taxus*, however are also reported to be synthesized by certain endophytic fungi, often associated with Taxol or taxoid (several closely to Taxol related natural products) producing *Taxus* species. Several genes of Taxol biosynthesis in *Taxus* plants have been already cloned and functionally identified (HEINIG and JENNEWEIN 2009), however much less information is available from the Taxol biosynthesis of endophytic fungi. Thus, the examination of the in evolutionary regard very interesting observation of the production of identical but extremely complex natural products by very distantly related organisms, in this case the yew tree *Taxus*, and its endophytic fungi may lead to fundamental new knowledge about the evolution of secondary metabolite pathways.

As mentioned above at the beginning of the project there was very little knowledge on the biosynthesis of Taxol or any other related taxanes in endophytic fungi, besides the radio-active precursor feeding experiments indicating independent biosynthesis of Taxol in the endophytic fungus *Taxomyces andreanae* (STIERLE *et al.* 1993). Fortunately, several biosynthetic steps of the Taxol biosynthesis in planta were already characterized and the underlying genes isolated. Initially the comparison for examination of the origin of the synthesis was not possible due to the in contrast to *Taxus* unknown biosynthesis in endophytes. Thus first the genes and enzymes of the pathway in endophytes had to be isolated and characterized.

IV.1 *Analysis of "taxane producing" endophytic fungi*

To confirm the results with respect to taxanes detection from endophytic fungi published earlier three endophytic fungi obtained from culture collection were analyzed phytochemically. This analysis of the published and patented strains did not led to the

detection of taxanes in the organic extracts in contrast to previous results (STROBEL *et al.* 1994; STROBEL *et al.* 2001; HOFFMAN 2003). Since no data on the proposed biosynthetic pathway in fungi was available the detection of taxanes from the organic fungal extracts was the only possible way to obtain an indication for to at least get a hint for a functional and hence complete pathway. For this reason the detection of one or more taxoids was regarded as an essential need before starting any molecular biological examination.

As already discussed in chapter III.2 one plausible explanation for the unexpected results was the scenario that the fungal species stopped synthesizing the natural products because of the cultivation under laboratory conditions which are fundamentally different to their natural environments. Loss of the ability of production in general could occur either due to silencing of the pathway or part of it, or however less probable because of loss of biosynthesis genes. As mentioned in the introduction for fungi it is known that secondary metabolites biosynthetic pathways are located mainly in the telomere regions of the genomes (WALTON 2000). Genetic instability in these regions can cause the loss of pathway genes or in case of a gene cluster the whole pathway. In general loss of pathways as a reason for loss of the ability for synthesis is described as a evolutionary scenario by analysis of genomes of organisms, normally from the same genus either shown to produce certain compounds or not able to synthesize the molecules. For different *Fusarium* species, gibberellin producers and species not synthesizing the compounds, it was shown that some parts of the biosynthesis gene cluster or the entire gene cluster was deleted (BÖMKE *et al.* 2008; PROCTOR *et al.* 2009). This kind of divergence during evolution might be due to changing requirements of the species during time, for example because of changing environmental conditions. In case of gibberellins this for example means that if a fungus does not act as a plant pathogen anymore there is no need for gibberellins biosynthesis. Hence there is no pressure for the organism to retain the ability for production and to possess the gene cluster. These considerations of course represent general evolutionary processes that can be assumed to happen over a long time period in parallel to divergence within the genus itself.

Such a phenomenon was never described in literature so far as a cause of cultivation in laboratory and it cannot be examined in the present case due to lack of information on the predicted fungal Taxol biosynthetic pathway.

Regarding the fungal species examined here classical evolutionary theories cannot be applied. Strains were cultivated under laboratory conditions which fundamentally differ from their

natural environment. At the time point of isolation all pressure for natural product synthesis for example for defense is removed as far as there is no interaction anymore with the multiple other inhabitants, meaning no stimulation neither negative or positive. The organisms might adapt to these new conditions quickly and stop all not needed biosynthetic efforts for example production of natural products, although this might not be the case for all secondary metabolite pathways.

Thus regarding the short time periods in which obviously taxanes cannot be detected anymore from the endophytes extracts it seems more likely that the biosynthetic pathway is silenced rather that the genes are completely deleted from the genome.

The example given in chapter III.2 that by using serial transfers nonaflatoxinogenic *A. flavus* strains can be created from previously aflatoxinogenic ones (CHANG *et al.* 2007) supports this theory. By only 20 serial transfers the fungi completely stop the synthesis of the compounds. The fungi obtained from culture collections analyzed in the present study were transferred more often over the years and most probably adapted completely to the standard laboratory cultivation media. Thereby it seems reasonable that natural product biosynthetic pathways still active directly after isolation are silent now.

By using of screening media also used in the past for cultivation for identification of taxanes from fungal extracts, mimicking the natural habitat in theory it should have been possible to re-induce production. But taking into account, that regulation of the predicted pathway and hence the entirety of factors needed for restoration of the assumed production was not known we could not be sure that this was possible at all. Furthermore by removing the evolutionary pressure for the predicted production also the need to retain the genes was lost. With regard to the known characteristics of Taxol biosynthetic pathway from yew, especially the specificity of the enzymes for complex substrates as well as products, it seems feasible that even minor changes of a single gene, for example through one mutation might lead to an inactive enzyme. This would most probable lead to a collapse of the entire pathway towards taxoid end-products, like Taxol or Baccatin III, without any possibility to restore the ability of production.

All these factors, known or being assumed to influence the proposed taxane production by the endophytes led to the conclusion that the influence of growth under laboratory conditions had to be minimized. This could only be achieved by isolation of new endophytes from their natural environment. These fungal species would be still adapted to the conditions there and

furthermore if taxanes would be detectable from their organic culture extracts it is possible to take this observation as a hint for an active, complete and not degenerated fungal Taxol biosynthetic pathway.

IV.2 *Isolation of endophytic fungi from Taxus species*

The isolation and characterization with regard to natural product synthesis of endophytic fungi from different plant species not only from *Taxus* species is well described in the literature (ZHOU *et al.* 2010; ALY *et al.* 2011). Although as already mentioned Taxane producing endophytes have been isolated from different plants such as hazelnut or *Cupressus* species (HOFFMAN 2003; KUMARAN *et al.* 2008b; ZHOU *et al.* 2010) (Table S1) besides the genus *Taxus,* in the work presented special attention was paid towards fungi isolated from *Taxus* in accordance to the hypothesis of a connected evolution of the biosynthetic pathways. The 26 of 34 new isolates characterized by BLAST of the 5.8 ITS rDNA sequences thereby were all ascomycetes belonging to the orders *Hypocreales, Pleosporales, Diaporthales* or *Xylariales.* The most abundant genus was *Phomopsis* of the order *Diaporthales* followed by different *Fusarium* strains. All of the orders and most of the species were already described as endophytes and in particular as *Taxus* endophytic fungi (SOCA-CHAFRE *et al.* 2011). Four of the isolates were most closely related to other fungal endophytes which were not annotated on taxonomic level in detail.

This diversity of fungi was rather low compared to other studies on endophytic fungi from plants and from *Taxus* (GANLEY *et al.* 2004; SOCA-CHAFRE *et al.* 2011). Nevertheless especially the most abundant species were not only described as endophytes but also as taxane producers before (CHAKRAVARTHI *et al.* 2008; KUMARAN and HUR 2009; KUMARAN *et al.* 2010b) (Table I-1). Since there are still many questions on the occurrence and function for example symbiosis of endophytes there are different explanations for the results. Puzzling in this context is the host specificity of the fungi for the host plant. Whereas some individual species seem to be highly specific for a host there are multiple genera that were isolated from a wide range of different and unrelated plant species. For example *Taxomyces andreanae,* the first predicted taxane producer ever was only found on a yew tree in Montana and not on any other yew species from other locations what might indicate a co-evolution of this endophyte

and yew in this special habitat. For most of the endophytes the case is less obvious. There are only tendencies of certain genera that are more common than others, for example the occurrence of *Pestalotiopsis* spp. on yew trees from subtropical locations (STROBEL *et al.* 1996; ALY *et al.* 2011), whereas endophytic *Penicillium* spp. seem to be rare in yew, although they occur in high frequencies on other plant species. Nevertheless there is a certain host specificity regarding plant families e.g. in conifers. The occurrence of dominating fungal communities in plants is correlated to the taxonomic relatedness of the hosts (SIEBER 2007) but also seems to be influenced by the geographic location and hence the environmental conditions of the plant/endophyte system. In case of the isolated fungi the predominant host species was *Taxus baccata* collected from different locations in Europe, growing under comparable climatic conditions. Other species like *Taxus media* were also collected in Europe, for example in botanical gardens. This might explain the occurrence of similar endophytic fungi isolated from these species because of either host-specificity or a preference due to the environment. The diversity of endophytes was furthermore shown to be dependent on the tissue of plant species indicating an influence of the micro-environment on the occurrence of fungal species (ALY *et al.* 2011). Comparison of endophytes from *Cupressus arizonica* leaves and woody material showed that most of endophytic species were found only in one of the plant parts and that the overlap of species was rather low (ARNOLD 2007). Furthermore the different tissues contain by it selves different numbers of endophytes (GANLEY and NEWCOMBE 2006). In the present study all endophytic fungi were isolated from surface sterilized inner-bark material by cultivation on solid agar followed by separation. This might have also led to a selection for faster growing fungi and to the loss of either over grown species or species that might not be cultivatable under the laboratory conditions applied (ARNOLD 2007). Hence, only the dominant *Taxus baccata* endophytes from one tissue of the plant, inner-bark, could be expected from the isolation procedure performed in this study. Because of these reasons the number and diversity of endophytic fungi had to be lower than shown in other studies.

However the aim of the project was the isolation of taxane producing endophytes and not the examination of the total number or whole diversity of fungal species within the collected plant material as for example shown by Soca-Chafr and co-workers (SOCA-CHAFRE *et al.* 2011).

The diversity obtained was estimated as high enough, due to the occurrence of several fungal species in which organic extracts already taxanes were detected (Table III-2, Table I-1) and

because of the number of different isolates obtained. In other studies performed for isolation of potential taxane producing endophytes in around 10 % of the fungal extracts taxanes were detected, for example 15 of 150 by Caruso and co-workers or 16 of 105 by the group of Soca-Chafre (CARUSO *et al.* 2000b; SOCA-CHAFRE *et al.* 2011). Hence, in around three extracts of the isolates taxanes should be detectable. These estimated number of potential taxane producers could be used for molecular biological examination of the proposed pathway.

IV.3 *Phytochemical examination of newly isolated endophytic fungi*

To identify endophytes able to produce Taxol or related taxanes, such as Baccatin III or 10-Deacetylbaccatin III the 34 fungal isolates were cultivated in liquid culture using either M1D, S7 or YM-6.3 medium. For detection of taxanes we used an immunological assay and LC/MS/MS. Two endophytes were identified as potential taxane producing fungi, EF0001 and EF0021.

For evaluation of the results the chosen setup for screening, including cultivation and analytics has to be critically discussed. Furthermore assuming that the detected taxanes were produced by the endophytes possible regulation mechanisms for secondary metabolite production have to be considered.

The immunoassay purchased by Cardax Pharmaceuticals (Hawaii) was aimed to be a fast and very sensitive method for detection of all taxanes in parallel from raw organic extracts. The amounts of taxanes in the two fungal giving a signal in this assay, EF0001 and EF0016, were calculated to 7.8 ng/L and 2.5 ng/L, respectively. In comparison to *Taxus baccata* extract used as a positive control these values were >10,000 fold lower but lied in the range of observed taxane concentrations shown in literature when using this assay for quantification (Table I-1). Although many already as potential taxane producers described endophytes yield also only amounts of ng/L these values had to be validated by using an extract of a tissue definitely taxane free as negative control. The results of the analysis of tobacco leaf organic extracts clearly showed a cross reactivity probably due to other natural products, most like phenolic compounds leading to a “taxane” concentration comparable to the values determined in the fungal samples. This background activity was not described in literature up to now. As negative controls only pure medium was used (SOCA-CHAFRE *et al.* 2011) or no negative

control at all was mentioned (literature Table I-1). It is not known which part of the molecule is bound specifically to the used polyclonal antibody. For *Taxus* extract which showed a 10,000 fold higher response, due to the much higher production level this background is negligible. Nevertheless these results raise questions about the reliability of the data presented in reports using only the immunological detection method for Taxol or taxane analysis from endophytic fungi's extracts (KIM *et al.* 1999; CARUSO *et al.* 2000b; METZ *et al.* 2000; MIAO *et al.* 2009a; STANIEK *et al.* 2010; SOCA-CHAFRE *et al.* 2011).

However these observations do not clearly ruled out the taxane production of the fungi found positive in the assay since the composition of the complex organic raw extracts obtained from newly isolated fungal species was not known. Thus it was necessary to control the results via a structure dependent analysis method as done in this thesis with LC/MS/MS.

As described in chapters III.1 and III.3.3 this analysis was performed using "multiple reaction monitoring" scan mode as a detection method in combination with HPLC separation.

This targeted approach of course has advantages as well as disadvantages. Whereas the immunological assay could be used for the detection of total taxanes the method developed here was selective to three defined compounds Taxol, Baccatin III and 10-Deacetyl-baccatin III. Regarding the existence of more than 350 known Taxoids (BALOGLU and KINGSTON 1999) this selection seems rather limited. But all publications up to date describing taxane detection in fungal extracts report the only Taxol and Baccatin III or 10-Deacetyl-baccatin III (Table I-1), although this of course does not mean that no other taxoid compounds are present. Furthermore the aim of the study was to clarify Taxol biosynthetic pathway in the first instance and not the discovery of biosynthetic routes towards other compounds like C14-hydroxy-taxoids. Thus it was reasonable to focus on Taxol itself and the two other late precursors of Taxol in the analysis using the highly selective triple quadrupol scan mode. The detection limits determined under optimal conditions for the pure substances were in fmol range what is corresponding to pico gram amounts of the compounds. Assuming at least ng/L culture medium yields of taxanes, as described as minimal yields in literature (Table I-1) the sensitivity was more than sufficient. For example EF0001 extract had a volume of 3 mL obtained by extraction of 10 L culture broth. According to immunological quantification, assuming that the signal was due to taxanes and not to cross reactivity would contain around 26 ng/mL taxanes. Hence, 10 µL of this extract, as injected in our experiments, would contain 260 pg taxanes. This value is still around 10 fold over the detection limits. Of course there

was still the possibility that other taxanes besides the three compounds were present in the extracts instead. These compounds were not detected by this LC/MS/MS analysis. A scan of total ions would have been necessary to achieve this. An approach like this would have been less sensitive and not selective at all. In contrast to selection of defined molecular ions in Q1 the whole spectrum of ionization products would pass this first "mass-filter". The amount of "desired ions" in this mixture would be lower compared to selection for definite ions in the same scan time. Furthermore the identification of unknown compounds out of crude extracts with a mass spectrometer like a QTrap3200 used here is very difficult due to the resolution of the machine. In contrast to "time of flight" mass spectrometers the mass/charge is determined only to the second position after the decimal point. This unit resolution does not allow the determination of compounds by calculation of molecular formula from the accurate mass. Another point to consider is the unknown ionization of these unknown taxoid compounds. Probable adducts are not known as well as the optimal ionization mode. Hence, an approach like this would more likely not lead to doubtless identification of taxoids, which are supposed to be minor secondary metabolites in the fungal extracts. The targeted analysis as performed here was very sensitive and highly selective, also by usage of three transitions instead of only one for confirmation of the compounds identity. Its advantages strongly outweighed the possible loss of other synthesized substances that might additionally have been present.

As shown in chapter III.3.3 in endophyte EF0001 extract Baccatin III was detected and in EF0021 extract a clear signal for 10-Deacetylbaccatin III occurred. In case of EF0001 these results corresponded to the initial taxane detection via immunoassay. At least part of the initially observed response observed was due to Baccatin III in the extract.

Regardless all analytical considerations compared to literature the amounts of the identified Taxoids were in the in the range of the lowest detected amounts described so far (Table I-1), although as initially discussed all values reported, obtained by HPLC-UV, are most likely over estimated due to background absorption from raw organic extracts. The reason for the taxane yields in ng/L range observed are probably the cultivation conditions which have the strongest impact on natural product formation.

Taking into account that the production levels in all so far isolated endophytic fungi are much lower than described for any *Taxus* species, there is no obvious pressure for the endophyte to produce the compound in its natural environment and even less reason to do so under laboratory conditions. As already discussed for the general occurrence of endophytes (I.3,

IV.2) the micro conditions, in this case inside the yew tree might play a major role for surviving, life cycle and the secondary metabolism of the microbes (ALY *et al.* 2011).
Hence, media composition, cultivation time and temperature play major roles for secondary metabolite yields. The media used in this study were taken either from literature as S7 and M1D (literature Table I-1), or represented a standard full medium for natural product screening of fungi. YM-6.3 was for example used for production and analysis of melleolides from different *Armillaria* spp. (ENGELS *et al.* 2011).
The two media S7 and M1D were used besides PDA liquid medium in all present studies for the cultivation of endophytes for taxane production (Table I-1). Thereby S7 was designed in order to mimic the natural conditions within the inner bark of the plant host *Taxus* (STIERLE *et al.* 1993). Cultivation times for the different described species varied from one to three weeks or were carried out in a two step procedure (SREEKANTH *et al.* 2009) at 20 to 25 °C with or without shaking. Besides these cultivation characteristics fungi were furthermore grown either in the dark completely or with defined changes between dark and light cycles (STIERLE *et al.* 1993; KUMARAN and HUR 2009). By comparing all these different methods described before, in this study the cultivation conditions were set to three weeks in the dark with constant agitation (160 rpm) at room temperature (22 °C). Regarding the diversity of endophytes isolated these conditions were probably not optimal for all the fungi. This was obviously one explanation for the detected amounts of taxanes.
Another important point in this context is the not completely known regulation of the pathway in *Taxus* as well as in the associated endophytes. For *Taxus* cell cultures Taxol biosynthesis can be influenced by addition of different plant growth regulators like indole-3-acedic acid (IAA) or 6-benzyladenine (BA) and choice of carbohydrate (KETCHUM and GIBSON 1996). Nevertheless this influence was found to be strongly dependent on the individual cell line. The best production levels were obtained by induction of *Taxus canadensis* cell cultures with jasmonate known to be involved in plant defense mechanisms regulation (MCCONN *et al.* 1997) by Ketchum and co-workers (KETCHUM *et al.* 1999). They were able to determine the kinetics of Taxol accumulation as well as the highest titers, showing the influence of age of induced culture, time of induction and time point of highest yields. Thereby it was at least possible to find a correlation between maximum yields and taxadiene synthase activity (HEZARI *et al.* 1997; KETCHUM *et al.* 1999) as far as the accumulation of Taxol was

highest inducing 7 day old cultures, where taxadiene synthase activity was found to be at its maximum.

