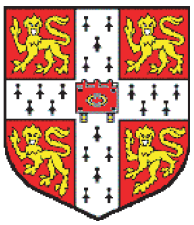
The role of microRNAs in neurons

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 LaTeX 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 *in vitro* development
miRNA	microRNA
***n*-mer**	an oligomer of a length n
nt	nucleotides
ref.	reference
RT-PCR	real-time PCR
qRT-PCR	quantitative real-time PCR
P	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, Gs form connections with Cs, while As pair with both Ts and Us. 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 *et al.* . . . 158
5.3.2 Comparison to targets derived from Makeyev *et al.* . . . 158
5.3.3 Comparison to targets identified by Chi *et al.* . . . 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 Deption 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 targeting . . . 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

Chapter 1

Introduction

In the first part of the Introduction I describe the discovery of miRNAs and how this class of molecules came into the scientific limelight. I also describe findings that demonstrate functions and roles of miRNAs, and their importance for the brain and neurons. The second part of the Introduction states the aims of this thesis, and introduces the experimental paradigm, that have been followed during the course of the thesis project.

1.1 Significance of miRNAs

1.1.1 Discovery of miRNAs

Nucleic acids were discovered in the nineteenth century, and a function for RNA was described in the fifties of the twentieth century, when it was identified to be a messenger between DNA and protein. The perception of RNA as a passive ancillary carrier of genetic information changed dramatically during the following six decades. Discoveries, including introns capable of self-splicing by Thomas Cech, the RNase P that cleaves tRNA using RNA at its core by Sidney Altman, and ultimately the discovery of Venkatraman Ramakrishnan, Thomas Steitz and Ada Yonath that synthesis of all proteins in the ribosome is catalyzed solely by its RNA component shifted RNA to the center of biology. The hypothesis of a primordial RNA-world became generally accepted, and few scientists doubt the importance of RNA-dependent mechanisms for evolution and the existence of life's complexity, such as retrotransposition driven genome rearrangements and alternative splicing.

Contrary to the original functional paradigm of RNAs as DNA-protein intermediates, more recently discovered regulatory RNAs do not encode proteins, but have a separate

functional significance of their own. One type of regulatory RNAs, called miRNAs, is the subject of this thesis. The first miRNA was discovered in 1993 by Victor Ambros and Gary Ruvkun (Lee et al., 1993; Wightman et al., 1993) in *Caenorhabditis elegans* (*C. elegans*) as a post-transcriptional regulator of gene expression. The discovered miRNA, called *lin-4*, was found to decrease expression of its target gene, *lin-14*, through interaction of the *lin-4* miRNA with the transcript of *lin-14*, leading to the decrease in levels of LIN-14 protein. Since their discovery in 1993, miRNAs have been identified in all multicellular animals and plants and also in some unicellular plants (Grimson et al., 2008). The official repository of information concerning metazoan miRNAs is miRBase (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006, 2008) (http://www.mirbase.org/). According to the current release (miRBase Release 16), there are 667 and 1,049 known genes for miRNAs in the mouse and human genomes respectively.

Curiously, after their original discovery, miRNAs received relatively little attention from the scientific community. However, studies of a different type of non-coding RNA, that were also performed in *C. elegans*, eventually brought miRNAs to the attention of scientist world wide. In 1995, an injection of antisense RNA into *C. elegans*, was shown to repress expression of a gene to which it was complementary, a phenomenon that was later called RNA interference (RNAi) (Guo and Kemphues, 1995). Mysteriously, the injection of the sense sequence also induced RNAi. Three years later, Andrew Fire and Craig Mello found an explanation by showing that double-stranded RNA (dsRNA) was the effective trigger of the phenomenon. Soon after the role of dsRNA was established, RNAi was identified in *Drosophila* (Kennerdell and Carthew, 1998), successfully used to silence a gene in *Xenopus* (Oelgeschläger et al., 2000) and was described in mice (Svoboda et al., 2000; Wianny and Zernicka-Goetz, 2000). Moreover, RNAi explained the enigmatic phenomenon of post-transcriptional silencing of endogenous genes by clones of homologous sequences, which was reported in plants (Napoli et al., 1990; van der Krol et al., 1990; Smith et al., 1990; de Carvalho et al., 1992) and fungi (Romano and Macino, 1992; Cogoni et al., 1996). As RNAi is triggered by tiny amounts of dsRNA, it suggested the existence of mechanisms that can propagate and sustain the RNA-mediated gene silencing (as opposed to purely stoichiometric sense-antisense interactions) in all major branches of eukaryotic tree of life. Very soon after the discovery of dsRNA as the trigger of RNAi, several groups showed that in both plants (Hamilton and Baulcombe, 1999) and animals (Hammond et al., 2000; Zamore et al., 2000; Elbashir et al., 2001) the effective dsRNA was converted to short functional oligomers (21 to 25 nucleotids long), that were termed short interfering RNAs (siRNAs). The discovery of siRNAs, which were chemically identical

to several then known miRNAs, suggested that RNAi and miRNA induced regulation of gene expression were related processes, with miRNAs being an endogenous form of siRNAs (siRNAs and miRNA and other short RNAs are together referred to as sRNAs). As RNAi was demonstrated in a variety of species, these insights prompted the cataloging miRNAs and other endogenous sRNAs in various species. The discoveries of new sRNAs followed soon. In the year 2001, Tomas Tuschl's laboratory were able to clone dozens of miRNAs from *Drosophila* embryos and human cell cultures (Lagos-Quintana et al., 2001), which confirmed that a pool of sRNAs existed naturally in different animals. Finally, 8 years after their discovery, miRNAs came into the scientific limelight.

In the ensuing decade, miRNAs and other sRNAs remained at the frontier of research world-wide. Entirely new classes of sRNAs were discovered in metazoans, including Piwi-interacting RNAs (piRNAs) (Aravin et al., 2007), recently indentified promoter-associated short RNAs (PASRs) (Taft et al., 2009) and splice-site RNAs (spliRNAs) (Taft et al., 2010). Two excellent reviews give a comprehensive summary of information about siRNAs and piRNAs (Carthew and Sontheimer, 2009; Malone and Hannon, 2009). There are several important distinctions between plant and animal miRNAs, despite general principles being similar in both kingdoms (Voinnet, 2009). The biogenesis and function of animal[1] miRNAs is described in the next section.

1.1.2 Biogenesis and molecular mechanisms of miRNA function

The main role of miRNAs in the cell is the regulation of expression of genes (miRNA targets) at a post-transcriptional level. The absence of direct miRNA-DNA interactions (Sharp, 2009) and demonstration of co-localisations of miRNAs and mRNAs to the cytoplasmic compartments (Liu et al., 2005) supported this view. As in the case of siRNAs, a majority of published reports showed miRNAs to have an inhibitory effect on expression of their targets. RNAi, a cleavage of the mRNA triggered by siRNAs (see section 1.1.1), is catalyzed by an enzymatic complex to which siRNAs are bound and which they direct to the mRNA targets. This protein complex is located in the cytoplasm, and is known as RNA-induced silencing complex or RISC (Hammond et al., 2000). Soon after the discovery of RISC being a catalytic machine of RNAi, it was demonstrated that miRNAs were also associated with RISC (Hutvágner and Zamore, 2002). Complementarity between ≈

[1]The work of this thesis is focused on miRNAs in the mouse, therefore the introduction to miRNA biology in animals is also focused on the mouse. Genes, transcripts and protein names will, by default, refer to those in the mouse, unless specified otherwise. Conventions for the mouse notations will be used (gene names are in italic with the first letter capitalised, while products of the genes (transcripts and proteins) are in a regular font with the first letter capitalised), unless specified otherwise.

7 bases located at the 5′-end of miRNA (called the seed region) and an mRNA transcript was found to be a good predictor of the transcript being targeted by the miRNA (Lewis et al., 2003, 2005). Properties of the seed region and the target sites (called the seed matching sites) are discussed in detail in Introduction, section 1.2.1. Here it should be noted that the requirement of only a partial complementarity between a miRNA and an mRNA transcripts enables, in principle, a single miRNA to target hundreds of mRNA transcripts (Enright et al., 2003; Stark et al., 2003; Farh et al., 2005; Lim et al., 2005; Giraldez et al., 2006; Baek et al., 2008; Selbach et al., 2008).

Biogenesis of miRNAs

The ≈ 22nt miRNAs that are incorporated into RISC and act as its guide are sometimes referred to as mature miRNAs. Mature miRNAs originate from longer transcripts, called primary-miRNAs (pri-miRNAs), that are produced by Pol II (Polymerase II) transcription[1], and are capped and polyadenylated (Lee et al., 2004; Cai et al., 2004). Frequently, pri-miRNA transcripts give rise to more than one mature miRNA (i.e. they are polycistronic), or, as was shown for ≈ 40% of human miRNAs, pri-miRNA transcripts can also encode a protein sequence, in which case mature miRNA sequences are usually located within introns and are called intragenic (Kim et al., 2009). Within a pri-miRNA, the sequence of a mature miRNA is within a secondary structure, a hairpin (Winter et al., 2009; Kim et al., 2009). These hairpins are recognized in the nucleus by the enzyme-complex that is sometimes referred to as the Microprocessor, the principal component of which is an RNase III type endonuclease, called Drosha (Lee et al., 2002). Drosha introduces a cut in the stem of the hairpin within the pri-miRNA releasing a shorter hairpin (with the stem ≈ 33 nt), called the precursor-miRNA (pre-miRNA). Characteristically for the RNase III type endonucleases, when Drosha cuts the base of the dsRNA hairpin, it leaves a ≈ 2nt overhang of the 3′RNA-end, and a phosphate at the 5′-end (Basyuk et al., 2003; Lee et al., 2003). The Drosha cut is important, as it produces one end of the mature miRNA. Interestingly, in the case of a few intronic miRNAs, called mirtrons, their splicing produces pre-miRNAs directly, thus bypassing the Microprocessor (Babiarz et al., 2008).

The next step, pre-miRNA is exported from the nucleus to the cytoplasm via the Exportin 5 complex (Yi et al., 2003; Bohnsack et al., 2004). Upon export to the cytoplasm, pre-miRNAs are recognized and cleaved by another protein complex, a principle

[1]Pol III transcription was reported to produce pri-miRNAs in case of Alu-element derived miRNAs (Borchert et al., 2006), but such cases are rare.

component of which is an RNase type III enzyme, called Dicer (encoded by a single gene in the mouse, *Dicer1*) (Grishok et al., 2001; Hutvágner et al., 2001; Bernstein et al., 2001; Ketting et al., 2001; Knight and Bass, 2001). Dicer introduces a cut into the pre-miRNA, removing the loop from the stem, and generates the second end of the mature miRNA. This end also has ≈ 2 nt overhang at the 3′-end and a phosphate at the 5′-end. The RNA, that is generated by Dicer, is a ≈ 22 nt double-stranded oligonucleotide, one strand of which is going to become a mature miRNA (called a guide strand). The duplex is thought to be loaded into RISC in a process dependant on the interaction of Dicer with the Argonaute component of RISC (Ago proteins (Höck and Meister, 2008; Joshua-Tor and Hannon, 2010)) and other ancillary proteins (Chendrimada et al., 2005; Maniataki and Mourelatos, 2005). However, miRNA duplexes introduced into a Dicer-null background have been reported to efficiently downregulate expression of the targets despite lacking the Dicer protein (Hanina et al., 2010). Therefore, at least for exogenously produced miRNA duplexes, RISC assembly can be independent of Dicer. To produce a functionally competent RISC one of the two strands of the duplex is degraded (passenger strand). The strand of the duplex with the weakest hydrogen bonds at its 5′-end is more likely to survive (Schwarz et al., 2003; Khvorova et al., 2003). Mechanism of degradation of the passenger strand is not enirely clear, as endonucleolytic activity of the Argonaute component is not essential for assembly of the miRNA RISC (Matranga et al., 2005; Leuschner et al., 2006). Recently, Dicer independent generation of one mature miRNA (miR-451) from pre-miRNA was reported (Cifuentes et al., 2010; Yang et al., 2010). However, as in the case of Microprocessor independent miRNAs, Dicer independent miRNAs are thought to be rare and currently only one such miRNA is known (Cifuentes et al., 2010; Yang et al., 2010).

Molecular mechanisms of miRNA function

Historically, the function of miRNAs was thought to be the inhibition of the translation of targeted mRNAs, while siRNAs were thought to trigger endonucleolytic cleavage of targeted mRNAs. However, the destabilization and subsequent degradation of target mRNA transcripts was subsequently observed (Lim et al., 2005; Giraldez et al., 2006). Subsequently, a significant correlation was detected between the regulatory effects of several miRNAs measured at mRNA and protein levels (Baek et al., 2008; Selbach et al., 2008). Moreover, according to recent reports, the inhibitory effect of several miRNAs on protein production is predominantly explained by the reduction in mRNA levels (Hendrickson et al., 2009; Guo et al., 2010). The ability of miRNAs to destabilize target mRNAs was

at the basis of the experimental paradigm in this thesis, where lowering of mRNA levels served as an indication of miRNA activity (see Introduction, section 1.2.2).

The regulatory function of miRNAs is not exerted by the miRNA itself. The same is true for siRNAs during RNAi. miRNAs bind the RNA induced silencing complex (RISC), which it guides to target transcripts (Fabian et al., 2010). The target recognition by miRNAs is observed in complementarity of miRNA sequences to the sequence of their target transcript. The rules of guiding itself (i.e. recognition of target transcripts by miRNAs) will be discussed in section 1.2.1, in conjunction with the use of these rules for the computational prediction of miRNA targets. Below is a brief overview of the mechanisms by which RISC destabilises the target transcript and/or represses its translation after it was brought to the transcript by the miRNA.

A protein of the Argonaute family is the component of RISC that directly binds the miRNA (Peters and Meister, 2007). There are four genes in the mammalian genome that encode members of Argonaute family (*Ago1*, *Ago2*, *Ago3* and *Ago4*), each of which can bind miRNAs (Azuma-Mukai et al., 2008; Landthaler et al., 2008). All four Ago proteins are thought to be functionally competent as RISC components, although only Ago2 possesses endonucleolytic activity, which catalyzes RNAi (Liu et al., 2004; Baillat and Shiekhattar, 2009). As a consequence, the main role of Ago proteins in animal miRNA mediated regulation is thought to be either in the repression of translation or in directing mRNA targets to components of a generic mRNA degradation machinery (Fabian et al., 2010). One of the proteins that binds to Ago and that is essential for both miRNA mediated mRNA destabilisation and translational repression is GW182 (Eulalio et al., 2008). In fact, both mRNA destabilisation and the repression of translation can be triggered by chemical tethering of GW182 to mRNA, in the absence of miRNAs and Ago (Pillai et al., 2004; Behm-Ansmant et al., 2006; Chekulaeva et al., 2009; Zipprich et al., 2009), which confirms that GW182 operates downstream of Ago. GW182 is thought to recruit the mRNA to the deadenylase complex, which triggers deadenylation of mRNA transcripts, followed by their decapping and subsequent degradation by exonucleases (Eulalio et al., 2009). The inhibition of translation occurs both during initiation and elongation stages. During initiation RISC can interfere with assembly of the initiation complex at the Cap of the mRNA (Pillai et al., 2005; Humphreys et al., 2005; Wang et al., 2006; Wakiyama et al., 2007; Thermann and Hentze, 2007; Mathonnet et al., 2007) or with assembly of the 80S ribosome (Wang et al., 2008a; Chendrimada et al., 2007). During the elongation stage RISC can cause ribosomes to stall and trigger them to drop-off the transcript (Gu et al., 2009; Olsen and Ambros, 1999; Nottrott et al., 2006; Maroney et al., 2006; Petersen et al.,

2006). More detailed information on molecular mechanism of miRNA mediated mRNA destabilisation and translational repression is described elsewhere (Fabian et al., 2010).

There are currently three reports of miRNAs activating gene expression. Two of these described miRNA targeting of mRNA transcripts through seed matching sites in 5′UTRs (which is unconventional for miRNA-mRNA interaction, see section 1.2.1). Subsequently the activation of translation of mRNA transcripts was observed (Ørom et al., 2008; Henke et al., 2008). In another report, the activation of translation of miRNA targets was observed in a human cell culture system upon the stress of serum withdrawal and cell-cycle arrest (Vasudevan et al., 2007). It is still unclear, however, how general and reproducible these observations are, and the activatory roles of miRNAs will not be further discussed in this Introduction[1].

1.1.3 miRNAs in brain development and neuronal function

Despite research of miRNA function outside of the *C. elegans* developmental paradigm commencing only a decade Ago, their significance for neuronal biology was already obvious by 2007, when this thesis project was designed. The evidence of the importance of miRNAs for brain development and neurogenesis came, primarily, from the ablation of components of the miRNA biogenesis pathways (e.g. *Dicer1* deletions) and simultaneous removal of nearly all miRNAs as a consequence. Further details of the roles of miRNAs in determining neuronal identity and neuronal function were revealed through experiments with individual miRNAs. Results of these experiments are discussed in the current section with the purpose of illustrating importance of miRNAs for brain development, the establishment of neuronal identity and neuronal function.

miRNAs in brain development

The depletion of all miRNAs by the disruption of miRNA biogenesis provided some of the first insights into the function of miRNAs in brain development. Dicer is an RNase III type enzyme that is indispensable for biogenesis of all but one known miRNA (Cifuentes et al.,

[1]Analysis of the results of experiments performed in this thesis project, where miRNAs were exogenously added to cultured neurons, were not consistent with activatory activity of overexpressed miRNAs (Chapters 4 and 5). However, a signature of activity of endogenous miRNAs in certain conditions, such as the stresses described in Chapter 6 (section 6.2.1) was consistent with the relief of miRNA mediated regulation during stresses, or, perhaps, a switch to an activatory mode. Before such conclusions can be drawn, additional experiments specifically designed to directly test this proposition are required to convincingly demonstrate that the modulation of miRNA regulation in primary neurons under stresses does take place (see Discussion, section 7.4).

2010). There is one copy of *Dicer1* gene in the mouse and human genomes and a stable mouse knockout (Dicer-null) was generated (Bernstein et al., 2003). Development of Dicer-null embryos did not proceed beyond 7.5 days of development, which is before formation of the body plan during gastrulation (Bernstein et al., 2003). Although this result is likely to mean that Dicer and miRNAs are essential for mouse development, it did not prove that Dicer and miRNAs are important for brain development *per se*, as developmental arrest occurred too early for this conclusion to be drawn. Ablation in mice of one of the four Ago encoding genes, *Ago2*, also lead to a severe developmental delay, however its onset was at a later time (E10.5). Interestingly, one of the most prominent developmental defects in Ago2-null mice was failure of the neural tube closure (Liu et al., 2004). Similar observations were made by Antonio Giraldez and colleagues in their experiments on *Danio rerio* (Giraldez et al., 2005). Giraldez created *D. rerio* Dicer-null zygotes that were lacking maternal Dicer (*MZdicer*). Interestingly, although *MZdicer* embryos were not viable, their development progressed further relative to the development of Dicer-null mouse embryos, and severe abnormalities in development of both neural and nonneural systems were uncovered. For example, formation of the neurocel and the midbrain-hindbrain boundary were severely undermined in *MZdicer* embryos. Importantly, it was possible to confirm that the phenotype of *MZdicer* embryos was triggered by the lack of mature miRNAs through rescue experiments. Indeed, injection of a dsRNA mimic of dre-miR-430, a highly abundant miRNA in early embryonic development, rescued many aspects of neural development (including formation of normal size brain ventricles and the midbrain-hindbrain boundary). This experiment showed that miRNAs were not only essential for early vertebrate development, but that they also played a significant role in development of the nervous system.

With improvement in gene targeting technologies, conditional knockout mice were created, which enables one to directly observe the consequences of disrupting miRNA biogenesis for mammalian brain development and for mature neurons. By deleting Dicer at a specific time and in a specific cell type it was possible to circumvent the requirement for Dicer in the early embryonic development of the mouse and study consequences of its loss for later developmental stages. Isolation of Dicer-null neural progenitors suggested that miRNAs are essential for commitment of neural progenitors to differentiation. Neural progenitors without detectably expressed miRNAs were obtained from the embryonic Dicer-null cerebral cortex and were shown to be incapable of differentiation (Andersson et al., 2010). Similarly, Dicer-null oligodenrocyte progenitors were also incapable of differentiation. This phenotype was partially rescued by ectopic expression of miR-219 and

miR-338 (Zhao et al., 2010). The requirement of Dicer and miRNAs for commitment of neural progenitors was in agreement with reports of miRNAs being essential for the differentiation of mouse embryonic stem (ES) cells (Kanellopoulou et al., 2005; Wang et al., 2007). In addition to the regulation of stem cells commitment, miRNAs were found to be important for the survival of differentiated cell types. Increased apoptosis was frequently reported upon deletion of Dicer in various cell types, including dopaminergic neurons (Kim et al., 2007), Purkinje cells (Schaefer et al., 2007), and forebrain neurons (Davis et al., 2008; Konopka et al., 2010; Hébert et al., 2010).

miRNAs in establishment of neuronal identity

Studying consequences of perturbation (downregulation and/or overexpression) of individual miRNAs provided further evidence of their functional significance in neuronal development, particularly in differentiation and acquisition of a cellular identity. The first evidence of the role of miRNAs in the establishment of differentiated cell types, including the neuronal cell type, came from an experiment on ectopic expression of two tissue specific miRNAs in HeLa cell culture (Lim et al., 2005). One of this miRNAs was miR-124, a miRNA highly conserved and highly expressed in the central nervous system (CNS) and specific to neurons (Lagos-Quintana et al., 2002; Landgraf et al., 2007; Cheng et al., 2009; Clark et al., 2010). Ectopic expression of miR-124 in HeLa cell culture[1] caused inhibition of a number of genes that were normally expressed at a low level in the brain (Lim et al., 2005). Therefore, expression of miR-124 transformed the gene expression of HeLa cells to be more like that of a neuron, which hinted at its role in the establishment of neuronal gene expression (neuronal state). Subsequently, this proposition was supported by the observation that introduction of miR-124 into dividing neural precursors caused them to cease division and undergo neuronal differentiation (Cheng et al., 2009). In addition to miR-124, other miRNAs, such as miR-9, miR-125b and miRNAs of let-7 family were shown to trigger premature differentiation upon overexpression in neuronal progenitors (Leucht et al., 2008; Le et al., 2009; Rybak et al., 2008).

