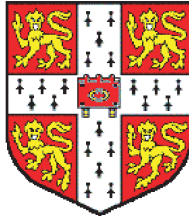


The role of microRNAs in neurons

Siarhei Manakou

Peterhouse



A dissertation submitted to the University of Cambridge
for the degree of Doctor of Philosophy

Wellcome Trust Sanger Institute,
Wellcome Trust Genome Campus,
Hinxton,
Cambridge.
CB10 1SA
UK

Email: siarheimanakov@gmail.com

June 1, 2011

To Alex, Leo, Marija,
Matias and Steve

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

This thesis does not exceed the specified length limit of 60,000 words as defined by the Biology Degree Committee.

This thesis has been typeset in 12pt font using L^AT_EX according to the specifications defined by the Board of Graduate Studies and the Biology Degree Committee.

Siarhei Manakou

“The role of microRNAs in neurons”

Abstract

Many individual functional microRNA (miRNA) targets have been identified in neurons, and their importance for neuronal differentiation is well established. However, with over 50% of genes in a mammalian genome being computationally predicted as miRNA targets, the global significance of the role of miRNAs in neurons is not yet fully understood. Using chemical transfection, I artificially overexpressed ten miRNAs in primary neuronal cultures. For six of them I identified hundreds of putative direct targets through analysis of the differential gene expression associated with the transfection experiments. Among these six miRNAs, there were two that are naturally enriched in the adult mouse brain (miR-124 and miR-434-3p), three miRNAs that were depleted from neurites (miR-143, miR-145 and miR-25) and one non-mouse miRNA (cel-miR-67). Analysis of the miRNA mediated effects on gene expression revealed that upon overexpression both miR-124 and miR-434-3p destabilised mRNA transcripts that are seen to be induced in stress conditions. The effect of overexpression of the other four miRNAs was found to be similar to that of miR-124 and miR-434-3p, although it was less significant. The ability of miRNAs to downregulate the inducibly expressed genes, and a widespread upregulation of these genes in stress conditions, implies that miRNAs normally act to prevent changes to equilibrium in the transcriptome. The results of this thesis also demonstrate that a repertoire of miRNA targets, including that of the neuron specific miR-124, is context-dependent. Given that the context can be influenced by a stress associated with experimental treatments, this work bears direct implications for future experiments aiming to ascribe particular functions to miRNAs.

Acknowledgements

I would like to start by acknowledging people who made it possible for me to study in Cambridge University: the people behind the Darwin Trust of Edinburgh, the Wellcome Trust, and EMBL. The Darwin Trust funded my final two years of the undergraduate education in Edinburgh University and then the first three years of my postgraduate education in Cambridge University. I would especially like to thank Sir Kenneth Murray and Noreen Murray of the Darwin Trust, who I was fortunate to know personally and who will remain an inspiration for me for the rest of my life. I am also grateful to The Wellcome Trust, which provided me (in addition to the Darwin Trust) with support through the first three years, and which was the sole source of my funding in the fourth year. A big thanks to EMBL, too, which paid me a stipend during the remainder of my stay in Cambridge.

I would like to thank people who organised and ran the PhD program at the Wellcome Trust Sanger Institute, and my two principle supervisors, Seth Grant and Anton Enright. I would also like to thank Derek Stemple and Frederick Livesey for additional supervision.

Apart from the mentoring provided by the supervisors, it would be impossible for me to go through the project without learning from my colleagues in the laboratories of Anton Enright and Seth Grant. In the laboratory of Anton Enright, I would like to especially thank Cei Abreu-Goodger and Stijn van Dongen. I learned all of my bioinformatic tricks from the two of them, and they provided me with day to day guidance in my work for the duration of the project. Incredibly, Cei Abreu-Goodger even found time to review my experimental designs and protocols, for which I am very grateful. Additionally, I would like to thank other people in the laboratory of Anton Enright, and in particular Mat Davis, Harpreet Saini and Nenad Bartonicek, whose support and help was very valuable. In the laboratory of Seth Grant, I learned the skills and techniques necessary to conduct the experiments of the thesis. For that, I am grateful to all people of the team, but I would especially like to express my gratitude to Rene Frank, Andrew Morton, Eric MacLaren, Ellie Tuck, Fei Zhu, Sharifah Syura, Alex Bayes and Tomas Ryan who took their valuable

time to train me and provide me with the required materials. Also, thanks to Meng Li for providing me with a plasmid and Elena Vigorito for helpful advice with RT-PCR.