For endophytic fungi so far no studies focused on determination of optimal growth conditions in correlation with proposed Taxol production, besides the work of Li and Tao co-fermenting *Taxus* cells with fungal endophytes in a membrane reactor proposing an influence on *Taxus* cell growth and Taxol production by endophytic fungi's metabolites like gibberellic acid (LI and TAO 2009; LI *et al.* 2009).

The great differences in detected taxane yields from ng/L to µg/L range under similar cultivation conditions, if not due to the analytical methods, might therefore be an effect of the endophytic fungal species itself, of course assuming that the compounds were really synthesized by the fungi. Through different adaption to the micro-environments isolated from and different regulatory mechanisms every fungal species might have its own preferences for growth, development and secondary metabolites production. For model systems the strong impact of environmental factors, as carbon or nitrogen source, temperature, light or pH was examined. These effects were shown to be typically transmitted through zinc-finger proteins (SHWAB and KELLER 2008) leading to positive or negative effects on natural product profile. Through standardization of cultivation procedure for taxane screening of course these differences were not examined in detail for every species. Secondary metabolite production is furthermore known to be dependent on developmental stage (CALVO *et al.* 2002) or global transcription factors like *LaeA* from *Aspergillus* spp., shown to be regulating entire gene clusters (BOK and KELLER 2004). Over-expression or deletion of the protein allowed the silencing and also the activation of formerly not expressed biosynthetic pathways leading to characterization of for example terrechinone gene cluster from *Aspergillus nidulans* (BOK *et al.* 2006).

Hence, many factors influencing natural product outcome in the natural environment and especially under laboratory conditions. Factors like expression of the proteins, global transcriptional regulation or possible stimulation of production by the host or by other endophytes as in the microhabitat the endophytes were isolated from could not be controlled and might have had a negative impact on natural product yields.

Independently from these considerations, all assuming that the endophytes are responsible for the detected taxanes the benefit of Taxane production itself for the endophytes is not obvious, even in their natural environment. The yew tree produces much higher amounts of the

compounds. For this reason fungal taxanes more likely play a role in the complex microhabitat e.g. for communication than in defending the potential microbial producers. Under laboratory conditions, cultivating a single fungus without its natural microbial partners and the host plant, there seems to be even less reason for production of trace amounts of highly complex natural products, like taxanes.

All these considerations deliver possible explanations for the experimental results observed. By changing the cultivation conditions, examining regulatory mechanisms, etc. it might have been possible to enhance the taxane yields in case of the present study. Nevertheless the low level but doubtless detection of taxanes via identification with LC/MS/MS, together with not detecting the compounds in other endophytes extracts isolated or purchased from culture collection and in the negative control could be regarded as sufficient indication that the fungi EF0001 and EF0021 were the best candidates for molecular biological examination.

Since the aim of isolation of genes from the proposed fungal Taxol biosynthesis was carried out on genomic level, production and hence transcription of genes or expression of proteins did not influenced the further molecular biological workflow.

IV.4 *Southern Blotting and phage library hybridization*

As a third screening method Southern Blot hybridization was chosen. Based on the proposed connection of evolution a certain homology of the biosynthetic genes was assumed that should lead to hybridization of probes designed from the sequences of known biosynthetic pathway genes from *Taxus* with fungal DNA. Due to the unknown degree of homology and furthermore the unknown genome sizes of the endophytic fungi, the method was first tested with genomic DNA from *Taxus baccata*. The hybridization conditions were optimized as well as the washing steps and exposure time. Examination of the newly isolated and the three organisms from culture collection led to the identification of signals for taxadiene synthase probe as well as for hydroxylase probes for EF0001 and EF0021. For none of the other fungi any signal was detected under these conditions. These results were in accordance to the phytochemical observations made by immunological and LC/MS/MS analysis of organic extracts in order to identify taxanes as described before. Nevertheless the intensity of the bands was rather low compared to the *Taxus* positive control leading to the conclusion that

there is only low homology between plant and fungal genes. With regard to the few existing reports, published during the time of the thesis on Taxol biosynthesis pathway genes from endophytes this result was rather surprising. In these works identification of genes of the predicted biosynthesis was done by amplification parts of genes of Taxol biosynthesis (from *Taxus*) from endophytes genomes. For three genes encoding for taxadiene synthase (TDS), 10-deacetylbaccatin III-10-*O*-acetyl transferase (DBAT) and C-13 phenylpropanoid side chain-CoA acyltransferase (BAPT) primers were designed on predicted conserved regions of the genes and used for PCR with genomic fungal DNA as a template (ZHANG *et al.* 2008; KUMARAN and HUR 2009; MIAO *et al.* 2009b). These approaches led to PCR products of 200 bp (DBAT), 530 bp (BAPT) and 632 bp (TDS), respectively what corresponded to the size of fragments homologous to *Taxus* cDNA. Sequence alignment of the TDS fragment amplified from *Mucor rouxianus* isolated from *Taxus chinensis* with TDS cDNA clone from *Taxus media* (U48796) showed 98 % identity of the sequences. Furthermore surprising was the observation of Staniek and co-workers who used the same primer set to amplify TDS and DBAT gene fragments from *Taxomyces andreanae* (STANIEK *et al.* 2009). The sizes of the PCR products here correspond to the sizes of the genomic clones of the *Taxus* genes, meaning that in contrast to the earlier published data the genes contain all introns present in plant genes. Besides these none consistent findings regarding gene structure with or without plant intron sequences a homology of 98 % should have led to signals of similar intensity in *Taxus* and endophytes Southern Blot analysis what was not the case. Furthermore in a previous study *Taxomyces andreanae* was subjected to genomic library screening.

Examination of 160,000 phage clones via hybridization with probes homologous to taxadiene synthase (TDS) from *Taxus* no positive clone, meaning a fungal TDS could be identified (HEINIG 2006). If the gene really was nearly identical in plant and fungus the approach should have led to positive results.

Despite these inconsistent findings leading to open questions the bands detected were defined and no signals were observed in the negative control (tobacco DNA). Additionally in both fungal extracts taxanes could be detected.

The construction of λ-phage libraries from EF0001 and EF0021 was done according to the Stratagene guidelines. Titers reached thereby were lower than expected, but nevertheless contained enough individual phage clones to be sure that the entire genome was packed. The most reasonable explanation for the observed library qualities were the incomplete digest in

combination with size selective packaging. Most probably in the DNA preparations the amount of smaller fragments was too high. Since smaller fragments can be ligated better than large ones. In contrast the phage packaging extracts are selective for large inserts. Hence all vectors containing small DNA fragments will not be packed effectively and do not occur in the final libraries. Regarding insert sizes of around 20 kb for EF0001 for example and screening of all clones of the libraries, over 300,000 for EF0001 with three different probes and 40,000 for T5H-oligonucleotide probes in case of EF0021 the probability to find a gene calculated as described in chapter III.4.2 was more than 99 % for all screening efforts.

Hybridization screening resulted in several positive clones for every probe used including TDS, T13H and T5H. Since all these clones were tested in a second screening round showing a clear enrichment of signals unspecific binding to the vector backbone could be excluded. All clones were analyzed by primer walking first, but the extensive sequencing efforts did not lead to the identification of a gene identical or similar to the hybridization probes. Whereas the sequencing from the vector backbone into the recombinant DNA inserts did not cover the entire fragments and hence there was still the possibility that the gene responsible for the hybridization signals lied in the non-sequenced region also sequencing of the complete phage vectors from two individual phage clones isolated from EF0001 genomic library, PC4 positive for T13H and PC9 positive for TDS did not led to the isolation of a gene of interest. Since hybridization results were highly reproducible these observations were unexpected and it is not easy to find an explanation for this phenomenon. Of course the hybridization and washing conditions were not carried out at maximum stringency, but as described before the method was verified by Southern Blot hybridization also with *Taxus* DNA and there was a clear difference between the samples in which bands were detected and the negative control. Furthermore the clones isolated represent only a very small percentage of the total number of examined clones. If there was a general problem with hybridization setup much more false positive signal would be expected, for example through binding to the phage DNA. The only explanation for the results is the binding of the probes to the very small parts in the inserts that showed homology (III.4.2.3), e.g. 16 bp in PC9 insert.

In summary these results did not confirm the hypothesis that genes of both Taxol biosynthetic pathways are identical as described by other groups (MIAO *et al.* 2009b; STANIEK *et al.* 2009) or are at least somehow homologous, due to a connected evolution and a possible gene transfer from one to the other species.

Thus the proposed evolutionary scenario of a connected evolution seemed to be incorrect. Although initially thought to be very improbable Taxol biosynthesis in fungi and *Taxus* spp. might have evolved convergently, as known for gibberellins pathway which is fundamental different in species from different kingdoms.

For P450 enzymes that scenario seems not too improbable since this enzyme class is not limited to secondary metabolism and hence these enzymes might have evolved through gene duplication and diversification from primary metabolism genes/enzymes separately in both species.

Especially in case of the key enzyme of Taxol biosynthesis the taxadiene synthase a scenario for convergent evolution cannot be built that easy. Although as mentioned in the introduction for gibberellins biosynthesis the enzymes responsible for *ent*-kaurene formation are not only dissimilar in sequence but also the reaction mechanism, catalyzed by two terpene synthases in plants whereas by one enzyme bifunctional enzyme in fungi is fundamental different (TUDZYNSKI and HÖLTER 1998; HEDDEN *et al.* 2002).

The results showing clearly that there is no homology between the genes of the two pathways, although throughout *Taxus* spp. for example taxadiene synthase gene is highly conserved (HEINIG 2006). Despite not being homologous on genetic level in case of taxadiene synthase the proteins have to share some characteristics. The enzyme from *Taxus* catalyzes the very complex reaction from geranylgeranyl diphosphate to taxadiene as described in detail in the introduction. Thereby the enzyme is highly specific for the substrate. Only the all-*trans* configuration is accepted. Furthermore only three isomers of taxadiene are formed differing in the position of the double bond at the C-ring of the compound (LIN *et al.* 1996; WILLIAMS *et al.* 2000a), due to elimination of a proton as final reaction step. These characteristics have to be conserved in a fungal taxadiene synthase, too.

The detection of Taxol and its two late precursors Baccatin III and 10-Deacetyl-baccatin III require the formation of the correct isomer of taxadiene to obtain the right stereochemistry in the end-products, assuming of course that as in *Taxus* the following hydroxylation steps are also highly substrate specific. This assumption is feasible because the specificity of the enzymes especially the first P450 hydroxylases represents a major difference to gibberellins biosyntheses where multiple hydroxylation reactions are catalyzed by the same enzyme, leading to the great variety of compounds through very early branching of the pathways, what is not the case in *Taxus* Taxol biosynthesis (FISCHBACH and CLARDY 2007; BÖMKE and

TUDZYNSKI 2009). One multifunctional cytochrome P450 dependent oxygenase for example catalyzes several steps in fungal gibberellins biosynthesis (TUDZYNSKI *et al.* 2002).

These known enzyme characteristics for the terpene synthase as well as for the hydroxylases involved in Taxol biosynthesis in *Taxus* led to the conclusion that despite being not homologous on genetic level, on protein level certain motives relevant for the reaction mechanism have to be conserved or in case of convergent evolution should have developed similarly. Therefore sequencing of the whole fungal genomes followed by analysis via comparison with protein sequences from Taxol biosynthesis was thought to be an alternative possibility to identify the fungal Taxol biosynthesis, even if the genes are not homologous as the results show.

IV.5 *Analysis of EF0021 and Taxomyces andreanae genomes*

In order to explore the possibility of an independent evolution of Taxol biosynthetic pathway in plants and fungi two endophytic fungi from *Taxus spp.* were selected for whole genome sequencing, isolate EF0021 and *Taxomyces andreanae*.

First EF0021 was sequenced and analyzed for identification of potential genes/enzymes that might be involved in Taxoid biosynthesis. Since none of the genes and predicted proteins identified could be clearly identified to be involved in the predicted fungal Taxol biosynthetic pathway, besides efforts for functional characterization of the only diterpene synthase 0021_TS_1762, *Taxomyces andreanae* genome was sequenced, too. Even if there is no homology between plant and fungal genes at least homology between genes of different fungi should be observable. Hence comparison of the results obtained from analysis of the genomes might provide the missing evidence.

It was thought to be very unlikely that not only the pathways towards the same compound in different kingdoms evolved convergently but that further also the genes of endophytes are fundamentally different.

With calculated sizes of 45.9 Mb and approximately 45 Mb for EF0021 and *Taxomyces andreanae* endophytes genomes lied in the expected range. Due to the high coverage it is very improbable that the information was not representing the complete genome

sequences what might call the results into question. If parts were missing it might be possible that these contained the genes of the proposed fungal Taxol biosynthesis.

Similarly to the choice of the organisms for sequencing also the strategy for sequencing analysis was highly selective and designed according to the aim of the project. The main focus thereby lied on the identification of a probable fungal taxadiene synthase which of course is essential as the key enzyme for a fungal Taxol biosynthesis. The results of genomic library screening indicated that the genes show no homology, but according to the unique features of taxadiene synthase from *Taxus* species (LIN *et al.* 1996; WILLIAMS *et al.* 2000a) especially with regard to substrate and product specificity, certain characteristics were assumed to be similar between plant and fungal enzyme.

With respect to identification of terpene synthase candidates, especially diterpene synthases from the genomes a data set for comparison was created. The choice thereby was done according to enzyme function on the one hand and the origin of the enzymes on the other hand. The set therefore contained a variety of diterpene synthases from plants, as (-)-abieta-7(8),13(14)-diene synthase from *Abies grandis* (Q38710) catalyzing reactions towards three ring terpenoid structures similarly to taxadiene synthase. Enzymes like for example casbene synthase forming fundamentally different products were not used. From plants furthermore copalyl diphosphate synthases and *ent*-kaurene synthases were included in order to have a direct comparison between enzymes involved in the formation of identical natural products known from plants as well as from fungi (Table II.9).

From the fungal kingdom, besides the enzymes involved in gibberellins biosynthesis all known diterpene synthases, as aphidicolan-16β-ol synthase from *Phoma betae* or fusicoccadiene synthase from *Phomopsis amygdali* were used. Furthermore the set contained a variety of sesquiterpene synthases from different fungi. At last also two fungal geranylgeranyl diphosphate synthases and a lanosterol synthase were included (Table II-9).

Thus in summary the data set used represented a repertory of enzyme sequences with which different information could be obtained in a single analysis. Of course all these enzymes contain the characteristic catalytic motives, like DDXXD or DDXXE leading to identification of all genes/enzymes in the genomes sharing these features. These enzymes were regarded as the total of possible terpene synthases in the respective organism. Furthermore by using enzymes derived from plants as well as from fungi the kingdom related differences between the enzymes were taken into account. Regarding the plant enzymes used these characteristics

are for example the position of the catalytical motives, plastidial leader sequences or family specific sequence domains (TRAPP and CROTEAU 2001). These characteristics do not occur or are different in fungal terpene synthases (TOYOMASU *et al.* 2007). Thus phylogenetic grouping of a predicted enzyme of the analyzed endophytes with the plant derived comparison sequences would give a hint for a possible trans-kingdom gene transfer. By using diterpene synthases as well as sesquiterpene synthases it was furthermore possible to sub-group the predicted terpene synthases according to a probable function either the formation of C15 or C20 terpenoids. The prenyltransferases and the lanosterol synthase were aimed to refine the analysis with regard to candidates that might not be secondary metabolite terpene synthases but involved in either prenylation or in primary metabolism meaning in sterol biosynthesis. The geranylgeranyl diphosphate synthases furthermore were described as "marker" genes/enzymes for gene clusters in gibberellins biosynthesis (BÖMKE and TUDZYNSKI 2009) and might therefore be useful also with respect to terpenoid gene cluster identification.
With regard to the rational selection due to the information available for terpene synthases the results obtained by this analysis surly represent the most substantive data.
The predicted terpene synthases, six from EF0021 and 20 from *Taxomyces andreanae* (Table III-7, Table III-8) were all found to be most closely related to terpene synthases from fungal origin. Two enzymes of EF0021 could furthermore clearly be identified as proteins involved in steroid biosynthesis. Of the remaining 24 enzymes 23 sequences were annotated as sesquiterpene synthases and only one, 0021_TS_1762 as diterpene synthase. Besides this initial analysis these sequences together with the terpene synthase comparison protein data set were used for construction of a phylogenetic tree in order to figure out if there are indications for either an enzyme similar to plant enzymes or with regard to the hypothesis of a convergent evolution if there are enzymes similar to each other from the endophytes. Analysis resulted in three major clades A, B and C (Figure III-14). Thereby clade B could be divided in additional 3 sub-clades.
Although so far there is no report on comparison of terpene synthases from different kingdoms, the enzyme classes in each kingdom are well examined. Plant terpene synthases thereby can be divided into seven sub-families dependent on either function or their origin (CHEN *et al.* 2011). The plant derived enzymes used in this study only represented a small number of all known synthases used recently for construction of phylogenetic tree and division of the enzymes into these sub-families. Nevertheless, the reference sequences are

grouped as described in literature (BOHLMANN *et al.* 1998; CHEN *et al.* 2011). Whereas the diterpene synthases from conifers like *Abies grandis* or *Taxus brevifolia* group together as expected, because they belong all to gymnosperm specific sub-family TPS-d, the copalyl diphosphate synthase from *Pisum sativum* belonging to TPS-c and the *ent*-kaurene synthase from *Oryza sativa* belonging to TPS-e/f were clearly separated (Figure III-14, B.I) (BOHLMANN *et al.* 1998; CHEN *et al.* 2011). Most closely related to the plant enzymes were all diterpene synthases from fungi including the fungal gibberellins biosynthesis proteins. This relationship might be best explained by mechanistic considerations. Terpene synthases from plants can be divided into class I and class II enzymes according to the reaction mechanism and products formed. Class II synthases thereby catalyze a cyclization towards copalyl diphosphate (CDP synthases) or to terpenoid end products similar to copalyl diphosphate over a bifunctional mechanism, like for example (-)-abieta-7(8),13(14)-diene synthase from *Abies grandis* (PETERS *et al.* 2000). Besides taxadiene synthase, being a typical class I synthase directly converting geranylgeranyl diphosphate to taxadiene all other diterpene synthases in the analysis were known to form products similar to abietadiene and hence to copalyl diphosphate. This is also true for all enzymes grouped in clade B.II. Fungal copalyl diphosphate/ *ent*-kaurene synthases of course form the same product as their plant pendants, although catalyzing the reaction in a bifunctional manner as described for abietadiene synthase. According to the product structure the same mechanism can be assumed for aphidicolan-16β-ol synthase from *Phoma betae* (TOYOMASU *et al.* 2004). All these enzymes belong to the clade in contrast to the mechanistically differing sesquiterpene synthases which are all in clades A and C. The third sub-clade contained the only two predicted enzymes from the sequenced endophytes in B, one prenyltransferase from *Taxomyces andreanae* and the only predicted protein annotated as diterpene synthase, 0021_TS_1762. Looking at the bootstrap support these proteins are phylogenetically less related to the other two sub-clades. Furthermore it is obvious that the reference proteins similar to the two predicted ones are prenyltransferases or fusicoccadiene synthases (TOYOMASU *et al.* 2007), representing bifunctional enzymes having a terpene synthase and a prenyltransferase domain. Hence the formation of B.III was most likely due to similarity of the prenyltransferase sequences rather than the terpene synthase part of fusicoccadiene synthase.