Experiments with individual miRNAs enabled the identification of genes regulated by these miRNAs (i.e. miRNA targets). In some cases targets of miRNAs themselves were shown to inhibit neuronal differentiation (i.e. they had an anti-neuronal activity), and therefore these miRNAs themselves can be said to have a pro-neuronal role. Perhaps the most studied example of a pro-neuronal miRNA is miR-124, and its targets were iden-

[1]HeLa was derived from a cervical carcinoma (Scherer et al., 1953) and, hence, is non-neuronal.

tified in several distinct pathways important for neuronal differentiation. For example, miR-124 was found to participate in a double negative feedback loop that involves REST, a transcription factor that has a gate-keeper role in the acquisition of the neuronal state (Conaco et al., 2006; Visvanathan et al., 2007). REST binds RE-1 elements in promoter regions of many neuronal genes. In non-neuronal cells it recruits co-repressors to promoters of these genes and causes their transcriptional inhibition (Ballas and Mandel, 2005; Ballas et al., 2005). One of these co-repressors, SCP1, is directly targeted by miR-124 and, thus miR-124 counteracts the anti-neuronal activity of REST (Visvanathan et al., 2007). Interestingly, in non-neuronal cells, REST was shown to inhibit expression of miR-124 (Conaco et al., 2006; Visvanathan et al., 2007), an interaction that completes the double negative feedback loop. Apart from being involved in inhibition of the function of the REST/SCP1 pathway, miR-124 was also implicated in the inhibition of at least three other anti-neuronal pathways. Its targets include PTBP1, a global inhibitor of the neuron specific alternative splicing (Makeyev et al., 2007), BAF53a, a neural-progenitor specific chromatin remodeling factor (Yoo et al., 2009) and Sox9, a transcription factor important for proliferation of neural progenitors (Cheng et al., 2009). In addition to miR-124, over a dozen different miRNAs were also shown to be regulators of neuronal differentiation through the inhibition of genes with anti-neuronal activity. An in depth description of the function of miR-124 and other miRNAs in neuronal differentiation can be found in an excellent review by Xuekun Li and Peng Jin (Li and Jin, 2010).

miRNAs in neuronal function

The first discovered miRNA, lin-4, was identified because of its essential role in regulation of *C. elegans* development (Lee et al., 1993; Wightman et al., 1993). Subsequently, as was described above, other miRNAs were also shown to have an important role in regulation of organ development and cell differentiation. At the same time, miRNAs, including miRNAs important for brain and neuronal development, were being discovered in adult organisms. This posed a challenge to identify the functions of miRNAs in mature neurons.

A seminal study in this field was published by Gerhard Schratt and colleagues, which described a role of miR-134 in regulation of synaptic morphology that was dependent on neuronal activity (Schratt et al., 2006). This miRNA was found to inhibit expression of *Limk1*, a gene that was shown to regulate actin filament dynamics. At the morphological level, miR-134 acted to decrease dendritic spine size. This work proposed the biological importance of miR-134 as an inhibitor of synaptic plasticity because activity of miR-134 itself was relieved by BDNF, a major stimulant of synaptic growth and function. The role

of miR-134 in neuronal function was further explored in a recent work from the laboratory of Li-Huei Tsai (Gao et al., 2010). Overexpression of miR-134 in the mouse hippocampus was shown to impair performance of the animals in a context fear-conditioning task and to abrogate induction of CA1-CA3 long-term potentiation (LTP). Inhibition of LTP and memory was suggested to be linked to miR-134 targeting Creb1 transcript that encodes a transcription factor important for induction of long-term synaptic plasticity (Gao et al., 2010).

Activity of other miRNAs was also shown to modulate the function of mature neurons, and their activity was frequently described as inhibitory to genes upregulated in neuronal plasticity (Schratt, 2009). Surprisingly, this mode of inhibition of synaptic plasticity was also shown for miRNAs that were previously demonstrated to promote neuronal differentiation (i.e. pro-neuronal miRNAs). For example, when miR-124 was injected into the cultured neurons of *Aplysia californica*, it significantly reduced long term facilitation of synaptic transmission (Rajasethupathy et al., 2009). Perhaps the most striking evidence of the inhibitory effect of miRNAs to neuronal plasticity came from demonstration that the loss of all miRNAs leads to enhancement of learning and memory in mice (Konopka et al., 2010). Using a transgenic mouse line in which ablation of Dicer could be induced in the adult forebrain neurons, it was possible to significantly deplete the pool of miRNAs in the mature neurons of a living animal. All mice eventually died, presumably due to the massive neurodegeneration that was observed after 14 weeks from inducing the deletion of *Dicer*, which was consistent with previously reported elevated apoptosis in Dicer-null backgrounds (see above). There was a time-window, however, at 12 weeks from the time of deletion, when the pool of miRNAs was significantly depleted, while the onset of apoptosis had not yet commenced. At this time, the mutant mice had increased performance in four different learning and memory tests and also displayed an elevated post-tetanic CA1-CA3 synaptic potentiation (Konopka et al., 2010).

In summary, miRNA mediated regulation of gene expression was revealed to be important for both differentiation of neuronal progenitors and for the function of mature neurons. Several miRNAs, including pro-neuronal miRNAs, were suggested to inhibit neuronal plasticity. Both development and plasticity are characterized by widespread changes in gene expression, and miRNA regulation of these processes is consistent with their proposed function as buffers of differential gene expression (Wu et al., 2009b). In the role as buffers, miRNAs are well suited to contribute to canalisation of developmental programs (Hornstein and Shomron, 2006) and confer robustness to gene expression

networks (Herranz and Cohen, 2010). The robustness may be particularly important in the face of stresses to biological systems and miRNAs were proposed to be important for adequate stress responses (Leung and Sharp, 2010). This aspect of miRNA function will be important for interpretation of experimental results obtained during the course of this thesis project, therefore the role of miRNAs in stress responses is reviewed in the next section.

1.1.4 The paradox of miRNAs and role of miRNAs in stress responses

A feature of miRNAs, that is sometimes referred to as a paradox of miRNAs, is their high degree of evolutionary conservation and yet apparently non-essential role in cell and organism viability. Some miRNAs are highly conserved between deeply branching metazoans (Pasquinelli et al., 2000; Sempere et al., 2006). Moreover, the very origin of metazoan organ systems, such as the central nervous system, the sensory tissue, the musculature and the gut, was found to coincide with the origin of tissue specific expression of certain miRNAs (Christodoulou et al., 2010). In the light of deep evolutionary conservation, finding that a majority of miRNAs was not essential for viability came as a surprise. For example, simultaneous removal of a majority of miRNAs from the ES cells through deletion of the *Dgcr8* gene[1], which is a key factor in miRNA biogenesis did not trigger their death (Wang et al., 2007). The individual deletion of a majority of genes encoding miRNAs or entire miRNA families from the nematode worm *Caenorhabditis elegans* genome did not induce significant phenotypic abnormalities (Miska et al., 2007; Alvarez-Saavedra and Horvitz, 2010). According to one report, inhibition of a highly expressed, evolutionarily conserved neuron-specific miRNA, miR-124, during development of the chick spinal cord did not affect neurogenesis (Cao et al., 2007). Additionally, both inhibition and overexpression of miRNAs was found to induce only subtle changes in the abundance and translation of the target transcripts (Baek et al., 2008; Selbach et al., 2008). To reconcile the seemingly conflicting evolutionary conservation and dispensability, miRNAs were suggested to act as tuners and buffers of gene expression, rather than major regulatory switches (Wu et al., 2009b). In this role, the function of miRNAs would be to ensure robustness of gene expression programs. In agreement with this proposition, the significance of miRNAs was identified for processes that shift equilibria of gene expression programs. The

[1]Together with Drosha, Dgcr8 is an essential component in the Microprocessor complex (see section 1.1.2) (Han et al., 2004).

role of miRNA mediated regulation in some of these processes, such as development and plasticity, had already been discussed in section 1.1.3. Significance of miRNAs in other equilibrium shifting processes (induced mutations and stresses) is described below.

The function of miRNAs in conferring robustness to biological systems was revealed when nematodes with deletions of miRNA genes, but no deleterious phenotype, developed significant abnormalities in the context of perturbations to the transcriptome (Brenner et al., 2010). In nematodes with stable deletions of seemingly non-essential miRNAs, Brenner and colleagues knocked down five hub nodes of the gene expression network, each of which was a component of several major signaling pathways (e.g. EGF, Wnt, Notch and etc.). Knock down of transcripts encoding these hub proteins was expected to perturb gene expression equilibrium. Wild type nematode worms could sustain these perturbations without developing significant abnormalities, however nematodes with miRNA deletions developed notable defects of germline development and a significant proportion of the worms was sterile. Such combinatorial interaction was identified for four non-essential for viability miRNAs (out of eleven included in the analysis) and three out of five genes involved in signalling (Brenner et al., 2010). Interestingly, the nature of the signalling pathways is such that expression of their members vary in different contexts, which makes them good candidate targets of miRNA mediated buffering. Indeed, regulation of signalling pathways was noted to be a feature of miRNA targeting (Inui et al., 2010).

Cellular stresses are known to shift gene expression from the state of homeostatic equilibrium, and several miRNAs have been implicated in conferring robustness to both developmental programs and the homeostatic state in the face of stresses (Herranz and Cohen, 2010; Leung and Sharp, 2010). As an extension to the example discussed above, of non-essential *C. elegans* miRNAs acting as buffers of perturbation in gene expression, several non-essential miRNAs in other organisms have been identified as key regulators of stress responses. Perhaps the best studied example of a miRNA imparting robustness to a developmental program under stress is that of miR-7 in sensory organ development of the common fruit fly, *Drosophila melanogaster* (Li et al., 2009). Despite miR-7 being perfectly conserved between protostomes and deuterostomes (Sempere et al., 2006), mutant fruit flies that lacked miR-7 developed normally. However, if the mutant larvae were subjected to temperature fluctuations, the two key transcription factors, Yan and Atonal, were abnormally expressed in the eye and antennal cells. Additionally, a sensory organ precursors (SOP), called arista, failed to develop and SOPs for coeloconic sensillae either failed to develop or were patterned abnormally. In eye and SOP development, miR-7 is

thought to act as a buffer against perturbation in a gene expression network through a feed forward inhibition (both direct and indirect) of the two aforementioned transcription factors. Another miRNA in *D. melanogaster*, miR-14, was identified as protective against stress (Xu et al., 2003). Fruit flies with deletion of *miR-14* were viable, but had reduced lifespan and were significantly more susceptible to salt stress. Cell death was increased in the mutants under the stress, and at the same time several pro-apoptotic genes were found to be inhibited by miR-14, which suggested miR-14 role in suppressing the apoptosis during the stress. The importance of miRNAs for an appropriate stress response was also reported in vertebrates. In the zebrafish *Danio rerio*, miRNAs of the miR-8 family were found to be expressed in the skin and kidneys and to be required for adaptation to osmotic fluctuations (Flynt et al., 2009). The development of fish, with miR-8 knocked down, was indistinguishable from the wild type under normal conditions. However, if the fish were placed into a high osmomolarity buffer and then transferred into distilled water, oedema was observed in the mutants with significantly greater frequency than in wild type. In the mouse, knock out of the gene for miR-208 did not reduce viability, but disrupted normal stress response to thoracic aortic banding (van Rooij et al., 2007).

Summary of section 1.1

In this section I described the discovery of miRNAs and the current understanding of the functions and roles of miRNAs in the brain and neurons. The ability of a single miRNA to regulate the expression of dozens to hundreds of genes and the involvement of these targets in a spectrum of key developmental and neurological processes, suggests a role of great significance in the nervous systems and neurons in particular. In this regard, dispensability of many individual miRNAs is paradoxical. Recent research into miRNA function showed that miRNAs can act as buffers of plasticity in gene expression (for example, during neuronal plasticity and during stresses), which suggested an explanation of the apparent dispensability of many miRNAs in standard, stable laboratory environments.

Although the roles of miRNAs as buffers of gene expression programs has began to emerge from the current research, the full extent of the functional significance of miRNA is likely not yet fully appreciated. The primary reason for this is a traditional "single gene approach" that was taken in many of the previous studies into miRNA biology, where some of the miRNA targets are studied in detail, while the effect of miRNAs on hundreds of other potential targets was frequently overlooked. This set the scene for this thesis project, where I used methods of whole transcriptome profiling for the identification of

hundreds of miRNA targets with the goal to better understand the roles of miRNAs in neurons.

1.2 Thesis aims and an experimental paradigm

Individual miRNAs have been shown to regulate hundreds of target mRNA transcripts (Stark et al., 2003; Enright et al., 2003; Farh et al., 2005; Lim et al., 2005; Giraldez et al., 2006; Baek et al., 2008; Selbach et al., 2008). The functions of even the most studied miRNAs are generally perceived through a prism of only a handful of validated miRNA-mRNA interactions. The significance of miRNAs for the development of metazoans has been recognised since their discovery, however relatively little is known about their function in committed cell types. Numerous published studies described targets of miRNAs important for development of the nervous system as a whole and neurons in particular. However, the function of miRNAs in mature neurons is less well understood. This incomplete understanding of the roles of miRNAs in differentiated cell types, in conjunction with the paradoxical dispensability of many individual miRNAs for animal viability (Miska et al., 2007; Brenner et al., 2010), made the study of the roles of miRNAs in differentiated neurons relevant.

Computational approaches for miRNA target analysis have been prevalent since 2003. Purely computational approaches have many pitfalls and a demonstrated over-prediction bias (Giraldez et al., 2006; Baek et al., 2008). In order to address this situation a number of approaches have been developed since. The three main methods are currently: mRNA profiling after miRNA perturbation, large-scale proteomic approaches after perturbation and direct sequencing of mRNA targets bound to RISC component enzymes. At the time of the design of this thesis project, only the first method of high throughput identification of putative miRNA targets had been established and received experimental validation. This method involves overexpression of miRNAs in cells by exogenous addition of miRNA mimics, and subsequent elucidation of miRNA targeting through microarray profiling of incurred changes in the transcriptome (Lim et al., 2005; Giraldez et al., 2006). This technique was suitable to study miRNAs in neurons, because neuronal gene expression can be studied in primary cultures (Valor et al., 2007), and these cultures can be efficiently transfected with miRNA mimics (Conaco et al., 2006). Additionally, the technology behind microarrays and protocols for their analysis are both mature and inexpensive. For these reasons I decided to assay miRNA targets in neuronal cultures using microarray profiling of differential expression triggered by transfection of miRNA mimics.

While the other two approaches were not available at the time of the design of this thesis project, it is still worth describing their strengths and weaknesses in comparison to the assay used. The first method is conceptually similar to the approach used during the thesis project. However, instead of profiling the effects of miRNA mimics at the transcript level, profiling is performed at the level of the proteome (Baek et al., 2008; Selbach et al., 2008). The advantage of this method is the possibility to identify miRNA targets that are not regulated by miRNAs at the level of mRNA stability. However, recent research shows that a majority of mammalian miRNA targets are indeed regulated predominantly at the level of mRNA stability (Guo et al., 2010). Additionally, these experiments are expensive and time-consuming.

The second method is based on sequencing of RNA co-precipitated with proteins of the RNA silencing machinery (Licatalosi et al., 2008). These methods, such as HITS-CLIP, par-CLIP and iCLIP are very promising for miRNA target research. Firstly, because the technique does not require perturbation of miRNA expression, it can potentially identify the *in vivo* targets of miRNAs in cognate tissue and organs. However, this is a new technology with great promise but many obstacles remain. Issues arise for a number of reasons including: availability and specificity of antibody used, technical variation and amplification biases from sequencing and finally poor capture of mRNA sequence compared to an abundance of miRNA sequence (Anton Enright, Eric Miska, Jernej Ule and Donál O'Carroll personal communication). Although this approach currently has many technical limitations that need to be addressed, it would appear, at least conceptually, to be the most promising approach for future work. A limited comparison between published miRNA targets defined by HITS-CLIP and by my own approach is given in Chapters 5 and 6.

Aims: To summarise, in this thesis I aimed to characterise the roles of miRNAs in neurons by identification and analysis of miRNA targets in primary neuronal cultures. The identification of targets was achieved through perturbation of levels of miRNA expression in the cultures, analysis of the incurred differential gene expression by microarrays and derivation of the lists of targets from the profiling data.

Below I will describe methods of chemical transfection as a way of inducing miRNA mediated perturbations of the transcriptome. I will also describe the seed enrichment analysis as an approach to derive putative direct targets from microarray profiling data. Finally, E17.5 mouse primary neuronal cultures will be introduced as a model system to study the biology of neurons.

1.2.1 Widespread effect of miRNAs on the transcriptome is guided by miRNA seed region

Complementarity between miRNA and mRNA sequences was suggested as a mechanism of targeting in the two articles reporting the discovery of the first known miRNA, lin-4, and its target, lin-14 mRNA (Lee et al., 1993; Wightman et al., 1993). Other features of target recognition by miRNAs were also noted in the original reports: only partial complementarity was required between miRNA and mRNA transcripts, and the sites in the mRNA that annealed to the miRNA were located in the 3′UTR of the transcript. Subsequently, another miRNA discovered in *C. elegans*, called let-7, was also shown to recognise the mRNA targets through basepairing with their 3′UTRs (Reinhart et al., 2000). These observations laid the foundation for the first computational sequence based miRNA target prediction algorithms (Stark et al., 2003; Enright et al., 2003; Lewis et al., 2003). These algorithms have one assumption in common: transcripts with one or more sites in their 3′UTRs, which are partially complementary to a miRNA, are more likely to be targeted by that miRNA than transcripts drawn at random. Around the time when the first target prediction algorithms were being developed, Eric Lai noted that the sequences between positions 2 to 8 at the 5′-end of several miRNAs were perfectly complementary to the 3′UTRs of a selection of post-transcriptionally regulated transcripts (Lai, 2002). Additionally, 5′-regions of fly and worm miRNAs were noted to be more evolutionarily conserved than the rest of the sequence (Lai, 2002; Lim et al., 2003). Based on this observation, and on analysis of the available validated miRNA-target pairs, Lewis and colleagues proposed that perfect complementarity in the positions from 2 to 8 at the 5′-end of miRNAs with the sequence of 3′UTRs was a key determinant in miRNA target recognition (Lewis et al., 2003, 2005). This important for target recognition region in the miRNA sequence received the name “miRNA seed region”, while the complementary sequence in the target sequence was called the seed matching site. Experiments with reporter constructs confirmed the importance of basepairing to the seed for targeting. Presence of a single seed matching site in a 3′UTR was shown to be sometimes sufficient for miRNA targeting of a transcript (Doench and Sharp, 2004; Lai et al., 2005), which further justified the use of seed matching sites for prediction of the targets.

Efforts from the laboratory of David Bartel greatly contributed to current understanding of the role of miRNA seed region in target recognition (Bartel, 2004, 2009). Through microarray profiling of changes in transcriptomes upon overexpression of miRNAs, Bartel and colleagues showed that complementarity of six to eight bases between a miRNA seed

region and 3′UTRs was frequently associated with destabilisation of these transcripts (Grimson et al., 2007). The efficiency of destabilisation was lower for transcripts that had sites with only six bases of complementarity than sites with seven and eight bases. Interestingly, sites with six bases of complementary to the seed region (positions 2 to 7) followed by an adenine nucleotide were significantly more effective at destabilisation than the six bases on their own. Sites with six or eight base complementarity will be referred to as 6(2)-type and 8(2)-type seed matching sites respectively, while the sites of seven base complementarity and six base complementarity followed by an adenine nucleotide will be referred to as 7(2)-type and 7(1A)-type seed matching sites.

On the basis of complementarity to the seed being sufficient to enable miRNA mediated regulation, individual miRNAs were predicted to potentially directly regulate expression of dozens to hundreds of genes (Brennecke et al., 2005; Grün et al., 2005; Lewis et al., 2005; Stark et al., 2005). Validation of a wide spread regulation of gene expression by miRNAs came from the whole transcriptome analysis of miRNA activity. Ectopic expression of miRNAs was found to induce wide spread changes in gene expression, where 3′UTRs of downregulated transcripts were significantly enriched in miRNA seed matching sites (Lim et al., 2005; Giraldez et al., 2006). Crucially, a significant fraction of the downregulated transcripts with the seed matching sites in their 3′UTRs were validated as direct miRNA targets using luciferase reporter assay (Lim et al., 2005). The validation of direct targeting served as a proof that a significant component of the widespread downregulation of gene expression was directly caused by miRNAs inhibiting their direct targets. The same conclusions were drawn from the reciprocal experiments, where individual miRNAs were removed from the system. In these experiments, 3′UTRs of transcripts upregulated were enriched in the seed matching sites (Krützfeldt et al., 2005; Rodriguez et al., 2007), which is consistent with the inhibitory role of miRNA. An elegant demonstration of the scale of direct miRNA mediated effects that can be identified through whole transcriptome profiling was made in the laboratories of Karen Steel and Anton Enright (Lewis et al., 2009). There, transcripts upregulated in a mouse mutant, which had a single substitution in the seed region of miR-96, were enriched in seed matching sites for the original miR-96 seed, while transcripts downregulated in the mutant were enriched in the sites complementary to the acquired miR-96 seed.

Regulation of gene expression by miRNAs on a large scale was also observed for endogenous miRNAs through profiling of transcription in specific tissues (Farh et al., 2005; Sood et al., 2006). Expression of some miRNAs was shown to be highly tissue specific (Wienholds et al., 2005; Landgraf et al., 2007). Therefore, given the multitude of targets

an individual miRNAs may have, the inhibitory activity of highly expressed and tissue specific miRNAs could shape gene expression of cognate tissues (Stark et al., 2005). Systematic analysis of the distribution of seed matching sites in transcriptomes of a wide range of tissues confirmed the broad impact of miRNA regulation on tissue specific gene expression (Farh et al., 2005; Sood et al., 2006). For example, the 3′UTRs of highly expressed in the liver transcripts were depleted of seed matching sites for liver specific miR-122, in the muscle – for muscle specific miR-1 and in the brain – for brain specific miR-124.

The biological reason behind the importance of basepairing between the seed region and the target sequence was understood when the crystal structure was solved for miRNA bound to a bacterial homologue of the metazoan Argonaute-component of RISC complex (Ago proteins) (Wang et al., 2008c). The edges of the bases in position 2 to 6 at the 5′-end of the miRNA were found to be readily available, and thought to nucleate annealing to the target. When the structure of the ternary complex of Ago-miRNA-target sequence was solved, it was found that bulges in positions 2 to 8 at the 5′-end of the miRNA were poorly accommodated (Wang et al., 2008b), which confirmed the importance of complementarity between the seed site and the target.

