



Universiteit
Leiden
The Netherlands

A functional genomics study of extracellular protease production by *Aspergillus niger*

Braaksma, M.

Citation

Braaksma, M. (2010, December 15). *A functional genomics study of extracellular protease production by Aspergillus niger*. Retrieved from <https://hdl.handle.net/1887/16246>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/16246>

Note: To cite this publication please use the final published version (if applicable).

A FUNCTIONAL GENOMICS STUDY OF EXTRACELLULAR
PROTEASE PRODUCTION BY *ASPERGILLUS NIGER*

Machtelt Braaksma

A functional genomics study of extracellular
protease production by *Aspergillus niger*

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. P.F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 15 december 2010
klokke 15:00 uur
door

Machtelt Braaksma

geboren te Stadskanaal
in 1977

PROMOTION COMMITTEE

Promotoren:	Prof. dr. P.J. Punt Prof. dr. C.A.M.J.J. van den Hondel
Co-promoter:	Dr. ir. M.J. van der Werf (DSM)
Other members:	Prof. dr. P.J.J. Hooykaas Prof. dr. J.H. de Winde (Technische Universiteit Delft) Prof. dr. A.K. Smilde (Universiteit van Amsterdam) Dr. M. Saloheimo (VTT Technical Research Centre of Finland)

The work described in this thesis was financially supported by the Kluyver Centre for Genomics of Industrial Fermentation and carried out at TNO Quality of Life.

Printed by: GVO drukkers & vormgevers B.V. | Ponsen & Looijen

CONTENTS

	Outline	7
Chapter 1	<i>Aspergillus</i> as a cell factory for protein production: controlling protease activity in fungal production	9
Chapter 2	The effect of environmental conditions on extracellular protease activity in controlled fermentations of <i>Aspergillus niger</i>	27
Chapter 3	An inventory of the <i>Aspergillus niger</i> secretome by combining <i>in silico</i> predictions with shotgun proteomics data	47
Chapter 4	Metabolomics as a tool for target identification in strain improvement: the influence of phenotype definition	67
Chapter 5	Identification of modules in <i>Aspergillus niger</i> by gene co-expression network analysis	89
Chapter 6	A top-down systems biology approach for the identification of targets for fungal strain and process development	115
	Summary	131
	Samenvatting	135
	References	141
	List of publications	157
	Curriculum vitae	159

OUTLINE

The filamentous fungus *Aspergillus niger* has a long track record as a highly efficient producer of a wide variety of enzymes. Already soon after the development of fungal transformation systems this species was acknowledged for its potential as a production host for heterologous proteins. However, the production of homologous and especially heterologous proteins is often limited by the high levels of proteases produced by this fungus as well. **Chapter 1** reviews the role that protease activity plays in strain and process development of *A. niger* and other aspergilli. It discusses several approaches and techniques that have been applied to generate strains with reduced protease activity. Furthermore, it provides an outlook on how new research approaches, such as the -omics techniques, may play a role in understanding the proteolytic system of aspergilli.

The objective of the project described in this thesis is to study the complex induction of extracellular proteases in *A. niger* using information gathered with functional genomics technologies. A special emphasis is given to the requirements for performing a successful systems biology study and addressing the challenges met in analyzing the large, information-rich data sets generated with functional genomics technologies.

Chapter 2 of this thesis describes a systematic study of the influence of several environmental factors on the production of extracellular proteases of *A. niger* in controlled batch cultivations. Using a change-one-factor-at-a-time approach, the effect of pH and various medium components on protease production was investigated. Subsequently, a full two-level factorial design was applied with four environmental factors selected from the screening experiments that affected the protease production the most. Six protease-related quantitative phenotypes were calculated from these samples to study the individual and interaction effects of the tested environmental factors on each of these phenotypes. Samples generated in this full factorial experimental design were used for analysis with different functional genomics technologies (Chapter 3 to 5).

Chapter 3 presents an improved list of potential signal peptide directed proteins encoded by the *A. niger* genome. For the compilation of this list, the signal peptide predictions from *A. niger* were compared to those of the best homologs of three neighbouring *Aspergillus* species. In addition, a shotgun proteomics approach was used to determine the *A. niger* secretome under different experimental conditions.

Based on this analysis the complexity of the repertoire of secreted proteases was confirmed.

The effect of different quantitative phenotypes related to product formation on the information content of a metabolomics data set is investigated in **Chapter 4**. For this purpose, besides the production of secreted proteases the production of another industrially relevant product by *A. niger* was evaluated, i.e. the enzyme glucoamylase. For both products, different quantitative phenotypes associated with activity and productivity were defined and for each phenotype the relation with metabolome data was investigated. Results showed that, depending on the product studied, different quantitative phenotypes had the highest information content in relation to the metabolomics data set.

Chapter 5 describes the clustering of co-expressed genes using two DNA microarray data sets; one of these data sets was derived from the experiments described in Chapter 2. A set of conserved genes was used to construct gene co-expression networks for both the individual and combined data sets. By comparative analysis the existence of modules was revealed, some of which are present in all three networks. Subsequently, all protein-coding *A. niger* genes, including hypothetical and poorly conserved genes, were integrated into the co-expression analysis. We have used this two-step approach to relate the genes encoding hypothetical proteins to the identified functional modules.

In top-down systems biology the information gathered with functional genomics technologies is analyzed with multivariate data analysis tools, and can be used as a method to achieve unbiased selection and ranking of targets for both strain improvement and bioprocess optimization. **Chapter 6** discusses the key factors for a successful top-down systems biology approach.