After spending some significant amount of time on an experimental project, it is very obvious that a great amount of work is undertaken outside of the lab, perhaps unnoticed, but without which the work in the lab would be impossible. I would like to thank the Wellcome Trust Sanger Institute Animal Research Facilities and also the Seth Grant's laboratory, for providing me with the animals for my experiments. Additionally, a massive amount of work has been done by the Institute's microarray facilities, and without the contribution of Peter Ellis, Naomi Hammond and Cordelia Langford my project would not be possible.

Finally, I would like to acknowledge my parents and my friends. My parents directly supported me through the duration of my project, and provided me with the shelter of their kind words and home. As for my friends, Marija Buljan, Alexandra Nica, Leopold Parts, Steve Pettitt and Matias Piipari... the margins of this thesis are too narrow to contain all my gratitude to them. What would the whole time and work in Cambridge be worth, if it was not for the friends like them?

Abbreviations and comments

Abbreviations:

Cat. no.	catalogue number
DIV	days of <i>in vitro</i> development
miRNA	microRNA
<i>n</i>-mer	an oligomer of a length <i>n</i>
nt	nucleotides
ref.	reference
RT-PCR	real-time PCR
qRT-PCR	quantitative real-time PCR
<i>P</i>	P-value
UTR	Untranslated region

Comments:

- Very small numbers are presented using “E notation” as an alternative to the standard decimal notation. In this notation a letter *e* is used to represent *times ten risen to the power of*. For example, 0.000000012 in “E notation” is presented as $1.2e - 8$ or $1.2e - 08$.
- DNA is a polymer consisting predominantly of four types of units (nucleotides) containing the following four bases: adenine (the corresponding nucleotide is commonly denoted as *A*), cytosine (*C*), guanine (*G*) and thymine (*T*). RNA is also a polymer, which predominantly consists of nucleotides containing adenine, cytosine, guanine and uracil (the corresponding nucleotide is denoted as *U*) bases. In conventional Watson-Crick double stranded forms of RNA, DNA or DNA-RNA heteroduplexes, *G*s form connections with *C*s, while *A*s pair with both *T*s and *U*s. Therefore, *U* is RNA’s equivalent of DNA’s *T*. For purposes of consistency, sequences of DNA and RNA are frequently stored in databases as a sequence of the four letters *A*, *G*, *C* and *T*, where *T* is understood to be *U* in case of RNA sequences. In this thesis, I

preserved this notation, and both DNA and RNA nucleotide words are represented as sequences of *A*, *T*, *G* and *C*.

- The research of miRNA function that is presented in this system was conducted in an *in vitro* cell culture system derived from mice (*Mus musculus*). Conventionally, names of genes that encode miRNAs and names of miRNAs themselves are preceded by a three letter prefix, which uniquely corresponds to the species of the origin. Mouse miRNAs are preceded by three letters “mmu” (as in mmu-miR-124 or mmu-let-7c), while, human miRNAs (*Homo sapiens*) are preceded by “hsa” (as in hsa-miR-124 or hsa-let-7c). For convenience the three letter prefix of mouse miRNAs is frequently omitted, therefore names miR-124 and let-7c mean mmu-miR-124 and mmu-let-7c. Prefixes for other species are not omitted.

Contents

Contents	13
List of Figures	17
List of Tables	21
1 Introduction	25
1.1 Significance of miRNAs	25
1.1.1 Discovery of miRNAs	25
1.1.2 Biogenesis and molecular mechanisms of miRNA function	27
1.1.3 miRNAs in brain development and neuronal function	31
1.1.4 The paradox of miRNAs and role of miRNAs in stress responses . .	36
1.2 Thesis aims and an experimental paradigm	39
1.2.1 Widespread effect of miRNAs on the transcriptome is guided by miRNA seed region	41
1.2.2 Transfections of miRNA mimics enables identification of direct miRNA targets	44
1.2.3 The use of seed enrichment analysis of direct miRNA effects and identification of targets	46
1.2.4 Primary neuronal cultures as a model system to study neuronal biology	52
2 Methods	55
2.1 Primary Neuronal Cultures	56
2.2 RNA extraction	61
2.3 Quantitative RT-PCR. mRNA	66
2.4 Quantitative RT-PCR. miRNA	71
2.5 Transfection protocol	73