Traditionally, functional seed matching sites were identified in 3′UTRs of the target transcripts (Lee et al., 1993; Wightman et al., 1993). Later 3′UTRs of miRNA targets were experimentally found to be significantly enriched in seed matching sites for that miRNA (Lim et al., 2005; Giraldez et al., 2006). This allowed the use of statistics of enrichment of seed matching sites in 3′UTRs of transcripts to discern miRNA mediated effects on gene expression and the compilation of lists of putative direct targets in this thesis (Chapter 5). However, using RNA precipitation and new generation sequencing (i.e. a method which is not reliant on seed enrichment statistics to identify putative direct targets) showed that a significant fraction of target sites are located outside of the 3′UTRs of putative targets (Chi et al., 2009). For example, in a recent study, of all sites in putative direct miRNA targets in embryonic stem cells $\approx 30\%$ were found to be located in coding region (Leung et al., 2011). Nevertheless, the greatest enrichment of the seed matching sites (per length of a sequence) in miRNA perturbation experiments was routinely observed in 3′UTRs of transcripts responding to the perturbations (Cei Goodger-Abreu, personal communication). Therefore, for the purpose of this thesis, miRNA targeting was considered in a traditional way and only 3′UTR sequences were used for seed matching site based compilation of putative miRNA targets (Chapter 5).

Two main conclusions can be drawn from the results of the experiments and analysis described in this section. Firstly, a single miRNA can directly inhibit the expression of dozens to hundreds of genes, a property that shapes expression of the whole transcriptome. Secondly, profiling of differential gene expression upon perturbation of individual miRNAs, in combination with the search for seed matching sites in 3′UTRs, enables identification of direct targets of miRNAs. These principles were used in this thesis to study function of miRNAs, which is described in more detail in the next section.

1.2.2 Transfections of miRNA mimics enables identification of direct miRNA targets

Cationic lipid transfections were developed over two decades ago as a method of introducing DNA molecules into cultured eukaryotic cells (Felgner et al., 1987). Efficiency of cationic lipid transfection was higher than that of some of the more traditional methods, including calcium phosphate transfections, and it became a popular tool that is now widely used for transfections of both DNA and RNA molecules into cells in culture. Transfections of mammalian cell cultures with miRNA mimics (double-stranded RNA, with one of the strands being equivalent to mature miRNA), was successfully used to overexpress miRNAs (Lim et al., 2005; Conaco et al., 2006; Baek et al., 2008; Selbach et al., 2008), and transfection with miRNA inhibitors (single-stranded chemically modified RNA complementary to mature miRNAs) was used to inhibit activity of endogenous miRNAs (Krützfeldt et al., 2005; Conaco et al., 2006; Selbach et al., 2008). Mechanism for the transfection of DNA was previously studied and reported in the literature (Zabner et al., 1995; Xu and Szoka, 1996), and the same mechanisms were suggested for transfection of RNA (Schroeder et al., 2010). Cationic lipids were suggested to form aggregates with nucleic acids that can be engulfed by cells through endocytosis (Zabner et al., 1995). After the nucleic acid/lipid complex is internalized, cationic lipids interact with endosomal membranes, which destabilises the aggregates and leads to the release of DNA from cationic lipids into the cytoplasm (Xu and Szoka, 1996). In the case of DNA transfections, the limiting step in efficiency of transfection was proposed to be the transfer of DNA from the cytoplasm into the nucleus (Zabner et al., 1995). This step was not considered to be an obstacle for miRNA perturbation experiments, because the biological activity of miRNA mimics and inhibitors takes place in the cytoplasm. Similarly high efficiency was expected in experiments with miRNA mimics and inhibitors, because efficiency of DNA

entry into the cytoplasm was shown to be high ($\approx$ 72.3% after 24 h incubation (Zabner et al., 1995)).

Stable knock out mice have been used in miRNA research (Rodriguez et al., 2007; Elia et al., 2009; Xin et al., 2009), however transfection of miRNAs had several advantages for studying miRNAs in mature neurons. The first advantage is the relative simplicity with which transfection experiments can be performed. In comparison to creation of knockout lines, transfections can be performed faster, thus, even within time constraints of the thesis projects, it was possible to study the functions of several miRNAs. The second advantage is that transfections are acute, meaning that the development of the transfected cells is no different from that of control cells prior to the experiment itself. The latter aspect was particularly important in this project on miRNAs in differentiated neurons, because miRNAs are involved in neuronal differentiation (see section 1.1.3). Because of this, analysis of mature neurons in stable knockout mutant lines would be problematic due to their development being affected prior to the experiment. This problem could have in principle been circumvented through creation of conditional knockout lines, however creation of such mice was not logistically possible within the time constraints of the project.

An important factor arguing for the use of transfections to study miRNAs was a report of successful detection of miRNA targets with this methodology in a seminal work by Lee Lim and colleagues (Lim et al., 2005). Two miRNAs, miR-1 and miR-124, were transfected into HeLa cell culture. Sets of genes significantly downregulated in the two experiments ($P < 0.001$) encoded transcripts that were enriched in seed matching sites for miR-1 and miR-124 ($P < 7.0e-27$ and $P < 1.1e-54$). This enrichment suggested that downregulation of a significant fraction of genes was directly caused by the inhibitory activity of the transfected miRNAs, and the downregulated transcripts with seed matching sites comprised lists of putative direct targets of the two miRNAs. This proposition was supported through a validation experiment: using a luciferase reporter system (Lewis et al., 2003), direct inhibition was confirmed for six out of ten selected targets. Further confirmation that targets identified in the transfection experiment by Lim and colleagues came from a different study conducted on primary neuronal cultures (Conaco et al., 2006). In that work, a mimic for miR-124 was transfected into primary neuronal cultures, and 17 out of 17 genes, which were inferred by Lim et al. as direct targets of miR-124, were significantly downregulated in neuronal cultures. Moreover, upon transfection of neuronal cultures with an inhibitor for miR-124, ten out of the 17 genes were significantly upregulated (Conaco et al., 2006). The experiments with miR-124 inhibition in neuronal

cultures showed that over half of genes, which were identified by transfection of mimics as putatively direct targets, were likely to have been the innate miR-124 targets in primary neurons.

Because of the logistical advantages and reported effectiveness of transfections of miRNA mimics for identification of direct miRNA targets, the transfection of miRNA mimics in primary neuronal cultures was chosen as the basis of the experimental approach in this thesis. A transfection protocol using a cationic lipid reagent, DharmaFECT 3, was optimised for the transfection of siRNAs into primary neuronal cultures by my colleague Dr. Erik MacLaren (Maclaren et al., 2011). This protocol was used for transfecting miRNA mimics and inhibitors (Methods, section 2.5), because miRNA mimics and inhibitors have a similar length to siRNAs.

1.2.3 The use of seed enrichment analysis of direct miRNA effects and identification of targets

Before the list of genes downregulated in transfection experiments could be used for the identification of putative direct miRNA targets, it was necessary to confirm that overexpressed miRNAs were likely to be the direct cause of differential gene expression. miRNAs can guide inhibitory RISC to the targets through complementarity of short sequences (six to eight nucleotides (Lewis et al., 2003, 2005)). Sequences that are complementary to a miRNA seed region can be encountered by chance in 3′UTRs, because the seed region is short. Therefore, even if a transfection of a miRNA mimic failed to induce miRNA mediated inhibition, it was possible to falsely identify downregulated transcripts with the seed matching sites as direct miRNA targets. To avoid such false positive results, it was important to develop a method that could discriminate between experiments where transfection of miRNA mimics failed to elicit miRNA mediated inhibition. A significant hypergeometric overrepresentation of seed matching sites for a particular miRNA in 3′UTRs of downregulated transcripts suggests that their downregulation was likely to be directly caused by the miRNA. Not all of the transcripts that are downregulated during miRNA overexpression experiments and bearing seed matching sites for the miRNA are real miRNA targets, however over 60% of these putative direct targets have been experimentally validated in previous studies (Lim et al., 2005; Giraldez et al., 2006). Furthermore, significant biases in the distribution of seed matchings sites have been observed in tissue specific expression profiles (Farh et al., 2005; Sood et al., 2006). These biases represent the depletion of the seed matching sites for highly expressed endogenous miR-

NAs from the 3′UTRs of highly expressed mRNA transcripts (Farh et al., 2005; Sood et al., 2006). The correlation of natural depletion signals with cognate miRNAs provided additional evidence for shifts in seed distribution as indicative of the biological activity of the miRNAs.

To identify signals of direct miRNA effects on gene expression profiles, it is possible to select genes based on a differential expression cutoff (e.g. $P < 0.05$), and to analyse composition of their 3′UTRs. Such an approach was used in the Lim et al. study, where significant enrichment of the seed matching sites for overexpressed miRNAs was detected in 3′UTRs of downregulated genes at the cutoff $P < 0.001$ (Lim et al., 2005). Although a simple hypergeometric test can produce useful information, there are several crucial drawbacks to this approach that prevent it from being used as the sole tool for discovery of miRNA mediated effects. First, 3′UTR length and composition biases cannot be easily accounted for with this traditional approach. Second, an arbitrary selection of the differential expression cutoff can artificially increase or decrease the size of the selection of putative direct miRNA targets. In order to account for length and composition biases and not to rely on arbitrary differential expression cutoffs, my colleagues, Stijn van Dongen and Cei Abreu-Goodger in the Enright laboratory, developed a method of nucleotide word enrichment analysis, called Sylamer (van Dongen et al., 2008).

Sylamer works by identifying occurrence biases of nucleotide words (one to 15 bases long) in a sorted list of sequences. For example, Sylamer can estimate enrichment or depletion of seed matching sites in a list of 3′UTRs of all transcripts that were detected by microarray transcriptome profiling, when this list is ranked from most downregulated to most upregulated. Assessment of the enrichment of a particular nucleotide word of a given length is done by calculation of hypergeometric enrichment P-value of that word in samples of sequences (or bins) from the list. Sampling of sequences from the ordered list is done from the most downregulated to the most upregulated and the size of a leading bin is iteratively incremented (i.e. at each step the leading bin includes all previously sampled sequences plus a certain number of new sequences). For example, if the size of the increment is 100 sequences, then the first bin includes 0 sequences, the second bin includes 100 sequences, the third bin includes 200 sequences and so on until the leading bin includes all sequences from the list. At each step, Sylamer calculates enrichment P-value of a particular nucleotide word by comparing its occurrence in the leading bin to its occurrence in all sequences of the list. Importantly, Sylamer operates on counts of nucleotide words of a given length per bin and per whole list, and the values of words-per-3′UTR are never a part of the equation, which automatically excludes the possibility

of a length bias. Additionally, correction for composition biases was incorporated into Sylamer, where hypergeometric statistics of a word of a given length can be adjusted to account for biases in underlying distributions of the related words of a shorter length (van Dongen et al., 2008).

The output of Sylamer can be visualised by plotting lines that correspond to occurrence biases of words of a given length in all bins from a sorted gene list. For example, Figures 1.1a and 1.1b show a cartoon diagram of a mock Sylamer line representing an occurrence bias of a single word. The x-axes of these plots corresponds to sorted 3′UTRs sequences. The y-axes values greater than 0 indicate enrichment and values less than 0 indicate depletion of the words in the bins when compared to the whole list. A shortcut to interpreting Sylamer plots is to consider the direction of the slope of the lines as indication of enrichment or depletion (shown as dashed arrows, Figures 1.1a and 1.1b). As shown in the cartoons, where the line goes upward there is an enrichment of the word in the underlying 3′UTRs, while if the line slopes downward there is a depletion.

In the analysis of the data presented in this thesis, the Sylamer plots illustrate occurrence biases of 876 distinct seven nucleotide seed matching sites[1], (7(2) and 7(1A) types) complementary to the seed regions of all known mature mouse miRNAs[2], which is 591 mature miRNAs according to miRBase Release 14 (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006, 2008) (Figure 1.1c and 1.1d). Two types of seven nucleotide seed matching sites (7(2) and 7(1A)-seed matching sites) were used here as predictors of miRNA targeting (Introduction, section 1.2.1). Therefore, in this thesis Sylamer analysis was always performed for nucleotide words of length seven. In the Sylamer plots, distribution of all words is shown as grey lines, apart from two to four manually selected words, which are shown in colors. As in the example cartoon diagrams, the x-axes of these plots correspond to sorted 3′UTRs sequences of genes, with one 3′UTR being selected per gene (Methods, section 2.7), and genes are usually sorted from most downregulated to most upregulated. The y-axes correspond to hypergeometric P-value of the words in each of the bins. The

[1] 6-mer and 8-mer seed matching sites can also be used as predictors of miRNA targeting, however 7-mers were reported to have a better predictive power for evolutionarily conserved miRNA targets (Bartel, 2009). In Sylamer analysis of distribution of nucleotide words and subsequent prediction of putative direct miRNA targets, peaks in occurrence biases of 7-mer sites (7(2) and 7(1A)-types) were seen to provide a higher signal to noise ratio for prediction of direct targets, than 6-mer and 8-mer sites (Cei Goodger-Abreu, personal communication).

[2] Three seed matching sites, *CAATAAA*, *TATTTAT* and *TCAATAA* that are similar to the *AATAAA*, a polyadenylation signal in 3′UTRs (Connelly and Manley, 1988), had to be excluded from Sylamer analyses, as biases in distribution of these words could not be attributed to miRNA mediated effects.

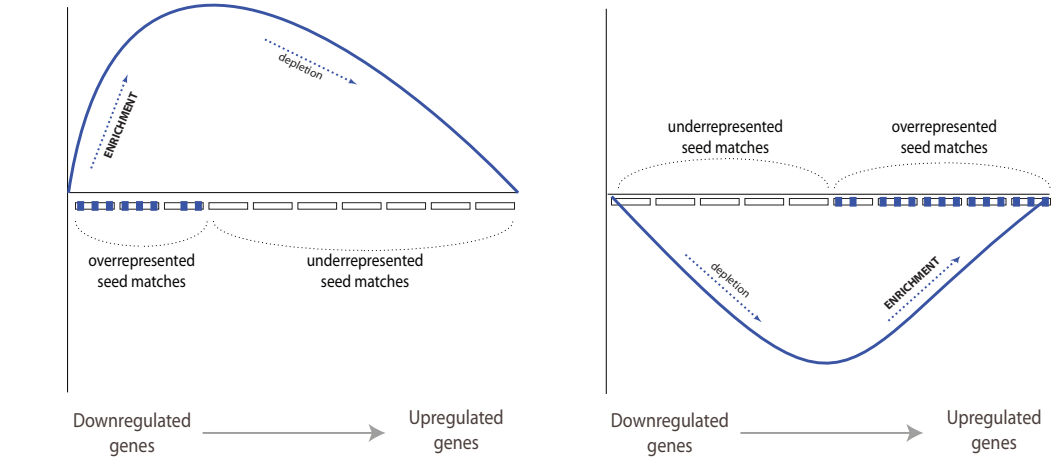
y-axes values above 0 indicate enrichment (equivalent to $-\log_{10}$ of the P-value) and values below 0 indicate depletion (equivalent to $\log_{10}$ of the P-value).

Visualisation of the results allows one to view the enrichment curve for a word or several words of interest in the context of the enrichment curves of all microRNA seed matching sites of the same length. This context defines a natural background, and a significant result should in all cases stand out among it. For example, the blue lines in Figures 1.1c and 1.1d show that occurrence bias of the word *GTGCCTT* is more significant than that of the background. This ability to instantly evaluate how unique the shifts in distribution of the words are, can qualitatively prove or disprove that selected miRNAs are specifically involved in the regulation of the gene expression patterns that are reflected in sorted 3′UTRs.

Sylamer plots can also help to select a threshold of differential expression that is most relevant to observed distributions, a choice that otherwise is arbitrary. For example, in Figure 1.1c genes are sorted by t-statistics from the most downregulated to the most upregulated upon overexpression of miR-124 (Chapter 5, section 5.1.2). Vertical lines show cutoffs of P-value of differential expression of 0.01 and 0.05 on both sides of the list (the x-axes). The steady rise of the colored line reflecting enrichment of a seed matching site (*GTGCCTT*, blue) peaks near P-value cutoff 0.01 at the left side of the plot (where downregulated genes are). Therefore, from this plot it follows that the most appropriate way to compile a list of candidate direct miR-124 targets is to select genes that were downregulated with differential expression P-value below 0.01, and subselect among them the genes that bare miR-124 seed matching sites in their 3′UTRs. Without this test, a researcher may be tempted to select a more relaxed cutoff (0.05) or a more stringent cutoff (0.001), which would artificially increase or decrease the selected list of candidate targets.

(a) Enrichment in downregulated genes

(b) Enrichment in upregulated genes



(c) Enrichment in downregulated genes

(d) Enrichment in upregulated genes

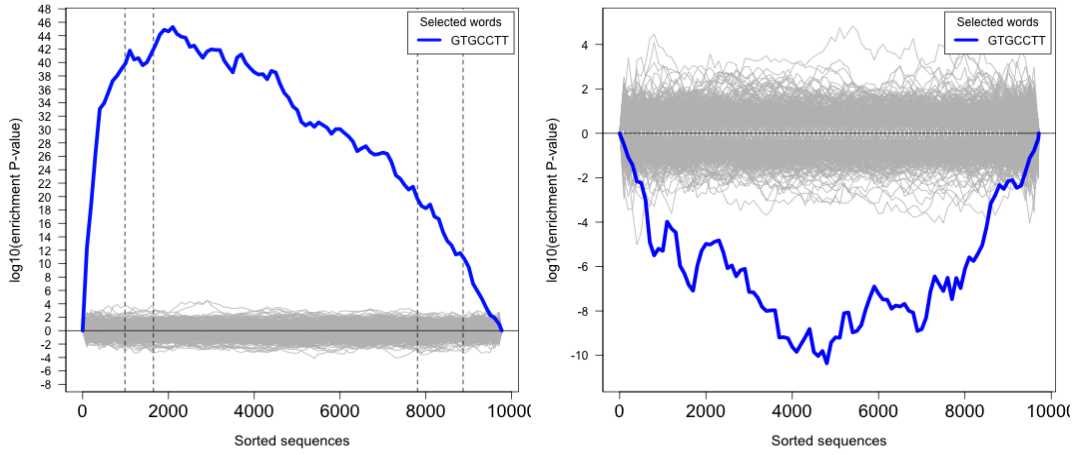


Figure 1.1: How to interpret Sylamer plots. *[The legend is on the next page]*

Figure 1.1: How to interpret Sylamer plots. *[The figure is on the previous page]*
The thick blue line in the cartoon Sylamer plots (Figures 1.1a and 1.1b) corresponds to the bias in the occurrence of one nucleotide word, which is represented by the small blue dashes in the ordered list of 3′UTR sequences (shown by the rectangles aligned under the x-axis). The 3′UTRs are ordered according to the change in expression of the corresponding genes: on the left side of the list there are downregulated genes, while on the right side there are upregulated genes. In Figure 1.1a the word is overrepresented in the left part of the list, where the 3′UTRs correspond to the downregulated genes. In Figure 1.1b the word is overrepresented in the right part of the list, where the 3′UTRs correspond to the upregulated genes. The y-axis qualitatively describes these biases: points on the line with the y-coordinate greater than 0 represent enrichment of the word in 3′UTRs preceding that point in the list, while the values less than 0 represent depletion of the word in the preceding 3′UTRs. The shortcut to understanding the Sylamer plots is to consider the direction of a slope of the line: the line that is going upward corresponds to the enrichment of the word in the underlying 3′UTRs, while the line going downward, corresponds to the depletion.
The lines in the Figures 1.1c and 1.1d correspond to the biases in occurrence of hundreds of different nucleotide words in thousands of ordered 3′UTR sequences (the x-axis). The 3′UTRs are ordered according to differential expression of the corresponding genes: on the left side of the list there are downregulated genes, while on the right side there are upregulated genes. The values on the y-axis correspond to the hypergeometric statistic for the enrichment/depletion of a word in the sequences that precede a point on the line in comparison to all sequences. The values greater than 0 correspond to the enrichment, while the values below zero correspond to the depletion. The occurrence of the majority of the words does not have significant biases (the grey lines). However, the distribution of one word ($GTGCCTT$), represented by the blue line, is empirically different from distribution of all other words. The blue line in the Figure 1.1c corresponds to the enrichment of the word in the 3′UTRs of the downregulated genes. In the Figure 1.1d the blue line shows enrichment in the 3′UTRs corresponding to the upregulated genes. The vertical dashed lines in Figure 1.1c show the P-value cutoffs (0.01 and 0.05 on the two ends of the x-axis) for the fold change t-statistic of the genes in the ordered lists. The dashed lines facilitate selection of the cutoff for identification of the lists of miRNA targets (see text), for example in Figure 1.1c the 0.01 cutoff on the left side corresponds almost exactly to the peak of the initial enrichment of the nucleotide word.

1.2.4 Primary neuronal cultures as a model system to study neuronal biology

In section 1.2.2 transfection of miRNA mimics was described as an efficient way to induce miRNA mediated changes in gene expression that can reveal the direct targets of miRNAs. Chemical transfections are usually performed on cells in culture, therefore to study the functions of miRNAs in neurons using transfections it was necessary to grow neurons in culture. Methods for culturing neurons were being actively developed for over a century (Nelson, 1975; Dichter, 1978). The cultures became a popular model for neuronal development and function, because they offered the researchers easy access to live cells, and made available nearly identical replicates for dose-response and timecourse experiments. One of these methods, known as dissociated primary neuronal culture, was used in this thesis to model neuronal growth and function and study the roles of miRNAs in differentiated neurons.

Dissociated primary neuronal cultures can be obtained from the brains of prenatal, neonatal and adult rats and mice (Brewer et al., 1993; Ahlemeyer and Baumgart-Vogt, 2005; Brewer, 1997; Brewer and Torricelli, 2007). Obtaining cultures from adult brains is more technically challenging, because dissociation of the brain cells is complicated by established adhesion between cell bodies and entanglement of mature neurites. Additionally, viability of cultured adult neurons was noted to be lower than that of prenatal and neonatal neurons and to require a constant supply of trophic factors (Brewer, 1997). Of the prenatal and neonatal brains, the former were selected as the source of neurons for this thesis project, as the use of embryonic material allowed the growth of more pure neuronal populations. The reason for this is due to the wave of embryonic neurogenesis preceding the wave of gliogenesis (Götz and Huttner, 2005; Freeman, 2010). Consequently it is possible to time the dissection of the embryonic brains in order to maximise proportion and number of neurons in the starting material. For example, it was shown that in primary cultures plated from the prenatal rat hippocampus (E18) less than 0.5% of cells were glial (Brewer et al., 1993), while in primary cultures from a neonatal mouse hippocampus the proportion of glial cells was approximately 7% (Ahlemeyer and Baumgart-Vogt, 2005), despite nearly identical isolation and culturing protocols.