The remaining predicted proteins initially annotated as sesquiterpene synthases from fungi also in phylogenetic analysis were clearly separated from all diterpene synthases and enzymes with plant origin. Since for fungal terpene synthases a detailed analysis and allocation into terpene synthase sub-families, as described for plant terpene synthases was not done so far interpretation of the data was done mainly by comparison of it with the few functionally characterized reference proteins in the respective clades. From the known end products of these enzymes mechanistic features can be predicted. The differences of these mechanisms were used as an explanation for the observed grouping of the endphytes' sesquiterpene synthases.

Clade A consists of all proteins similar to Cop6 from *Coprinopsis cinereus* and furthermore contains the two trichodiene synthases included in the terpene synthase data set. Cop 6 is characterized as a α-cuprenene synthase and was also found in previous phylogenetic analysis to be most closely related to trichodiene synthases (HOHN and BEREMAND 1989; AGGER *et al.* 2009). Comparing the chemical structures of trichodiene and α-cuprenene it can be assumed that both molecules are synthesized by similar mechanisms, starting with an isomerization of the precursor molecule farnesyl diphosphate (FPP) from 2,3-*trans* to 2,3-*cis* configuration leading in the end to the observed terpenoids. According to these mechanistic considerations all predicted proteins in A are most likely terpene synthases sharing these mechanistic characteristics, especially the capability of FPP isomerization leading to similarities to the mentioned reference sequences.

Similar considerations deliver an explanation for clade C. Besides the described sesquiterpene synthases forming trichodiene like compounds many others convert FPP in its all-*trans* configuration to terpenoid products like germacrene A, aristolochene and others as protoilludene or illudanes (PROCTOR and HOHN 1993; ABRAHAM 2001; AGGER *et al.* 2009). Analysis of sesquiterpene synthases from *Coprinopsis cinereus* led to the identification of three putative germacrene A synthases (Cop1-3) and one δ cadinene synthase (Cop4) (AGGER *et al.* 2009). All predicted terpene synthases not member of A and B cluster with these enzyme sequences. Thereby six synthases predicted from genes located on one single contig in *Taxomyces andreanae* genome were found to be most closely related to Cop4 and one to Cop1. The physical linkage and the close sequence relationship show that these genes most probable represent duplication products. Although Cop4 is also supposed to catalyze a cyclization starting from *cis*-FPP the enzymes show more similarity to proteins of Cop1-3

type. Some sesquiterpene synthases do not show a definite similarity to any reference protein, what was not surprising as far as in the data set only a limited number of proteins was included. As already shown in the initial annotation using blastp the best protein matches were predicted proteins for example from *Postia placenta* and not functional characterized terpene synthases. A selection of the most important structures mentioned here is shown in Figure VI-1.

In summary the results showed that the strategy itself was highly useful for the analysis of the newly predicted proteins. By comparison with the chosen reference data set clear differences according to origin of terpene synthases on the one hand and mechanistic features on the other hand were observed. The results thereby clearly show that there is no indication for a terpene synthase from the endophytic fungi showing plant protein or taxadiene synthase characteristics, giving hint for a connection of evolution of Taxol biosynthesis in plants and fungi on protein level. This was assumed according to the unique mechanistic features of plant enzyme. An enzyme having these mechanistically relevant features would have been expected to group at least in clade B.II with other fungal diterpene synthases that were found to be most closely related to plant enzymes.

Of course fungal taxadiene synthase still could be totally different from any similarity expectations. If so at least comparison of the enzymes from the two endophytes should have led to an obvious similarity because it seems impossible that a biosynthetic pathway towards a compound like Taxol evolved independently not only in different kingdoms but also in every single endophyte. Such a similarity was not found.

The other part of the targeted analysis approach was done using reference protein data sets for identification of potential enzymes modifying the terpenoid backbone after cyclization. This sets on the one hand contained 14 known and functionally characterized enzyme sequences from *Taxus* Taxol biosynthesis and 34 sequences of cytochrome P450 dependent oxygenases from plants and fungi on the other hand.

Whereas the Taxol sequence set is self explaining, in case of cytochrome P450 dependent oxygenase enzymes a variety of proteins was selected from plants as well as from fungi that are most likely involved in secondary metabolism. Besides proteins having a known function, like *ent*-kaurenoic acid hydroxylase from *Arabidopsis thaliana* (AAK11564) (HELLIWELL *et al.* 2001) or trichothecene C-8 hydroxylase from *Fusarium sporotrichioides* (BROWN *et al.* 2003) others were only annotated as P450ox in genome projects and it is not known in

which biosynthetic pathway they are involved. For two P450ox from *Coprinopsis cinereus* a function was predicted according to the location of the genes next to terpene synthase Cop6, assuming clustering of the biosynthesis (AGGER *et al.* 2009).

Thus by this analysis in contrast to terpene synthase search the results were not limited to terpene biosynthesis and contain probably also P450ox not involved in secondary metabolism at all. Therefore as described in chapter III.4.6 the results from analysis using Taxol biosynthesis data set and P450ox data set were directly compared in order to identify the P450ox that are most similar to taxoid hydroxylases and to examine if as observed for *Taxus* P450ox multiple enzymes, in case of *Taxus* known to belong to the synthesis are more similar to each other that to other P450ox (JENNEWEIN and CROTEAU 2001).

This analysis led to the same conclusion as made for terpene synthases. There was only weak indication for a trans-kingdom gene transfer, for one predicted protein of each endophyte and no hint for a family of P450ox proteins either similar to each other in one endophyte or similar in both endophytic species, as assumed. For other fungal biosynthetic pathways it is known that P450ox having functions in the same biosynthesis are very heterogeneous. In example the five known P450ox involved in aflatoxin formation in *Aspergillus* all belong to different P450ox classes and catalyze different types of P450 mediated reactions like hydroxylation, desaturation or oxidation (PAYNE and BROWN 1998; CRESNAR and PETRIC 2011).

For taxoid hydroxylases this was not expected since all characterized P450ox characterized so far were mono hydroxylases (HEINIG and JENNEWEIN 2009). For the steps like oxidation of hydroxyl-group at C9 or the oxetane ring formations the enzymes are still unknown (I.2).

Of course the analysis only resulted in a very limited view of the whole P450ox diversity, according to the aim of the project. By further analysis using other reference proteins for example from fatty acid biosynthesis or triterpene biosynthesis or even enzymes catalyzing unusual reactions (MIZUTANI and SATO 2011) the endophytes P450ox could be analyzed in more detail. Nevertheless this could not be done in this study since the aim was the identification of a fungal Taxol biosynthetic pathway.

Summarizing all these results of genomes analyses with regard to terpene synthases and additional search for potentially terpene backbone modifying enzymes like P450ox the approach clearly showed that there is obviously no Taxol biosynthesis in endophytic fungi's genomes. Intensive analysis of terpene synthase candidates delivered only one diterpene

synthase from EF0021 genome, which was found to be clearly of fungal origin since chimeric enzymes like this are very unusual in general, were found only in fungi so far and the sequence was clearly separated from not only other plant terpene synthases but furthermore from other known fungal diterpene synthases, too. Although it cannot be ruled out completely that this enzyme is able to catalyze the reaction towards taxadiene without a functional proof the absence of an ortholog in *Taxomyces andreanae* is a strong indication for another function, for example the described one the formation of fusicoccadiene (TOYOMASU *et al.* 2007). In accordance to genomic library screening results, finding no homology to *Taxus* biosynthesis genes also alignment with whole EST library sequences from induced *Taxus* cell culture did not led to homology between any gene might be involved in Taxol biosynthesis or in secondary metabolism in general. Altogether this eliminated the initially theory of a predicted trans-kingdom gene transfer in evolution of Taxol biosynthesis. Neither on genetic level nor on protein level was any conserved sequence, due to gene transfer or conservation of mechanistically relevant protein characteristics observed. By showing furthermore that there were no similarities between predicted proteins of the two individual fungal species the only possible conclusion was that the endophytic fungi do not have the ability of Taxol biosynthesis. Hence, detection of taxanes from endophytes organic extracts must be due to another phenomenon.

IV.6 *Analysis of diterpene synthase 0021_TS_1762*

The only remaining potential fungal taxadiene synthase after all analyses was 0021_TS_1762. Although no similar enzyme was found in *Taxomyces andreanae* what led to the conclusion that the gene most probably does not encode a fungal taxadiene synthase. Despite this fact the gene and enwas further analyzed on basis of curiosity yme were analyzed in more detail. Since amplification from a cDNA RACE library was not successful, a synthetic gene was designed according to the cDNA prediction made with FGENESH software. The absence of transcript of the gene was again a hint that the enzyme might not be the fungal taxadiene synthase, but also could be due to selected harvesting time point. The cultivation conditions were not chosen with regard to transcription of certain genes but for production and detection of taxoids according to previously reports (Table I-1).

The enzyme obtained by expression in *E. coli* after confirmation by Western Blot analysis was tested by *in vitro* activity assay but neither taxadiene nor other products were detected. The reasons for an inactive protein were thought to be most probably due to an incorrect initial prediction of the gene. The splicing mechanisms of the organism found to be most closely related to *Phialocephala fortinii* (Table III-2) were of course not known. Therefore it was not sure if the initial intron/exon prediction done with *Aspergillus* mode resulted in the correct gene. In case of the basidiomycete *Armillaria* it has been shown that the usage of different prediction modes in FGENESH developed according to known splicing mechanisms from different fungi resulted in fundamentally heterogeneous output (MISIEK and HOFFMEISTER 2008).

Re-examination with regard to catalytic motives revealed that one essential motive for terpene synthase activity was missing, the DDXXD or DDXXE motive responsible for co-factor binding (WILLIAMS *et al.* 2000b; TRAPP and CROTEAU 2001; FELICETTI and CANE 2004). Since the position where the whole BLAST protein domain prediction was interrupted was in the region of intron1 this part of the gene was manually re-analyzed.

By enhancing the size of intron1 from initial 93 bp to 165 bp, due to lack of BLAST support of this fragment variant 0021_TS_1762 was constructed which now contained a DDXXE motive (Figure III-15). BLAST domain search now did not show interruption of the terpene synthase domain anymore. However, also this protein was inactive as well as the 35 further analyzed variants arose from the other specialty of intron1 which is divisible by three, resulting in 56 theoretical cDNA sequences that do not interrupt the reading frame.

Thus although the terpene synthase could be expressed all examined enzymes were active and therefore it was not possible to determine the function and product.

Reason for these results were most likely due to still incorrect prediction leading to lack or addition of amino acids somewhere in the sequence. Besides the considerations regarding the catalytic motive, which is only present in the variant 0021_TS_1762 and absent in all other predicted variants what obviously might lead to loss of activity, also other even minor changes in the sequence can lead to changes in tertiary structure of the protein. Geometry of the active site of terpene synthases is sensitive to even slight modifications that either lead to different product outcome or completely inactive enzymes (SEEMANN *et al.* 2002; FELICETTI and CANE 2004). To find the reason for this inactivation, which might be due to only one amino acid exchange, would have taken large additional experimental work. With

regard to the overall aim this was not subject to the current investigation although 0021_TS_1762 as only diterpene synthase obtained from genome analysis still is the most promising candidate for a fungal taxadiene synthase.

IV.7 *Conclusions and suggestion of an alternative reason for taxane detection from fungi*

Since the data on Taxol or taxane detection in endophytic fungi prior to the start of the work was very heterogeneous with regard to plant host species, diversity of fungal species and natural product yields there was no feasible explanation for this phenomenon. Multiple reports described the phytochemical examination of endophytic fungi and taxanes were detected in organic extracts of these. From these observations it was concluded that the fungi themselves produce the compounds and hence have a Taxol biosynthetic pathway.

In summary analysis and interpretation of all data of this study led to the conclusion that there is no indication for a Taxol biosynthetic pathway in endophytic fungi. Search for genes homologous to Taxol biosynthesis genes known from *Taxus* were unsuccessful. Hence, the possibility of a divergent evolution as discussed in the introduction could be excluded. Further the analysis of two individual endophytic fungal genomes with regard to potential genes/enzymes involved in terpene synthesis revealed that there were no candidate genes either supporting the theory of a trans-kingdom gene transfer s. Another strong indication for the now stated conclusion that fungi do not have the pathway was that there were also no similarities between the two fungal species enzymes besides generally highly conserved primary metabolism genes. Thus taking these data together there was also no evidence for a convergent evolutionary scenario as discussed initially as a second possible way of development of the pathway towards Taxol in the two different kingdoms and comparable to gibberellins biosynthesis (BÖMKE and TUDZYNSKI 2009).

Finally the only result obtained within this thesis that supports taxane production by endophytic fungi was the detection of Baccatin III and 10-Deacetyl-baccatin III from extracts of fungal cultures in accordance to literature (Table I-1).

By proofing the absence of the biosynthesis the only possible source of the compounds detected is *Taxus* and the phenomenon of identical natural products in extracts from only

distantly related species is most probably not due to an evolutionary event on molecular biological level and biosynthesis in both organisms.

Hence, the taxoid molecules detected in previous studies (Table I-1) and in this study had to be transferred from plant to fungi. Furthermore all the analyses were performed after isolation of the fungi. Hence the microbes were not associated with the host anymore, when the compounds were detected. If, as proposed now as an alternative explanation for the phenomenon, taxanes are of plant origin a mechanism for uptake and short or mid time storage of lipophilic compounds in the fungal species is required. For existence of both processes evidence can be found in literature. Due to the impact of Taxol as anti-cancer drug (WANI *et al.* 1971; SUFFNESS and WALL 1995; GOLDSPIEL 1997) the interactions of the molecule with cells and in particular with membranes were studied intensively (BALASUBRAMANIAN and STRAUBINGER 1994; SHARMA and STRAUBINGER 1994; BERNSDORFF *et al.* 1999; CROSASSO *et al.* 2000; ZHAO and FENG 2004). Of course in these cases the motivation for the research was to understand the behavior of Taxol when used in therapy. Nevertheless the data obtained by analysis of interaction with artificially created membranes can be well used to create a model for Taxol uptake and accumulation out of the natural environment, the inner bark of *Taxus* into the membranes or as discussed later into compartments of the endophytes, like vesicles. By this theory it is at least possible to create a model that explains Taxol detection in *Taxus* derived fungi. The scenario of uptake from the plant does not explain detection of taxanes from endophytes from the great variety of other plants that do not produce the compound (Table I-1). Examination of Taxol membrane interaction included several factors that might play a role in the proposed mechanism. Some of the analyses like the influence on membrane fluidity or the maximum amounts of Taxol possible to incorporate into lipid layers or liposomes are more related to the usage of Taxol as a drug. Other factors examined like stability of Taxol-membrane complexes and effects of the compound in very low concentration ranges might be indications that it is possible for fungi to uptake the compounds and also that they could accumulate the compounds up to a certain degree without a negative effect on the membrane (BALASUBRAMANIAN and STRAUBINGER 1994; SHARMA and STRAUBINGER 1994; ZHAO and FENG 2004). The stability of lipid-Taxol constructs was found to be highly dependent on the composition of the membranes. Some constructs were not stable at whereas others could be stored up to several months (SHARMA and STRAUBINGER 1994;

CROSASSO *et al.* 2000). Furthermore in low concentration ranges (10^{-5} M) the physical effects on the membrane structure were also low (ZHAO and FENG 2004). With respect to the system *Taxus*/ *Taxus* endophyte these factors might explain on the one hand that by the low amount in ng/L range the fungi are not affected negatively and on the other hand why within the observed total diversity of endophytic fungi examined from one individual *Taxus* species in only around 10 % of the organisms extracts taxanes were detected. Due to the diversity it seems probable that also membrane composition of the species was heterogeneous. Hence not all fungi might be able form lipid-Taxane complexes and to store the compounds from the point of isolation, over time of cultivation until phytochemical analysis. Besides these experimental data for Taxol interaction with artificially constructed membranes there is also at least an example for the uptake and storage of highly lipophilic compounds by fungi. It was shown that *Fusarium solani* is able to accumulate polyaromatic molecules by a passive transport (VERDIN *et al.* 2005). In addition to the degradation ability for the compounds of some strains it was furthermore shown that all examined ones were able to accumulate the molecules in intracellular lipid vesicles. It could also been shown that the presence of the lipophilic compounds did not affect the growth (VERDIN *et al.* 2005). Together with the general features discussed regarding the interaction of Taxol with non polar lipids this supports the theory of an uptake and storage of taxanes from yew tree leading to the detection of the ng/L amounts of the compounds in fungal extracts.

Altogether this theory explains some point of the phenomenon but also cannot be applied to whole heterogeneity of the problem. As mentioned above the scenario only fits to endophytic fungi from yew. For all others no taxanes can be provided by the plant host (Table I-1).

In case of the heterogeneity of fungal species observed to contain the identical natural product the theory might be the most reasonable explanation as far as it is not dependent on a certain genus but only on composition of membranes or the existence of lipophilic compartments for storage of the molecules. Furthermore the low yields in comparison to *Taxus* which led to the question about the driving force for the initially predicted production, are logical with respect to the physical scenario. Taxanes as non polar molecules will tend to stick to any also lipophilic molecules in the whole environment, hence of course also and maybe preferentially to the plant cells. Therefore only minor amounts migrate to the endophytes membranes. Since the transfer is supposed to be passive, only by diffusion the probability that the compounds are accumulated in the plant producers' membranes is higher (VERDIN *et al.* 2005). If taxane

detection is really due to this incorporation effect proposed as an alternative this might be also the reason that no taxanes were detected in extracts of the three organisms described as "taxane producers" earlier (III.2). During the time storage and cultivation in laboratory including multiple serial transfers the supposed Taxane-lipid complexes were not stable anymore and the anyway trace amounts of taxanes present after isolation were lost.