2.6	Microscopy	76
2.7	Microarray profiling of mRNA and miRNA expression	80
2.8	Seed enrichment analysis	87
2.9	Neuron specific genes	89
2.10	Enrichment of GO and KEGG terms	90
3	A model of neuronal development	91
3.1	A model of developmental gene expression	92
3.1.1	Gene expression changes in development of both hippocampal and forebrain primary cultures were highly correlated	92
3.1.2	The reciprocal trends of gene expression in development of hippocampal and forebrain primary cultures	93
3.1.3	Cell growth, not proliferation, was a predominant ongoing process in development of primary cultures	96
3.2	A model of miRNA activity in neurons	102
3.2.1	Identification of three categories of differentially expressed miRNAs in forebrain cultures development	102
3.2.2	miRNA expression in cultures development was similar to that in the brain and neurons	103
3.2.3	miRNAs were active in primary cultures: miR-124 and let-7 miRNAs shaped gene expression	105
3.3	Selection of miRNAs for functional experiments	110
4	A system to study microRNA function	113
4.1	Efficient transfection of neurons	114
4.1.1	Microscopy confirmed efficient transfection of neurons in primary forebrain cultures	114
4.1.2	Transfection did not cause neuronal death	117
4.1.3	Transfection of miRNA mimics consistently induced miRNA mediated changes in gene expression	123
4.2	Improving detection of miRNA targets	126
4.2.1	The use of miRNA inhibition instead of mock transfection improved detection of putative direct targets	126
4.2.2	Selection of an optimal incubation time and cell plating density	130
5	Results of transfection experiments	135

5.1	Unidirectional overexpression experiments	136
5.1.1	The effect of cel-miR-67, miR-143 and miR-145 overexpression was maximal at 3DIV or 4DIV	136
5.1.2	The effect of miR-124 overexpression was maximal at 6DIV	143
5.1.3	Endogenous miR-124 constrained gene expression in more mature primary neurons (6DIV)	143
5.2	Bidirectional perturbation experiments	146
5.2.1	Identification of targets for a steady state expressed miRNA (miR-124)	146
5.2.2	Identification of targets for downregulated miRNAs in development (miR-143, miR-145 and miR-25) and of a non-mouse miRNA (cel-miR-67)	149
5.2.3	Identification of targets for an upregulated miRNA in development (miR-434-3p)	154
5.3	Validation of the methodology	157
5.3.1	Comparison to targets identified by Lim <i>et al.</i>	158
5.3.2	Comparison to targets derived from Makeyev <i>et al.</i>	158
5.3.3	Comparison to targets identified by Chi <i>et al.</i>	162
6	Analysis of miRNA function in neurons	163
6.1	Characterisation of identified miRNA targets	164
6.1.1	Significant intersections between targets of different miRNAs	164
6.1.2	Explaining the intersection: A hypothesis of a pool of transcripts primed for miRNA mediated regulation	165
6.1.3	Context-dependent nature of the pool: Over 20% of targets were induced by the transfection procedure itself	168
6.1.4	Context dependent nature of published miR-124 targets	174
6.1.5	Recurrent enrichment of GO terms was in agreement with the pool of targets hypothesis	176
6.1.6	Enrichment of GO terms highlighted the importance of miR-124 and miR-434-3p	179
6.2	The function of miRNAs in neurons and the brain	182
6.2.1	In transfection experiments miRNAs downregulated stress inducible genes	182

6.2.2	Synaptic genes linked to neurological disease were enriched in miR-124 targets and in stress induced genes	188
6.2.3	miR-124 in development: Reduction of variability in gene expression	191
7	Discussion	197
7.1	Characterising the experimental system	199
7.2	Identification of miRNA targets	201
7.3	Context dependent function of miRNAs	204
7.4	Directions for future work	207
7.5	Conclusion	210
A	Supplementary Data	211
B	Publications and presentations of this work	277
	Bibliography	279

List of Figures

1.1	How to interpret Sylamer plots	50
2.1	Example of an output of Bioanalyzer	65
3.1	Correlation of forebrain and hippocampal cultures development	93
3.2	Gene expression clusters of down- and upregulated genes	95
3.3	Intersection of gene expression clusters in hippocampal and forebrain cultures	95
3.4	Enrichment of Gene Ontology (GO) terms in differentially expressed genes during primary cultures development	98
3.5	Neurons in primary cultures	99
3.6	Viability of cells in primary neuronal cultures	100
3.7	A gene expression switch point in development near 4DIV	101
3.8	Validation of miRNA expression profiles with qRT-PCR	103
3.9	Signature of miR-124 and let-7 regulation of the level of gene expression in primary cultures	107
3.10	Signature of miR-124 regulation of differential gene expression in primary cultures	108
4.1	Transfection of primary cultures with GFP-expressing plasmid	115
4.2	Transfection of primary cultures with AlexaFlour 488 labelled oligo at 3DIV and 6DIV	116
4.3	Visual inspection of transfected cultures	118
4.4	Differential regulation of neuron specific genes	119
4.5	The total number of spikes recorded in cultures after transfections	121
4.6	Viability of cells in transfected cultures	122
4.7	Correlation of cel-miR-67 experiments at 4DIV and 6DIV	124
4.8	Differential gene expression in miR-124 transfection experiments at 3DIV and 6DIV	128