In this thesis, in order to maximise the neuronal content of the plating material for primary neuronal cultures, mouse hippocampal and whole forebrain cultures were plated from E17.5 mouse embryonic brains, which approximately coincided with the end of embryonic neurogenesis (Götz and Huttner, 2005). To further enrich cultures for neurons,

the plated cells were cultivated in B27 supplemented Neurobasal (Methods, section 2.1), which is the media specifically developed to enhance neuronal and inhibit glial survival and growth (Brewer et al., 1993). Additionally, particular attention was paid to the control of the concentration of glutamine in the media, as it can deaminate to glutamate, an excitotoxic amino acid, and also because glial growth was suggested to be enhanced at higher glutamine concentrations (Brewer et al., 1993).

The experiments of this thesis project were designed based on the assumption that primary cultures, plated and incubated in conditions maximising neuronal content and survival, were a good model system to study neuronal biology. This assumption was supported by a body of existing evidence, where the cultures, similar to the ones used in this thesis, were shown to be a suitable system to study both neuronal physiology and gene expression. For example, some of the early studies on dissociated primary neuronal cultures showed that morphological characteristics of the cultured neurons were similar to the neurons in the brain (Dichter, 1978; Kriegstein and Dichter, 1983). Later it was demonstrated that neurons in the cultures were electro-physiologically active and developed functional synaptic connections (Bading et al., 1995; Hardingham et al., 2001). Moreover, it was also shown that it was possible to modify the strength of these connections, which meant that primary neuronal cultures may be used to study neuronal plasticity (Arnold et al., 2005).

Apart from studying the morphology and physiology of neurons, primary neuronal cultures can be a model system of choice to study the gene expression program of growing and functioning neurons. An advantage, that primary neuronal cultures provide, is the ease with which neuronal gene expression can be profiled: Because the cultures are enriched in neurons, neuronal gene expression can be studied simply by profiling the extract of total RNA. For example, in a study conducted in Seth Grant's laboratory, by profiling gene expression during development of cultured primary neurons, it was possible to identify hundreds of genes whose upregulation in cultures preceded the morphological and electrophysiological appearance of the synapses (Valor et al., 2007).

And finally, transfection of cultures with miRNA mimics and inhibitors, was shown to be suitable for the identification of miRNA targets (Conaco et al., 2006). Transfection of primary cortical cultures with the miR-124 mimic was shown to downregulate genes, while transfection of the inhibitor upregulates genes that were previously identified and validated as targets of miR-124 (Conaco et al., 2006).

Summary of section 1.2

Advances in our understanding of the mechanisms of miRNA target recognition, in conjunction with whole transcriptome analysis of gene expression, has enabled researchers to conduct experiments, which can identify the whole spectrum of putative direct miRNA targets at the same time. This approach is based on perturbation of expression of individual miRNAs followed by analysis of the sequences of 3′UTRs of differential expressed genes. 3′UTRs of transcripts targeted by a miRNA were shown be enriched in seed matching sites for that miRNA. Therefore identification of such transcripts, which also respond to perturbation of the miRNA expression levels, allows the compilation of lists of putative direct miRNA targets. Such an approach has been previously used, and lists of identified putative targets were significantly enriched in the validated direct miRNA targets (Lim et al., 2005; Giraldez et al., 2006). In this thesis I will describe the use of these experimental and computational techniques to identify putative direct targets of several miRNAs in primary neuronal cultures. Subsequent analysis of the lists of these targets allowed me to suggest explanations for some of the previous observation made in the published literature, and formulate testable hypotheses of miRNA functions in neurons.

Chapter 2

Methods

Equipment in continuous use

- Sparkfree laboratory refrigerator (+4 °C). Thermo Electron Corporation.
- Sparkfree laboratory freezer (-20 °C). Thermo Electron Corporation.
- Milli-Q® Gradient +A10 (Water purification system). EMD Millipore Corporation. Cat. no. QGARD00R1.
- Pipettes:
  - Gilson PIPETMAN® P2 (0.2 - 2 μl). Anachem Ltd. Cat. no. F144801
  - Gilson PIPETMAN® P10 (1 - 10 μl). Anachem Ltd. Cat. no. F144802
  - Gilson PIPETMAN® P20 (2 - 20 μl). Anachem Ltd. Cat. no. F123600
  - Gilson PIPETMAN® P100 (10 - 100 μl). Anachem Ltd. Cat. no. F123615
  - Gilson PIPETMAN® P200 (20 - 200 μl). Anachem Ltd. Cat. no. F123601
  - Gilson PIPETMAN® P1000 (200 - 1000 μl). Anachem Ltd. Cat. no. F123602
- Pippete tips (Ranin Aerosol Resistant Tips):
  - Capacity 10 μl. Anachem Ltd. Cat. no. RT-10F
  - Capacity 20 μl. Anachem Ltd. Cat. no. RT-20F
  - Capacity 100 μl. Anachem Ltd. Cat. no. RT-100F
  - Capacity 200 μl. Anachem Ltd. Cat. no. RT-200F
  - Capacity 1000 μl. Anachem Ltd. Cat. no. RT-1000F

2.1 Primary Neuronal Cultures

All mice were treated in accordance with the U.K. Animals Scientific Procedures Act of 1986, and all procedures were approved through the British Home Office Inspectorate.

Materials

Reagents

- C57BL/6 c/c mice at 18 or 19 days of gestation. Supplied on site.
- Fetal Calf Serum (FCS). Supplied on site.
- Dulbecco's Modified Eagle Medium (DMEM). Invitrogen[1]. Cat. no. 31330-038
- B-27 Supplement (optimised medium supplement for neurons (Brewer et al., 1993)). Invitrogen. Cat. no. 17504-044
- L-Glutamine 200 mM. Invitrogen. Cat. no. 31330-038
- Natural Mouse Laminin. Invitrogen. Cat. no. 23017-015
- Papain Vial. Worthington Biochemical Corprotation. Cat. no. PAP2
- Neurobasal Medium. Invitrogen. Cat. no. 21103-049.
- Penicillin Streptomycin (PenStrep). Invitrogen. Cat. no. 15140-122
- Dulbecco's Phosphate Buffered Saline 1× (1× DPBS). Invitrogen. Cat. no. 14190-094
- Poly-D-lysine hydrobromide (PDL). Sigma-Aldrich Corporation. Cat. no. 1000689047
- Ethanol 99.7 - 100% v/v. VWR International. Cat. no. 101707 2.5LT

Equipment

Dissection:

- 35 mm Petri dish. Corning Inc. Cat. no. 430588
- 55 mm Petri dish. Sterilin Ltd. Cat. no. PF55
- 100 mm Petri dish. Corning Inc. Cat. no. 430167
- 140 mm Petri dish. Sterilin Ltd. Cat. no. 501V
- 6-well cell culture plate. Corning Inc. Cat. no. 3156
- Pasteur pipette. Alpha Laboratories Ltd. Cat. no. LW4070
- Squirt bottle. Supplied on site.

[1]Invitrogen Corporation is a part of Life Technologies

- Leica MZ9.5 Stereomicroscope (binocular dissection microscope) with Leica CLS 150× light source. Meyer Instruments Inc.
- Thermo Scientific Holten HV Mini Laminar (sterile hood for dissections). Thermo Fisher Scientific. Cat. no. 54250130
- Scissors. Supplied on site.
- Dumont (curved) forceps. Fine Science Tools Inc. Cat. no. 11295-20
- Narrow Pattern Forceps, 2 pairs. Fine Science Tools Inc. Cat. no. 11002-12
- Dumont #5 Mirror Finish Forceps - Inox Biologie. Fine Science Tools Inc. Cat. no. 11252-23
- Iris Spatula - Slight Curve. Fine Science Tools Inc. Cat. no. 10093-13

Plating:

- 12 Well Cell Culture Cluster (12 well cell culture plate). Corning Inc. Cat. no. 3512
- 15 ml centrifuge tubes. Becton, Dickinson and Company. Cat. no. 4-2097-8
- AC1000 Improved Neubauer (cell counting chamber). Hawksley. Cat. no. AC1000
- Galaxy R CO_2 incubator. Wolf Laboratories Ltd
- Axiovert 200 (inverted microscope) with temperature control. Carl Zeiss AG
- 90 mm filter unit. Nalgene Nunc International. Cat. no. 450-0020
- Vacuum pump. Manufactured on site.
- Microbiological Safety Cabinet Class II. Holten.
- Sterile syringe filter, 0.2 μm. Nalgene Nunc International. Cat. no. 190-2520
- Swinging bucket centrifuge for 15-ml tubes. Eppendorf. Model: Centrifuge 5702
- Water bath. Grant Instruments (Cambridge) Ltd.
- Wide-bore glass tips. Manufactured on site.

Reagent and equipment setup

L-Glutamine solution L-Glutamine solution was aliquoted (1ml) and stored at –20 °C. *IMPORTANT:* Upon repeated freezing and thawing glutamine can spontaneously deaminate into excitotoxic glutamate, therefore re-freezing had to be avoided.

Laminin Laminin solution was thawed on ice, aliquoted (50 μl) and stored at -20 °C. For coating plates an aliquot of laminin was thawed on ice and mixed with 8 ml of ice-cold 1× DPBS. 600 μl of the mixture was used to coat one well of 12-well plate ($\sim$0.02 μg

$\cdot$ mm^{-2}). *IMPORTANT:* Upon rapid thawing or repeated freezing and thawing laminin may polymerize and form a gel, thus laminin solution was always handled on ice.

NeurobasalFull On the day of preparing the media, L-Glutamine solution (1ml aliquot) and B-27 (10ml) were added to 490 ml of Neurobasal. The mixture was filtered using the 90 mm filter unit and the vacuum pump. NeurobasalFull was stored at 4 °C and it was used until a change in its color was detected, which indicated a change in pH (approximately one month from the time of preparation).

DMEM + FCS An aliquot (50ml) of FCS was added to 450 ml of DMEM and filtered through 90 mm filter unit with vacuum pump. DMEM + FCS was used solely to terminate papain treatment and wash off the digested material. Similar to NeurobasalFull, DMEM + FCS was stored at 4 °C until change in pH was observed (approximately one month).

Papain Papain was used to partially digest dissected tissue, with the aim to dissolve extracellular matrix and preserve cell viability. A dry stock of papain was dissolved in 12 ml of 1× DPBS and stored at 4 °C. Maximal amount of the material that could be efficiently treated by a 1 ml papain aliquot was equal to that of two forebrains. An attempt to digest more material prevented an efficient digestion of extracellular matrix. Papain was filtered through a syringe 0.2 μm filter before use.

PDL Lyophylized PDL (5 mg) was dissolved in 50 ml of 1× DPBS to obtain 2× stock. Aliquots of 10 ml were stored at -20 °C. When defrosted, 10 ml of 1× DPBS were added to obtained 1× PDL. 1× PDL solution was stored at 4 °C for up to 2 weeks.

70% ethanol In order to decrease the risk of contamination of cells during dissection and plating, all work surfaces were treated with 70% ethanol (prepared in the squirt-bottle before the dissection procedure from 100% ethanol and Milli-Q water).

Coating cell culture plates 2 ml of 1× DPBS was added to the 10 outer wells of a 12 well cell culture plate in order to maintain humidity. 450 μl of 1× PDL solution was added into the two central wells and incubated for minimum 1 h (or overnight) at 37 °C in the incubator. During the incubation, a 50 μl aliquot of laminin was thawed on ice and diluted in 8 ml of ice cold 1× DPBS. After the incubation PDL was aspirated and wells were washed once with 1 ml of 1× DPBS. 600 μl of laminin solution was added to the two central wells and incubated for minimum 2 h (or overnight). After the incubation

laminin solution was aspirated and 950 μl of NeurobasalFull was added to the wells (this was always done within 2 h of plating cells). The cell culture plates were subsequently left in the 37 °C incubator to warm up the media.

Prewarming solutions Prior to dissections, papain (aliquots of 1 ml in 15 ml Falcon tube), DMEM+FCS and NeurobasalFull were prewarmed on the waterbath at 37 °C. This was necessary to eliminate an additional cold shock to the cells during plating.

Preparing of chilled 1× DPBS (5% PenStrep) Before dissections 1× DPBS with 5% PenStrep was prepared and chilled on ice. Concentration of antibiotics in 1× DPBS (5% PenStrep) is approximately 500 units of penicillin and 500 μg of streptomycin per 1 ml. 1× DPBS (5% PenStrep) was used to store dissected foetuses and heads prior to dissection. 140 mm, 100 mm, 35 mm Petri dishes and a 6-well plate were placed on ice and filled with ice-cold 1× DPBS (5% PenStrep).

Procedure

Dissection:

- A pregnant mouse (17.5 days post coitus) was killed by cervical dislocation. The abdomen was disinfected by a squirt of 70% ethanol. The uterus with fetuses was dissected, using scissors, and narrow forceps. After a quick wash with 70% ethanol, the uterus with fetuses was placed into a 140 mm Petri dish with 1× DPBS (5% PenStrep), and the fetuses were killed by cooling.
- The fetuses were dissected from the uterus and decapitated. Holding one head at a time with curved forceps, the heads were quickly washed in a 100 mm Petri dish prefilled with ice-cold 1× DPBS (5% PenStrep). Afterwards the heads were distributed among the wells of a 6-well plate prefilled with DPBS (5% PenStrep).
- One head at a time was placed into the upper lid of a 55 mm Petri dish prefilled with ice-cold DPBS (5% PenStrep) and placed under a dissection binocular microscope. To maintain the sterile environment, the dissection was performed in a mini laminar.
- Using mirror finish forceps, the skin and the calvarium were removed from the head.
- The cranial nerves were severed with a spatula, and the brain was removed from the skull.
- The forebrains were dissected using mirror finished forceps and a spatula.

- The meningi were removed from each hemisphere, using mirror finished forceps. If necessary, the hippocampi were dissected at this stage using mirror finish forceps.
- After removing the meningi, either the whole forebrains or the hippocampi were transferred with a Pasteur pipette into a 35 mm Petri dish prefilled with ice-cooled 1× DPBS (5% PenStrep). If the material appeared intact after the transfer, then it was further shredded either with scissors or by trituration through a Pasteur pipette (shredding facilitated the papain digestion stage). The dissected material was kept on ice in 1× DPBS (5% PenStrep) until the end of the dissections.

Plating cells:

- The dissected tissue was taken out 1× DPBS (5% PenStrep) and placed into 1 ml of prewarmed papain solution (in 15 ml centrifuge tube), carrying over as little PBS as possible. The material was incubated in papain solution for 25 min in 37 °C waterbath. *IMPORTANT:* To achieve better disruption of extracellular matrix, the material equivalent to at most two embryonic forebrains was digested in one tube at a time.
- After the papain treatment, all further manipulations were performed in a Microbiological Safety Cabinet Class II. As much as possible papain solution was quickly removed using a wide-bore glass pipette and 1 ml of prewarmed DMEM+FCS was added to the tissue, which terminated the papain lysis.
- The tissue was macerated through a wide-bore glass pipette and P1000 tip (narrow-bore). Further DMEM+FCS was added to bring the total volume to 5 ml.
- After making sure that no clumps of the undisrupted tissue remained, the obtained cell suspension was centrifuged at 400 g (in a swinging bucket centrifuge) for 3 min 30 sec.
- As much as possible DMEM+FCS was removed. 1 ml of NeurobasalFull was added and the pellet was triturated through P1000 tip.
- The cell suspension was centrifuged at 400 g for 3 min 30 sec.
- The supernatant was removed and approximately 2 ml of NeurobasalFull was added to the pellet. *IMPORTANT:* If several papain digestions were carried out in parallel, all the pellets were mixed before counting at this stage.
- After trituration of the pellet (with P1000), the cells were counted using a cell counting chamber and an inverted microscope.

- After counting, NeurobasalFull was added to the cells so that the desired plating volume was approximately 50 μl if possible (as it made total $1,000\mu$l of the growth media per well).

Throughout this work, cells were plated at two densities: $1,850$ cells $\cdot$ mm^{-2} and 790 cells $\cdot$ mm^{-2} (referred to as high and low densities). Cells were cultured in a humidified incubator, with CO_2 concentration held at 5%, and temperature at 37 °C.

2.2 RNA extraction

Materials

Reagents

Extraction:

- Dulbecco's Phosphate Buffered Saline 1× (1× DPBS). Invitrogen. Cat. no. 14190-094
- miRNeasy®Mini Kit. Qiagen N.V. Cat. no. 217004
- QIAzol®. Qiagen N.V. Cat. no. 79306
- RNase-Free DNase Set. Qiagen N.V. Cat. no. 79254. Components:
  - DNase I (solid)
  - RDD buffer
  - Nuclease free water
- Liquid nitrogen. Supplied on site.
- Ethanol 99.7 - 100% v/v. VWR International, LLC. Cat. no. 101707 2.5LT
- Chlorophorm, $\geq$ 99%. Sigma-Aldrich Corporation. Cat. no. C2432-500ML
- Nuclease-Free Water. Ambion Inc.[1] Cat. no. AM9937
- RNase Zap® wipes. Ambion Inc. Cat. no. AM9786
- Azowipe®. Vernon-Carus Ltd. Cat. no. 81103

Quality Control:

- Agilent RNA 6000 Nano Kit. Agilent Technologies. Cat. no. 5067-1511

[1] Ambion Inc. is a part of Applied Biosystems Inc. (Life Technologies)

Equipment

Extraction:

- RNase-Free 1.5 ml Microfuge Tubes. Ambion Inc. Cat. no. 12400
- -90 °C freezer. SANYO Electric Co. Limited. Model: MDF-450V
- Table top centrifuge. Eppendorf AG. Model: Centrifuge 5415D
- Centrifuge with thermocontrol. Eppendorf AG. Model: 5417R
- Vortex. Fisons Scientific Equipment. Cat. no. SGP-202-0109

Quality Control:

- NanoDrop Spectrophotometer. Thermo Fisher Scientific. Model: ND-1000
- RNA Nano Chips for use with Agilent 2100 Bioanalyzer. Agilent Technologies. Cat. no. 5067-1511
- Agilent Technologies 2100 Bioanalyzer. Agilent Technologies
- Agilent 2100 Expert Software. Agilent Technologies

Reagent and equipment setup

Good laboratory practice of working with RNA

Performing biological experiments always requires following a set of rules that to prevent detrimental contamination of samples. This is especially important when working with samples of RNA, as any contamination can become detrimental to integrity of RNA samples due to the abundance of RNase in the environment. Thus, it was important to treat all work surfaces, gloves and pipettes with RNase inhibitor containing solution (such as RNase Zap) and to subsequently clean the surfaces with a fast evaporating alcohol liquid (in Azowipes). RNA was stored in water, thus it was important to use nuclease free water. Overtime, contamination with RNase is perhaps inevitable, thus RNA samples were stored in -90 °C freezer and always handled on ice when out of the freezer. There is a widely held view that RNA can lose its integrity due to mechanical shearing by ice crystals during freezing and thawing cycles. Therefore the number of these cycles was minimized when possible.

Preparation of DNase I

Before the RNA extraction procedure, DNase I was dissolved in 550 μl of water (can be stored at 4 °C for one month) and mixed with RDD buffer (in 1:7 ratio).

Preparation of the centrifuge

Prior to the beginning of extraction a centrifuge with a thermocontrol must be cooled down to 4 °C.

Procedure

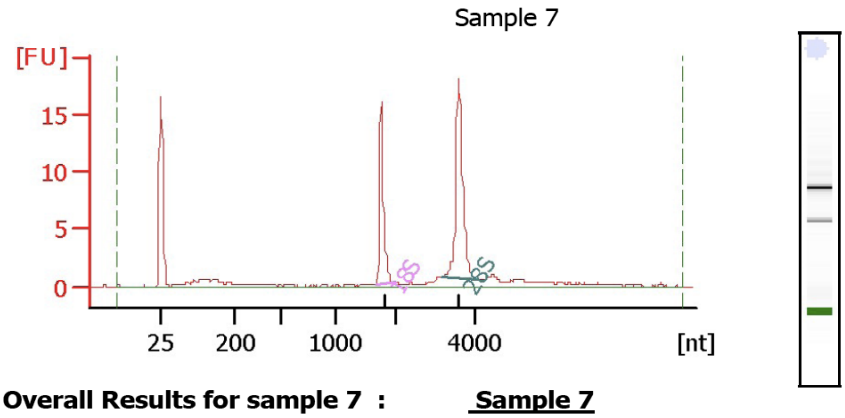
RNA extraction with RNeasy RNA extraction kit (Qiagen)

- Cultures from which RNA was to be extracted were transferred from the incubator onto the bench, the growth media was removed and the cultures were quickly washed once with 1× DPBS. After DPBS was removed, 700 μl of Qiazol was added per well and cells were scrapped off the bottom of the well using a P1000 tip. The suspensions were transferred into 1.5 μl microfuge tubes. *IMPORTANT:* It was possible that the stress stress associated with the removal from the controlled environment of the incubator could perturb gene expression in the cells. Therefore it was important to minimize the time between transportation and addition of Qiazol (usually, no more than 4 cultures were dealt with at a time).
- At this stage, it was possible to either snap-freeze the suspensions with liquid nitrogen and store at -90 °C or to immediately proceed with the rest of the extraction protocol. Freezing of the suspension was frequently more convenient when many cultures had to be dealt with at one time. For example, if 16 cultures were to be extracted, that would involve four rounds of Qiazol lysis (as only four cultures were taken out of the incubator at one time). Thus the lysates from the first batch of cultures would have remained significantly longer at the ambient temperature than the lysates from the last batch, which was not desirable.
- If the suspension was freshly obtained then it was incubated on a benchtop for 5 min. If the suspension was taken out of -90 °C freezer, it was incubated on benchtop until the lysate defrosted and appeared clear. Gentle shaking was found to increase the speed of defrosting.
- 140 μl of chloroform was added to the suspension and it was thoroughly vortexed for 15 s. Afterwards, the suspension was incubated on benchtop from 2 to 3 minutes.