The only experimental approach at least mimicking the natural *Taxus*/*Taxus* endophyte system to a certain degree is the co-cultivation approach of plant cells and endophytic fungi in a co-bioreactor (LI and TAO 2009; LI *et al.* 2009). The authors conclude from their data that the increase in Taxol concentration is a result of Taxol production of the endophyte in combination with positive effects of other compounds of endophyte origin enhancing the Taxol yield of the plant cells. However the production of taxanes by the plant cells lead to formation of a concentration gradient. According to the newly proposed reason for taxane detection in fungal extracts based on the results of this thesis the explanation for enhanced taxane yields in this case could also be transfer of the taxoid compounds by diffusion triggered by the concentration gradient to the other compartment of the reactor. Due to the limited solubility of the compounds in aqueous solutions like the medium they will accumulate than most likely in the endophytic fungus' membranes or vesicles.

A potential future perspective would therefore be the re-examination of the fungi with regard to this newly proposed scenario. This could be done using for example the strains described to accumulate polyaromatic compounds in vesicle as described before (VERDIN *et al.* 2005). By cultivation on *Taxus* biomaterial, recovery of the fungi and analysis it should be possible to find out if not only Polyaromatic compounds accumulate in the fungal vesicles, but also other lipophilic compounds, in this special case taxanes. Also the still unsolved question, observation of taxanes in extracts of endophytes isolated from other plants than *Taxus* might need some further investigations. Since for these fungi there is even less indication for a reason for production on the one hand and an acquirement of the biosynthesis from the plant producer of the natural products *Taxus* the only possibility would be the examination of a convergent evolution by whole genome sequencing as described for the *Taxus* endophytes within this study.

In summary by the work presented here it was possible to eliminate several predictions for the observation of identical natural products the taxanes in yew and a great variety of endophytic fungi. All initially made assumptions and scenarios regarding evolution of a fungal Taxol

biosynthetic pathway either divergent by a trans-kingdom gene transfer or convergent in the different species were examined and could be excluded. The data shows that there most probably is no fungal Taxol biosynthesis and hence the conclusion made according to the detection of the natural product in organic extracts of the cultures that these microbes produce the detected compounds was not confirmed. With having no biosynthesis and hence cannot produce taxanes in the end the here proposed alternative scenario of an uptake and accumulation of the compounds by the endophytic fungi seems to be the only and most probable explanation.

After 20 years of controversial discussion about the phenomenon and hundreds of publications about Taxol detection from endophytes, finally the answer might be a physical mechanism and not a complex evolutionary event.

V Summary

Focus of the present PhD thesis was the examination of the interesting and only poorly understood phenomenon of detection of identical structurally elaborate natural products in only distantly related organisms with regard to secondary metabolite pathway evolution.

Due to the impact of Taxol as one of the most important anti-cancer drugs today intensive research efforts were made with respect to development and improvement of production systems in order to overcome the supply problems. During the search for alternative production systems in the early 1990s for the first time trace amounts of Taxol were found in the endophytic fungus *Taxomyces andreanae* isolated from *Taxus brevifolia*. Besides the potential impact as an alternative microbial production system for Taxol this was especially interesting with respect to evolution of the thereby predicted biosynthesis towards identical natural products in organisms belonging different kingdoms in general.

Since the isolation of *Taxomyces andreanae* many more endophytic species were isolated from *Taxus* spp. but also from other plants where taxanes were detected in their organic extracts. Thereby the plant species described as hosts of potential "taxane producing" fungi did not follow a certain pattern. Furthermore the variety of endophytic species isolated was found to be highly diverse. Though the Taxol biosynthetic pathway was well examined in yew, no information was available on genes or enzymes involved in the predicted fungal Taxol biosynthesis.

In order to find answers to these questions arising through heterogeneity of the data available at the time of project start the work presented here was based on considerations about imaginable evolutionary scenarios of the pathway in the different species either divergent or convergent. Thereby a scenario of divergent evolution is based on the hypothesis that a complex biosynthetic pathway is assumed to consist of at least 19 enzymatic steps most likely did not evolve completely independently, supporting a convergent evolution in species from different kingdoms. Thus, assuming that a trans-kingdom horizontal gene transfer was involved, the genes and enzymes were predicted to be similar.

In this thesis both predicted scenarios were examined. Since there was no information on the potential fungal Taxol biosynthesis three fungal species obtained from culture collection were analyzed as described in literature in order to detect taxoid molecules. Since surprisingly no taxanes were detected in these extracts further 34 new endophytes were isolated from

Taxus spp. and phytochemically examined. Thereby taxanes were detected in two fungal extracts. These fungi named EF0001 and EF0021 were chosen for molecular biological experiments. To investigate a divergent evolution first more than 300,000 clones of genomic λ–phage libraries were screened via hybridization using probes homologous to genes of Taxol biosynthetic pathway from yew. By this approach several potential positive clones were obtained. However none of the recombinant fungal DNA inserts showed any homology to genes of Taxol biosynthesis from yew. This led to the conclusion that the predicted fungal pathway, if present at all, evolved by convergent evolution. To figure out if this scenario was right two endophytic fungi's genomes were sequenced, the genome of the newly isolated fungus EF0021 and that of *Taxomyces andreanae*, the first described fungus in which extracts taxanes were detected. Analysis of the sequences was done by a "targeted approach" for identification of all candidate genes and hence prediction of proteins that might be involved in the proposed taxoid synthesis. This led to the prediction of 24 fungal terpene synthases. Using phylogenetic analysis two main results were obtained. First a horizontal gene transfer of the pathway from yew to fungus or in the opposite direction could be excluded in accordance to hybridization results. Second phylogenetic analysis of terpene synthases, analysis of other possible enzyme candidates like cytochrome P450 oxygenases and the alignment of both genomes to each other did not reveal an indication for a fungal Taxol biosynthesis. Only one diterpene synthase was predicted from EF0021 genome, but had no ortholog in *Taxomyces andreanae*. Thus also a convergent scenario of pathway evolution independently in each kingdom could be eliminated. All results obtained in this thesis taken together led to the conclusion that there is no fungal Taxol biosynthesis. Data presented thereby represent the first detailed molecular biological investigation of this phenomenon.

Due to these findings, contradicting the initial hypothesis which was based on the assumption of the existence of a Taxol biosynthesis in endophytic fungi as well as in *Taxus*, an alternative scenario was proposed based on a transfer of the natural product from yew tree, the producer, to the fungus. This scenario is supported by known mechanisms for Taxol membrane interaction as well as the ability of fungi to store lipophilic compounds. Many observations in context of taxane detection from fungi can be explained nicely by this theory and the molecular biological results of this thesis doubtlessly exclude any evolutionary reason for the phytochemical results in the past but also during this study. Therefore it is probable for this thesis to represent the final answer to a controversy hold up for the last 20 years.

Zusammenfassung in deutscher Sprache

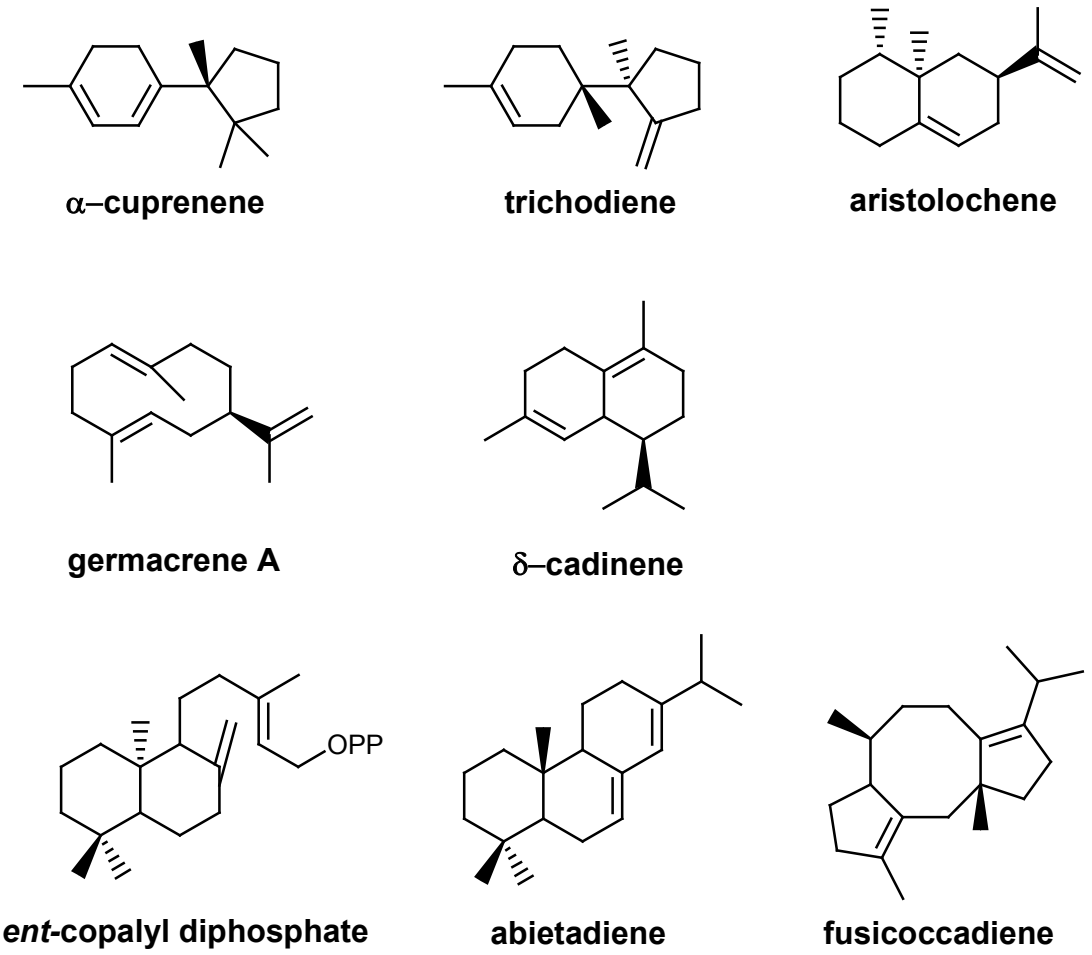
Thema dieser Doktorarbeit waren Untersuchungen des interessanten und bisher nahezu nicht verstandenen Phänomens des Nachweises von identischen strukturell hochkomplexen Naturstoffen in Organismen die nur sehr weitläufig miteinander verwandt sind im Hinblick auf die Evolution von Sekundärstoffbiosynthesewegen.

Taxol zählt heutzutage zu den bedeutendsten Krebsmedikamenten. Um den daraus resultierenden hohen Bedarf zu decken, wurden vor allem intensive Anstrengungen hinsichtlich der Entwicklung und Verbesserung von neuen und bereits bestehenden Produktionssystemen unternommen. Bei der Suche nach alternativen Produzenten in den frühen 90er Jahren wurde der endophytische Pilz *Taxomyces andreanae* aus der pazifischen Eibe *Taxus brevifolia* isoliert. Die phytochemische Analyse des Naturstoffextrakts dieses Pilzes führte zur ersten Detektion von Taxol in einem mikrobiellen Extrakt. Neben der Bedeutung dieser Entdeckung als mögliches alternatives mikrobielles Produktionssystem für Taxol, ergaben sich vor allem Fragen hinsichtlich des Vorkommens und des Ursprungs einer anhand der Analyseergebnisse vorhergesagten Biosynthese hin zu identischen Naturstoffen wie im bekannten Produzent der Eibe. Seit dieser ersten Entdeckung wurden bei zahlreichen weiteren endophytischen Pilzen Taxol nachgewiesen. Dabei beschränkte sich die Herkunft der Organismen nicht auf *Taxus* Spezies. Auch von verschiedenen anderen Pflanzen konnten derartige Pilze isoliert werden. Der Biosyntheseweg ist in *Taxus* gut untersucht, wohingegen nichts über Gene oder Enzyme bekannt war, die am vorhergesagten pilzlichen Weg beteiligt sind. In Anbetracht der sich stark unterscheidenden publizierten Daten zu Anfang des Projekts, wurde eine Strategie basierend auf evolutiven Überlegungen gewählt. Die Evolution des Biosynthesewegs kann entweder divergent oder konvergent abgelaufen sein. Legt man ein divergentes Szenario zugrunde, basierend auf der Überlegung dass sich ein Biosyntheseweg der ca. 19 enzymatische Schritte umfasst wahrscheinlich nicht konvergent entwickelt hat, muss ein horizontaler Gentransfer stattgefunden haben und demzufolge kann angenommen werden, dass eine gewisse Ähnlichkeit zwischen Genen und Enzymen in den verschiedenen Spezies besteht. Beide möglichen Evolutionsszenarien wurden im Rahmen dieser Arbeit untersucht. Aufgrund der nicht vorhandenen Informationen zur pilzlichen Biosynthese wurden zunächst drei in der Literatur als potentielle Taxol Produzenten beschriebene endophytische Pilze phytochemisch analysiert. In diesen Analysen konnten jedoch keine Taxane nachgewiesen werden, woraufhin weitere 34 neu aus *Taxus* Spezies isolierte Pilze

analog auf das Vorkommen von Taxanen hin untersucht wurden. Die führte zur Detektion der von taxoiden Verbindungen in zwei Pilzextrakten, von EF0021 und EF0001. Diese wurden für die molekularbiologischen Experimente ausgewählt. Dabei wurden zunächst, ausgehend von der Theorie eines möglichen Gentransfers >300.000 Klone aus genomischen λ–Phagen Bibliotheken mittels Hybridisierung mit Sonden homolog zu Taxol-Biosynthese Genen aus *Taxus* durchmustert. Da bei diesem Ansatz kein Gen oder eine Homologie zu den Pflanzensequenzen identifiziert werden konnte, lag die Vermutung nahe, dass möglicherweise obwohl zunächst anders angenommen eine konvergente Entwicklung stattgefunden hat. Um diese Hypothese zu überprüfen wurden zwei verschiedene Pilzgenome sequenziert, das von EF0021 und von *Taxomyces andreanae.* Die Analyse der Genome wurde zielgerichtet im Hinblick auf die Identifizierung von Genen und Enzymen die an Terpen Biosynthesen beteiligt sein könnten durchgeführt. Im Zuge dessen wurden 24 Terpen synthasen identifiziert, von denen allerdings nur ein Enzym vorhergesagt wurde das hohe Ähnlichkeit zu einer Diterpen Synthase aufwies. Durch phylogenetische Analyse dieser Proteine im Vergleich mit einer repräsentativen Auswahl von bekannten Terpen Synthasen aus Pflanzen und Pilzen konnte dabei gezeigt werden, dass zum einen keines der Enzyme Characteristika aufwies die auf eine Herkunft aus dem Pflanzenreich schließen ließ und zum anderen waren auch keine Gemeinsamkeiten zwischen den Enzymen der beiden Pilz Spezies ersichtlich. In Kombination mit vergleichbaren Resultaten weiterer Analysen, beispielsweise der Suche nach Cytochrom P450 Hydroxylasen wurde letztlich die Schlussfolgerung gezogen, dass kein Taxol Biosyntheseweg in endophytischen Pilzen vorliegt und die Organismen somit auch nicht in der Lage sind eigenständig die Verbindungen zu synthetisieren. Diesem Schluss folgend blieb als einzige Erklärung für die Herkunft der detektierten Verbindungen die Eibe. Da anhand der molekularbiologischen Daten in dieser Arbeit ein biologischer Zusammenhang also ein Biosyntheseweg in beiden nur weitläufig miteinander verwandten Organismen ausgeschlossen werden konnte, bleibt als Alternative nur eine Aufnahme und Akkumulation der taxoiden Verbindungen von den mit dem pflanzlichen Wirt assoziierten Pilzen. Da diese Theorie zum einen von Daten bezüglich Taxol Interaktion mit Membranen oder Einlagerung von lipophilen Molekülen in pilzlichen Vesikeln unterstützt wird und weiterhin zumindest ein plausibel das Vorkommen von Taxanen in *Taxus* Endophyten erklärt, liefert dieser physikalische Prozess möglicherweise die Antwort auf die Fragen dieses Phänomen betreffend das seit 20 Jahren überaus kontrovers diskutiert wird.