4.9	Sylamer analysis of biases in distributions of seed matching sites in miR-124 transfection experiments at 3DIV and 6DIV	129
4.10	qRT-PCR profiling of the effect of miR-124 transfection in a high cell plating density timecourse at 4DIV	133
4.11	qRT-PCR profiling of the effect of miR-124 transfection in a low cell plating density timecourse at 6DIV	134
5.1	Scatterplot of differential gene expression upon unidirectional over-expression of miR-143, miR-145 and cel-miR-67	140
5.2	Sylamer analysis of biases in distribution of seed matching sites upon unidirectional over-expression of miR-143, miR-145 and cel-miR-67	142
5.3	Differential gene expression and seed matching site enrichment in miR-103 transfection experiments at 4DIV.	147
5.4	Sylamer analysis of biases in distributions of seed enrichment in miR-124 and miR-103 transfection experiments	148
5.5	Differential gene expression and seed matching site enrichment in miR-143, miR-145 and miR-25 transfections experiments at 4DIV	152
5.6	Sylamer analysis of biases in distributions of seed matching sites in miR-143, miR-145 and miR-25 transfection experiments at 4DIV	153
5.7	Differential gene expression and seed matching site enrichment in miR-434-3p transfection experiments at 6DIV	155
5.8	Sylamer analysis of distribution biases in distributions of seed matching sites in miR-434-3p, miR-370, miR-410 and miR-551b transfection experiments at 6DIV	156
5.9	A significant intersection of miR-124 targets with those identified by Lim et al.	159
5.10	Differential expression and shifts in seed matching site distributions induced by miR-124 in CAD cells	160
5.11	A significant intersection of miR-124 targets with those derived from Makeyev et al.	161
5.12	A significant intersection of miR-124 targets with those identified by Chi et al.	161
6.1	Intersections of the lists of putative direct miRNA targets	167
6.2	Intersection of genes induced by the transfection reagent and miRNA targets	171
6.3	Recurrence in top 25 most enriched KEGG terms	172

6.4	Intersection of published miR-124 targets and genes induced by the transfection reagent	175
6.5	Enrichment of Gene Ontology (GO) terms in miRNA targets	178
6.6	Counts of GO categories (“Biological process”, size > 5 genes) enriched in miRNA targets	181
6.7	The innate miRNA seed matching site distribution biases in transfection and kainate stresses	183
6.8	miR-124 and miR-434-3p downregulated stress induced genes	186
6.9	Significant intersections of genes induced in three stresses of neurons and the brain	187
6.10	Variability in expression of stress induced genes increased with progression of development	192
6.11	Variability in expression of stress induced genes was reduced by miR-124 overexpression	193
A.1	Pairwise correlation of raw mRNA microarray probe intensities in profiles of hippocampal cultures	215
A.2	Pairwise correlation of raw mRNA microarray probe intensities in profiles of hippocampal cultures	216
A.3	Sample relation between raw mRNA microarray profiles between replicates of hippocampal and forebrain cultures	217
A.4	Sample relation between normalised mRNA microarray profiles between replicates of hippocampal and forebrain cultures	218
A.5	Pairwise correlation of raw miRNA microarray probe intensities in profiles of forebrain cultures	226
A.6	Relationship between replicate raw miRNA microarray profiles	227
A.7	Regulation of stress induced genes by miR-25, miR-143 and miR-145	271
A.8	Correlation of expression and standard deviation in hippocampal primary cultures	272
A.9	Correlation of expression and standard deviation in forebrain primary cultures	273
A.10	Correlation of expression and standard deviation in miR-124 overexpression and inhibition	274
A.11	Correlation of expression and standard deviation in mock transfection and cel-miR-67 overexpression	275