- The suspension was centrifuged at 12,000 g for 15 min at 4 °C.
- Aqueous upper phase was collected into a new tube and 1.5 volume 100% ethanol was added to it. Usually it was possible to collect approximately 300 μl of supernatant without risking contamination from the bottom phase.
- The solution was vortexed for 3 sec and immediately transferred onto a spin column (a part of RNeasy kit).
- The colum was centrifuged for 15 sec at 10,000 g and the flow through was discarded.
- 350 μl of RWT buffer (part of miRNeasy kit) was added to the spin column. The column was centrifuged for 15 sec at 10,000 g and the flow through discarded.
- 80 μl of prepared DNase I solution (see section 2.2) was added to the column. The column was incubated with DNase I for 15 min on a benchtop.
- 350 μl of RWT buffer was added to the column. The column was centrifuged for 15 sec at 10,000 g and the flow-through discarded.
- 500 μl of RPE buffer (part of miRNeasy it) was added to the column. After 15 sec centrifugation at 10,000 g the flow-through was discarded.
- 500 μl of RPE buffer (part of miRNeasy it) was added to the column. After 2 min centrifugation at 10,000 g the flow-through was discarded.
- The column was carefully placed into a new 1.5 ml tube and centrifuged at full speed for 2 min. This step was described as an "optional" in miRNeasy manual, but I found it to be important for the removal of the residual solvent.
- The column was transfered into a new 1.5 ml and RNA was eluted with 30 μl of nuclease free water (upon 1 min centrifugation at 10,000 g). This step was repeated with the flow-through solution to obtain higher final concentration of RNA.
- The solution of RNA in water was stored at -90 °C (either immediately following the extraction, or after measuring RNA concentration and quality control using Nanodrop).

Measuring concentration of RNA and quality control

Concentration of RNA was determined using Nandrop following the manufacturer's protocol. RNA integrity was assessed using the Bioanalyzer machine and Bioanalyzer 6000 Nano kit, following the manufacturer's protocol. The analysis of total RNA on bioanalyzer enables visual assessment of the integrity of peaks corresponding to 18S and 28S ribosomal RNA. No significant degradation was ever detected in any of the samples (a representative Figure 2.1).



Overall Results for sample 7 : Sample 7

RNA Area:	67.6
RNA Concentration:	40 ng/µl
rRNA Ratio [28s / 18s]:	1.8
RNA Integrity Number (RIN):	10.0 (B.02.05)
Result Flagging Color:	
Result Flagging Label:	RIN:10

Fragment table for sample 7 : Sample 7

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
18S	1,694	1,983	14.6	21.5
28S	3,156	4,061	25.5	37.7

Figure 2.1: Example of an output of Bioanalyzer.

2.3 Quantitative RT-PCR. mRNA

Materials

Reagents

- Oligo(dT)$_{12-18}$ Primer. Invitrogen. Cat. no. 18418-012
- Random Primers (Nanomers). New England Biolabs Inc. Cat. no. S1254S
- Random Decamers RETROscript®. Ambion Inc. Cat. no. AM5722G
- 10 mM dNTP Mix (deoxyribonucleotides mix). Invitrogen. Cat. no. 18427-013
- SuperScript® II Reverse Transcriptase Kit. Invitrogen. Cat. no. 18064-014
- TaqMan® Universal PCR Master Mix, No AmpErase UNG (2×). Applied Biosystems Inc. Cat. no. 4324018
- Primers and probes mix for real-time PCR (TaqMan® Gene Expression Assays). Applied Biosystems Inc. Custom design
- Nuclease-Free Water. Ambion Inc. Cat. no. AM9937

Equipment

- 96-well plate (Thermowell 96 Well Plate Model (M)), Corning Inc. Cat. no. 6511
- 96-well plate cover (Microseal A film). MJ Research Inc. Cat. no. MSA-5001
- Optical 96-well plate (MicroAmp Optical 96-Well Reaction Plate). Applied Biosystems Inc. Cat. no. N8010560
- MicroAmp Optical Adhesive Film. Applied Biosystems Inc. Cat. no. 4311971
- Centrifuge with a 96-well plate rotor. DJB Labcare Ltd. Model: Heraeus Biofuge Stratos
- Peltier Thermal Cycler. MJ Research Incorporated
- NanoDrop Spectrophotometer. Thermo Fisher Scientific. Model: ND-1000
- 7500 Real Time PCR System. Applied Biosystems Inc.
- 7500 Real Time PCR System Sequence Detection Software v1.2.2. Applied Biosystems Inc.
- Primer Express® Software v3.0. Applied Biosystems Inc.

Reagent and equipment setup

Primers

Usage of internal controls is important for RT-PCR experiments, as the method is very sensitive to the amount of input cDNA. Thus in addition to profiling of the differentially expressed genes (Acta2 and Lass2), it was important to identify a stably expressed control gene. Microarray profiling of gene expression upon miR-124 over-expression and inhibition at 3DIV and 6DIV (Chapter 5, section 5.2.1) showed that some of the components of splicing machinery and translation factors were stably expressed. Two genes were picked from these categories as potential control genes: Sfrs7 (splicing factor, arginine/serine-rich 7) and Eif2b4 (eukaryotic translation initiation factor 2B, subunit 4 delta).

Primers and probes for the real-time PCR were designed to span two neighbouring constitutive exons (i.e. exons that were present in all annotated transcript of the gene, Ensembl version 56 (Flicek et al., 2008)). Primers were designed using Primer Express® Software v3.0 (Applied Biosystems Inc.) according to the manufacturer's protocol. Preference was given to neighbouring constitutively present exons at the 3′-end of the genes, as it maximized the chances of the regions of interest to be reverse transcribed with the oligo(dT) primer. The BLAST analysis (Altschul et al., 1990) confirmed that sequences of all primers were uniquely present in the mouse genome.

	Acta2	**Lass2**
Ensembl Gene ID	ENSMUSG00000035783	ENSMUSG00000015714
ExonA ID	4 ENSMUSE00000545192	3 ENSMUSE00000253057
ExonB ID	5 ENSMUSE00000545191	4 ENSMUSE00000253053
Forward Primer	CCCAGATTATGTTTGAGACCTTCAA	CAGACCAGCGGCAAGCA
Probe	TCCCCGCCATGTATGT	CCCAAGCAGGTGGAG
Reverse Primer	GGACAGCACAGCCTGAATAGC	CTCTGCCGTGACAAAAGGTCTA
	Eif2b4	**Sfrs7**
Ensembl Gene ID	ENSMUSG00000029145	ENSMUSG00000024097
ExonA ID	4 ENSMUSE00000186164	3 ENSMUSE00000138340
ExonB ID	5 ENSMUSE00000186166	4 ENSMUSE00000138342
Forward Primer	CAACAGGTTCCTACACGAAAGGA	CATCGCTATAGCCGACGAAGA
Probe	TACGGATCCAAAGTCA	AAGCAGGTCACGATCT
Reverse Primer	TGAGGCAGGTGGGAGAAGAG	CCCTGGATCGGGAATGG

Table 2.1: Real-time PCR primers and probes.

In a trial RT-PCR detection of Sfrs7 had lower standard deviation between and within treatments than Eif2b4. Therefore Sfrs7 was chosen as the control gene for the analysis of differential expression of Acta2 and Lass2.

Procedure

Reverse transcription: Generation of first strand cDNA

- Total RNA was mixed with dNTP (Table 2.2).

Component	per 12 μl mix
total RNA (40 ng $\cdot$ μl^{-1})	9.00
Random primers	1.00
Oligo dT	1.00
dNTP	1.00
TOTAL	12

Table 2.2: mRNA reverse transcription starting mix.

- The starting mix was transferred into 96-well PCR plate and incubated at 65 °C for 5 min in Peltier Thermal Cycler.
- The plate was transferred on ice. 4 μl of 5× 1 st strand buffer and 2 μl of DTT (0.1 M dithiothreitol from SuperScript® II Reverse Transcriptase Kit) were added to the samples.
- The plate was incubated at 42 °C for 2 min in Peltier Thermal Cycler.
- 1 μl of SuperScript II reverse transcriptase was added to the samples.
- The sample was incubated for 50 min at 42 °C, followed by 15 min incubation at 70 °C for 15 min.
- The product of the reaction (cDNA) could be stored at 4 °C overnight.

Real time PCR

Concentration of cDNA in all samples was normalized to 67.5 ng·ml^{-1}. Initially, several dilutions were tested and 67.5 ng· ml^{-1} was found to consistently produce the Ct values (see below for the definition of Ct) between 23 and 30 for amplification of Acta2, Lass2 and Sfrs7 transcripts, which is within the optimal range of 7500 Real Time PCR Sequence Detection System.

- 15 μl of 2× TaqMan Universal RT-PCR Mix was pipetted into the wells of optical 96-well plates. *IMPORTANT:* Pipetting errors are a very common source of variability in RT-PCR results. In order to minimize pipetting errors, a new pipette tip for each

well of the 96-well reaction plate was used for addition of PCR components, as I found that variable amounts of liquid may adhere to the walls of the tip and lead to inconsistencies.

- 10 μl of cDNA (at 67.5 ng $\cdot$ μl^{-1}), 5 μl of appropriately diluted Assay Mix (see above) was added to the wells to make 30 μl of total reaction volume. *IMPORTANT:* Pipetting of small volumes was found to be relatively imprecise and lead to additional variability in RT-PCR results. This was especially important for addition of primers-probe mix as the amount of the probe determined the detection of real time PCR progression. Thus the dilutions were scaled to obtain the volumes that could be pipetted relatively accurately.
- The plate was sealed with optical plate cover and centrifuged briefly in a centrifuge with a rotor for 96-well plates in order to bring all liquid to the bottom of the wells and remove large air bubbles.
- The real time PCR was conducted in the thermal cycler using the default 40 cycle program (Table 2.3).

HOLD	**HOLD**	**CYCLE (40 cycles)**	
2 min	10 min	15 sec	60 sec
50 °C	95 °C	95 °C	60 °C

Table 2.3: mRNA real time PCR program.

Analysis of expression by RT-PCR

TaqMan probes were used to evaluate the rate of PCR in real time. The probes were designed to anneal to the internal part of the amplification product. The probes had a fluorescent reporter dye at the 5′ and a quencher at the 3′-end. During the amplification, DNA polymerase hydrolysed the probe with its 5′ nuclease activity. This decoupled the reporter from the quencher and lead to the increase in the fluorescence with each cycle. During the initial stages of a typical correctly set-up PCR, the fluorescence stays at a baseline. As PCR progresses and the amount of product approximately doubles every cycle, the increase in the fluorescences eventually becomes detectable and to raise exponentially. Eventually, the pool of the available probe depletes and the fluorescence levels at a plateau.

If primers and a probe are designed correctly, the stage of PCR at which the fluorescence starts to grow exponentially correlates with the amount of the starting material,

i.e. it correlates with the abundance of transcripts of the gene in question. Even though this estimate is not immediately indicative of the absolute level of the transcript, it is straightforward to make a comparison across several samples and draw conclusions on their relative abundance.

It is difficult to identify the exact cycle at which the fluorescence starts to grow exponentially. Instead, it is more reliable to estimate the fractional cycle number at which the fluorescence reaches a threshold value above the baseline. The fractional cycle number at which the fluorescence reaches a threshold value is called the Ct value. The rate of amplification of the same transcript is approximately identical between different samples, even if the starting amounts of the transcript are different. Thus the slopes of the curves describing the exponential growth in the fluorescence are approximately parallel for the same transcripts. Therefore, the threshold can be chosen at any level of the fluorescence for relative comparison, as long as it is above the baseline and below the plateau.

The fluorescence threshold is chosen arbitrarily and it cannot be immediately used to describe absolute levels of the transcript in question. However, the difference between the Ct values of different samples is indicative of the relative difference in abundance of the transcript. To make a comparison between different samples, the within sample differences in Ct values (ΔCt) between a control transcript and a transcript of interest are usually compared (instead of comparing actual Ct values between different samples). This is done because PCR progression is dependent on the amount of the starting material, and having an internal control is a way of normalizing differences of input RNA across the different samples. Because the difference of the differences is compared across the samples, the estimate is called $\Delta\Delta$Ct.

The Ct values for each of the primers-probe mixes was obtained separately with the threshold automatically defined by the software. Therefore, one threshold was applied to all samples probed with one probe, but the threshold was different for samples probed with a different set of primers and probes. For example, at 36 h of incubation, the threshold of 0.24 was used to define the Ct values analysis of Lass2 (including samples that were transected with mimics and inhibitors of miR-124), but a different threshold of 0.42 was used to define the Ct value of for Sfrs7 in the same samples. Such an approach optimised the detection of relative differences in gene expression between samples treated with mimics and inhibitors. The disadvantage of this method was that it did not allow for estimation of absolute expression levels.

2.4 Quantitative RT-PCR. miRNA

Materials

Reagents

- Reverse transcription primers and real time PCR primers and probes:
  - snoRNA202 (mouse). Applied Biosystems Inc. Cat. no. 4380914
  - hsa-miR-143[2]. Applied Biosystems Inc. Cat. no. 4373134
  - hsa-miR-let7c. Applied Biosystems Inc. Cat. no. 4373167
  - hsa-miR-370. Applied Biosystems Inc. Cat. no. 4373031
- TaqMan® MicroRNA Reverse Transcription Kit. Applied Biosystems Inc. Cat. no. 4366596. Components of the kit:
  - 100 mM dNTPs (deoxyribonucleotides)
  - MultiScribe Reverse Transcriptase, 50U/μl
  - 10 x Reverse Transcriptase Buffer
  - RNase Inhibitor, 20U/μl
- Nuclease-Free Water. Ambion Inc. Cat. no. AM9937
- TaqMan® Universal PCR Master Mix, No AmpErase UNG (2×). Applied Biosystems Inc. Cat. no. 4324018

Equipment

The same equipment as used for mRNA RT-PCR (see section 2.3).

Reagent and equipment setup

Primers

Primers for synthesis of the first strand cDNA and primers and probes for real time PCR were purchased from Applied Biosystems Inc. (the sequences were proprietary). For the reasons discussed in section 2.3 it was important to use an internal control. Measurement of expression of snoRNA202 was chosen as the internal control, as its use was previously reported in miRNA quantification experiments (Bak et al., 2008; Elia et al., 2009; Judson et al., 2009; Quintavalle et al., 2010).

[2]The same primers were sold for profiling of mouse and human miR-143, miR-let7c and miR-370

Procedure

Reverse transcription: Generation of cDNA

- The components of the kit were thawed on ice
- The master mix for reverse transcription was produced according to Table 2.4.

Component	per 15 μl reaction
100mM dNTPs	0.15
MultiScribe Reverse Transcriptase, 50U/μl	1.00
10x Reverse Transcriptase Buffer	1.50
RNase Inhibitor, 20U/μl	0.19
Nuclease-free water	4.16
TOTAL	7.00

Table 2.4: miRNA reverse transcription master mix.

- The reverse transcription primers were thawed on ice and mixed with the reverse transcription master mix and total RNA (20 ng/μl) in 3:7:5 ratio. The mixture was gently mixed and centrifuged to bring down the droplets. The total volume for reverse transcription was 20 μl and it was performed in a 96-well plate. *IMPORTANT:* Centrifugation speed at this stage must not exceed 400 g.
- The 96-well plate with reaction mixes was incubated on ice for at least 5 min before the start of the reverse transcription. The reverse transcription was performed in Peltier Thermal Cycler according to the program in Table 2.5.

Step Type	Time(min)	Temperature(°C)
HOLD	30	16
HOLD	30	42
HOLD	5	85
HOLD	∞	4

Table 2.5: miRNA reverse transcription program.

Real time PCR

- The product of the reverse transcription reaction (cDNA) was diluted with nuclease free water in 1:15 ratio.
- Real time PCR reactions were set according to Table 2.6.

Component	per 20 μl reaction
primers-probe mix for real time PCR	1.00
appropriately diluted cDNA	1.33
TaqMan 2X Universal PCR MasterMix	10.00
Nuclease Free Water	7.67
TOTAL	20

Table 2.6: miRNA real time PCR master mix.

- The real time PCR (20 μl reaction volume) was conducted in optical 96-well plates using 7500 Real Time PCR System (see section 2.3) with the program described in Table 2.7.

HOLD	CYCLE (40 cycles)	
10 min	15 sec	60 sec
95°C	95°C	60°C

Table 2.7: miRNA real time PCR program.

Analysis of expression by RT-PCR

As in section 2.3.

2.5 Transfection protocol

Primary forebrain cultures for transfection experiments were obtained and cultured as described in section 2.1.

Materials

Reagents

- DharmaFECT 3 siRNA Transfection Reagent. Dharmacon[1]. Cat. no. T-2003-01
- miRNA mimics:
  - mmu-miR-143. Dharmacon, cat. no. MI0000257 / MIMAT0000247 or Qiagen N.V. Cat. no. MSY0000247
  - mmu-miR-145. Dharmacon, cat. no. MI0000169 / MIMAT0000157 or Qiagen N.V. Cat. no. MSY0000157

[1]Dharmacon is a part of Thermo Fisher Scientific

  - mmu-miR-451. Dharmacon Cat. no. MI0001730 / MIMAT0001632
  - cel-miR-67. miRIDIAN microRNA Mimic Negative Control #1). Dharmacon. Cat. no. CN-001000-01-05
  - mmu-miR-25. Qiagen N.V. Cat. no. MSY0000652
  - mmu-miR-410. Qiagen N.V. Cat. no. MSY0001091
  - mmu-miR-551b. Qiagen N.V. Cat. no. MSY0003890
  - mmu-miR-370. Qiagen N.V. Cat. no. MSY0001095
  - mmu-miR-434-3p. Qiagen N.V. Cat. no. MSY0001422
  - mmu-miR-124. Qiagen N.V. Cat. no. MSY0000134
  - mmu-miR-103. Qiagen N.V. Cat. no. MSY0000546

- miRNA inhibitors:

  - Anti-mmu-miR-124. Qiagen N.V. Cat. no. MIN0000134
  - Anti-mmu-miR-434-3p. Qiagen N.V. Cat. no. MIN0001422
  - Anti-mmu-miR-145. Qiagen N.V. Cat. no. MIN000157
  - Anti-mmu-miR-103. Qiagen N.V. Cat. no. MIN0000546
  - Anti-mmu-miR-551b. Qiagen N.V. Cat. no. MIN0003890
  - Anti-mmu-miR-370. Qiagen N.V. Cat. no. MIN0001095
  - Anti-mmu-miR-410. Qiagen N.V. Cat. no. MIN0001091
  - Anti-mmu-miR-25. Qiagen N.V. Cat. no. MIN0000652
  - Anti-mmu-miR-143. Qiagen N.V. Cat. no. MIN0000247

Procedure

The transfection protocol described here was developed by Eric MacLaren (Maclaren et al., 2011) for transfection of siRNA into primary neuronal cultures. This protocol was developed in order to reduce levels of several genes encoding components of the post-synaptic density and to study if such pertrubations affected electrical activity of neurons.

For the first round of miRNA transfections, MacLaren's protocol was followed precisely. The cultures were plated at a high density ($\approx$ 1,850 cells $\cdot$ mm^{-2}) and the incubation time after transfection was 48 h. After a round of test transfections (Chapter 4, section 4.2.2) lower plating density ($\approx$ 790 cells$\cdot$mm^{-2}) and shorter incubation time (36h) were used in some of the experiments (the exact settings for each of the experiments conducted in this thesis project are listed in Table 2.8). The steps of the transfection procedure are listed below:

- Before starting the transfection procedure, the growth media was removed from a cell culture well to leave only 400 μl.
- 3.5 μl of 20 μM stock RNA was diluted in 98 μl of Neurobasal in Tube 1. In parallel, 2.4 μl of DharmaFECT 3 were added to 12 μl of Neurobasal in Tube 2. The tubes were incubated on a benchtop for 5 minutes. For 20 mM stock of RNA this produced approximately 115 nM concentration of the mimic in the final volume of approximately 600 μl. In order to achieve a different concentration in the final volume, a different amount of RNA could be added (adjusting the amount of Neurobasal appropriately). For mock transfection no RNA was added.
- Contents of Tube 1 and Tube 2 were combined in Tube 3 and gently mixed by pipetting up and down. Tube 3 was incubated on benchtop for 20 minutes.
- 80 μl of Neurobasal was added to Tube 3, the content was gently mixed and transferred drop by drop to the cell culture well (evenly distributing the content across the culture well).
- After the addition of the reaction mixture, the culture was transferred back to the incubator (set to 37 °C).
- After 36 h or 48 h incubation time (see Table 2.8) total RNA was extracted from cultures as described in Methods, section 2.2. On all occasions transfection experiments were carried out in four biological replicates.

miRNA ID	**DIV**	**Mimic**	**Inhib**	**Mock**	**Dens**	**Inc**	**M**	**Batch**
cel-miR-67	3DIV	115nM		available	1,850	48h	D	i
cel-miR-67	4DIV	115nM		available	1,850	48h	D	ii
cel-miR-67	4DIV	115nM		available	1,850	48h	D	iii
cel-miR-67	6DIV	115nM		available	1,850	48h	D	iv
cel-miR-67	6DIV	115nM		available	1,850	48h	D	v
miR-143	2DIV	250nM	250nM	available	790	36h	Q	vi
miR-143	3DIV	115nM		available	1,850	48h	D	i
miR-143	4DIV	250nM	250nM	available	790	36h	Q	vii
miR-143	6DIV	115nM		available	1,850	48h	D	iv
miR-145	3DIV	115nM		available	1,850	48h	D	i
miR-145	4DIV	115nM		available	1,850	48h	D	ii
miR-145	4DIV	250nM	250nM	available	790	36h	Q	vii
miR-145	6DIV	115nM		available	1,850	48h	D	iv

Continued on the next page

miRNA ID	DIV	Mimic	Inhib	Mock	Dens	Inc	M	Batch
miR-25	4DIV	250nM	250nM	available	790	36h	Q	vii
miR-103	4DIV	250nM	250nM	available	790	36h	Q	viii
miR-124	3DIV	115nM	230nM	available	1,850	48h	Q	ix
miR-124	4DIV	115nM	230nM		1,850	48h	Q	x
miR-124	6DIV	115nM	230nM	available	1,850	48h	Q	xi
miR-370	6DIV	250nM	250nM		790	36h	Q	xii
miR-410	6DIV	250nM	250nM		790	36h	Q	xii
miR-551b	6DIV	250nM	250nM		790	36h	Q	xii
miR-434-3p	6DIV	250nM	250nM	available	790	36h	Q	xiii
Mock only	4DIV			available	790	36h	na	vii

Table 2.8: Parameters of transfection experiments.
miRNA ID - miRBase Release 14 mature miRNA identifiers (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006, 2008); *DIV* - the timepoint at which the experiment was conducted; *Mimic* - the concentration of a miRNA mimic in the transfection mixture; *Inhib* - the concentration of a miRNA inhibitor in the transfection mixture; *Mock* - was the mock transfection *available* or not for the unidirectional contrasts (in case of the *Mock only* experiment, mock transfected cultures were contrasted with untransfected; *Dens* - the cell density that was used for plating of cultures for the experiments; *Inc* - the post-transfection incubation time (i.e. time from the transfection until cells were killed during miRNA extraction); *M* - the manufacturer of miRNA mimics and inhibitors (*D* - Dharmacon, *Q* - Qiagen N.V.); *Batch* - the batch identifier (cultures for experiments that have the same batch numbers were extracted from the same set of mice on the same day).