VI Appendix

VI.1 *Supplementary structures*



2,3-trans-farnesyl diphosphate *2,3-cis*-farnesyl diphosphate

Figure VI-1: Chemical structures of molecules mentioned in chapter IV.5, including several products of sesquiterpene synthases as well as of diterpene synthases and the structures of *trans*- and *cis*-farnesyldiphosphate

VI.2 *List of abbreviations*

Abbreviation	Description
[α-^{32}P]dATP	Desoxy-adenosin triphosphate labeled with radioactive phosphor isotope at alpha phosphate
[γ-^{32}P]dATP	Desoxy-adenosin triphosphate labeled with radioactive phosphor isotope at gamma phosphate
Δ	deletion
µg	micro gram
µL	micro liter
10-DAB	10-deacetylbaccatin III
2YT medium	2x yeast extract trypton medium
A	adenine
α	alpha
A. flavus	*Aspergillus flavus*
aa	amino acid
ACN	acetonitrile
AIDS	acquired immunodeficiency syndrome
AP	alkaline phospatase buffer
β	beta
B/B_0	corrected absorption at 405 nm
B_0	measured absorption at 405 nm
BA	6-benzyladenine
bapt	C-13 phenylpropanoid side chain-CoA acyltransferase
BLAST	Basic Local Alignment Search Tool
bp	base pair
BSA	bovine serum albumin
C	cyrosine
C. helicacabum	*Cardiospermum halicacabum*
CAD	collision-activated dissociation
CD	compact disk
cDNA	complementary *DNA*
CE	collision energy
CIEIA	competitive inhibition enzyme immuno assay
cit-primer	citrate synthase primer
cm	centimeter
Cop	sesquiterpene synthases from *Coprinopsis cinereus*
cps	counts per second
CPS	*ent*-copalyl diphosphate synthase
CTAB	cetyltrimethylammonium bromide
CXP	collision cell exit potential.

d	day
δ	delta
dbat	10-deacetylbaccatin III-10-*O*-acetyl transferase
dd	double distillation
DDT	dichlorodiphenyltrichloroethane
DEPC	diethylpyrocarbonate
DMAPP	dimethylallyl diphosphat
DNA	deoxyribonucleic acid
DP	declustering potential
DXP	1-deoxy-D-xylulose-5-phosphate
E. coli	*Escherichia coli*
e.g.	*exempli gratia*
EDTA	ethylenediaminetetraacetic acid
EF	endophyte fungus
ELISA	enzyme-linked immunosorbent assay
EP	entrance potential
ESI	electrospray ionization
EST	expressed sequence tag
et.al.	*et alii*
E-Value	expected value
FCDS	fusicoccadine synthase
FDA	Food and Drug Administration
fmol	femtomol
FPP	farnesyl diphosphate
g	gram
G	guanine
GA	gibberellic acids
GGPP	geranylgeranyl diphosphate
GGPPS	geranylgeranyl diphosphate synthase
GmbH	Gesellschaft mit beschränkter Haftung - company with limited liability
h	hour
HIS-tag	poly histidine-tag
HPLC	high performance liquid chromatography
IAA	indole-3-acedic acid
IME	Institute for Molecular Biology and Applied Ecology
Inc.	Incorporation
IPP	isopentenyl diphosphate
IPS	impulse per second
IPTG	isopropyl β-D-1-thiogalactopyranoside
IR	infrared spectroscopy
IST	internal transcribed spacer
kb	kilo base pair
K_m	Michaelis constant
KS	*ent*-kauren synthase

kV	kilo volt
L	liter
λ	lamda
LB-medium	Luria Bertani-Medium
LC/MS/MS	liquid chromatography coupled with mass spectrometry
Lib	library
M	molar
M. speciosus	*Maguireothamnus speciosus*
m/z	mass-to-charge ratio
MA	Massachusetts
Mb	mega base pairs
MCS	multiple cloning site
mg	milligram
MIC	minimal inhibitory concentration assay
min	minute
mL	milliliter
mM	milli molar
mm	millimeter
MRM	multiple reaction monitoring
MS	mass spectrometry
msec	milliseconds
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
N. tabacum	*Nicotiana tabacum*
NBT/BCIP	nitro-blue tetrazolium chloride/5-bromo-4-chloro-3'-indolyphosphate
NCBI	National Center for Biotechnology Information
NCI	National Cancer Institute
nd	not determined
Neg.	negative
ng	nano gram
nm	nano meter
NMR	nuclear magnetic resonance spectroscopy
No.	number
NZY	NZ amine yeast extrct medium
OD	optical density
orf	open reading frame
P450ox	cytochrome P450 dependent monooxygenases
PBS	phosphate buffered saline
PBS-T	phosphate buffered saline with tween
PC	positive clone
PCR	polymerase chain reaction
PDA	potato dextrose agar
pfu/ml	plaque forming unit per milliliter
pH	negative decimal logarithm of the hydrogen ion activity in a solution
PhD	*philosophiae doctor*

Pki-primer	pyruvate kinase primer
PKS	polyketide synthases
PNK Buffer	polynucleotide kinase reaction buffer
psi	pound per square inch
QX	quadrupole X
RACE	rapid amplification of cDNA-ends
radio-TLC	radioactive thin layer chromatography
RNA	ribonucleic acid
rpm	rounds per minute
RT	room temperatur
S.cerevisiae	*Saccharomyces cerevisiae*
SDS	sodium dodecyl sulfate
sec	seconds
SOC	super optical broth with catabolic repressor
sp.	specie
spp.	species
SSC	saline-sodium citrate buffer
ssDNA	single strange DNA
Su–	nonsuppressing
T	thymine
T. baccata	*Taxus baccata*
T. celebica	*Taxus celebica*
T. chinensis	*Taxus chinensis*
T. globosa	*Taxus globosa*
T. wallichiana	*Taxus wallichiana*
T. baccata	*Taxus baccata*
T. brevifolia	*Taxus brevifolia*
T. canadensis	*Taxus canadensis*
T. canadensis	*Taxus canadensis*
T. cuspidata	*Taxus cuspidata*
T. hicksii	*Taxus hicksii*
T. mairei	*Taxus mairei*
T. x media	*Taxus x media*
T. yunnanensis	*Taxus yunnanensis*
T13H	taxane-13a-hydroxylase
T5H	taxane-5a-hydroxylase
TBE buffer	tris/borate/EDTA buffer
TBS	tris-buffered saline
TDS	taxadiene synthase
TE buffer	tris EDTA buffer
TEMED	Tetramethylethylenediamine
TES	tris EDTA sodium cloride
Tet^r	tetracycline resistant
Tm	melting temperture

TPS	terpene synthase
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol
U	units
U	uracil
UPGMA	unweighted pair group method with arithmetic means
USA	United States of America
USDA	United States Department of Agriculture
UV	ultraviolet
V	volt
v/v	volume per volume
w/v	weight per volume
WT	wild typ
*x*g	gravitational acceleration
YM	yeast and malt

VI.3 *List of sequence data (CD)*

- Genome data of EF0021 (multi fasta file of all contigs)
- Genome data of *Taxomyces andreanae* (multi fasta file of all contigs)
- Predicted protein sequences of terpene synthases from EF0021 and *T. andreanae* (.pro files)
- Insert sequences of PC4 and PC9 phage clones from EF0001
- Consensus sequences of phage library clones summarized in Table III-5

VI.4 *List of figures*

VI.5 *List of tables*

VIIReferences

ABRAHAM, W. R. (2001): Bioactive sesquiterpenes produced by fungi: are they useful for humans as well? *Curr Med Chem*, **8**(6): 583-606.

AGGER, S., LOPEZ-GALLEGO, F. and SCHMIDT-DANNERT, C. (2009): Diversity of sesquiterpene synthases in the basidiomycete Coprinus cinereus. *Mol Microbiol*, **72**(5): 1181-1195.

AJIKUMAR, P. K., XIAO, W. H., TYO, K. E. J., WANG, Y., SIMEON, F., LEONARD, E., MUCHA, O., PHON, T. H., PFEIFER, B. and STEPHANOPOULOS, G. (2010): Isoprenoid Pathway Optimization for Taxol Precursor Overproduction in Escherichia coli. *Science*, **330**(6000): 70-74.

ALY, A. H., DEBBAB, A. and PROKSCH, P. (2011): Fungal endophytes: unique plant inhabitants with great promises. *Appl Microbiol Biotechnol*, **90**(6): 1829-1845.

ANTEROLA, A., SHANLE, E., PERROUD, P. F. and QUATRANO, R. (2009): Production of taxa-4(5),11(12)-diene by transgenic Physcomitrella patens. *Transgenic Research*, **18**(4): 655-660.

ARNOLD, A. E. (2007): Understanding the diversity of foliar endophytic fungi: progress, challenges, and frontiers. *Fungal Biology Reviews*, **21**(2-3): 51-66.

BACON, C. W., WHITE, J. F. and (EDS) (2000): Microbial Endophytes. New York, Marcel Dekker Inc.

BALASUBRAMANIAN, S. V. and STRAUBINGER, R. M. (1994): Taxol-lipid interactions: taxol-dependent effects on the physical properties of model membranes. *Biochemistry*, **33**(30): 8941-8947.

BALOGLU, E. and KINGSTON, D. G. (1999): The taxane diterpenoids. *J Nat Prod*, **62**(10): 1448-1472.

BERNSDORFF, C., RESZKA, R. and WINTER, R. (1999): Interaction of the anticancer agent Taxol (paclitaxel) with phospholipid bilayers. *J Biomed Mater Res*, **46**(2): 141-149.

BESUMBES, O., SAURET-GUETO, S., PHILLIPS, M. A., IMPERIAL, S., RODRIGUEZ-CONCEPCION, M. and BORONAT, A. (2004): Metabolic engineering of isoprenoid biosynthesis in Arabidopsis for the production of taxadiene, the first committed precursor of Taxol. *Biotechnology and Bioengineering*, **88**(2): 168-175.

BLANKENSHIP, J. D., SPIERING, M. J., WILKINSON, H. H., FANNIN, F. F., BUSH, L. P. and SCHARDL, C. L. (2001): Production of loline alkaloids by the grass endophyte, Neotyphodium uncinatum, in defined media. *Phytochemistry*, **58**(3): 395-401.

BOHLMANN, J., MEYER-GAUEN, G. and CROTEAU, R. (1998): Plant terpenoid synthases: molecular biology and phylogenetic analysis. *Proc Natl Acad Sci U S A*, **95**(8): 4126-4133.

BOK, J. W., HOFFMEISTER, D., MAGGIO-HALL, L. A., MURILLO, R., GLASNER, J. D. and KELLER, N. P. (2006): Genomic mining for *Aspergillus* natural products. *Chem Biol*, **13**(1): 31-37.

BOK, J. W. and KELLER, N. P. (2004): LaeA, a regulator of secondary metabolism in *Aspergillus spp. Eukaryot Cell*, **3**(2): 527-535.

BÖMKE, C., ROJAS, M. C., HEDDEN, P. and TUDZYNSKI, B. (2008): Loss of gibberellin production in *Fusarium verticillioides* (Gibberella fujikuroi MP-A) is due to a deletion in the gibberellic acid gene cluster. *Appl Environ Microbiol*, **74**(24): 7790-7801.

BÖMKE, C. and TUDZYNSKI, B. (2009): Diversity, regulation, and evolution of the gibberellin biosynthetic pathway in fungi compared to plants and bacteria. *Phytochemistry*, **70**(15-16): 1876-1893.

BROWN, D. W., PROCTOR, R. H., DYER, R. B. and PLATTNER, R. D. (2003): Characterization of a fusarium 2-gene cluster involved in trichothecene C-8 modification. *J Agric Food Chem*, **51**(27): 7936-7944.

CALVO, A. M., WILSON, R. A., BOK, J. W. and KELLER, N. P. (2002): Relationship between secondary metabolism and fungal development. *Microbiol Mol Biol Rev*, **66**(3): 447-459, table of contents.

CARUSO, M., COLOMBO, A. L., CRESPI-PERELLINO, N., FEDELI, L., MALYSZKO, J., PAVESI, A., QUARONI, S., SARACCHI, M. and VENTRELLA, G. (2000a): Studies on a strain of *Kitasatospora sp*. paclitaxel producer. *Ann. Microbiol.*, **50**(2): 89-102.

CARUSO, M., COLOMBO, A. L., FEDELI, L., PAVESI, A., QUARONI, S., SARACCHI, M. and VENTRELLA, G. (2000b): Isolation of endophytic fungi and actinomycetes taxane producers. *Annals of Microbiology*, **50**(1): 3-13.

CHAKRAVARTHI, B. V., DAS, P., SURENDRANATH, K., KARANDE, A. A. and JAYABASKARAN, C. (2008): Production of paclitaxel by *Fusarium solani* isolated from *Taxus celebica*. *J Biosci*, **33**(2): 259-267.

CHANG, M. C. Y. and KEASLING, J. D. (2006): Production of isoprenoid pharmaceuticals by engineered microbes. *Nature Chemical Biology*, **2**(12): 674-681.

CHANG, P. K., WILKINSON, J. R., HORN, B. W., YU, J., BHATNAGAR, D. and CLEVELAND, T. E. (2007): Genes differentially expressed by *Aspergillus flavus* strains after loss of aflatoxin production by serial transfers. *Appl Microbiol Biotechnol*, **77**(4): 917-925.

CHAU, M. and CROTEAU, R. (2004): Molecular cloning and characterization of a cytochrome P450 taxoid 2a-hydroxylase involved in Taxol biosynthesis. *Archives of Biochemistry and Biophysics*, **427**(1): 48-57.

CHAU, M., JENNEWEIN, S., WALKER, K. and CROTEAU, R. (2004a): Taxol biosynthesis: Molecular cloning and characterization of a cytochrome P450 taxoid 7 beta-hydroxylase. *Chem Biol*, **11**(5): 663-672.

CHAU, M., WALKER, K., LONG, R. and CROTEAU, R. (2004b): Regioselectivity of taxoid-O-acetyltransferases: heterologous expression and characterization of a new taxadien-5alpha-ol-O-acetyltransferase. *Archives of Biochemistry and Biophysics*, **430**(2): 237-246.

CHEN, F., THOLL, D., BOHLMANN, J. and PICHERSKY, E. (2011): The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant J*, **66**(1): 212-229.

CHI, Y., ZHAO, D.-L. and ZHOU, D.-P. (2008): Identification of taxol biosynthesis stage-enriched transcripts in Nodulisporium sylviforme, using suppression subtractive hybridization. *World J. Microbiol. Biotechnol.*, **24**(11): 2601-2605.

CHU, H. Y., WEGEL, E. and OSBOURN, A. (2011): From hormones to secondary metabolism: the emergence of metabolic gene clusters in plants. *Plant J*, **66**(1): 66-79.

COATES, R. M., JIN, Y. H., WILLIAMS, D. C. and CROTEAU, R. (2005): Taxadiene synthase-catalyzed cyclization of 6-fluorogeranylgeranyl diphosphate to 7-fluoroverticillenes. *Journal of the American Chemical Society*, **127**(21): 7834-7842.

CRAGG, G. M., SCHEPARTZ, S. A., SUFFNESS, M. and GREVER, M. R. (1993): The Taxol Supply Crisis - New Nci Policies for Handling the Large-Scale Production of Novel Natural Product Anticancer and Anti-Hiv Agents. *Journal of Natural Products*, **56**(10): 1657-1668.

CRESNAR, B. and PETRIC, S. (2011): Cytochrome P450 enzymes in the fungal kingdom. *Biochim Biophys Acta*, **1814**(1): 29-35.

CROOM, E. M. (1995): Taxus for taxol and taxoids. *Taxol science and applications*: 37-70.

CROSASSO, P., CERUTI, M., BRUSA, P., ARPICCO, S., DOSIO, F. and CATTEL, L. (2000): Preparation, characterization and properties of sterically stabilized paclitaxel-containing liposomes. *J Control Release*, **63**(1-2): 19-30.

CROTEAU, R., KETCHUM, R., LONG, R., KASPERA, R. and WILDUNG, M. (2006): Taxol Biosynthesis and Molecular Genetics. *Phytochemistry Reviews*, **5**(1): 75-97.

CROTEAU, R., WALKER, K., KETCHUM, R. E. B., HEZARI, M., GATFIELD, D., GOLENIOWSKI, M. and BARTHOL, A. (1999): Partial purification and characterization of acetyl coenzyme A: Taxa-4(20),11(12)-dien-5 alpha-ol O-acetyl transferase that catalyzes the first acylation step of Taxol biosynthesis. *Archives of Biochemistry and Biophysics*, **364**(2): 273-279.

DAHIYA, J. S. (1996): Fermentation for Taxol Production. **WO 96/32490**. Canada.

DANIEWSKI, W. M., GUMULKA, M., ANCZEWSKI, W., MASNYK, M., BLOSZYK, E. and GUPTA, K. K. (1998): Why the yew tree (*Taxus baccata*) is not attacked by insects. *Phytochemistry*, **49**(5): 1279-1282.

DANISHEFSKY, S. J., MASTERS, J. J., YOUNG, W. B., LINK, J. T., SNYDER, L. B., MAGEE, T. V., JUNG, D. K., ISAACS, R. C. A., BORNMANN, W. G., C.A., A., C.A., C. and M.J., D. G. (1996): Total Synthesis of Baccatin III and Taxol. *Journal of the American Chemical Society*, **118**(12): 2843-2859.

DEJONG, J. H. M., LIU, Y., BOLLON, A. P., LONG, R. M., JENNEWEIN, S., WILLIAMS, D. and CROTEAU, R. B. (2006): Genetic engineering of taxol biosynthetic genes in *Saccharomyces cerevisiae. Biotechnology and Bioengineering*, **93**(2): 212-224.

DENG, B. W., LIU, K. H., CHEN, W. Q., DING, X. W. and XIE, X. C. (2009): *Fusarium solani,* Tax-3, a new endophytic taxol-producing fungus from *Taxus chinensis*. *World J. Microbiol. Biotechnol.*, **25**(1): 139-143.

DEUS-NEUMANN, B. and ZENK, M. H. (1984): Instability of indole alkaloid production in *Catharanthus roseus* cell suspension cultures. *Planta Medica*, **50**(5): 427-431.

DREYFUSS, M. M. and CHAPELA, I. H. (1994): Potential of fungi in the discovery of novel, low-molecular weight pharmaceuticals. *Biotechnology*, **26**: 49-80.

EISENREICH, W., MENHARD, B., HYLANDS, P. J., ZENK, M. H. and BACHER, A. (1996): Studies on the biosynthesis of taxol: the taxane carbon skeleton is not of mevalonoid origin. *Proc Natl Acad Sci U S A*, **93**(13): 6431-6436.

EISENREICH, W., ROHDICH, F. and BACHER, A. (2001): Deoxyxylulose phosphate pathway to terpenoids. *Trends Plant Sci*, **6**(2): 78-84.

EISENREICH, W., SCHWARZ, M., CARTAYRADE, A., ARIGONI, D., ZENK, M. H. and BACHER, A. (1998): The deoxyxylulose phosphate pathway of terpenoid biosynthesis in plants and microorganisms. *Chem Biol*, **5**(9): R221-233.

EL, J. M. and DIALLO, B. (2000): Taxane and taxine derivatives from Penicillium sp.

ELMER, W. H., MATTINA, M. J. I. and MACEACHERN, G. J. (1994): Sensitivity of Plant-Pathogenic Fungi to Taxane Extracts from Ornamental Yews. *Phytopathology*, **84**(10): 1179-1185.

ENGELS, B., DAHM, P. and JENNEWEIN, S. (2008): Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards Taxol (Paclitaxel) production. *Metabolic Engineering*, **10**(3-4): 201-206.

ENGELS, B., HEINIG, U., GROTHE, T., STADLER, M. and JENNEWEIN, S. (2011): Cloning and characterization of an *Armillaria gallica* cDNA encoding protoilludene synthase, which catalyzes the first committed step in the synthesis of antimicrobial melleolides. *J Biol Chem*, **286**(9): 6871-6878.

FELDMANN, K. A., CHOE, S. W., KIM, H. B. and PARK, J. H. (2002): Functional genomics of cytochromes P450 in plants. *Phytochemistry in the Genomics and Post-Genomics Eras*, **36**: 125-143.

FELICETTI, B. and CANE, D. E. (2004): Aristolochene synthase: mechanistic analysis of active site residues by site-directed mutagenesis. *J Am Chem Soc*, **126**(23): 7212-7221.

FIELD, B., FISTON-LAVIER, A. S., KEMEN, A., GEISLER, K., QUESNEVILLE, H. and OSBOURN, A. E. (2011): Formation of plant metabolic gene clusters within dynamic chromosomal regions. *Proc Natl Acad Sci U S A*, **108**(38): 16116-16121.

FIELD, B. and OSBOURN, A. E. (2008): Metabolic diversification--independent assembly of operon-like gene clusters in different plants. *Science*, **320**(5875): 543-547.

