

DISS. ETH NO. 16410

**MOLECULAR IDENTIFICATION OF FUNGAL PATHOGENS  
USING THE ITS REGION**

A dissertation submitted to the  
SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZÜRICH  
for the degree of  
Doctor of natural sciences

presented by  
DIANA EVA CIARDO  
Dipl. Natw. ETH

born 4. 10. 1977  
citizen of Zug (ZG), Seon (AG) and Depressa (LE, Italy)

accepted on the recommendation of  
Prof. Dr. Markus Aebi, examiner  
Prof. Dr. Bruce McDonald, co-examiner  
Prof. Dr. Martin Altwegg, co-examiner  
Dr. Philipp Bosshard, co-examiner

Zürich 2006

## Abstract

Fungi are gaining importance as infectious agents due to an increasing number of severely ill and immunocompromised patients. Identification of isolates is essential for proper treatment, especially because certain species show an inherent resistance to antimycotic agents. Traditional phenotypic methods may be ambiguous and occasionally result in misidentification, whereas molecular methods promise more rapid and reliable identification. In this project, we have chosen sequence analyses of the internal transcribed spacer (ITS) region, located between the 18S and the 28S rRNA genes and including the highly variable ITS1 and ITS2 regions and the 5.8S rRNA gene, as identification procedure. Because mislabeled as well as incomplete sequences and sequences of low quality may be found in public sequence databases, we decided to build our own in-house database, including most pathogenic fungi and common contaminants occurring in the medical diagnostic laboratory. Only confirmed sequences were included, and extensive analyses were performed to define sequence homologies between different species. For evaluation of this database, two studies were designed. First, ITS based sequence analysis was compared to identification by biochemical properties for yeasts and second, molecular identification was compared to phenotypic identification of moulds. In both instances, sequence analysis resulted in a more accurate identification. In the diagnostic laboratory, the use of sequencing should be limited due to costs and extensive hands-on time required for this procedure. Therefore, algorithms were developed for reasonable application of the available identification methods in the diagnostic laboratory. To further molecular identification, DNA extraction methods and amplification were optimized and hybridization assays for the most important pathogenic species were developed.

## Zusammenfassung

Mit der Zunahme von schwer kranken und immunkompromittierten Patienten wächst auch die Bedeutung von Pilzen als Infektionserreger. Eine Identifizierung der Isolate ist essentiell für die Auswahl der optimalen Behandlung, vor allem da verschiedene Spezies unterschiedliche Empfindlichkeit gegen die gebräuchlichen Antimykotika zeigen. Traditionelle phänotypische Methoden sind manchmal unklar und können zu Falschidentifizierungen führen. Molekulare Methoden versprechen eine schnellere und zuverlässigere Identifizierung. In diesem Projekt haben wir die Sequenzanalyse des internen transkribierten Zwischenstücks (ITS) als Identifizierungsmethode gewählt. Die ITS Region liegt zwischen den 18S und 28S rRNA Genen und umfasst die hochvariablen Abschnitte ITS1 und ITS2 und das 5.8S rRNA Gen. Weil falsch benannte, unvollständige und qualitativ schlechte Sequenzen in öffentlichen Sequenzdatenbanken hinterlegt sind, haben wir eine eigene institutsinterne Datenbank aufgebaut. Diese umfasst die meisten pathogenen Pilze und die häufigsten Kontaminanten, die im medizinischen Diagnostiklabor vorkommen. Es wurden ausschliesslich bestätigte Sequenzen eingeschlossen, und ausgedehnte Analysen ermöglichten Sequenzhomologien zwischen den einzelnen Spezies festzulegen. Um diese Datenbank zu evaluieren, wurden zwei Studien durchgeführt. Identifizierung mittels Analyse der ITS-Sequenz wurde mit der phänotypischen Identifizierung von Schimmelpilzen und mit der biochemischen Identifizierung von Hefen verglichen. In beiden Fällen führte die Sequenzierung zu einer präziseren Identifizierung. Im diagnostischen Labor sollte die Sequenzanalyse aus Aufwand- und Kostengründen nicht zu häufig eingesetzt werden. Deshalb wurden Algorithmen zur sinnvollen Verwendung der vorhandenen Identifizierungsmethoden entwickelt. Zur weiteren Verbesserung der molekularen Identifizierung wurden nicht nur DNA-Extraktionsmethode und Amplifikation optimiert, sondern auch Hybridisierungsassays für die Detektion der häufigsten pathogenen Pilze entwickelt.

## Riassunto

Funghi acquisiscono sempre più importanza come agenti patogeni a causa dell'aumento di pazienti con malattie gravi e con sistema immunitario compromesso. L'identificazione del fungo patogeno è essenziale per la scelta del trattamento appropriato, particolarmente siccome certe specie dimostrano una resistenza inerente ai trattamenti antimicotici. Il metodo tradizionale basato sulla fenotipia è ambiguo e può condurre ad un'identificazione falsa, mentre metodi molecolari permettono un'identificazione univoca e attendibile. In questo progetto abbiamo scelto come metodo d'identificazione l'analisi della sequenza della regione interna trascritta spaziatrice (ITS), situata tra i geni di 18S e 28S rRNA e contenente le regioni altamente variabili ITS1 e ITS2 e il gene conservato di 5.8S rRNA. Nelle banche dati di sequenze pubbliche è possibile trovare sequenze denominate falsamente, incomplete e di qualità inferiore. Per questa ragione abbiamo costruito una banca dati propria contenente sequenze della regione ITS della maggioranza dei funghi patogeni e di specie contaminanti trovate frequentemente nel laboratorio diagnostico. Solo sequenze confermate furono incluse, e analisi dettagliate furono condotte per definire l'omologia di sequenze tra specie diverse. Per valutare la nostra banca dati furono eseguiti due studi. L'analisi di sequenze della regione ITS fu paragonata all'identificazione basata su proprietà biochimiche di feccia e all'identificazione fenotipica di muffa. In ambedue i casi, con l'analisi di sequenze fu ottenuta un'identificazione più precisa. Nel laboratorio diagnostico è consigliabile di limitare l'uso dell'analisi di sequenze a causa del costo elevato e dell'impiego di tempo. Per questo furono sviluppati degli algoritmi per l'uso sensato dei metodi esistenti nel laboratorio diagnostico. Per migliorare l'identificazione molecolare fu ottimata l'estrazione e l'amplificazione del DNA e fu sviluppato un metodo di detezione dei funghi patogeni più importanti basato sull'ibridazione.