2.6 Microscopy

Materials

Reagents

- Dulbecco's Phosphate Buffered Saline 1× (1× DPBS). Invitrogen. Cat. no. 14190-094
- 4% Paraformaldehyde (PFA). Supplied on site.
- Methanol. VWR. Cat. no. 20847.320
- Albumin, from bovine serum (BSA). Sigma-Aldrich Corporation. Cat. no. A2153-100G
- Triton® X-100. Sigma-Aldrich Corporation. Cat. no. 93443-100ML
- Dulbecco's Phosphate Buffered Saline 1× (1× DPBS). Invitrogen. Cat. no. 14190-094

- Nuclease-Free Water. Ambion Inc. Cat. no. AM9937
- Prolong® Gold antifade reagent. Invitrogen. Cat. no. P36934
- Prolong® Gold antifade reagent with DAPI (4′,6-diamidino-2-phenylindole). Invitrogen. Cat. no. P36935
- Enhanced GFP (eGFP) expressing plasmid (pML40-CAG), 410 ng·ml^{-1} (obtained from Meng Li, personal communication)
- Fluorescently labelled oligonucleotide (AllStars Neg. siRNA AF 488). Qiagen N.V. Cat. no. 1027284
- Trypan blue stain 0.4%. Invitrogen. Cat. no. 15250-061
- β3-tubulin rabbit polyclonal antibody (primary antibody). Synaptic Systems. Cat. no. 302 302g
- Goat anti-rabbit IgG (H+L) AlexaFlour 633 (secondary antibody). Molecular Probes. Cat. no. A21070

Equipment

- Super Frost® Plus Slides. VWR International. Cat. no. 631-0108
- Coverslips. Supplied on site.
- Digital camera (AxioCam MRm). Carl Zeiss Ltd. Cat. no. 000445-554
- Light microscope (Axioplan 2 Imaging). Carl Zeiss Ltd.
- Imaging software (AxioVision Release 4.6). Carl Zeiss Ltd.
- Adobe Photoshop, CS4 Extended V11.0. Adobe Systems Inc.

Reagent and equipment setup

eGFP expressing plasmid The plasmid pML40-CAG expressed enhanced GFP (eGFP) (Zhang et al., 1996) under CAGGs promoter (cytomegalovirus immediate-early enhancer sequence connected to a modified (AG) chicken β-actin promoter) (Alexopoulou et al., 2008).

Fluorescently labelled oligonucleotide Fluorescently labelled oligonucleotide (AllStars Negative control) was recommended by Qiagen N.V. as a way of controlling the transfection efficiency. This oligonucleotide had a proprietary sequence and it was labelled by AlexaFlour 488 fluorophore at the 3′-end. AllStars Neg. control was diluted in nuclease-free water to obtain 20 μM stock, and stored at -20 °C.

Growing cultures Cultures for microscopy were grown as described in section 2.1, with the only difference that prior to addition of PDL (see section 2.1) one glass coverslip was placed into culture wells. For transfections (see below) cultures were grown at relatively low density ($\sim$ 790 cells $\cdot$ mm^{-2}), and for the immunostaining and the trypan blue assay cultures were grown at intermediate density ($\sim$ 1,250 cells $\cdot$ mm^{-2}).

Transfection of cultures Cultures were transfected as described in section 2.5. The concentration of AllStar Neg. control in the final transfection volume was 115 nM. Approximately 1.5 ng of the plasmid was used for the transfection.

The blocking solution To reduce non-specific binding of antibodies, and to dilute antibodies to appropriate concentrations, I used the following blocking solution: 0.2% triton X-100 and 3% BSA (in 1$\times$ DPBS).

Procedure

Imaging cultures transfected with the eGFP expressing plasmid or the AlexaFlour 488 labelled oligonucleotide

- The growth media was removed from culture wells and replaced with 1$\times$ DPBS.
- 1$\times$ DPBS was aspirated and replaced with 0.5 ml of 4% PFA solution for 20 min.
- 4% PFA solution was aspirated, coverslips removed from the culture wells. Remainder of 4% PFA was removed from the coverslips by daubing off.
- The coverslips were rinsed by 5 dips into Milli-Q water.
- Small drops of antifade mounting reagent were placed on a microscope slide. The coverslips were carefully placed on top of the drops. The slides were stored either at 4 °C or -20 °C.

Slides were visualised using Axioplan 2 Imaging microscope 495 nm light emission or differential interference contrast (DIC) settings. For DIC images, all parameters were set automatically with the microscope software. The fluorescence of both eGFP and AlexaFlour 488 transfected cultures was detected with 495 nm light for excitation of the fluorophores. The exposure time was controlled either manually or automatically (see figure legends). For imaging of cultures transfected with the plasmid, the exposure was set to the level that produced a maximum contrast between fluorescence of the cells and the background. For imaging of cultures transfected with the AlexaFlour 488 labelled oligonucleotide, the exposure was manually set in all cases to 220 ms to make images

comparable. For better visual contrast, the spectrum was inverted in all photographs of eGFP and the labelled oligonucleotide transfections using Photoshop. (No other parameters of the original images were changed).

Immunostaining β3-tubulin

Staining of a neuronal marker, β3-tubulin (Lee et al., 1990), was performed to evaluate abundance of neurons in primary cultures. Below are the steps of the procedure that was taken for the immunostaining:

- Coverslips were removed into 1× DPBS at room temperature.
- Coverslips were placed in a dish (pre-cooled on ice) with methanol (at -20 °C) for 7 min incubation.
- Coverslips were rehydrated in 1× DPBS briefly (for less than 1 min).
- Coverslips were drained and 100 μl of the blocking solution was added for 1 h.
- Coverslips were washed briefly in 1× DPBS and the primary antibody (100 μl of the supplier stock diluted at 1:1000 in the blocking solution) was added for 1 h.
- Coverslips were washed in 2 ml 1× DPBS for 5 minutes (this was repeated three times).
- Coverslips were drained and the secondary antibody (100 μl of of the supplier stock diluted at 1:1000 in the blocking solution) was added for 20 min (incubation in the dark).
- Coverslips were washed in 2 ml 1× DPBS for 5 minutes (this was repeated three times).
- Coverslips were washed briefly in distilled water.
- A drop of the antifade reagent (with DAPI) was placed onto slides and coverslips were placed on top (cells down). The slides were left overnight at 4 °C before imaging.
- Slides were visualised using Axioplan 2 Imaging microscope 358 nm (for DAPI) and 632 nm (for β3-tubulin staining) light emission with default settings.
- False colors were added to photographs using Photoshop.

Counts of DAPI stained nuclei were assumed to correspond to the number of all cells in a culture, counts of cells stained for β3-tubulin were assumed to correspond to the number of neurons. The percentage of neurons was estimated in the same fashion as viability in Trypan assay (see below).

Trypan assay

Dead cells were visualised in cultures with Trypan blue stain (Altman et al., 1993). Below are the steps of the procedure that was taken for Trypan blue staining:

- 100 μl of Trypan blue stain (0.4% solution, as supplied) was added per cell culture (1000 μl of a growth media, see section 2.1) for 6min.
- All liquid was aspirated and replaced with 0.5 ml of 4% PFA solution for 20 min.
- Coverslips were washed briefly in distilled water.
- A drop of the antifade reagent (with DAPI) was placed onto slides and coverslips were placed on top (cells down). The slides were left overnight at 4 °C before imaging.
- Slides were visualised using Axioplan 2 Imaging microscope 358 nm light (for DAPI) and white light for Trypan with default settings.
- False colors were added to photographs using Photoshop.

Counts of DAPI stained nuclei were assumed to correspond to the number N of all cells in a culture, counts of Trypan blue stained cells were assumed to correspond to the number D of dead cells. Photographs obtained at a low magnification (10× objective) were used for counting ($\sim$ 500 cells per photograph), counts from three non-overlapping images were averaged per each treatment (see text). Viability was defined by the following formula:

$$\frac{N - D}{N} \cdot 100\%$$

2.7 Microarray profiling of mRNA and miRNA expression

Acquisition and analysis of two types of microarray data are described in this section: the in-house data and the external data. The in-house data was generated by microarray facilities at the Wellcome Trust Sanger Institute to profile mRNA or miRNA abundances in total RNA samples that I extracted as a part of experiments described in Chapters 3, 4, 5 and 6. The external data was generated elsewhere in independent experiments.

In-house microarray data: Experimental procedures

Microarray profiling of mRNA and miRNA abundances in total RNA (extracted as described in Methods, section 2.2) was performed by the staff of the microarray facility at the Wellcome Trust Sanger Institute (Naomi Hammond [nh4@sanger.ac.uk], Peter Ellis [pde@sanger.ac.uk] and Cordelia Langford [cfl@sanger.ac.uk]). All procedures were carried out according to the standard Illumina protocols (http://www.illumina.com/support/literature.ilmn). Below is a summary of the procedures.

Reagents

- Illumina® TotalPrep RNA Amplification Kits. Illumina Inc. AMIL1791

Equipment

- Nanodrop
- BeadArray reader
- BeadStudio software
- For mRNA expression profiling: Illumina Sentrix BeadChip Array Mouse-WG6_v1.1 (used in the hippocampal and forebrain developmental timecourse experiments, Chapters 3) or Illumina Sentrix BeadChip Array Mouse-WG6_v2 (used in all miRNA perturbation experiments, Chapters 4, 5 and 6).
- For miRNA expression profiling: Illumina Universal Sentrix Array Matrix (used in the forebrain developmental timecourse experiments, Chapter 3)
- Thermal cycler
- Hybridisation oven

Illumina mRNA microarray profiling assay

- The total RNA was reverse-transcribed with oligo(dT) primers. The oligo(dT) primers had a T7 RNA polymerase binding site (promoter) at the 5′-ends, which was necessary for the *in vitro* transcription step (see below).
- The RNA was digested with RNaseH.
- The cDNA was converted to double-stranded cDNA with a DNA polymerase.
- The purified double-stranded cDNA was incubated with the T7 RNA polymerase and rNTPs (including biotin-tagged rUTP) to produce biotinylated single-stranded anti-sense RNA (called aRNA or cRNA). This step was equivalent to *in vitro* transcription, and amplification was achieved at this step (Gelder et al., 1990).

- The cRNA was purified, quantitated (using Nanodrop), mixed with the hybridisation buffer and applied to the array slides.
- The slides were washed and labelled with streptavidin-Cy3.
- The arrays were scanned using BeadArray reader and the image data was processed by BeadStudio.
- Quality control of loading and hybridisation efficiencies was performed using sample dependent and sample independent control measurements.
- The raw data output (not normalized) of BeadStudio was used for the next stage (**Data processing**).

Illumina miRNA microarray profiling assay

- The 3′-ends of total RNA were polyadenylated and the total RNA was reverse-transcribed using biotin-tagged oligo(dT) primers. The oligo(dT) primer had a universal sequence at the 5′-end which was necessary for PCR step (see below).
- The cDNA was attached to streptavidin beads and hybridised to miRNA-specific oligos.
- The miRNA-specific oligos were extended using a DNA polymerase.
- The extended products were eluted and PCR was performed using fluorescently labelled primers.
- Single-stranded PCR products (ssDNA) were prepared, quantitated (using Nanodrop) and hybridised to the arrays.
- The arrays were scanned using BeadArray reader and the image data was processed by BeadStudio.
- Quality control of loading and hybridisation efficiencies was performed using sample dependent and sample independent control measurements.
- The raw data output (not normalized) of BeadStudio was used for the next stage (**Data processing**).

In-house microarray data: Data processing

Data processing and normalization. mRNA arrays

Analysis of the array data was performed in R environment (RTeam, 2008) with Bioconductor packages (Gentleman et al., 2004). The output of BeadStudio was imported into R using *lumi* package functions (Du et al., 2008). In addition to the essential quality

control steps performed by the microarray facility (see above), additional quality control steps were performed using *lumi* package functions. First, correlation of probe intensities levels was examined between pairs of biological replicates. Second, all samples were clustered hierarchically to examine their relation. In case of mRNA profiling, pairwise correlations of all replicate samples was above 0.99, and clustering of samples corresponded perfectly to the design of the experiments (Supplementary Data Figure A.2, A.1 and A.4). Therefore all replicate samples were used for further steps of the analysis. Detection call P-values were obtained with the *detectionCall* method, which is available via the *lumi* package. If the P-value was < 0.01 (the default threshold), the probe was considered "Present", otherwise it was considered "Absent". The "Absent" probes were removed, and expression values of the "Present" probes were transformed with the Variance Stabilizing Transformation (VST) method and normalized with the robust spline normalization (RSN) method (Lin et al., 2007), both available in the *lumi* package (Du et al., 2008).

Data processing and normalization. miRNA arrays

The output of BeadStudio was imported into R using *lumi* package functions (Du et al., 2008). Furthermore, using methods in the *lumi* package, the correlation of probe intensity values was assessed between the samples as a way of QC (to complement the QC steps performed by the microarray facility, see above). Some miRNA array samples stood out as poorly correlated with the rest of the samples (Supplementary Data Figure A.5 and A.6a). When raw probe intensities were visualised, it became apparent that probe intensities in the poorly correlated samples displayed a global downward shift (Supplementary Data Figure A.6b). In order to minimise the biases to the subsequent steps of analysis, seven poorly correlated samples were removed from further analysis (specified in the legend to Supplementary Data Figure A.6b). After the removal of these samples, there remained 4 samples for 1DIV, 2 samples for 2DIV, 4 samples for 4DIV and 5 samples for 8DIV timepoints (see Chapter 3, section 3.2). Raw values of the remaining samples were transformed using log_2 transformation and normalized using quantile normalization, as these methods of transformation and normalisation were shown to be optimal for Illumina miRNA microarray data (Rao et al., 2008).

Mapping the Illumina probes. mRNA arrays

Mapping of Illumina mRNA probes to gene and transcript identifiers was performed by my colleague, Cei Goodger-Abreu [cei@langebio.cinvestav.mx]. A summary of the steps involved is given below.

Illumina microarray probe sequences were taken from the Illumina mRNA array annotation files (BGX files) available from the Illumina website[1] http://www.illumina.com/support/annotation_files.ilmn. Illumina mRNA array probes were aligned to the complete set of the full-length Ensembl v56 mouse transcripts (Hubbard et al., 2009) with SSAHA2 (Ning et al., 2001). All categories of Ensembl transcripts, were retrieved using the Ensembl Perl API, which enabled access to Core, Vega and OtherFeatures mouse transcripts (Hubbard et al., 2009). A transcript from the highest scoring SSAHA2 alignment was chosen for each probe (at least 30 perfect consecutive matches were required). If more than one alignment had an equally high score, manually curated Vega transcripts were preferred. In order to resolve ambiguity of multiple transcripts from the same source aligning equally well, the "biotype" annotation was considered (protein coding transcripts were preferred to pseudogenes, and nonsense-mediated decay had the lowest preference). If the ambiguity was still not resolved, transcripts with the longest 3′UTR or cDNA were selected. When probes did not align to any of the Ensembl transcripts, mapping of the probes to RefSeq 38 (Pruitt et al., 2009) transcript identifiers was taken from the Illumina BGX files. After probes were uniquely mapped to the best transcript identifiers (either from Ensembl or from the Illumina BGX files), the corresponding GeneBank gene identifiers (in this thesis they are referred to as Entrez gene IDs) were matched to the probes. For probes mapped to Ensembl or Vega transcripts, the Entrez gene IDs were retrieved using Ensembl API, and for probes mapped to RefSeq IDs via the BGX annotation files, the Entrez gene IDs were taken from the BGX file.

Mapping the Illumina probes. miRNA arrays

Sequences of miRNA microarray probes were taken from the Illumina annotation file (the BGX file) available from the Illumina website[2] http://www.illumina.com/support/annotation_files.ilmn. Illumina miRNA array probes were aligned to the full set of mature miRBase Release 13 miRNA sequences (Griffiths-Jones et al., 2008, 2006; Griffiths-Jones, 2004) using SSAHA2 (Ning et al., 2001). Full length perfect matching to mature sequences

[1]The name of the annotation file for v1.1 beadarrays: MouseWG-6_V1_1_R4_11234304_A, for v2 beadarrays: MouseWG-6_V2_0_R2_11278593_A

[2]The name of the annotation file for the miRNA array matrix: mouseMI_V1_R0_XS0000127-MAP

was required. This requirement alone resolved all ambiguity and resulted in mapping of 362 Illumina miRNA probes to unique mature mouse miRNAs in the miRBase.

Analysis of differential expression

The differential expression was estimated for all probes (VST transformed and VSN normalised probe intensity values, see above) using the R package *limma* (Smyth, 2004). A linear model was fitted and coefficients were estimated for every probe according to the design of the experiments. Moderated t-statistics for each probe were evaluated with the empirical Bayes method. The P-values associated with the t-statistics were multiple-test corrected with Benjamini and Hochberg method (Benjamini and Hochberg, 1995). Each probe was paired up with a transcript and Entrez gene ID using probe mapping procedure that was described above. Probes that could not be mapped were excluded at this stage. Ambiguity in the mapping was resolved by keeping one probe (with the best adjusted P-value of differential expression) per gene. The remaining probe set usually contained from 9,000 to 11,000 probes, each uniquely corresponding to one Entrez gene ID. Fold-changes (in log_2 scale), t-statistics, P-value and adjusted P-value for the probes were used to describe differential expression of the corresponding genes.

For the miRNA arrays, the analysis of differential expression for all probes (log_2 transformed and quantile normalised values) that mapped to miRBase Release 13 mature miRNA identifiers (see above) was performed with *limma* (Smyth, 2004). Fold-changes (in log_2 scale), t-statistics, P-value and adjusted P-value for the probes were used to describe differential expression of the corresponding miRNAs.

Clustering of genes according to gene expression trends

The probes that were uniquely mapped to genes and which were differentially expressed in the timecourse experiments between any pair of consecutive timepoints or between the first and the last timepoint[3] (adjusted $P < 0.1$) were clustered according to their expression. The clustering was done using the Markov Cluster Algorithm (MCL) software package[4] (van Dongen, 2000; Freeman et al., 2007). Probes were described by measurements on the four timepoints, with each measurement defined as the median of the intensity values taken over the replicates. For each pair of the probes, their similarity was computed as the Pearson correlation coefficient over their intensity values. A graph was defined where

[3] four developmental timepoints, see Chapter 3, section 3.1
[4] freely available for download at http://www.micans.org/

the nodes are probes, and two nodes (probes) are connected if the correlation between them was at least 0.9, with the weight of the edge set to that correlation value. This graph was processed with the MCL algorithm, which naturally partitions the graph into separate clusters. The inflation parameter was set to 3.

External microarray data

Data retrieval

External microarray data (not normalised) was obtained directly from Gene Expression Omnibus (GEO) website http://www.ncbi.nlm.nih.gov/geo/. Accession identifiers for three analysed external data sets were GSM210760 (Makeyev et al., 2007), GSE6388 (Akahoshi et al., 2007) and GSE10246 (Lattin et al., 2008).

Normalisation, differential expression analysis and mapping

The external data sets were generated using Affymetrix microarray platforms. A method that was specifically designed for transformation and normalisation of Affimetrix data, called RMA (Irizarry et al., 2003a,b), was used in analysis of the external data. RMA was implemented via *affy* Bioconductor package (Gautier et al., 2004),

For the external datasets, differential expression was estimated for all probes using *limma* (Smyth, 2004) in the same way as was described for the in-house data (see above). As a result, each probe was assigned fold changes (on the RMA scale, wich is approximately equal to log_2) moderated t-statistics and corresponding P-values (both unadjusted and adjusted with Benjamini and Hochberg method).

Bioconductor annotation libraries (Gentleman et al., 2004) of microarray platforms were used in analysis of the external data ("mouse4302.bd" (Makeyev et al., 2007; Lattin et al., 2008) and "mgu74bv2.db" (Akahoshi et al., 2007)). These libraries provided mapping of microarray probes to RefSeq transcript and Entrez gene identifiers. For RefSeq transcript IDs, 3′UTR sequences were obtained from Ensembl, using Ensembl API (Hubbard et al., 2009). The length of 3′UTRs was used to resolve the ambiguity in mapping of probes to RefSeq transcript IDs: when the probes were mapped to more than one RefSeq transcript ID, the ID corresponding to the transcript with the longest 3′UTR was selected. In the Bioconductor annotation files, each RefSeq transcript ID corresponded to one Entrez gene ID. Therefore, by uniquely matching the probes to RefSeq transcript IDs, each probe was also uniquely matched to an Entrez gene ID. The adjusted P-values of differential expression were used for selection of the best probe per Entrez gene ID (i.e.

a probe with the most significant adjusted P-value). By selecting one probe per Entrez gene ID, the total number of probes was reduced to the total number of genes represented on the platform.

2.8 Seed enrichment analysis

Obtaining sequences

3′UTRs

The 3′UTR sequence for the transcripts mapped to each of the microarray platforms (see section 2.7) were obtained from Ensembl v56 with the Ensembl API (Hubbard et al., 2009). Three FASTA files were created using the retrieved sequences (referred here to as raw sequences): one for the two versions of the in-house Illumina mRNA arrays[1], and one file for the external data (see section 2.7). For Sylamer analysis (see below) it was recommended to remove regions of low complexity and repetitive (redundant) sequences, therefore the FASTA files with raw sequences were processed as previously described (van Dongen et al., 2008). The low complexity regions were masked out using DUST (Tatusov R.L. and Lipman D.J., personal communication) and redundant sequences were masked out using purge-sequence from the RSA-tools (Thomas-Chollier et al., 2008). The processed sequences are referred in this thesis to as dusted/purged sequences.

Seed matching sites for miRNAs

The sequences of all mature mouse miRNAs were downloaded directly from miRBase Release 14 (Griffiths-Jones et al., 2008, 2006; Griffiths-Jones, 2004). For each miRNA, two sequences complementary to the seed region (i.e. the seed matching sites) were produced: the sequence complementary to bases 2-8 from the 5′-end of the miRNA (7(2)-type seed matching site), and to bases 1-7 with an A opposite to position 1 (7(1A)-type). This resulted in 876 distinct seed matching oligonucleotide words. The seed matching sites were stored as a flat file and used for the enrichment analyses (see below).