FIRN, R. D. and JONES, C. G. (2003): Natural products--a simple model to explain chemical diversity. *Nat Prod Rep*, **20**(4): 382-391.

FISCHBACH, M. A. and CLARDY, J. (2007): One pathway, many products. *Nature Chemical Biology*, **3**(7): 353-355.

FITZPATRICK, D. A. (2011): Horizontal Gene Transfer in Fungi. *FEMS Microbiol Lett*, **"accepted article", 10.1111/j.1574-6968.2011.02465.x**: 1-8.

FLEMING, P. E., KNAGGS, A. R., HE, X. G., MOCEK, U. and FLOSS, H. G. (1994): Biosynthesis of Taxoids - Mode of Attachment of the Taxol Side-Chain. *Journal of the American Chemical Society*, **116**(9): 4137-4138.

FLOSS, H. G. and MOCEK, U. (1995): Biosynthesis of taxol. Taxol: Science and Application. M, S. Boca Raton, CRC Press**:** 191-208.

FREY, M., CHOMET, P., GLAWISCHNIG, E., STETTNER, C., GRUN, S., WINKLMAIR, A., EISENREICH, W., BACHER, A., MEELEY, R. B., BRIGGS, S. P., SIMCOX, K. and GIERL, A. (1997): Analysis of a chemical plant defense mechanism in grasses. *Science*, **277**(5326): 696-699.

FREY, M., HUBER, K., PARK, W. J., SICKER, D., LINDBERG, P., MEELEY, R. B., SIMMONS, C. R., YALPANI, N. and GIERL, A. (2003): A 2-oxoglutarate-dependent dioxygenase is integrated in DIMBOA-biosynthesis. *Phytochemistry*, **62**(3): 371-376.

GANGADEVI, V., MURUGAN, M. and MUTHUMARY, J. (2008): Taxol determination from *Pestalotiopsis pauciseta*, a fungal endophyte of a medicinal plant. *Shengwu Gongcheng Xuebao*, **24**(8): 1433-1438.

GANGADEVI, V. and MUTHUMARY, J. (2008a): A simple and rapid method for the determination of taxol produced by fungal endophytes from medicinal plants using high performance thin layer chromatography. *Sepu*, **26**(1): 50-55.

GANGADEVI, V. and MUTHUMARY, J. (2008b): Taxol, an anticancer drug produced by an endophytic fungus *Bartalinia robillardoides* Tassi, isolated from a medicinal plant, Aegle marmelos Correa ex Roxb. *World J. Microbiol. Biotechnol.*, **24**(5): 717-724.

GANGADEVI, V. and MUTHUMARY, J. (2009a): A novel endophytic Taxol-producing fungus Chaetomella raphigera isolated from a medicinal plant, Terminalia arjuna. *Appl. Biochem. Biotechnol.*, **158**(3): 675-684.

GANGADEVI, V. and MUTHUMARY, J. (2009b): Taxol production by *Pestalotiopsis terminaliae*, an endophytic fungus of *Terminalia arjuna* (arjun tree). *Biotechnology and Applied Biochemistry*, **52**: 9-15.

GANLEY, R. J., BRUNSFELD, S. J. and NEWCOMBE, G. (2004): A community of unknown, endophytic fungi in western white pine. *Proc Natl Acad Sci U S A*, **101**(27): 10107-10112.

GANLEY, R. J. and NEWCOMBE, G. (2006): Fungal endophytes in seeds and needles of Pinus monticola. *Mycol Res*, **110**(Pt 3): 318-327.

GARDES, M. and BRUNS, T. D. (1993): ITS primers with enhanced specificity for basidiomycetes--application to the identification of mycorrhizae and rusts. *Mol Ecol*, **2**(2): 113-118.

GIBSON, D. M., KETCHUM, R. E. B., VANCE, N. V. and CHRISTEN, A. A. (1993): Initiation and growth of cell lines of *Taxus brevifolia* (Pacific yew). *Plant Cell Reports*, **12**(9): 479-482.

GOGOI, D. K., BORUAH, H. P. D., SAIKIA, R. and BORA, T. C. (2008): Optimization of process parameters for improved production of bioactive metabolite by a novel endophytic fungus *Fusarium sp.* DF2 isolated from *Taxus wallichiana* of north east India. *World J. Microbiol. Biotechnol.*, **24**(1): 79-87.

GOLDSPIEL, B. R. (1997): Clinical overview of the taxanes. *Pharmacotherapy*, **17**(5): S110-S125.

GOODIN, S. (2008): Novel cytotoxic agents: epothilones. *Am J Health Syst Pharm*, **65**(10 Suppl 3): S10-15.

GOODWIN, S. B., M'BAREK S, B., DHILLON, B., WITTENBERG, A. H., CRANE, C. F., HANE, J. K., FOSTER, A. J., VAN DER LEE, T. A., GRIMWOOD, J., AERTS, A., ANTONIW, J., BAILEY, A., BLUHM, B., BOWLER, J., BRISTOW, J., VAN DER BURGT, A., CANTO-CANCHE, B., CHURCHILL, A. C., CONDE-FERRAEZ, L., COOLS, H. J., COUTINHO, P. M., CSUKAI, M., DEHAL, P., DE WIT, P., DONZELLI, B., VAN DE GEEST, H. C., VAN HAM, R. C., HAMMOND-KOSACK, K. E., HENRISSAT, B., KILIAN, A., KOBAYASHI, A. K., KOOPMANN, E., KOURMPETIS, Y., KUZNIAR, A., LINDQUIST, E., LOMBARD, V., MALIEPAARD, C., MARTINS, N., MEHRABI, R., NAP, J. P., PONOMARENKO, A., RUDD, J. J., SALAMOV, A., SCHMUTZ, J., SCHOUTEN, H. J., SHAPIRO, H., STERGIOPOULOS, I., TORRIANI, S. F., TU, H., DE VRIES, R. P., WAALWIJK, C., WARE, S. B., WIEBENGA, A., ZWIERS, L. H., OLIVER, R. P., GRIGORIEV, I. V. and KEMA, G. H. (2011): Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. *PLoS Genet*, **7**(6): e1002070.

GUERITTE-VOEGELEIN, F., GUENARD, D. and POTIER, P. (1987): Taxol and derivatives: a biogenetic hypothesis. *J Nat Prod*, **50**(1): 9-18.

GUO, B. H., WANG, Y. C., ZHOU, X. W., HU, K., TAN, F., MIAO, Z. Q. and TANG, K. X. (2006): An endophytic taxol-producing fungus BT2 isolated from *Taxus chinensis var. mairei*. *African Journal of Biotechnology*, **5**(10): 875-877.

HALL, R. D. and YEOMAN, M. M. (1987): Intercellular and intercultural heterogeneity in secondary metabolite accumulation in cultures of Catharanthus roseus following cell line selection. *Journal of Experimental Botany*, **38**(193): 1391-1398.

HARTZELL, M. (1991): The yew tree, a thousend whispers. Hulogosi, Eugene, Ore.

HEDDEN, P., PHILLIPS, A. L., ROJAS, M. C., CARRERA, E. and TUDZYNSKI, B. (2002): Gibberellin Biosynthesis in Plants and Fungi: A Case of Convergent Evolution? *J Plant Growth Regul*, **20**(4): 319-331.

HEFNER, J., KETCHUM, R. E. and CROTEAU, R. (1998): Cloning and functional expression of a cDNA encoding geranylgeranyl diphosphate synthase from *Taxus canadensis* and assessment of the role of this prenyltransferase in cells induced for taxol production. *Arch Biochem Biophys*, **360**(1): 62-74.

HEINIG, U. (2006). Klonierung der Gene der Taxolbiosynthese aus *Taxomyces andreanae*. Darmstadt, Technical University of Darmstadt.

HEINIG, U. and JENNEWEIN, S. (2009): Taxol: A complex diterpenoid natural product with an evolutionarily obscure origin. *African Journal of Biotechnology*, **8**(8): 1370-1385.

HEINIG, U., SCHOLZ, S., DAHM, P., GRABOWY, U. and JENNEWEIN, S. (2010): Development of carbon plasma-coated multiwell plates for high-throughput mass spectrometric analysis of highly lipophilic fermentation products. *Analytical Biochemistry*, **403**(1-2): 108-113.

HELLIWELL, C. A., CHANDLER, P. M., POOLE, A., DENNIS, E. S. and PEACOCK, W. J. (2001): The CYP88A cytochrome P450, ent-kaurenoic acid oxidase, catalyzes three steps of the gibberellin biosynthesis pathway. *Proc Natl Acad Sci U S A*, **98**(4): 2065-2070.

HEZARI, M., KETCHUM, R. E. B., GIBSON, D. M. and CROTEAU, R. (1997): Taxol production and taxadiene synthase activity in *Taxus canadensis* cell suspension cultures. *Archives of Biochemistry and Biophysics*, **337**(2): 185-190.

HEZARI, M., LEWIS, N. G. and CROTEAU, R. (1995): Purification and Characterization of Taxa-4(5),11(12)-Diene Synthase from Pacific Yew (Taxus-Brevifolia) That Catalyzes the First Committed Step of Taxol Biosynthesis. *Archives of Biochemistry and Biophysics*, **322**(2): 437-444.

HOFFMAN, A. (2003): Methods for obtaining Taxanes. **US6638742 B1**. USA.University of Portland.

HOFFMAN, A., KHAN, W., WORAPONG, J., STROBEL, G., GRIFFIN, D., ARBOGAST, B., BAROFSKY, D., BOONE, R. B., NING, L., ZHENG, P. and DALEY, L. (1998): Bioprospecting for taxol in angiosperm plant extracts. *Spectroscopy (Eugene, Oreg.)*, **13**(6): 22-32.

HOFFMEISTER, D. and KELLER, N. P. (2007): Natural products of filamentous fungi: enzymes, genes, and their regulation. *Nat Prod Rep*, **24**(2): 393-416.

HOHN, T. M. and BEREMAND, P. D. (1989): Isolation and nucleotide sequence of a sesquiterpene cyclase gene from the trichothecene-producing fungus *Fusarium* sporotrichioides. *Gene*, **79**(1): 131-138.

HOHN, T. M., MCCORMICK, S. P. and DESJARDINS, A. E. (1993): Evidence for a gene cluster involving trichothecene-pathway biosynthetic genes in Fusarium *sporotrichioides*. *Curr Genet*, **24**(4): 291-295.

HOLTON, R. A., BIEDIGER, R. J. and BOATMAN, D. (1995): Semisynthesis of taxol and taxotere. Taxol: Science and Applications. SUFFNESS, M. Boca Raton, CRC Press**:** 97-121.

HOLTON, R. A., KIM, H. B., SOMOZA, C., LIANG, F., BIEDIGER, R. J., BOATMAN, P. D., SHINDO, M., SMITH, C. C., KIM, S., NADIZADEH, H., SUZUKI, Y., TAO, C. L., VU, P., TANG, S. H., ZHANG, P. S., MURTHI, K. K., GENTILE, L. N. and LIU, J. H. (1994a): First total synthesis of taxol. 2. Completion of the C and D rings. *Journal of the American Chemical Society*, **116**(4): 1599-1600.

HOLTON, R. A., SOMOZA, C., KIM, H. B., LIANG, F., BIEDIGER, R. J., BOATMAN, P. D., SHINDO, M., SMITH, C. C., KIM, S., NADIZADEH, H., SUZUKI, Y., TAO, C. L., VU, P., TANG, S. H., ZHANG, P. S., MURTHI, K. K., GENTILE, L. N. and LIU, J. H. (1994b): First total synthesis of taxol. 1. Functionalization of the B ring. *Journal of the American Chemical Society*, **116**(4): 1597-1598.

HOLTON, T. A. and LESTER, D. R. (1996): Cloning of novel cytochrome P450 gene sequences via polymerase chain reaction amplification. *Cytochrome P450, Pt B*, **272**: 275-283.

HTAY, T. and LIU, M. W. (2005): Drug-eluting stent: a review and update. *Vasc Health Risk Manag*, **1**(4): 263-276.

HUANG, E. X., HUANG, Q. L., WILDUNG, M. R., CROTEAU, R. and SCOTT, A. I. (1998): Overproduction, in *Escherichia coli*, of soluble taxadiene synthase, a key enzyme in the taxol biosynthetic pathway. *Protein Expression and Purification*, **13**(1): 90-96.

HUANG, Y., WANG, J., LI, G., ZHENG, Z. and SU, W. (2001): Antitumor and antifungal activities in endophytic fungi isolated from pharmaceutical plants *Taxus mairei*, *Cephalataxus fortunei* and *Torreya grandis*. *FEMS Immunol. Med. Microbiol.*, **31**(2): 163-167.

ITOKAWA, H. (2003): Introduction. Taxus. The Genus Taxus. H. ITOKAWA, K.-H. L. London, New York Taylor and Francis**:** 1-18.

JENNEWEIN, S. and CROTEAU, R. (2001): Taxol: biosynthesis, molecular genetics, and biotechnological applications. *Appl Microbiol Biotechnol*, **57**(1-2): 13-19.

JENNEWEIN, S., LONG, R. M., WILLIAMS, R. M. and CROTEAU, R. (2004a): Cytochrome p450 taxadiene 5alpha-hydroxylase, a mechanistically unusual monooxygenase catalyzing the first oxygenation step of taxol biosynthesis. *Chem Biol*, **11**(3): 379-387.

JENNEWEIN, S., PARK, H., DEJONG, J. M., LONG, R. M., BOLLON, A. P. and CROTEAU, R. B. (2005): Coexpression in yeast of *Taxus* cytochrome P450 reductase with cytochrome P450 oxygenases involved in Taxol biosynthesis. *Biotechnology and Bioengineering*, **89**(5): 588-598.

JENNEWEIN, S., RITHNER, C. D., WILLIAMS, R. M. and CROTEAU, R. (2003): Taxoid metabolism: Taxoid 14 beta-hydroxylase is a cytochrome P450-dependent monooxygenase. *Archives of Biochemistry and Biophysics*, **413**(2): 262-270.

JENNEWEIN, S., RITHNER, C. D., WILLIAMS, R. M. and CROTEAU, R. B. (2001): Taxol biosynthesis: taxane 13a-hydroxylase is a cytochrome P450-dependent monooxygenase. *Proceedings of the National Academy of Sciences of the United States of America*, **98**(24): 13595-13600.

JENNEWEIN, S., WILDUNG, M. R., CHAU, M., WALKER, K. and CROTEAU, R. (2004b): Random sequencing of an induced *Taxus* cell cDNA library for identification of clones involved in Taxol biosynthesis. *Proc Natl Acad Sci U S A*, **101**(24): 9149-9154.

JIN, Q. W., WILLIAMS, D. C., HEZARI, M., CROTEAU, R. and COATES, R. M. (2005): Stereochemistry of the macrocyclization and elimination steps in taxadiene biosynthesis through deuterium labeling. *Journal of Organic Chemistry*, **70**(12): 4667-4675.

JONES, C. G. and FIRN, R. D. (1991): The Screening Hypothesis - an evolutionary model to explain the chemical diversity of Natural Products. *Phil Trans Royal Soc*, **333**: 273-280.

KASPERA, R. and CROTEAU, R. (2006): Cytochrome P450 oxygenases of Taxol biosynthesis. *Phytochemistry Reviews*, **5**(2): 433-444.

KEELING, P. J. and PALMER, J. D. (2008): Horizontal gene transfer in eukaryotic evolution. *Nat Rev Genet*, **9**(8): 605-618.

KELLER, N. P. and HOHN, T. M. (1997): Metabolic Pathway Gene Clusters in Filamentous Fungi. *Fungal Genet Biol*, **21**(1): 17-29.

KELLER, N. P., TURNER, G. and BENNETT, J. W. (2005): Fungal secondary metabolism - from biochemistry to genomics. *Nat Rev Microbiol*, **3**(12): 937-947.

KETCHUM, R. and CROTEAU, R. (2006): The *Taxus* Metabolome and the Elucidation of the Taxol® Biosynthetic Pathway in Cell Suspension Cultures
Plant Metabolomics. SAITO, K., DIXON, R. A. and WILLMITZER, L. Berlin, Heidelberg, Springer **57:** 291-309.

KETCHUM, R. E. B. and GIBSON, D. M. (1996): Paclitaxel production in suspension cell cultures of Taxus. *Plant Cell, Tissue and Organ Culture*, **46**(1): 9-16.

KETCHUM, R. E. B., GIBSON, D. M., CROTEAU, R. B. and SHULER, M. L. (1999): The kinetics of taxoid accumulation in cell suspension cultures of *Taxus* following elicitation with methyl jasmonate. *Biotechnology and Bioengineering*, **62**(1): 97-105.

KETCHUM, R. E. B., HORIGUCHI, T., QIU, D., WILLIAMS, R. M. and CROTEAU, R. B. (2007): Administering cultured *Taxus* cells with early precursors reveals bifurcations in the taxoid biosynthetic pathway. *Phytochemistry (Elsevier)*, **68**(3): 335-341.

KETCHUM, R. E. B., RITHNER, C. D., QIU, D., KIM, Y. S., WILLIAMS, R. M. and CROTEAU, R. B. (2003): *Taxus* metabolomics: methyl jasmonate preferentially induces production of taxoids oxygenated at C-13 in *Taxus x media* cell cultures. *Phytochemistry (Elsevier)*, **62**(6): 901-909.

KHALDI, N., COLLEMARE, J., LEBRUN, M. H. and WOLFE, K. H. (2008): Evidence for horizontal transfer of a secondary metabolite gene cluster between fungi. *Genome Biol*, **9**(1): R18.

KIM, G. T. and TSUKAYA, H. (2002): Regulation of the biosynthesis of plant hormones by cytochrome P450s. *Journal of Plant Research*, **115**(1119): 169-177.

KIM, S.-U., STROBEL, G. and FORD, E. (1999): Screening of taxol-producing endophytic fungi from *Ginkgo biloba* and *Taxus cuspidata* in Korea. *Agric. Chem. Biotechnol. (Engl. Ed.)*, **42**(2): 97-99.

KINGSTON, D. G. I., JAGTAP, P. G., BALOGLU, E., BARRON, D. M. and BANE, S. (2002): Design and synthesis of a combinatorial chemistry library of 7-acyl, 10-acyl, and 7,10-diacyl analogues of paclitaxel (Taxol) using solid phase synthesis. *Journal of Natural Products*, **65**(8): 1136-1142.

KOEPP, A. E., HEZARI, M., ZAJICEK, J., VOGEL, B. S., LAFEVER, R. E., LEWIS, N. G. and CROTEAU, R. (1995): Cyclization of Geranylgeranyl Diphosphate to Taxa-4(5),11(12)-Diene Is the Committed Step of Taxol Biosynthesis in Pacific Yew. *Journal of Biological Chemistry*, **270**(15): 8686-8690.

