# Probabilistic Models for Gene Silencing Data

Florian Markowetz

Dezember 2005

Dissertation zur Erlangung des Grades
eines Doktors der Naturwissenschaften (Dr. rer. nat.)
am Fachbereich Mathematik und Informatik
der Freien Universität Berlin

1. Referent: Prof. Dr. Martin Vingron
2. Referent: Prof. Dr. Klaus-Robert Müller

Tag der Promotion: 26. April 2006

# Preface

**Acknowledgements** This work was carried out in the *Computational Diagnostics* group of the Department of Computational Molecular Biology at the Max Planck Institute for Molecular Genetics in Berlin. I thank all past and present colleagues for the good working atmosphere and the scientific—and sometimes maybe not so scientific—discussions.

Especially, I am grateful to my supervisor *Rainer Spang* for suggesting the topic, his scientific support, and the opportunity to write this thesis under his guidance. I thank *Michael Boutros* for providing the expression data and for introducing me to the world of RNAi when I visited his lab at the DKFZ in Heidelberg. I thank *Anja von Heydebreck*, *Jörg Schultz*, and *Martin Vingron* for their advice and counsel as members of my PhD commitee.

During the time I worked on this thesis, I enjoyed fruitful discussions with many people. In particular, I gratefully acknowledge *Jacques Bloch*, *Steffen Grossmann*, *Achim Tresch*, and *Chen-Hsiang Yeang* for their contributions. Special thanks go to *Viola Gesellchen*, *Britta Koch*, *Stefanie Scheid*, *Stefan Bentink*, *Stefan Haas*, *Dennis Kostka*, and *Stefan Röpcke*, who read drafts of this thesis and greatly improved it by their comments.

**Publications** Parts of this thesis have been published before. Chapter 2 grew out of lectures I gave in 2005 at the *Instiute for Theoretical Physics and Mathematics* (IPM) in Tehran, Iran, and at the *German Conference on Bioinformatics* (GCB), Hamburg, Germany. I thank Prof. Mehrdad Shahshahani (Tehran) and Prof. Stefan Kurtz (Hamburg) for inviting me. Chapter 3 gathers results of two conference papers: the first one at the *Workshop on Distributed Statistical Computing* (DSC 2003) in Vienna, Austria [82], and the second one at the *Conference on Artifical Intelligence and Statistics* (AISTATS), in Barbados, 2005 [81]. Parts of chapter 4 were previously published in the journal *Bioinformatics* [80].

**Figures** This thesis reproduces three figures from other publications. I thank Prof. Danny Reinberg and Prof. Jules Hoffmann for the friendly permissions to reproduce Fig. 1.1 and Fig. 1.2, respectively. Fig. 1.3 is reproduced with permission from `www.ambion.com`.

Florian Markowetz Berlin, December 2005

# Contents