[1] Retrieval of sequences for the in house arrays was done by Cei Goodger-Abreu [cei@langebio.cinvestav.mx]

Seed matching site enrichment

Simple hypergeometric test of enrichment

The hypergeometric test was used to evaluate the enrichment of transcripts with seed matching sites for a particular miRNA among the 3′UTRs of transcripts[2] differentially expressed beyond a certain threshold. The urn model is a popular way to describe the hypergeometric test. In this model an urn contains N balls of which K are black and $N-K$ are white. We draw a sample of n balls from the urn without replacement and observe k black balls. The probability of such an event follows the hypergeometric distribution, and is given by

$$\frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

The one-sided test of enrichment asks for the cumulative probability of finding at least k black balls in a sample of size n randomly drawn from the urn according to the formula given above, by summing all probabilities over the range $k \ldots K$. This test, computed using standard R functions (RTeam, 2008), was used to test enrichment of seed matching site containing transcripts by equating the set of all transcripts with the set of all balls, by equating transcripts that contain at least one seed-matching site with black balls (and all other transcripts with white balls), and by considering a sample to be a set of transcripts differentially expressed beyond a certain threshold T. Denoting the sample size by n, when observing k transcripts containing at least one seed-matching site in the sample, the one-sided test thus gives the probability of observing as least as many as k transcripts containing seed-matching sites.

Sylamer

Sylamer tests for nucleotide word occurrence biases in a sorted list of sequences using hypergeometric test (van Dongen et al., 2008). In this thesis, Sylamer was applied to test for miRNA effects by searching sorted lists of 3′UTRs[3] for enrichment or depletion of miRNA seed matching sites. The mechanism of Sylamer is described in the Introduction (section 1.2.3). Below is the description of the internal parameters that were used for Sylamer analyses in this thesis:

[2]The matching of 7(2) and 7(1A)-type seed sites to the raw 3′UTR sequences was done using Perl.
[3]Transcripts without 3′UTR sequence were excluded from the test.

- Sequences were sorted either by t-statistic (see section 2.7) or by the intensity values (i.e. the level of expression) of corresponding microarray probes (see the figure legends).
- In each test distribution of 876 seed matching sites of length 7 bases corresponding to mouse mature miRNAs[4] (see above) was assessed.
- The sample size (bin size) of selected sequences was incremented by 100 at each step.
- The level of Markov-correction was set to 4.
- Sequences of 3′UTRs were dusted/purged (see above).

2.9 Neuron specific genes

Neuron specific genes were defined as genes with expression significantly higher in the mouse brain than in other organs, and which at the same time encoded proteins of post-synaptic density (PSD). These two characteristics were combined, because on its own they did not guarantee specificity of expression in neurons, while selecting the intersection between the two types of genes increased the likelihood of the specificity. The list of 1,634 mouse PSD genes was obtained experimentally by my colleague Alex Bayes (personal communication). The list of genes with expression significantly higher in the mouse cortex relative to other tissues and organs was obtained through analysis of "mouse gene expression tissue atlas" from "BioGPS" gene annotation portal (http://biogps.gnf.org/) (Lattin et al., 2008; Wu et al., 2009a) of the Genomic Institute of the Novartis Research Foundation.

The raw microarray data that comprised the "mouse gene expression tissue atlas" was downloaded from GEO (Barrett et al., 2007), accession number GSE10246 (Lattin et al., 2008)). This dataset consisted of a microarray gene expression profile of 91 mouse tissues and cell cultures. The raw microarray data was normalized as described in section 2.7. Of the 91 tissues in the data set, there were three that represented the cortex ("cerebral cortex", "cerebral cortex prefrontal", "hippocampus"). These tissues were labelled as "cortex". The expression in these cortical structures was contrasted to the expression in a selection of unrelated to the cortex tissues ("bone", "bone-marrow", "epidermis", "heart", "intestine large", "intestine small", "kidney", "lens", "liver", "lung", "lymph

[4]Additional two seed matching sites were included in analysis of experiments where the mimic of a *Caenorhabditis elegans* miRNA, cel-miR-67, was transfected. These words corresponded to 7(2) and 7(1A)-type seed matching sites for cel-miR-67 (*GGTTGTG* and *GTTGTGA*)

nodes", "mammary gland (lact)", "mast cells", "NK cells", "ovary", "pancreas", "placenta", "prostate", "salivary gland", "spleen", "stomach", "testis", "umbelical cord", "uterus"). These tissues were labelled as "non-cortex".

The tissues with the label "cortex" were contrasted against samples with the label "non-cortex" and differentially expressed probes were identified using the *limma* R package (Smyth, 2004), as described in section 2.7. Subsequently, the probes were uniquely mapped to Entrez gene and RefSeq transcript identifiers as described in section 2.7.

Genes that were upregulated in the cortex by more than two fold (with adjusted P-value of differential expression below 0.05) comprised the list of 2,732 genes with expression enriched in the cortex. The intersection of these genes with the PSD genes (Alex Bayes, personal communication) produced the list of 732 genes, which were putatively neuron-specific (referred in this thesis to as neuron-specific). Of these genes, 725 were represented on Illumina Mouse-WG6 v2 microarray platform that was utilised in functional miRNA experiments in this thesis (Methods, section 2.7). These putatively neuron-specific genes are listed in Supplementary Data (Table A.1).

2.10 Enrichment of GO and KEGG terms

Analysis of the enrichment of GO (Ashburner et al., 2000) and KEGG (Kanehisa et al., 2008, 2000) terms was performed in R environment (RTeam, 2008) with Bioconductor packages (Gentleman et al., 2004). Annotation of the Entrez gene IDs to GO and KEGG terms was obtained from "illuminaMousev2.db" library, available via the Bioconductor website http://www.bioconductor.org/packages/2.6/data/annotation/. The test of enrichment of GO and KEGG terms in the selected categories of genes, in comparison to the gene universes, was performed using the *GOstats* package (Falcon and Gentleman, 2007).

Chapter 3

A model of neuronal development

Dissociated primary cultures are a popular model of neuronal development and function (Introduction, section 1.2.4) and it was used in this thesis to investigate the role of miRNAs in neurons. The cultures were studied in a time-window centered around 4 days of *in vitro* development (DIV): from **1DIV** to **8DIV**. The significance of the 4DIV timepoint as a switch point in development of primary neuronal cultures was previously reported (Valor et al., 2007). Valor *et al.* showed that before 4DIV the ratio of average abundances of neuritic to somatic transcripts was below one, while after 4DIV it was above one, which is similar to that of mature neurons. Appearance of early synapses at around 4DIV and detection of early events of electrical activity in the cultures at around 6DIV (Valor et al., 2007), also supported the proposition that neurons in primary neuronal cultures after 4DIV were similar to mature neurons. Studying miRNAs in the cultures between 1DIV and 8DIV, could, therefore, highlight functions of miRNAs in both immature and mature primary neurons.

In this chapter, I describe profiling of mRNA and miRNA abundances in 1DIV to 8DIV time-window in development of E17.5 primary cultures. This profiling demonstrated that primary forebrain cultures were a good model to study mRNA and miRNA expression during growth of committed (differentiated) neurons. The profiling of mRNA and miRNA expression in developing primary forebrain cultures was published (Manakov et al., 2009). Additionally, I found evidence that endogenous miRNAs shaped gene expression in primary forebrain cultures. Apart from being a novel observation on its own, this finding further supported the cultures to be a suitable model system to study roles of miRNAs.

Based on miRNA profiling results, nine mouse miRNAs and a control non-mouse miRNA were selected for functional perturbation experiments. The selection is described at the end of the current chapter.

3.1 A model of developmental gene expression

3.1.1 Gene expression changes in development of both hippocampal and forebrain primary cultures were highly correlated

To establish if gene expression in E17.5 primary forebrain cultures was sufficiently similar to that of better characterized hippocampal cultures (Introduction, section 1.2.4), gene expression was profiled in development of both forebrain and hippocampal cultures. Microarrays were used to profile gene expression changes throughout a timecourse of E17.5 primary cultures development. The cultures were grown as described in Methods (section 2.1). At four timepoints, 1DIV, 2DIV, 4DIV and 8DIV, total RNA was extracted and profiled on microarrays (Methods, sections 2.2 and 2.7).

To obtain comparable measurements across four developmental timepoints it was necessary to have all biological replicates derived from a single batch of cultures. For profiling of mRNA expression in hippocampal cultures, a batch of 12 cultures was plated (Methods, section 2.1), producing three biological replicates for each of the 4 timepoints. Forebrain cultures were plated in a separate experiment, i.e. dissociated forebrains were obtained from a set of embryos collected from a different group of pregnant mice. Because forebrains are much larger, it was possible to plate a batch of 23 cultures. This produced five biological replicates for 1DIV, and six replicates for each of the timepoints 2DIV, 4DIV and 8DIV. Total RNA was extracted (Methods, section 2.2) and profiled on mRNA microarrays (Methods, section 2.7).

Results of microarray profiling of development of cultures were highly consistent. Pearson correlation of within-timepoint biological replicates prior to normalisation was 0.99 or higher for both hippocampal and forebrain cultures at all timepoints (Supplementary Data Figures A.2 and A.1). Transformation and normalisation of raw data was performed as described in Methods (section 2.7). Hierarchical clustering of normalized expression values revealed that gene expression detected by microarrays was more similar between biological replicates within any of the timepoints than between different timepoints. At the same time, gene expression between consecutive timepoints was relatively closely related, with pairs of timepoints 1DIV, 2DIV and 4DIV, 8DIV, forming distinct outgroups. This consistent trend was true for profiles of both hippocampal and forebrain cultures (Supplementary Data Figure A.4).

Approximately 10,000 genes were detected with high confidence in primary cultures. Of 46,628 probes on the microarray platform, 16,408 probes in hippocampal culture ex-

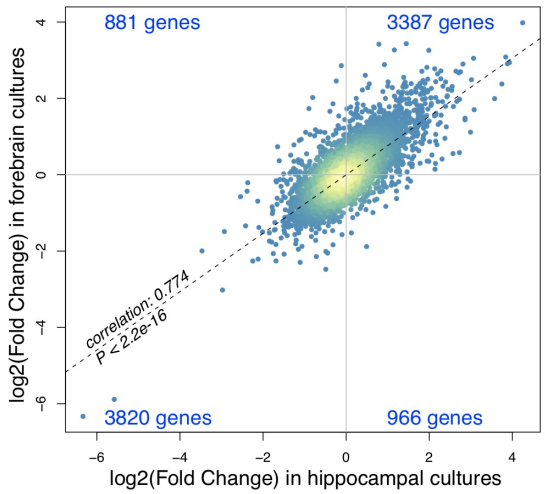


Figure 3.1: Correlation of forebrain and hippocampal cultures development. The points correspond to 9,054 genes detected in both the development of hippocampal and forebrain cultures (the x-axis corresponds to log_2 of the expression fold change between 1DIV and 8DIV in development of hippocampal cultures, the y-axis – to forebrain cultures). The colors of the points depend on the density of the points in a given region of the plot (yellow – highest, blue – lowest). The text in blue corresponds to gene counts in each quadrant of the plot, the italic in black gives Pearson correlation of the fold changes and P-value of the correlation. The dashed line is a linear model fitted through the points.

periment and 16,003 probes in forebrain culture experiment were reliably detected (using the standard Illumina detection call $P < 0.01$). The detected probes were mapped to 10,067 and 9,826 genes respectively. Detection, normalization and mapping are described in Methods (section 2.7).

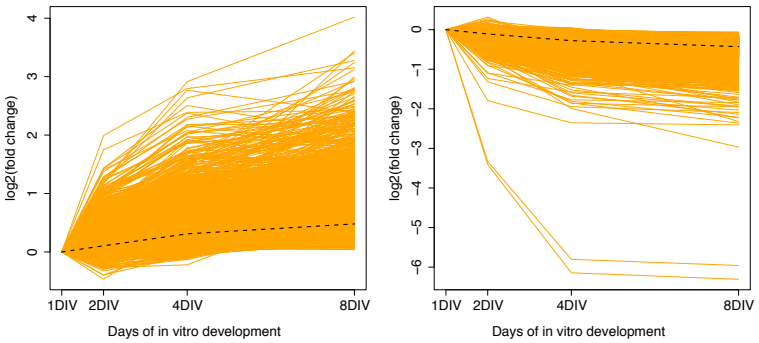
Comparison of differential expression between whole forebrain and hippocampal cultures, showed that gene expression trends in the two types of cultures were similar. Of 9,826 genes, whose expression was detected in forebrain cultures, 9,054 were also detected in hippocampal cultures development. When expression fold changes between 1DIV and 8DIV were compared between the two experiments, a Pearson correlation of 0.774 ($P < 2.2e-16$) was observed (Figure 3.1). This observation indicated that global trends in mRNA gene expression in development of forebrain cultures was similar to that of primary hippocampal cultures, an established model of neuronal development and function (Introduction, section 1.2.4).

3.1.2 The reciprocal trends of gene expression in development of hippocampal and forebrain primary cultures

Almost 90% of genes detected by microarrays in primary cultures were differentially expressed during development. Of all genes detected in hippocampal and forebrain culture development, 8,999 and 9,040 genes were differentially expressed between any of the two consecutive timepoints (adjusted $P < 0.1$, Methods, section 2.7). Between both experiments, the intersection of differentially expressed genes was 7,646 genes.

To identify trends of differential expression, differentially expressed genes (adjusted $P < 0.1$) were clustered using MCL (van Dongen, 2000) (see Methods, section 2.7). Clustering of gene expression trends identified 32 distinct clusters in hippocampal and 28 in forebrain culture experiments. The two largest clusters of genes encompassed a majority of all differentially expressed genes in both experiments (66.67% in hippocampal and 72.62% in forebrain cultures). The third largest cluster in both cases included less than 5% of genes. Median expression trends of the two largest clusters were approximately inverse, with median trends of expression being gradual upregulation and downregulation (Figure 3.2). A significant overlap was observed between genes in the two major clusters of upregulated and downregulated genes between hippocampal and forebrain cultures. Of genes that were expressed both in hippocampal and forebrain cultures, approximately 67.40% and 71.68% of hippocampal genes in upregulated and downregulated clusters were in the respective forebrain clusters (Figure 3.3). The P-value for these intersections exceeded the precision limit for the hypergeometric test as implemented in the R *stats* package (equivalent to $P < 1e - 45$) (RTeam, 2008).

(a) Upregulated cluster (HP) **(b)** Downregulated cluster (HP)



(c) Upregulated cluster (FB) **(d)** Downregulated cluster (FB)

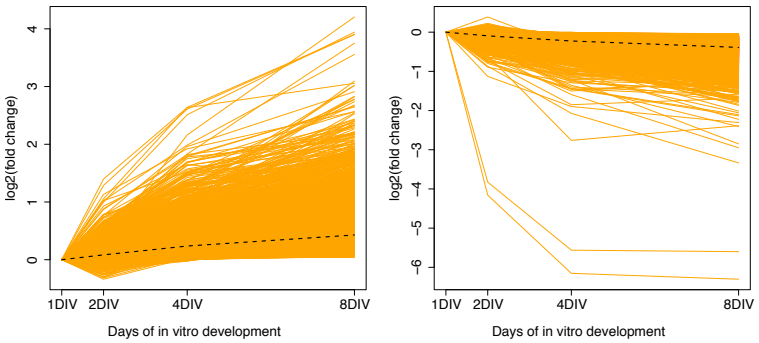


Figure 3.2: Gene expression clusters of down- and upregulated genes.
The orange lines correspond to the fold change of gene expression (log_2) starting from 1DIV and across the other three developmental timepoints (2DIV, 4DIV and 8DIV). The x-axis shows time (DIV, days of *in vitro* development), the y-axis shows fold change in gene expression (log_2)). The subfigures show: 3.2a – trends in the biggest gene expression cluster in developing hippocampal cultures (dashed line – median trend); 3.2b – trends in the second biggest expression cluster; 3.2d – trends in the biggest gene expression cluster in developing forebrain cultures; 3.2c – in the second biggest expression cluster. Abbreviations: *HP* – hippocampal cultures; *FB* – forebrain cultures. Clustering is described in Methods (section 2.7).

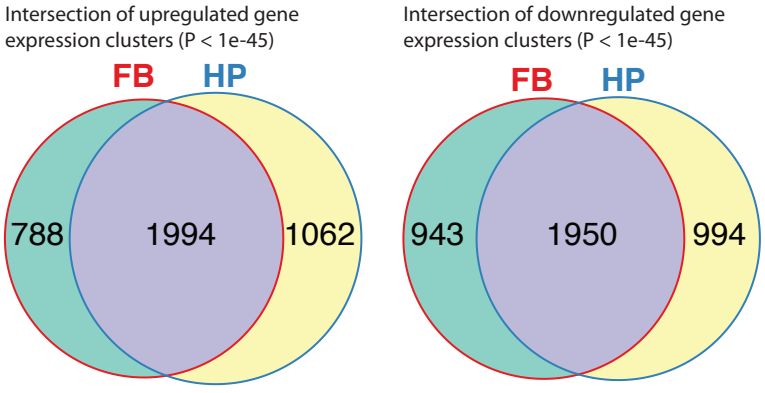


Figure 3.3: Intersection of gene expression clusters in hippocampal and forebrain cultures development.
FB – forebrain cultures; *HP* – hippocampal cultures. The genes detected as differentially expressed in hippocampal cultures (8,999 genes) were used as a gene universe for calculation of the hypergeometric P-values.

3.1.3 Cell growth, not proliferation, was a predominant ongoing process in development of primary cultures

Enrichment analysis of Gene Ontology (GO) (Ashburner et al., 2000) and KEGG terms (Kanehisa et al., 2008, 2000) was used to determine whether development of primary forebrain cultures was predominantly characterized by cell growth and neuronal activity, or by cell proliferation. If gene expression trends in development of cultures were consistent with increasing cell growth, then such observations would have been consistent with the growth of neurites being a predominant process in development of cultures. On the other hand, if gene expression trends were consistent with ongoing cell proliferation, then such a result would be suggestive of proliferating non-neuronal cell types being a dominant component of cultures. To test this, significantly differentially expressed genes (adjusted $P < 0.1$, Methods, section 2.7) between 1DIV and 8DIV were separated into two groups: down and upregulated in development. The separation of differentially expressed genes in two groups was done solely based on the direction of change during development without considering magnitude of change. Based on this criterion, 4,426 genes were defined as downregulated and 4,098 as upregulated in forebrain cultures development. Enrichment of GO terms of "Biological Process" and "Cellular compartment" types and of KEGG terms (also known as KEGG pathways) was then evaluated to describe function and localization of proteins encoded by down- and upregulated genes.

Analysis of GO term enrichment implied that in development of primary cultures there was an increase in cell growth, and not in proliferation. GO terms describing nuclear localization and biological processes taking place in the nucleus were enriched in downregulated genes (Figure 3.4). At the same time, terms describing extracellular, plasma membrane and other non-nuclear localizations were enriched in upregulated genes. Terms relating to biological processes not taking place in the nucleus were also overrepresented in upregulated genes. Importantly, some of the terms specifically related to neuronal biology and activity (e.g. "synapse" and "neurological system process") were among enriched terms in upregulated genes. Gene counts of down and upregulated genes in a representative selection of 40 most enriched GO terms is shown in Figure 3.4 (the full lists of the top 40 most enriched GO terms is in Supplementary Data Tables A.2 to A.5).

In agreement with the GO enrichment results, KEGG pathway analysis showed that biological processes taking place in the nucleus and cell cycle related pathways were downregulated while processes taking place not in the nucleus, as well as specifically neuronal pathways, were upregulated. For example, among the top 10 most enriched pathways

in downregulated genes were "spliceosome" ($P < 3.04e-20$), "DNA replication" ($P < 8.61e-10$) and "Cell cycle" ($P < 1.14e-09$). At the same time pathways "Neuroactive ligand-receptor interaction" ($P < 9.14e-05$), "Long-term potentiation" ($P < 0.00134$), "Calcium signalling function" ($P < 8.5e-05$) and "Cell adhesion molecules (CAMs)" ($P < 0.00134$) were in top 10 pathways enriched in upregulated genes. A complete list of the top 25 most enriched KEGG pathways is in Supplementary Data, Tables A.6 and A.7.

Abundant neurite outgrowth during the developmental timecourse was also evident upon visual inspection of primary cultures. On the day of plating (0DIV) cells almost completely lacked any appendages (Figure 3.5a) while by 3DIV the outgrowth was already evident (Figure 3.5b). To confirm that cultured cells were indeed neurons, I immunostained a neuronal marker, β3-tubulin (Lee et al., 1990), at 3DIV and 8DIV (Methods, section 2.6). By combining this immunostaining with visualisation of all nuclei (DAPI staining, see Methods, section 2.6) it was established that throughout the developmental timecourse the population of cells was comprised almost entirely of neurons (Figure 3.5 and Table 3.1). Additionally, I confirmed viability of cells in primary neuronal cultures with Trypan assay (Altman et al., 1993) at 3DIV (Figures 3.6a and 3.6b, Table 3.2) and at 8DIV (Figures 4.6c and 4.6d, Table 3.2), see for a comparison cultures treated with sodium azide (Figures 4.6e and 4.6f). Trypan assay is described in Methods (section 2.6).

As described in the Introduction (section 1.2.4) in the work by Valor *et al.*, which investigated the gene expression program of developing primary E17.5 hippocampal cultures, the 4DIV timepoint was identified as a switch point in maturation of the cultures (Valor et al., 2007). A similar effect was observed in developing primary E17.5 forebrain cultures: if expression trends of the upregulated genes were overlaid on the same plot with the downregulated genes, their median trends intersected at almost exactly the 4DIV timepoint (Figure 3.7). Therefore, in this thesis 4DIV was treated as a developmental switch point, when gene expression of neuritic genes became, on average, higher than that of nuclear genes (i.e. many of the somatic genes). In other words, after 4DIV the ratio of neuritic and somatic genes in developing cultures was similar to that of mature neurons. This central position of the 4DIV timepoint in gene expression program of cultures was important for selection of experimental timepoints for transfection experiments that are described in Chapters 4 and 5.

