

**FUNCTIONAL GENOMICS STUDY TO UNDERSTAND THE ROLE
OF SEROTONIN IN MOUSE EMBRYONIC STEM CELLS**

Anusha Nagari

Submitted to the faculty of the School of Informatics
in partial fulfillment of the requirements
for the degree of
Master of Science in Bioinformatics,
Indiana University

May 2011

Accepted by the Faculty of Indiana University,
in partial fulfillment of the requirements for the degree of Master of Science
in Bioinformatics

**Master's Thesis
Committee**

Narayanan B. Perumal, Ph.D., Chair

Mitradas M. Panicker, Ph.D.

Yaoqi Zhou, Ph.D.

Meeta Pradhan, Ph.D.

© 2011

Anusha Nagari

ALL RIGHTS RESERVED

Dedicated to my family

TABLE OF CONTENTS

	Page
LIST OF TABLES.....	vii
LIST OF FIGURES.....	viii
ACKNOWLEDGEMENTS.....	ix
ABSTRACT.....	x
CHAPTER ONE: INTRODUCTION.....	1
1.1 Cell.....	1
1.1.1 Cell division.....	1
1.2 Stem cells.....	2
1.2.1 Biological role and properties of stem cells.....	3
1.2.1.1 Stem cells are self-renewing.....	3
1.2.1.2 Stem cells are unspecialized and differentiate into more specialized cells.....	3
1.2.2 Origin of stem cells.....	4
1.2.3 Types of stem cells.....	4
1.3 Embryonic stem cells.....	6
1.4 Multipotent stem cells.....	7
1.4.1 Neural stem cells.....	7
1.5 Stem cells in therapy and research.....	8
1.6 Neurotransmitters involved in stem cell development.....	9
1.7 Gene regulation.....	10
1.8 Trends and challenges in transcriptional regulation.....	13
1.9 Transcriptional regulation studies.....	14
1.10 Techniques in functional genomics studies.....	16
1.11 Computational advances in transcriptional regulation studies.....	19
CHAPTER TWO: BACKGROUND.....	21
2.1 Role of serotonin in ES cell differentiation.....	21
2.5 Serotonin in early development.....	21

2.3 High throughput studies in ES cells.....	23
2.4 Knowledge gap and motivation.....	24
2.5 Goals and objectives.....	25
CHAPTER THREE: METHODS.....	26
3.1 Microarray experiment.....	26
3.2 Quality assessment.....	29
3.2.1 Exploratory data analysis.....	29
3.2.1.1 Box plots.....	29
3.2.1.2 MA plots.....	30
3.3 Normalization of the data.....	30
3.4 Selection of differentially expressed genes.....	31
3.5 Sequence retrieval and repeat masking.....	33
3.6 Identification of TFBS.....	34
3.6.1 MEME.....	35
3.6.2 STAMP.....	37
3.6.3 MotifScanner.....	38
3.6.4 Motif enrichment analysis.....	39
3.7 Literature validation and comparison with public domain data.....	40
3.8 Functional annotation of 5-HT responsive genes.....	41
CHAPTER FOUR: RESULTS.....	43
4.1 Exploratory data analysis.....	43
4.2 Selection of differentially expressed 5-HT responsive genes.....	47
4.3 Identification of TFBS using motif prediction tools.....	51
4.4 Motif enrichment analysis.....	54
4.5 Literature validation and comparison with public domain data.....	61
4.6 Gene ontology analysis.....	64
CHAPTER FIVE: DISCUSSION.....	66
CHAPTER SIX: CONCLUSIONS.....	71
REFERENCES.....	73

LIST OF TABLES

4.1 Summary of ANOVA results.....	48
4.2 Summary of Volcano plot results.....	49
4.3 Summary of differentially expressed 5-HT responsive genelists.....	50
4.4 List of differentially expressed 5HT induced and suppressed genes after stringent filtering.....	50
4.5 List of top 15 hits from MotifScanner for genelist1 sequences.....	55
4.6 List of top 15 hits from MotifScanner for genelist3 sequences.....	55
4.7 Top 15 hits from MotifScanner for genelist2 sequences.....	56
4.8 Top 15 hits from MotifScanner for genelist4 sequences.....	57
4.9 List of 5HT induced genes having TFBS predicted by MotifScanner.....	58
4.10 List of 5-HT suppressed genes having the enriched TFs predicted by MotifScanner.....	58
4.11 List of genes common in our dataset, PDD1 and PDD2.....	64
4.12 Functional annotation of 5-HT responsive genes.....	65