KOKSAL, M., JIN, Y. H., COATES, R. M., CROTEAU, R. and CHRISTIANSON, D. W. (2011): Taxadiene synthase structure and evolution of modular architecture in terpene biosynthesis. *Nature*, **469**(7328): 116-138.

KROKEN, S., GLASS, N. L., TAYLOR, J. W., YODER, O. C. and TURGEON, B. G. (2003): Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proc Natl Acad Sci U S A*, **100**(26): 15670-15675.

KUMARAN, R., MUTHUMARY, J., KIM, E. K. and HUR, B. K. (2009a): Production of taxol from Phyllosticta dioscoreae, a leaf spot fungus isolated from Hibiscus rosa-sinensis. *Biotechnology and Bioprocess Engineering*, **14**(1): 76-83.

KUMARAN, R. S. and HUR, B.-K. (2009): Screening of species of the endophytic fungus *Phomopsis* for the production of the anticancer drug taxol. *Biotechnol. Appl. Biochem.*, **54**(1): 21-30.

KUMARAN, R. S., KIM, H. J. and HUI, B.-K. (2010a): Taxol-producing fungal endophyte, *Pestalotiopsis* species isolated from *Taxus cuspidata*. *J. Biosci. Bioeng.*, **110**(5): 541-546.

KUMARAN, R. S., MUTHUMARY, J. and HUR, B. K. (2008a): Isolation and Identification of Taxol, an Anticancer Drug from Phyllosticta melochiae Yates, an Endophytic Fungus of Melochia corchorifolia L. *Food Science and Biotechnology*, **17**(6): 1246-1253.

KUMARAN, R. S., MUTHUMARY, J. and HUR, B. K. (2008b): Production of taxol from *Phyllosticta spinarum*, an endophytic fungus of *Cupressus sp*. *Engineering in Life Sciences*, **8**(4): 438-446.

KUMARAN, R. S., MUTHUMARY, J. and HUR, B. K. (2008c): Taxol from Phyllosticta citricarpa, a leaf spot fungus of the angiosperm *Citrus medica*. *Journal of Bioscience and Bioengineering*, **106**(1): 103-106.

KUMARAN, R. S., MUTHUMARY, J. and HUR, B. K. (2009b): Isolation and identification of an anticancer drug, taxol from *Phyllosticta tabernaemontanae*, a leaf spot fungus of an angiosperm, *Wrightia tinctoria*. *Journal of Microbiology*, **47**(1): 40-49.

KUSARI, S., LAMSHOFT, M. and SPITELLER, M. (2009a): Aspergillus fumigatus Fresenius, an endophytic fungus from *Juniperus communis L. Horstmann* as a novel source of the anticancer pro-drug deoxypodophyllotoxin. *J Appl Microbiol*, **107**(3): 1019-1030.

KUSARI, S., ZUHLKE, S. and SPITELLER, M. (2009b): An endophytic fungus from *Camptotheca acuminata* that produces camptothecin and analogues. *J Nat Prod*, **72**(1): 2-7.

KUZUYAMA, T. and SETO, H. (2003): Diversity of the biosynthesis of the isoprene units. *Natural Product Reports*, **20**(2): 171-183.

LASALA, J. M., STONE, G. W., DAWKINS, K. D., SERRUYS, P. W., COLOMBO, A., GRUBE, E., KOGLIN, J. and ELLIS, S. (2006): An overview of the *TAXUS* Express, paclitaxel-eluting stent clinical trial program. *Journal of Interventional Cardiology*, **19**(5): 422-431.

LASKARIS, G., VAN DER HEIJDEN, R. and VERPOORTE, R. (2000): Purification and partial characterisation of geranylgeranyl diphosphate synthase, from *Taxus baccata* cell cultures. An enzyme that regulates taxane biosynthesis. *Plant Science*, **153**(2): 97-105.

LAWRENCE, D. P., KROKEN, S., PRYOR, B. M. and ARNOLD, A. E. (2011): Interkingdom Gene Transfer of a Hybrid NPS/PKS from Bacteria to Filamentous Ascomycota. *PLoS One*, **6**(11): e28231.

LEETE, E. and BODEM, G. B. (1966): Biosynthesis of 3-Simethylamino-3-Phenyl-Propanoic Acid in Yew. *Tetrahedron Letters*(33): 3925-3927.

LI, J.-Y., STROBEL, G., SIDHU, R., HESS, W. M. and FORD, E. J. (1996): Endophytic taxol-producing fungi from bald cypress, *Taxodium distichum*. *Microbiology (Reading, U. K.)*, **142**(8): 2223-2226.

LI, J. Y., SIDHU, R. S., FORD, E. J., LONG, D. M. and STROBEL, G. A. (1998): The induction of taxol production in the endophytic fungus *Periconia sp* from *Torreya grandifolia*. *J. Ind. Microbiol. Biotechnol.*, **20**(5): 259-264.

LI, T., ZHANG, Z.-J., ZHANG, P., WANG, C.-I., LIU, B., LIU, T.-T., FU, C.-H. and YU, L.-J. (2011): Isolation and identification of a taxol-producing endophytic fungus identified from *Taxus media*. *Agric. Sci. Technol.*, **11**(5): 38-40,68.

LI, Y.-C. and TAO, W.-Y. (2009): Interactions of Taxol-producing endophytic fungus with its host (*Taxus spp*.) during Taxol accumulation. *Cell Biol. Int.*, **33**(1): 106-112.

LI, Y.-C., TAO, W.-Y. and CHENG, L. (2009): Paclitaxel production using co-culture of Taxus suspension cells and paclitaxel-producing endophytic fungi in a co-bioreactor. *Appl. Microbiol. Biotechnol.*, **83**(2): 233-239.

LIN, X. Y., HEZARI, M., KOEPP, A. E., FLOSS, H. G. and CROTEAU, R. (1996): Mechanism of taxadiene synthase, a diterpene cyclase that catalyzes the first step of taxol biosynthesis in Pacific yew. *Biochemistry*, **35**(9): 2968-2977.

LIU, K., DING, X., DENG, B. and CHEN, W. (2009): Isolation and characterization of endophytic taxol-producing fungi from *Taxus chinensis*. *J. Ind. Microbiol. Biotechnol.*, **36**(9): 1171-1177.

LONG, R. M. and CROTEAU, R. (2005): Preliminary assessment of the C13-side chain 2 '-hydroxylase involved in Taxol biosynthesis. *Biochemical and Biophysical Research Communications*, **338**(1): 410-417.

LONG, R. M., LAGISETTI, C., COATES, R. M. and CROTEAU, R. B. (2008): Specificity of the N-benzoyl transferase responsible for the last step of Taxol biosynthesis. *Archives of Biochemistry and Biophysics*, **477**(2): 384-389.

MALPARTIDA, F. and HOPWOOD, D. A. (1986): Physical and genetic characterisation of the gene cluster for the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2). *Mol Gen Genet*, **205**(1): 66-73.

MCCONN, M., CREELMAN, R. A., BELL, E., MULLET, J. E. and BROWSE, J. (1997): Jasmonate is essential for insect defense in *Arabidopsis*. *Proc Natl Acad Sci U S A*, **94**(10): 5473-5477.

MEHRABI, R., BAHKALI, A. H., ABD-ELSALAM, K. A., MOSLEM, M., BEN M'BAREK, S., GOHARI, A. M., JASHNI, M. K., STERGIOPOULOS, I., KEMA, G. H. and DE WIT, P. J. (2011): Horizontal gene and chromosome transfer in plant pathogenic fungi affecting host range. *FEMS Microbiol Rev*, **35**(3): 542-554.

METZ, A. M., HADDAD, A., WORAPONG, J., LONG, D. M., FORD, E. J., HESS, W. M. and STROBEL, G. A. (2000): Induction of the sexual stage of *Pestalotiopsis microspora*, a taxol-producing fungus. *Microbiology*, **146 (Pt 8)**: 2079-2089.

MIAO, Z., WANG, Y., YU, X., GUO, B. and TANG, K. (2009a): A new endophytic taxane-production fungus from *Taxus chinensis*. *Prikl Biokhim Mikrobiol*, **45**(1): 92-96.

MIAO, Z., WANG, Y., YU, X., GUO, B. and TANG, K. (2009b): A new endophytic taxane production fungus from *Taxus chinensis*. *Applied Biochemistry and Microbiology*, **45**(1): 81-86.

MISIEK, M. and HOFFMEISTER, D. (2008): Processing sites involved in intron splicing of Armillaria natural product genes. *Mycol Res*, **112**(Pt 2): 216-224.

MIZUTANI, M. and SATO, F. (2011): Unusual P450 reactions in plant secondary metabolism. *Arch Biochem Biophys*, **507**(1): 194-203.

MORRIS, P., RUDGE, K., CRESSWELL, R. and FOWLER, M. W. (1989): Regulation of product synthesis in cell cultures of *Catharanthus roseus*. V. Long-term maintenance of cells on a production medium. *Plant Cell, Tissue and Organ Culture*, **17**(2): 79-90.

NADEEM, M., RIKHARI, H. C., KUMAR, A., PALNI, L. M. and NANDI, S. K. (2002): Taxol content in the bark of Himalayan Yew in relation to tree age and sex. *Phytochemistry*, **60**(6): 627-631.

NICOLAOU, K. C., YANG, Z., LIU, J. J., UENO, H., NANTERMET, P. G., GUY, R. K., CLAIBORNE, C. F., RENAUD, J., COULADOUROS, E. A., PAULVANNAN, K. and SORENSEN, E. J. (1994): Total synthesis of taxol. *Nature (London, United Kingdom)*, **367**(6464): 630-634.

OBERLIES, N. H. and KROLL, D. J. (2004): Camptothecin and taxol: historic achievements in natural products research. *J Nat Prod*, **67**(2): 129-135.

OGDEN, L. (1988): *Taxus* (Yews) - a Highly Toxic Plant. *Veterinary and Human Toxicology*, **30**(6): 563-564.

OLSZEWSKI, N., SUN, T. P. and GUBLER, F. (2002): Gibberellin signaling: biosynthesis, catabolism, and response pathways. *Plant Cell*, **14 Suppl**: S61-80.

OSBOURN, A. E. and FIELD, B. (2009): Operons. *Cell Mol Life Sci*, **66**(23): 3755-3775.

PANDI, M., KUMARAN, R. S., CHOI, Y.-K., KIM, H. J. and MUTHUMARY, J. (2011): Isolation and detection of taxol, an anticancer drug produced from *Lasiodiplodia theobromae*, an endophytic fungus of the medicinal plant *Morinda citrifolia*. *African Journal of Biotechnology*, **10**(8): 1428-1435.

PARR, A. J., PAYNE, J., EAGLES, J., CHAPMAN, B. T., ROBINS, R. J. and RHODES, M. J. C. (1990): Variation in tropane alkaloid accumulation within the *Solanaceae* and strategies for its exploitation. *Phytochemistry*, **29**(8): 2545-2550.

PAYNE, G. A. and BROWN, M. P. (1998): Genetics and physiology of aflatoxin biosynthesis. *Annu Rev Phytopathol*, **36**: 329-362.

PETERS, R. J., FLORY, J. E., JETTER, R., RAVN, M. M., LEE, H. J., COATES, R. M. and CROTEAU, R. B. (2000): Abietadiene synthase from grand fir (*Abies grandis*): characterization and mechanism of action of the "pseudomature" recombinant enzyme. *Biochemistry*, **39**(50): 15592-15602.

PICHERSKY, E. and GANG, D. R. (2000): Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. *Trends in Plant Science*, **5**(10): 439-445.

PLATT, R. V., OPIE, C. T. and HASLAM, E. (1984): Plant Proanthocyanidins .8. Biosynthesis of Flavan-3-Ols and Other Secondary Plant-Products from (2s)-Phenylalanine. *Phytochemistry*, **23**(10): 2211-2217.

PROCTOR, R. H. and HOHN, T. M. (1993): Aristolochene synthase. Isolation, characterization, and bacterial expression of a sesquiterpenoid biosynthetic gene (Ari1) from *Penicillium roqueforti*. *J Biol Chem*, **268**(6): 4543-4548.

PROCTOR, R. H., MCCORMICK, S. P., ALEXANDER, N. J. and DESJARDINS, A. E. (2009): Evidence that a secondary metabolic biosynthetic gene cluster has grown by gene relocation during evolution of the filamentous fungus Fusarium. *Mol Microbiol*, **74**(5): 1128-1142.

PURI, S. C., VERMA, V., AMNA, T., QAZI, G. N. and SPITELLER, M. (2005): An endophytic fungus from *Nothapodytes foetida* that produces camptothecin. *Journal of Natural Products*, **68**(12): 1717-1719.

QI, X., BAKHT, S., LEGGETT, M., MAXWELL, C., MELTON, R. and OSBOURN, A. (2004): A gene cluster for secondary metabolism in oat: implications for the evolution of metabolic diversity in plants. *Proc Natl Acad Sci U S A*, **101**(21): 8233-8238.

RAJA, S. G. (2006): Drug-eluting stents and the future of coronary artery bypass surgery: facts and fiction. *Ann Thorac Surg*, **81**(3): 1162-1171.

RICHARDS, T. A., SOANES, D. M., FOSTER, P. G., LEONARD, G., THORNTON, C. R. and TALBOT, N. J. (2009): Phylogenomic analysis demonstrates a pattern of rare and ancient horizontal gene transfer between plants and fungi. *Plant Cell*, **21**(7): 1897-1911.

RITTER, S. K. (2004): Green innovations. *Chemical & Engineering News*, **82**(28): 25-30.

RO, D. K., ARIMURA, G. L., LAU, S. Y. W., PIERS, E. and BOHLMANN, J. (2005): Loblolly pine abietadienol/abietadienal oxidase PtAO (CYP720B1) is a multifunctional, multisubstrate cytochrome P450 monooxygenase. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(22): 8060-8065.

ROHMER, M. (1999): The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. *Natural Product Reports*, **16**(5): 565-574.

SAMBROOK, J. and RUSSELL, D. W. (2001): Molekular Cloning – A Laboratory Manual New York, Cold Spring Harbor Laboratory Press

SCHIFF, P. B., FANT, J. and HORWITZ, S. B. (1979): Promotion of microtubule assembly in vitro by taxol. *Nature*, **277**(5698): 665-667.

SCHOENDORF, A., RITHNER, C. D., WILLIAMS, R. M. and CROTEAU, R. B. (2001): Molecular cloning of a cytochrome P450 taxane 10β-hydroxylase cDNA from *Taxus* and functional expression in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, **98**(4): 1501-1506.

SCHOPFER, C. R. and EBEL, J. (1998): Identification of elicitor-induced cytochrome P450s of soybean (Glycine max L.) using differential display of mRNA. *Mol Gen Genet*, **258**(4): 315-322.

SCHRIPSEMA, J. and VERPOORTE, R. (1992): Search for factors related to the indole alkaloid production in cell suspension cultures of *Tabernaemontana divaricata*. *Planta Medica*, **58**(3): 245-249.

SCHULER, M. A. (1996): Plant cytochrome P450 monooxygenases. *Critical Reviews in Plant Sciences*, **15**(3): 235-284.

SEEMANN, M., ZHAI, G., DE KRAKER, J. W., PASCHALL, C. M., CHRISTIANSON, D. W. and CANE, D. E. (2002): Pentalenene synthase. Analysis of active site residues by site-directed mutagenesis. *J Am Chem Soc*, **124**(26): 7681-7689.

SHAABAN, M., PALMER, J. M., EL-NAGGAR, W. A., EL-SOKKARY, M. A., HABIB EL, S. E. and KELLER, N. P. (2010): Involvement of transposon-like elements in penicillin gene cluster regulation. *Fungal Genet Biol*, **47**(5): 423-432.

SHARMA, A. and STRAUBINGER, R. M. (1994): Novel taxol formulations: preparation and characterization of taxol-containing liposomes. *Pharm Res*, **11**(6): 889-896.

SHIMURA, K., OKADA, A., OKADA, K., JIKUMARU, Y., KO, K. W., TOYOMASU, T., SASSA, T., HASEGAWA, M., KODAMA, O., SHIBUYA, N., KOGA, J., NOJIRI, H. and YAMANE, H. (2007): Identification of a biosynthetic gene cluster in rice for momilactones. *J Biol Chem*, **282**(47): 34013-34018.

SHIU, P. K., RAJU, N. B., ZICKLER, D. and METZENBERG, R. L. (2001): Meiotic silencing by unpaired DNA. *Cell*, **107**(7): 905-916.

SHRESTHA, K., STROBEL, G. A., SHRIVASTAVA, S. P. and GEWALI, M. B. (2001): Evidence for paclitaxel from three new endophytic fungi of Himalayan yew of Nepal. *Planta Medica*, **67**(4): 374-376.

SHWAB, E. K. and KELLER, N. P. (2008): Regulation of secondary metabolite production in filamentous ascomycetes. *Mycol Res*, **112**(Pt 2): 225-230.

SIEBER, T. N. (2007): Endophytic fungi in forest trees: are they mutualists? *Fungal Biology Reviews*, **21**(2-3): 75-89.

SILVERMAN, R. (2000): Monooxygenation. The Organic Chemistry of Enzyme-Catalyzed Reactions. San Diego, Academic Press**:** 175-226.

SIM, J. H., KHOO, C. H., LEE, L. H. and CHEAH, Y. K. (2010): Molecular diversity of fungal endophytes isolated from *Garcinia mangostana* and *Garcinia parvifolia*. *J Microbiol Biotechnol*, **20**(4): 651-658.

SLOT, J. C. and ROKAS, A. (2011): Horizontal transfer of a large and highly toxic secondary metabolic gene cluster between fungi. *Curr Biol*, **21**(2): 134-139.

SMITH, D. J., BURNHAM, M. K., EDWARDS, J., EARL, A. J. and TURNER, G. (1990): Cloning and heterologous expression of the penicillin biosynthetic gene cluster from *penicillum chrysogenum*. *Biotechnology (N Y)*, **8**(1): 39-41.

SMITH, M. E., DOUHAN, G. W. and RIZZO, D. M. (2007): Intra-specific and intra-sporocarp ITS variation of ectomycorrhizal fungi as assessed by rDNA sequencing of sporocarps and pooled ectomycorrhizal roots from a Quercus woodland. *Mycorrhiza*, **18**(1): 15-22.

SOCA-CHAFRE, G., RIVERA-ORDUNA, F. N., HIDALGO-LARA, M. E., HERNANDEZ-RODRIGUEZ, C., MARSCH, R. and FLORES-COTERA, L. B. (2011): Molecular phylogeny and paclitaxel screening of fungal endophytes from *Taxus globosa*. *Fungal Biol*, **115**(2): 143-156.