List of Tables

2.1	Real-time PCR primers and probes	67
2.2	mRNA reverse transcription starting mix	68
2.3	mRNA real time PCR program	69
2.4	miRNA reverse transcription master mix	72
2.5	miRNA reverse transcription program	72
2.6	miRNA real time PCR master mix	73
2.7	miRNA real time PCR program	73
2.8	Parameters of transfection experiments	76
3.1	Neurons in primary neuronal cultures.	99
3.2	Viability of cells in primary neuronal cultures	100
3.3	Ranking of expression of proneuronal miRNAs	104
3.4	Depletion of downregulated miRNAs from synaptosomes	105
3.5	Selection of down- and upregulated miRNAs	111
4.1	Viability of cells in transfected cultures	123
5.1	Number of genes downregulated ($P < 1e - 06$) in miRNA overexpression experiments	137
5.2	Previously published genome-wide studies of miR-124 targeting	157
6.1	A list of the miRNA targets found within the “Pathways in cancer” KEGG pathway	173
6.2	miRNA targeting of genes induced in three stress types	185
6.3	Synaptic genes linked to neurological diseases in stresses and miR-124 tar- geting	189
A.1	The list of putative neuron-specific genes	212

A.2	Top 40 most enriched GO terms (“BP”) in developmentally downregulated genes	219
A.3	Top 40 most enriched GO terms (“BP”) in developmentally upregulated genes	220
A.4	Top 40 most enriched GO terms (“CC”) in developmentally downregulated genes	221
A.5	Top 40 most enriched GO terms (“CC”) in developmentally upregulated genes	222
A.6	Top 25 most enriched KEGG terms in developmentally downregulated genes	223
A.7	Top 25 most enriched KEGG terms in developmentally upregulated genes .	224
A.8	Three categories of miRNAs in development of primary forebrain cultures .	228
A.9	Putative direct targets of miR-124	231
A.10	Putative direct targets of miR-143	232
A.11	Putative direct targets of miR-145	233
A.12	Putative direct targets of miR-25	234
A.13	Putative direct targets of cel-miR-67	235
A.14	Putative direct targets of miR-434-3p	237
A.15	Top 25 most enriched KEGG terms in targets of cel-miR-67	237
A.16	Top 25 most enriched KEGG term in targets of miR-124	238
A.17	Top 25 most enriched KEGG terms in targets of miR-143	239
A.18	Top 25 most enriched KEGG terms in targets of miR-145	240
A.19	Top 25 most enriched KEGG terms in targets of miR-25	240
A.20	Top 25 most enriched KEGG terms in targets of miR-434-3p	241
A.21	Top 25 most enriched KEGG terms in the induced by transfection set . . .	242
A.22	Top 25 most enriched KEGG terms in the Ago HITS-CLIP set	243
A.23	Top 40 most enriched GO terms (“BP”) in targets of cel-miR-67	244
A.24	Top 40 most enriched GO terms (“BP”) in targets of miR-124	245
A.25	Top 40 most enriched GO terms (“BP”) in targets of miR-143	246
A.26	Top 40 most enriched GO terms (“BP”) in targets of miR-145	247
A.27	Top 40 most enriched GO terms (“BP”) in targets of miR-25	248
A.28	Top 40 most enriched GO terms (“BP”) in targets of miR-434-3p	250
A.29	Top 40 most enriched GO terms (“BP”) in the induced by transfection set	251
A.30	Top 40 most enriched GO terms (“BP”) in the Ago HITS-CLIP set	252
A.31	Top 40 most enriched GO terms (“CC”) in targets of cel-miR-67	253
A.32	Top 40 most enriched GO terms (“CC”) in targets of miR-124	254

A.33 Top 40 most enriched GO terms (“CC”) in targets of miR-143	256
A.34 Top 40 most enriched GO terms (“CC”) in targets of miR-145	257
A.35 Top 40 most enriched GO terms (“CC”) in targets of miR-25	258
A.36 Top 40 most enriched GO terms (“CC”) in targets of miR-434-3p	259
A.37 Top 40 most enriched GO terms (“CC”) in the Ago HITS-CLIP set	260
A.38 Top 40 most enriched GO terms (“CC”) in the induced by transfection set	262
A.39 miRNA targets within “ECM-receptor interaction” KEGG pathway	263
A.40 miRNA targets within “Gap junction” KEGG pathway	264
A.41 miRNA targets within “ErbB signaling pathway” KEGG pathway	265
A.42 miRNA targets within “Tight junction” KEGG pathway	266
A.43 miRNA targets within “p53 signaling pathway” KEGG pathway	266
A.44 miRNA targets within “Regulation of actin cytoskeleton” KEGG pathway	267
A.45 miRNA targets within “Focal adhesion” KEGG pathway	268
A.46 miRNA targets within “MAPK signaling pathway” KEGG pathway	269
A.47 miRNA targets within “VEGF signaling pathway” KEGG pathway	270
A.48 miRNA targets within “Toll-like receptor signaling pathway” KEGG pathway	270