LIST OF FIGURES

1.1 Stem cell hierarchy.....	5
3.1 Workflow of experimental analysis to understand the role of serotonin responsive genes in ES cell differentiation.....	28
4.1 Box plot of four normalized sample conditions.....	44
4.2 MA plot of the duplicates of four samples before normalization.....	45
4.3 MA plot for normalized samples.....	46
4.4 Volcano plots of Control Vs 5HT and P+C Vs P+C+5HT comparisons.....	48
4.5 Stamp results for MEME motifs enriched in 5HT induced gene sequences.....	52
4.6 STAMP results for MEME motifs enriched in 5HT-supressed gene sequences.....	53
4.7 MotifScanner results of enriched TFs in 5HT responsive genes.....	59
4.8 Enrichment of TFAP2A in FG of 5-HT induced gene lists.....	60
4.9 Enrichment of TFAP2A in FG of 5-HT suppressed gene lists.....	61
4.10 Comparison of 1.5 fold enriched genes in our dataset and Mikkelsen et al study.....	62
4.11 Comparison between 1.5 fold up regulated genes from our results and Ramalho-Santos et al study.....	63
4.12 Venn diagram comparing the common genes across our genotypic analysis PDD1, and PDD2.....	63

ACKNOWLEDGEMENTS

I would like to take the opportunity to acknowledge some of the people who made my graduate studies a memorable experience and made this thesis possible. Words are really inadequate to express the deep sense of my gratitude and regards towards my advisor Dr. Narayanan B. Perumal. His valuable advice, motivation, close supervision, feed back, and encouragement inspired me to complete this work successfully.

I would like to thank Dr. Mitradas M. Panicker, Dr. Yaoqi Zhou and Dr. Meeta Pradhan for their time, encouragement, insightful comments, and critical feedback. I greatly appreciate their efforts in reviewing my work. In spite of their tight work schedules, they have helped me unconditionally.

I owe my sincere thanks to Indiana University for providing the prestigious Dean's award during my Master's program. I would like to extend my sincere thanks to Mark McCreary, Dr. Pedro Romero, Dr. Malika Mahoui, and other faculty and staff of the department of Bioinformatics for their continuous support throughout the Masters program.

Above all, I express my deep gratitude and affection to my parents and my loving sister whose persistent love, warmth and their cooperation was an encouragement for successful completion of my thesis. I would like to extend my loving thanks to Mr. Rahul Krishna Kollipara for his support and guidance throughout my Master's program. I would also like to thank my friends, Sriya, Madhan, Satish, Keerthi chandra, Abhinita, Keerthi Kosuri, Rini, Gokhul Krishna, and Sushanth for their support and all the fun we have had in the last two years.

ABSTRACT

FUNCTIONAL GENOMICS STUDY TO UNDERSTAND THE ROLE OF SEROTONIN IN MOUSE EMBRYONIC STEM CELLS

Serotonin (5-hydroxytryptamine, 5-HT) is a monoamine neurotransmitter that is synthesized from the amino acid L-tryptophan and is reported to localize in mitochondria of embryonic stem cells. Even before its role as a neurotransmitter in mature brain was discovered, 5-HT has been shown to play an important role in regulating brain development. However, there is a lack of knowledge about the downstream target genes regulated by serotonin in embryonic stem (ES) cells. Towards this end, our study helps in understanding transcriptional regulatory mechanisms of 5-HT responsive genes in ES cells. By combining the gene expression data with motif prediction algorithms, literature validation and comparison with public domain data, gene targets specific to endogenous or exogenous 5-HT in ES cells were identified. By performing one-way ANOVA, and volcano plots using GeneSpring software, we identified 44 5-HT induced and 29 5-HT suppressed genes, likely to be transcriptionally regulated by 4 & 2 TFs respectively. Motif enrichment analysis on these target genes using MotifScanner revealed that the transcription factor TFAP2A plays a key role in regulating the expression of 5-HT responsive genes. Furthermore, by comparing our dataset with published expression profiles of ES cells, we observed a number of 5-HT responsive target genes showing enrichment in ES cells. Genes such as *Nanog*, *Slc38a5*, *Hoxb1* and *Eif2s1* from this analysis have been observed to be components of 'stemness' phenotypes reported in literature. Functional annotation of the 5-HT responsive genes identified gene ontologies such as regulation of translation in response to stress and energy derivation by oxidation, suggesting a regulatory role for 5-HT in mitochondrial functions of ES cells. Additionally, enrichment of other biological process terms such as development of various parts of nervous system, cell adhesion, and apoptosis suggests that 5-HT target genes may play an important role in ES cell differentiation. Our study implemented a new combinatorial approach for identifying gene regulatory mechanisms involved in 5-HT responsive genes and proposed potential mediatory role for serotonin in ES cell differentiation and growth. Thus, this study provides potential 5-HT target genes in ES cells for biological validation.