SREEKANTH, D., SYED, A., SARKAR, S., SARKAR, D., SANTHAKUMARI, B., AHMAD, A. and KHAN, M. I. (2009): Production, purification, and characterization of taxol and 10-DABIII from a new endophytic fungus *Gliocladium sp*. isolated from the Indian yew tree, *Taxus baccata*. *Journal of Microbiology and Biotechnology*, **19**(11): 1342-1347.

SRINIVASAN, V., PESTCHANKER, L., MOSER, S., HIRASUMA, T. J., TATICEK, R. A. and SHULER, M. L. (1995): Taxol production in bioreactors: kinetics of biomass accumulation, nutrient uptake, and taxol production by cell suspension of *Taxus baccata*. *Biotechnology and Bioengineering*, **47**(6): 666-676.

STANIEK, A., WOERDENBAG, H. J. and KAYSER, O. (2009): Taxomyces andreanae: A Presumed Paclitaxel Producer Demystified? *Planta Medica*, **75**(15): 1561-1566.

STANIEK, A., WOERDENBAG, H. J. and KAYSER, O. (2010): Screening the endophytic flora of Wollemia nobilis for alternative paclitaxel sources. *J. Plant Interact.*, **5**(3): 189-195.

STAUNTON, J. and WEISSMAN, K. J. (2001): Polyketide biosynthesis: a millennium review. *Nat Prod Rep*, **18**(4): 380-416.

STIERLE, A., STIERLE, D. and STROBEL, G. (2000): Taxol production by a microbe. **US 6013493**.

STIERLE, A., STROBEL, G. and STIERLE, D. (1993): Taxol and taxane production by *Taxomyces andreanae*, an endophytic fungus of Pacific yew. *Science*, **260**(5105): 214-216.

STIERLE, A., STROBEL, G., STIERLE, D., GROTHAUS, P. and BIGNAMI, G. (1995): The search for a taxol-producing microorganisms among the endophytic fungi of the Pacific yew, *Taxus brevifolia*. *Journal of Natural Products*, **58**(9): 1315-1324.

STROBEL, G. and DAISY, B. (2003): Bioprospecting for microbial endophytes and their natural products. *Microbiology and Molecular Biology Reviews*, **67**(4): 491-502.

STROBEL, G., DAISY, B., CASTILLO, U. and HARPER, J. (2004): Natural products from endophytic microorganisms. *Journal of Natural Products*, **67**(2): 257-268.

STROBEL, G., STIERLE, A. and STIERLE, D. (1994): Taxol production by *Taxomyces andreanae*. **US 5322779**.

STROBEL, G., STIERLE, A. and STIERLE, D. (2001): Taxol production by a microbe. **US 6013493B1**.

STROBEL, G., YANG, X., SEARS, J., KRAMER, R., SIDHU, R. S. and HESS, W. M. (1996): Taxol from *Pestalotiopsis microspora*, an endophytic fungus of *Taxus wallachiana*. *Microbiology (Reading, United Kingdom)*, **142**(2): 435-440.

STROBEL, G. A. (2002): Rainforest endophytes and bioactive products. *Crit. Rev. Biotechnol.*, **22**(4): 315-333.

STROBEL, G. A., FORD, E., LI, J. Y., SEARS, J., SIDHU, R. S. and HESS, W. M. (1999): *Seimatoantlerium tepuiense* gen. nov., a unique epiphytic fungus producing taxol from the Venezuelan Guyana. *Systematic and Applied Microbiology*, **22**(3): 426-433.

STROBEL, G. A., HESS, W. M., LI, J.-Y., FORD, E., SEARS, J., SIDHU, R. S. and SUMMERELL, B. (1997): *Pestalotiopsis guepinii*, a taxol-producing endophyte of the wollemi pine, *Wollemia nobilis*. *Aust. J. Bot.*, **45**(6): 1073-1082.

SUFFNESS, M. and WALL, M. E. (1995): Discovery and developement of taxol. Taxol science and applications. SUFFNESS, M. Boca Raton, CRC Press**:** 1-25.

TANIMOTO, S., DAEMEN, J. and SERRUYS PATRICK, W. (2007): Update on stents: recent studies on the *TAXUS* stent system in small vessels. *Vasc Health Risk Manag FIELD Full Journal Title:Vascular health and risk management*, **3**(4): 481-490.

TOYOMASU, T., NAKAMINAMI, K., TOSHIMA, H., MIE, T., WATANABE, K., ITO, H., MATSUI, H., MITSUHASHI, W., SASSA, T. and OIKAWA, H. (2004): Cloning of a gene cluster responsible for the biosynthesis of diterpene aphidicolin, a specific inhibitor of DNA polymerase alpha. *Bioscience Biotechnology and Biochemistry*, **68**(1): 146-152.

TOYOMASU, T., NUDA, R., KENMOKU, H., KANNO, Y., MIURA, S., NAKANO, C., SHIONO, Y., MITSUHASHI, W., TOSHIMA, H., OIKAWA, H., HOSHINO, T., DAIRI, T., KATO, N. and SASSA, T. (2008): Identification of diterpene biosynthetic gene clusters and functional analysis of labdane-related diterpene cyclases in *Phomopsis amygdali*. *Bioscience Biotechnology and Biochemistry*, **72**(4): 1038-1047.

TOYOMASU, T., TSUKAHARA, M., KANEKO, A., NIIDA, R., MITSUHASHI, W., DAIRI, T., KATO, N. and SASSA, T. (2007): Fusicoccins are biosynthesized by an unusual chimera diterpene synthase in fungi. *Proceedings of the National Academy of Sciences of the United States of America*, **104**(9): 3084-3088.

TRAPP, S. C. and CROTEAU, R. B. (2001): Genomic organization of plant terpene synthases and molecular evolutionary implications. *Genetics*, **158**(2): 811-832.

TUDZYNSKI, B. and HÖLTER, K. (1998): Gibberellin biosynthetic pathway in *Gibberella fujikuroi*: Evidence for a gene cluster. *Fungal Genetics and Biology*, **25**(3): 157-170.

TUDZYNSKI, B., ROJAS, M. C., GASKIN, P. and HEDDEN, P. (2002): The gibberellin 20-oxidase of *Gibberella fujikuroi* is a multifunctional monooxygenase. *J Biol Chem*, **277**(24): 21246-21253.

VERDIN, A., LOUNES-HADJ SAHRAOUI, A., NEWSAM, R., ROBINSON, G. and DURAND, R. (2005): Polycyclic aromatic hydrocarbons storage by *Fusarium solani* in intracellular lipid vesicles. *Environ Pollut*, **133**(2): 283-291.

WALKER, K. and CROTEAU, R. (2000a): Molecular cloning of a 10-deacetylbaccatin III-10-O-acetyl transferase cDNA from *Taxus* and functional expression in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, **97**(2): 583-587.

WALKER, K. and CROTEAU, R. (2000b): Taxol biosynthesis: Molecular cloning of a benzoyl-CoA : taxane 2 alpha-O-benzoyltransferase cDNA from Taxus and functional expression in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, **97**(25): 13591-13596.

WALKER, K., FUJISAKI, S., LONG, R. and CROTEAU, R. (2002a): Molecular cloning and heterologous expression of the C-13 phenylpropanoid side chain-CoA acyltransferase that functions in Taxol biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, **99**(20): 12715-12720.

WALKER, K., LONG, R. and CROTEAU, R. (2002b): The final acylation step in Taxol biosynthesis: Cloning of the taxoid C13-side-chain N-benzoyltransferase from *Taxus*. *Proceedings of the National Academy of Sciences of the United States of America*, **99**(14): 9166-9171.

WALKER, K., SCHOENDORF, A. and CROTEAU, R. (2000): Molecular cloning of a taxa-4(20),11(12)-dien-5 alpha-ol-O-acetyl transferase cDNA from *Taxus* and functional expression in Escherichia coli. *Archives of Biochemistry and Biophysics*, **374**(2): 371-380.

WALKER, K. D. and FLOSS, H. G. (1998): Detection of a phenylalanine aminomutase in cell-free extracts of Taxus brevifolia and preliminary characterization of its reaction. *Journal of the American Chemical Society*, **120**(21): 5333-5334.

WALKER, K. D., KLETTKE, K., AKIYAMA, T. and CROTEAU, R. (2004): Cloning, heterologous expression, and characterization of a phenylalanine aminomutase involved in Taxol biosynthesis. *Journal of Biological Chemistry*, **279**(52): 53947-53954.

WALTON, J. D. (2000): Horizontal gene transfer and the evolution of secondary metabolite gene clusters in fungi: an hypothesis. *Fungal Genet Biol*, **30**(3): 167-171.

WANG, J. F., LI, G. L., LU, H. Y., ZHENG, Z. H., HUANG, X. J. and SU, W. J. (2000): Taxol from *Tubercularia sp* strain TF5, an endophytic fungus of *Taxus mairei*. *Fems Microbiology Letters*, **193**(2): 249-253.

WANG, Y. and TANG, K. (2011): A new endophytic taxol- and baccatin III-producing fungus isolated from *Taxus chinensis var*. *mairei*. *African Journal of Biotechnology*, **10**(72): 16379-16386.

WANI, M. C., TAYLOR, H. L., WALL, M. E., COGGON, P. and MCPHAIL, A. T. (1971): Plant antitumor agents. VI. The isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. *J Am Chem Soc*, **93**(9): 2325-2327.

WEIGEL, B. J., BURGETT, S. G., CHEN, V. J., SKATRUD, P. L., FROLIK, C. A., QUEENER, S. W. and INGOLIA, T. D. (1988): Cloning and expression in *Escherichia coli* of isopenicillin N synthetase genes from *Streptomyces lipmanii* and Aspergillus nidulans. *J Bacteriol*, **170**(9): 3817-3826.

WHEELER, A. L., LONG, R. M., KETCHUM, R. E. B., RITHNER, C. D., WILLIAMS, R. M. and CROTEAU, R. (2001): Taxol biosynthesis: Differential transformations of taxadien-5 alpha-ol and its acetate ester by cytochrome P450 hydroxylases from Taxus suspension cells. *Archives of Biochemistry and Biophysics*, **390**(2): 265-278.

WILDUNG, M. R. and CROTEAU, R. (1996): A cDNA clone for taxadiene synthase, the diterpene cyclase that catalyzes the committed step of taxol biosynthesis. *Journal of Biological Chemistry*, **271**(16): 9201-9204.

WILLIAMS, D. C., CARROLL, B. J., JIN, Q., RITHNER, C. D., LENGER, S. R., FLOSS, H. G., COATES, R. M., WILLIAMS, R. M. and CROTEAU, R. (2000a): Intramolecular proton transfer in the cyclization of geranylgeranyl diphosphate to the

taxadiene precursor of taxol catalyzed by recombinant taxadiene synthase. *Chem Biol*, **7**(12): 969-977.

WILLIAMS, D. C., WILDUNG, M. R., JIN, A. Q., DALAL, D., OLIVER, J. S., COATES, R. M. and CROTEAU, R. (2000b): Heterologous expression and characterization of a "Pseudomature" form of taxadiene synthase involved in paclitaxel (Taxol) biosynthesis and evaluation of a potential intermediate and inhibitors of the multistep diterpene cyclization reaction. *Arch Biochem Biophys*, **379**(1): 137-146.

WINK, M. (2008): Plant secondary metabolism: diversity, function and its evolution. *Natural Product Communications*, **3**: 1205-1216.

WONG, C., SRIDHARA, S., BARDWELL, J. C. and JAKOB, U. (2000): Heating greatly speeds Coomassie blue staining and destaining. *Biotechniques*, **28**(3): 426-428, 430, 432.

WU, Y., HILLWIG, M. L., WANG, Q. and PETERS, R. J. (2011): Parsing a multifunctional biosynthetic gene cluster from rice: Biochemical characterization of CYP71Z6 & 7. *FEBS Lett*, **585**(21): 3446-3451.

XIAO, Z., ITOKAWA, H. and LEE, K. H. (2003): Total synthesis of Taxol. *The Genus Taxus*: 245-297.

YOUNG, D. H., MICHELOTTI, E. L., SWINDELL, C. S. and KRAUSS, N. E. (1992): Antifungal Properties of Taxol and Various Analogs. *Experientia*, **48**(9): 882-885.

YUKIMUNE, Y., TABATA, H., HIGASHI, Y. and HARA, Y. (1996): Methyl jasmonate-induced overproduction of paclitaxel and baccatin III in *Taxus* cell suspension cultures. *Nat Biotechnol*, **14**(9): 1129-1132.

ZHANG, P., ZHOU, P.-P., JIANG, C., YU, H. and YU, L.-J. (2008): Screening of Taxol-producing fungi based on PCR amplification from *Taxus*. *Biotechnol. Lett.*, **30**(12): 2119-2123.

ZHANG, P., ZHOU, P.-P. and YU, L.-J. (2009a): An Endophytic Taxol-Producing Fungus from *Taxus media*, *Cladosporium cladosporioides* MD2. *Curr. Microbiol.*, **59**(3): 227-232.

ZHANG, P., ZHOU, P. P. and YU, L. J. (2009b): An endophytic taxol-producing fungus from *Taxus x media*, *Aspergillus candidus* MD3. *Fems Microbiology Letters*, **293**(2): 155-159.

ZHANG, S., MONAHAN, B. J., TKACZ, J. S. and SCOTT, B. (2004): Indole-diterpene gene cluster from *Aspergillus flavus*. *Applied and Environmental Microbiology*, **70**(11): 6875-6883.

ZHAO, K., PING, W., LI, Q., HAO, S., ZHAO, L., GAO, T. and ZHOU, D. (2009): Aspergillus niger var. taxi, a new species variant of taxol-producing fungus isolated from *Taxus cuspidata* in China. *J. Appl. Microbiol.*, **107**(4): 1202-1207.

ZHAO, L. and FENG, S. S. (2004): Effects of lipid chain length on molecular interactions between paclitaxel and phospholipid within model biomembranes. *J Colloid Interface Sci*, **274**(1): 55-68.

ZHOU, X., WANG, Z., JIANG, K., WEI, Y., LIN, J., SUN, X. and TANG, K. (2007): Screening of taxol-producing endophytic fungi from *Taxus chinensis var. mairei*. *Applied Biochemistry and Microbiology*, **43**(4): 439-443.

ZHOU, X., ZHU, H., LIU, L., LIN, J. and TANG, K. (2010): A review: recent advances and future prospects of taxol-producing endophytic fungi. *Appl Microbiol Biotechnol*, **86**(6): 1707-1717.

VIII Curriculum vitae & publications

VIII.1 *Curriculum vitae*

Name: Uwe Herbert Heinig
Date of birth: 02.09.1979
Place of birth: Freital, Germany
Nationality: German

Qualifications:

2007 – 2012 Ph.D. thesis at RWTH Aachen, Germany, Research performed at Fraunhofer institute of Molecular Biology and Applied Ecology, Aachen, Germany, Department industrial biotechnology
Supervisor: Dr. S. Jennewein (Fraunhofer IME, Aachen), Prof. Dr. R. Fischer (RWTH, Aachen)
Thesis title: "Studies on the evolution of complex natural products biosynthetic pathways on the basis of Taxol-biosynthesis in plants and endophytic fungi"

2000 – 2006 Graduate engineer, chemistry, major: biochemistry, at Technical University Darmstadt, Germany, Department of organic chemistry
Supervisor: Prof. Dr. W.-D. Fessner, Dr. S. Jennewein
Thesis title: "Klonierung der Gene der Taxol®-Biosynthese aus *Taxomyces andreanae*"

1999 Abitur, Hermann-Staudinger Gymnasium, Erlenbach, Germany

VIII.2 *Publications & presentations*

Related to the thesis:

- **Uwe Heinig**, Stefan Jennewein; Taxol: A complex diterpenoid natural product with an evolutionarily obscure origin; *African Journal of Biotechnology; 8(8);* **2009**; 1370-1385.

- **Uwe Heinig**, Susanne Scholz, Pia Dahm, Udo Grabowy, Stefan Jennewein; Development of carbon plasma-coated multiwell plates for high-throughput mass spectrometric analysis of highly lipophilic fermentation products; *Analytical Biochemistry*; *403 (1-2)*; **2010**; 108-113.

- **Uwe Heinig**, Susanne Scholz, Stefan Jennewein; Getting to the bottom of Taxol biosynthesis by fungi; *Fungal Diversity*; **2012**; in preparation.

- Oral presentation at Volkswagen Foundation status symposium **2009**
  Title: "Taxol: A complex diterpenoid natural product with an evolutionarily obscure origin"

Others:

- Benedikt Engels, **Uwe Heinig**, Torsten Grothe, Marc Stadler, Stefan Jennewein; Cloning and characterization of an *Armillaria gallica* cDNA encoding protoilludene synthase, which catalyzes the first committed step in the synthesis of antimicrobial melleolides; *Journal of Biological Chemistry; 286(9)*; **2011**; 6871-6878.

- Poster presentation as a co-author at 9th VAAM symposium "Molecular Biology of Fungi" **2009**, *Title:* "Purification of protoilludene synthase from *Armillaria gallica"*

IX Acknowledgement

I want to thank Prof. Dr. Rainer Fischer for giving me the chance to work as a PhD student at Fraunhofer IME, providing an excellent environment for me as a young scientist.

I further thank Prof. Dr. Marc Stadler for reviewing my thesis as well as Prof. Dr. Ralph Panstruga and Prof. Dr. Jan Schirawski for participation in the exam comitee.

Next I want to thank the Volkswagen Stiftung, Hannover for funding this work in the initiative “evolutionary biology” not only giving me the financial support for my thesis, but furthermore providing me the opportunity to stay in the lab of Prof. Asaph Aharoni at Weizmann Institute of Science, Israel during the PhD. This was an outstanding experience scientifically as well as personally.

Of course big thanks to Dr. Stefan Jennewein. We first met at Darmstadt nearly 10 years ago. Since than we went together on the “path of science” and with his help I was able to develop from an undergraduate student of chemistry to the PhD I am now. Despite of all difficulties I have found my place in the world of science. Thanks for this Stefan.

Furthermore I want to thank the IME people, many of them going the entire five years in Aachen with me. Thank you for all the support with respect to work and much more important becoming real friends Ben, Pia, Raph, Claudi, Kläuschen, Suse, Ossi, Burk, Jens, Bernd, Gustl, Eike, Daniel, Tanja, Markus and many more. Without you all it would not have been possible! It was an amazing time in Aachen.

Last but not least I want to thank my parents who always believed in me and supported me all the way from starting studying up to now. Thank you so much.