# COVER PAGE

# TITLE PAGE

Possible titles:

* Investigation of long RNAs using a novel ligation-based approach
* Investigation of long RNAs using RNA templated DNA ligation
* Long RNAs: biology, technology, and perspective
* Examination of dynamic long RNAs

A Dissertation Presented

By

Christian Knauf Roy

Submitted to the Faculty of the

University of Massachusetts Graduate School of Biomedical Sciences, Worcester

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSPHY

(MONTH, DAY, YEAR)

BIOCHEMISTRY

# SIGNATURE PAGE

A Dissertation Presented

By

Christian Knauf Roy

The signatures of the Dissertation Defense Committee signify completion and approval as to style and content of the Dissertation

The signatures of the Dissertation Defense Committee signify completion and approval as to style and content of the Dissertation

(Signature) (Name Typed),

Thesis Advisor (Signature) (Name Typed),

Thesis Advisor (Signature) (Name Typed),

Member of Committee

(Signature) (Name Typed),

Member of Committee

(Signature) (Name Typed),

Member of Committee

(Signature) (Name Typed),

The signature of the Chair of the Committee signifies that the written dissertation meets the requirements of the Dissertation Committee

(Signature) (Name Typed),

Chair of Committee The signature of the Dean of the Graduate School of Biomedical Sciences signifies that the student has met all graduation requirements of the school. (Signature)

Anthony Carruthers, Ph.D., Dean of the Graduate School of Biomedical Sciences Program (Typed)

Month, Day and Year (Typed)

# Front Matter

## Dedication

## Acknowledgements

1. Talk about going to Grad School
2. Advisors
3. MJM
4. PDZ
5. Committee members
6. Zhiping
7. Scot
8. Job
9. Lab Members
10. Aaron
11. Alper
12. Amrit
13. Eric and Erin
14. Collaborators
15. Dave Weaver
16. Muro
17. Graveley
18. Heinrich
19. Anna
20. Ogo
21. Family

Table of Contents

Front Matter 3

Dedication 3

Acknowledgements 4

Table of Contents 6

Abstract 7

List of Tables 7

List of Figures 8

List of Symbols, Abbreviations or Nomenclature (optional) 9

Preface 10

Body Matter 11

Introduction 11

Research Chapters 11

Back Matter 12

Appendices 12

Biblography 12

## Abstract

## List of Tables

## List of Figures

## List of Symbols, Abbreviations or Nomenclature (optional)

### Abbreviations

|  |  |
| --- | --- |
| AS | Alternative Splicing |
| DNA | Deoxyribonucleic acid |
| RNA | Ribonucleic acid |
| ChIP-Seq | Chromatin Immunoprecipitation followed by sequencing |
| HTS | High-throughput sequencing (see also *NGS*) |
| NGS | Next-generation sequencing |
|  |  |

### Symbols

|  |  |
| --- | --- |
| 5´ |  |
| 3´ |  |
| μ |  |
|  |  |

### Definitions

RNA-Seq

## Preface

*Here is where you will put wording discussing previous publications*

*Here is example wording from the Gupta thesis*

The work reported in this dissertation has been published in the following articles.

Chapter II has been published previously as: Ankit Gupta, Xiangdong Meng, Lihua J. Zhu, Nathan D. Lawson, and Scot A. Wolfe (2011). Zinc finger protein-dependent and -independent contributions to the in vivo off-target activity of zinc finger nucleases. Nucleic acids research 39, 381-392

Chapter III has been published previously as: Christensen RG, Gupta A, Zuo Z, Schriefer LA, Wolfe SA, Stormo GD (2011)

A modified bacterial one-hybrid system yields improved quantitative models of transcription factor specificity. Nucleic acids research 39, e83

Contents of Chapter IV have been accepted for publication Ankit Gupta, Ryan G. Christensen, Amy L. Rayla, Abirami Lakshmanan,

Gary D. Stormo, Scot A. Wolfe (2012)

An optimized two‐finger archive for ZFN‐mediated gene targeting, Nat. Methods, (Manuscript accepted).

1. A

# Body Matter

## Introduction

#### Movitation

Type your text under subheadings like this in paragraph form.

1. Topics that need to be discussed
2. Coordination in splicing
3. Introduction to piRNAs
4. Historical Review of the literature
   1. RNA Expression
   2. Splicing
   3. Alternative Splicing
   4. SR Proteins
   5. RNA-Processing
   6. mRNPs
5. piRNAs
   1. Small RNA
   2. Tissue specific RNAs
6. Technicalical limitations
   1. Sequencing history
   2. Deep sequencing history
   3. The Isoform problem
   4. lncRNAs
   5. Computaitonal tools
7. State of Current literature
   1. ENCODE
   2. Integration of different datasets for more complete transcirptonal pictures
      1. Mention lincRNAs
   3. Single Cell RNA-Seq
      1. 1. Shalek AK, Satija R, Adiconis X, et al. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature*. 2013:1–5. doi:10.1038/nature12172.
8. Relationship between alternative splicing and protein domains
9. 1. Light, S. & Elofsson, A. The impact of splicing on protein domain architecture. *Current opinion in structural biology* 1–8 (2013).doi:10.1016/j.sbi.2013.02.013

## Research Chapters

### CHAPTER I : SeqZip as a methodology

1. Nature methods paper will do 90% of the work here
2. SeqZip Development
   1. Discovery of novel enzyme activity
3. Search for long and complicated RNAs
   1. CD45
   2. FN1
4. You should tie together the fact that Dscam ALSO has Fibronectin sections!
5. DSCAM

### CHAPTER II : SeqZip applied to HIV and piRNA precursors

1. piRNA precursors
2. HIV Transcript integrity

### CHAPTER III : piRNA Chapter and Molecular Cell paper

## Perspective / Final Summary and Conclusions

Things to remember about this section:

1. Multiple people have stated that this can be highly speculative
2. This is my chance to show that I am a scientist
3. If I were to die tomorrow, what would be the experiments to conduct
4. Discuss long range RNA secondary structure and implecations of regulating alternative splicing (See Li, S., & Breaker, R. R. (2013). Eukaryotic TPP riboswitch regulation of alternative splicing involving long-distance base pairing. Nucleic acids research, 41(5), 3022–31. doi:10.1093/nar/gkt057) and Reg of AS by long RANGE SS folder in Mendeley
5. Full length analysis of mRNAs
6. Implications for discrination past one's DNA as it is the actual PRODUCT of the DNA and the actual biology (or at least closer to the functional biology) that is going on inside of every person
7. Catalogue of every possible mRNA isoform
8. Seperate the signal from the noise of transcription
9. For out there comments:
10. Most high cited pnas papers are methodology papers
11. www.pnas.org/reports/most-cited

# Back Matter

## Appendices

### Methods

#### SeqZip

#### Computational Scripts

#### Computational Approaches

##### Git

##### Perl

##### IGV

##### MySQL

#### ImageQuant

#### LOD of Radioactivity

### Oligo Database

### Important MySQL data tables

## Biblography