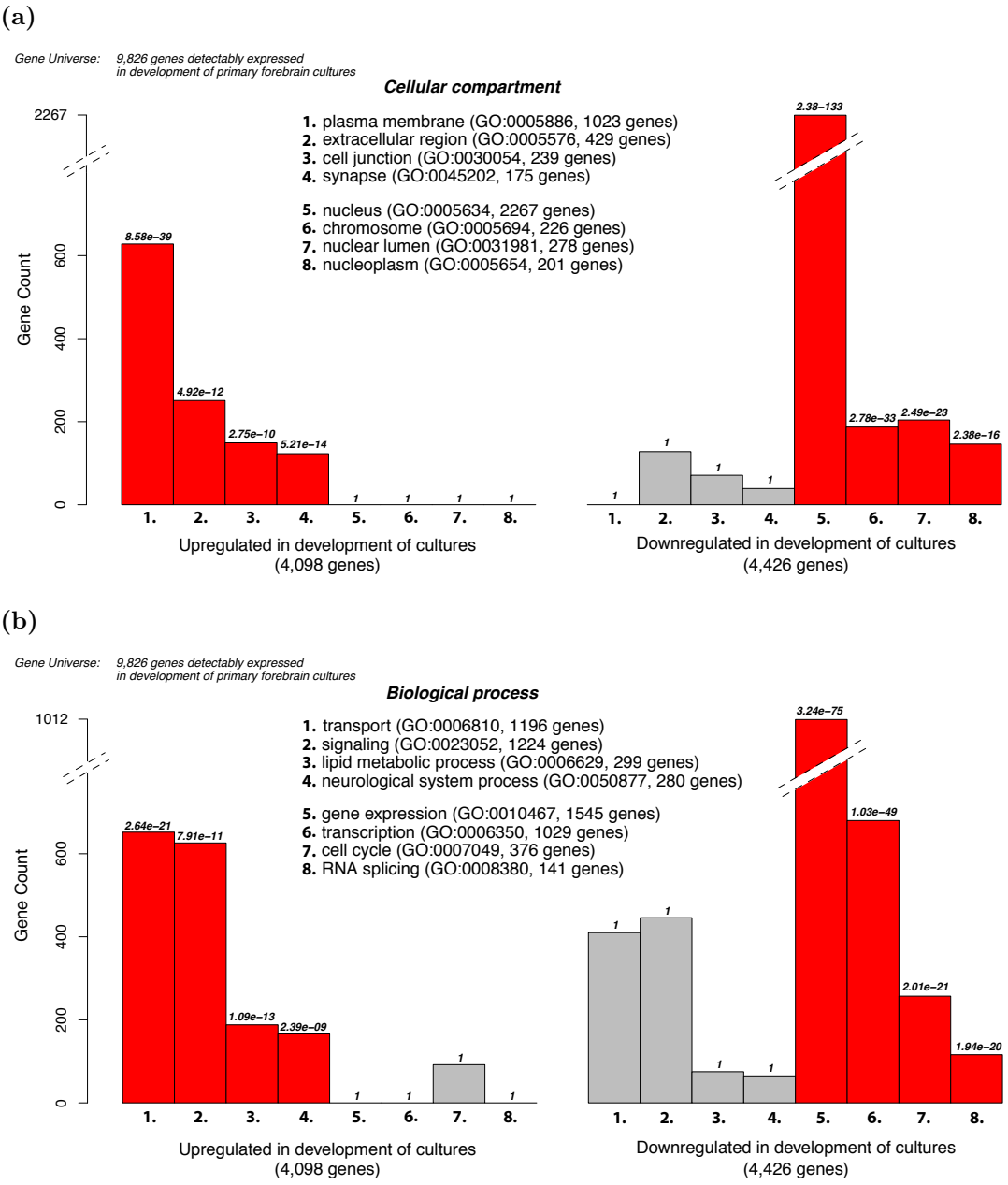
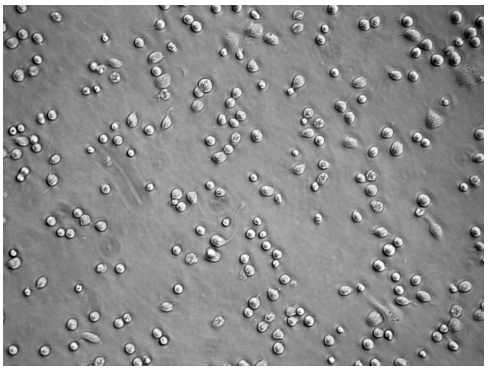


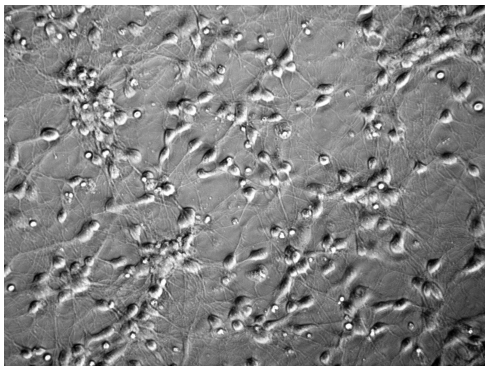
Figure 3.4: Enrichment of Gene Ontology (GO) terms in differentially expressed genes during primary cultures development.

The y-axes show the number of genes from a GO category that were identified among the up- or down-regulated genes. The enrichment P-values for each of the terms is given at the top of each bar (Methods, section 2.10). The numbers on the x-axes correspond to the GO terms listed in the plot areas. The bars for the significantly enriched GO terms is shown in red and not significantly enriched – in grey.

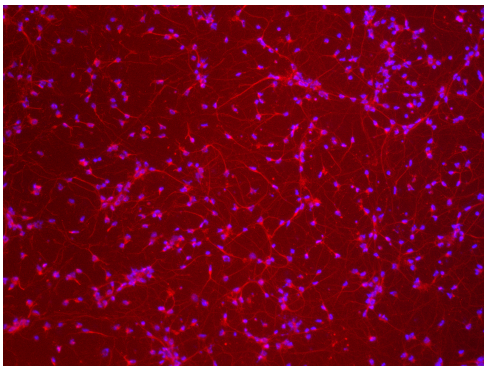
(a) 0DIV (DIC)



(b) 3DIV (DIC)



(c) 3DIV (β3-tubulin and DAPI)



(d) 8DIV (β3-tubulin and DAPI)

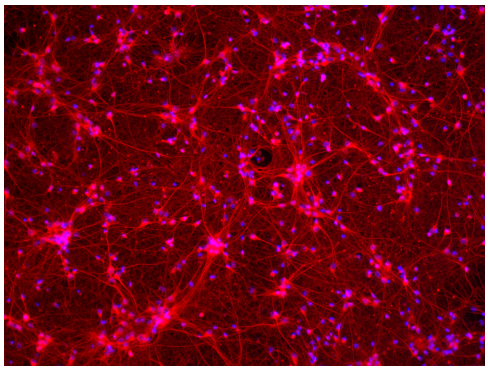
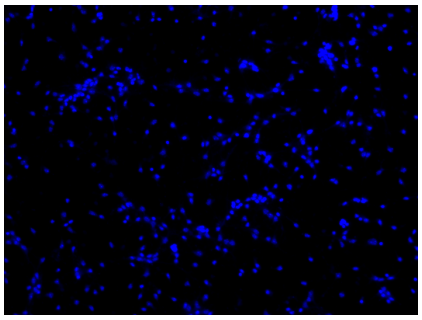


Figure 3.5: Neurons in primary cultures.
Cells in primary cultures were visualised using differential interference contrast (DIC) settings at 0DIV (Figure 3.5a) and 3DIV (Figure 3.5b). Immunostaining of a neuronal marker, β3-tubulin, is shown in red and DAPI staining is shown in blue (at 3DIV in Figure 3.5c and at 8DIV in Figure 3.5d). See Methods (section 2.6) for details.

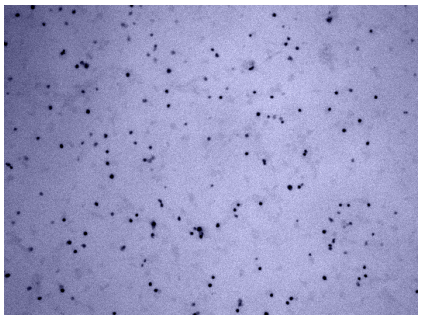
Days *in vitro*	**Percent of neurons**	**Standard deviation**
3DIV	99.7%	± 0.24%
8DIV	99.5%	± 0.34%

Table 3.1: Neurons in primary neuronal cultures.
Percent of neurons was estimated based on three non-overlapping 10× objective images (∼ 500 cells per image, e.g. Figures 3.5c and 3.5d). Total numbers of cells were estimated by counting DAPI stained nuclei, neurons – by counting β3-tubulin positive cells (Methods, section 2.6).

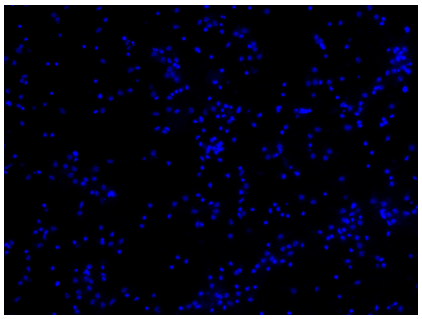
(a) 3DIV (DAPI)



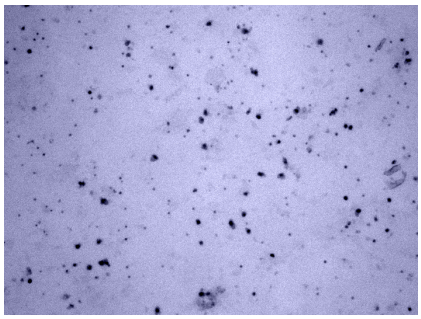
(b) 3DIV (Trypan)



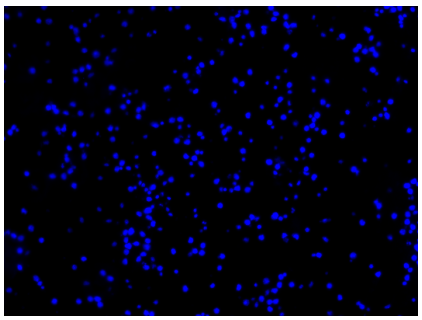
(c) 8DIV (DAPI)



(d) 8DIV (Trypan)



(e) killed by sodium azide (DAPI)



(f) killed by sodium azide (Trypan)

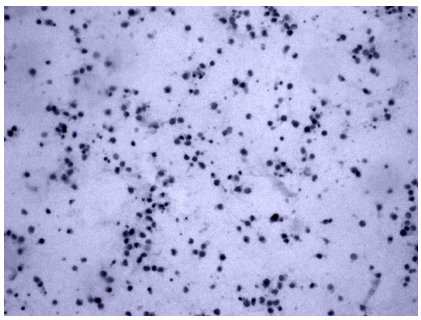


Figure 3.6: Viability of cells in primary neuronal cultures.
Pairs of figures show the same areas of cultures stained with DAPI or Trypan blue (see titles of the subfigures). Figures 3.6a and 3.6b show a culture at 3DIV; Figures 3.6c and 3.6d show a culture at 8DIV; Figures 3.6e and 3.6f show a culture treated at 8DIV with sodium azide (0.03%, 24 h incubation). See Methods (section 2.6) for details.

Days *in vitro*	**Viability**	**Standard deviation**
3DIV	68.8%	± 5.58%
8DIV	66.4%	± 8.05%

Table 3.2: Viability of cells in primary neuronal cultures.
Viability was estimated based on three non-overlapping 10× objective images (∼ 500 cells per image, see for example Figure 3.6). Total numbers of cells were estimated by counting DAPI stained nuclei, numbers of dead cells – by counting Trypan stained cells (Methods, section 2.6).

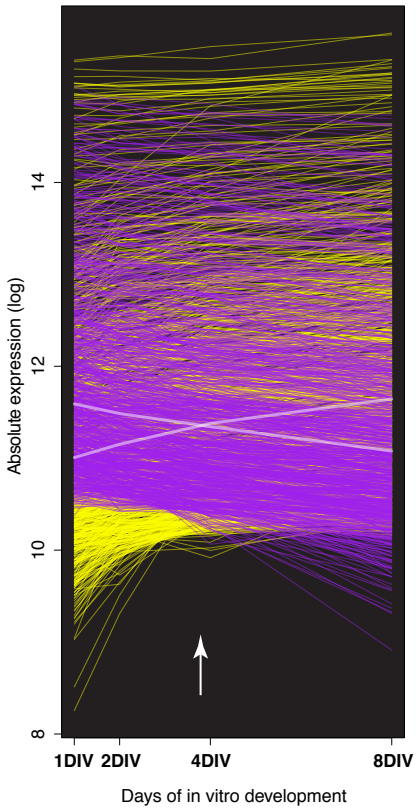


Figure 3.7: A gene expression switch point in development near 4DIV.
The thin lines represent trends of expression (median between replicates) of 2,000 most highly expressed genes from the downregulated (purple lines) and upregulated (yellow lines) categories. The y-axis shows *log* transformed and normalized absolute expression values (Methods, section 2.7), the x-axis shows time (DIV, days of *in vitro* development). The thick white lines are equivalent to the median trends of the 1,000 genes in each of the two categories. The arrow points to the crossing of the median trends (near 4DIV).

Summary of section 3.1

Analysis of GO and KEGG enrichment characterised the development of E17.5 primary forebrain cultures (1DIV to 8DIV time-window) in terms related to cell growth and neuronal activity, and not related to cell division (section 3.1.3). Genes associated with nuclear localisation and function (i.e. many of the somatic genes) appeared to be downregulated in the development of cultures, while genes associated with presumably neuritic localisation (e.g. plasma membrane and synaptic GO terms) were upregulated. After 4DIV the average abundance of the somatic genes remained lower than that of the neuritic genes, which indicated the importance of the 4DIV timepoint a switch timepoint in maturation of primary neurons. Also, these observations meant that a contribution of mRNA from proliferating secondary cell types (e.g. fibroblasts, endothelial cells and etc.) to the overall gene expression profile of the cultures was relatively small. These findings validated E17.5 primary neuronal cultures to be a suitable model to study gene expression in growing neurons. Additionally, I found that profiles of gene expression programs of E17.5 primary hippocampal and forebrain cultures were very similar (section

3.1.1 and 3.1.2). Therefore, primary forebrain cultures could be used to study neuronal gene expression in a way similar to primary hippocampal cultures.

3.2 A model of miRNA activity in neurons

3.2.1 Identification of three categories of differentially expressed miRNAs in forebrain cultures development

In addition to mRNA profiling (section 3.1), the samples of total RNA extracted from developing forebrain primary cultures were used to profile miRNA abundance using the *Illumina Universal Sentrix Array Matrix*. Micorarray analysis and mapping of array probes to official miRBase Release 13 miRNA symbols (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006, 2008) was performed as described in Methods (section 2.7). Overall, expression of 362 miRNAs was assessed in primary cultures, and 204 miRNAs were found to be differentially expressed between the first and last developmental timepoints (adjusted $P < 0.1$). As in the case of mRNA coding genes analysis (see section 3.1.3), all differentially expressed miRNAs were separated into two categories: Downregulated (99 miRNAs) and upregulated (105 miRNAs). Interestingly, of the 30 most highly expressed miRNAs (based on average expression between 1DIV and 8DIV), only 4 were differentially expressed, which was approximately 4.2 times less than expected by chance alone (hypergeometric $P < 5.04e-07$). Therefore, in addition to down- and upregulated categories, a third category was singled out, which was named as the steady state highly expressed. In summary, differential expression analysis identified three non-overlapping groups or categories of miRNAs with distinct patterns of expression in development of primary cultures:

- Steady state highly expressed (26 miRNAs)
- Downregulated (99 miRNAs)
- Upregulated (105 miRNAs)

A full listing of miRNAs per expression category, together with the ranks of their expression at the final developmental timepoint (8DIV), is in Supplementary Data (Table A.8).

To validate profiling of miRNA expression by microarrays, one miRNA was selected from each category (let-7c from the steady state, miR-143 from the downregulated and miR-370 from the upregulated categories) and their expression at 1DIV and 8DIV was

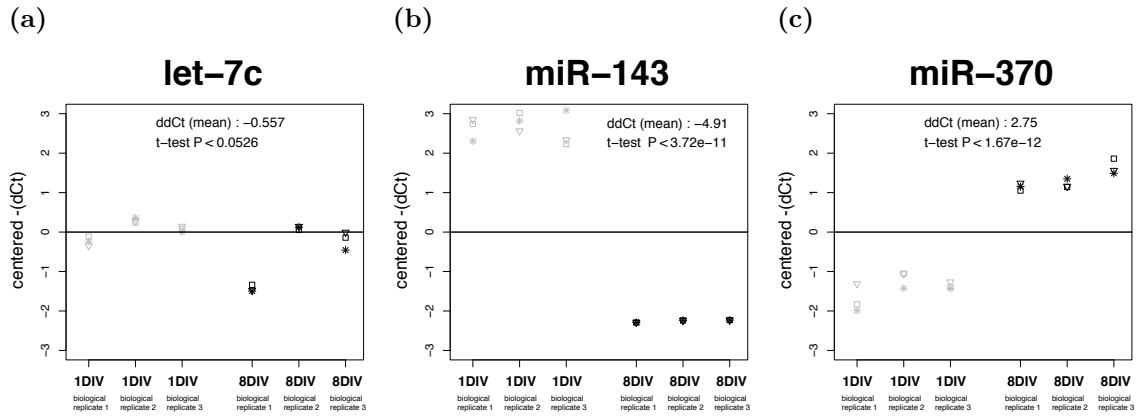


Figure 3.8: Validation of miRNA expression profiles with qRT-PCR.
The x-axes correspond to biological replicates (three per each of the two timepoints, 1DIV and 8DIV), the y-axes correspond to inverse ΔCt values centered around experimental medians (*"centered -(dCt)"*). The styled points correspond to ΔCt values for technical replicates per one biological replicate. The text gives provides the following information: 1) The mean of ΔΔCt values; 2) The t-test P-value for the differential expression. ΔΔCt method is described in Methods (section 2.3). The styled points correspond to ΔCt values of technical replicates for each biological replicate.

assessed with qRT-PCR (Methods, section 2.4). Results of qRT-PCR analysis were consistent with microarray results for all the three miRNAs (Figure 3.8).

3.2.2 miRNA expression in cultures development was similar to that in the brain and neurons

The dynamics of mRNA abundance, as described by gene expression analysis, were consistent with the development of primary E17.5 embryonic forebrain cultures being a model of neuronal development (section 3.1.3). It remained unknown, however, if miRNA expression and activity in primary cultures was similar to that in neuronal development. This section describes the analysis of miRNAs observed in primary cultures in the light of previously published research of neuronal miRNAs, which showed that expression of miRNAs in cultures was similar to that in the brain and neurons.

Several miRNAs that are known to be involved in inhibition of neural progenitor proliferation and promotion of neuronal differentiation were found to be among the top 20 most highly expressed miRNAs in primary cultures. These were six let-7 miRNAs, miR-124, miR-125b-5p, miR-9 and miR-137. Expression of these miRNAs along with references to relevant literature on their function is summarised in Table 3.3. Interestingly, all of these miRNAs belonged to the steady state highly expressed category (see section 3.2.1) and were highly expressed starting from 1DIV. Their high expression both at 1DIV and 8DIV

was consistent with E17.5 cultures to be comprised predominantly of developing differentiated neurons, as in published literature these miRNAs were shown to be implicated in neuronal differentiation.

miRNA	**#1DIV**	**#8DIV**	**Function**	**Reference**
miR-9	1st	1st	Promotes neurogenesis in the MH	(Leucht et al., 2008)
			Promotes differentiation of NSCs	(Zhao et al., 2009)
			Promotes differentiation of NPGs	(Shibata et al., 2008)
let-7 family	2nd	3rd	Inhibits Lin28, a plurepotency factor	(Rybak et al., 2008)
miR-125b-5p	11th	4th	Promotes neuronal differentiation of neuroblastoma	(Le et al., 2009)
miR-137	12th	9th	Promotes neuronal differentiation of glioblastoma	(Silber et al., 2008)
miR-9*	9th	14th	Inhibits BAF53a, a chromatin remodelling factor of NPGs	(Yoo et al., 2009)
miR-124	17th	10th	Inhibits SCP1, a partner of REST	(Visvanathan et al., 2007)
			Inhibits PTBP1, a repressor of neuronal splicing	(Makeyev et al., 2007)
			Promotes neuronal differentiation of glioblastoma	(Silber et al., 2008)
			Promotes neuronal differentiation of adult NPGs	(Cheng et al., 2009)
			Inhibits BAF53a, a chromatin remodelling factor of NPGs	(Yoo et al., 2009)

Table 3.3: Ranking of expression of pro-neuronal miRNAs.
miRNA - miRNA identifier as of miRBase Release 13; *#1DIV* - rank of expression at 1DIV; *#8DIV* - rank of expression at 8DIV; *Function* - published function in neurogenesis and/or establishing of neuronal identity; *Ref* - references to corresponding literature. Abbreviations: "NPGs" - neural progenitors (*in vivo*); "NSCs" - neural stem cells (*in vitro*); "MH" - midbrain-hindbrain domain. For let-7 family expression rank of let-7a is provided. Additional five let-7 miRNAs (let-7b, let-7d, let-7g, let-7c and let-7f) were among top 20 most highly expressed miRNAs at both of the timepoints (for 8DIV expression see Supplementary Data, Table A.8).

Additionally, the miRNAs in the downregulated category were in agreement with their reported depletion from the synaptic fraction in the adult mouse forebrain (Table 3.4) (Lugli et al., 2008; Siegel et al., 2009). Strikingly, the four miRNAs identified by Lugli *et al.* as the most depleted from the synaptic fraction of the mouse forebrain (in comparison to the whole forebrain homogenate) were exactly the same four miRNAs that were the most strongly downregulated in cultures development (Lugli et al., 2008). Additionally, Siegel and colleagues also identified four miRNAs that were significantly depleted from forebrain synaptic fraction, all of which were among 12 most downregulated miRNAs in cultures (Table 3.4).

miRNA	**# depletion** (Lugli et al., 2008)	**# depletion** (Siegel et al., 2009)	**# DR**	**× FC**	**adj. *P* downreg in cultures**
miR-143	1st	1st	1st	× 9.65	$P < 8.49e-14$
miR-451	2nd	*na*	4th	× 4.03	$P < 3.36e-09$
miR-150	3rd	2nd	3rd	× 7.36	$P < 5.11e-10$
miR-145	4th	3rd	2nd	× 7.27	$P < 2.35e-13$
miR-301	5th	*na*	38th	× 1.51	$P < 1.29e-04$
miR-153	6th	*na*	*na*	*na*	*not sign.*
miR-126-5p	7th	4th	21st	× 1.89	$P < 6.67e-07$
miR-126-3p	8th	*na*	6th	× 2.79	$P < 2.77e-10$

Table 3.4: Deption of downregulated miRNAs from synaptosomes.
miRNA - miRNA identifier as of miRBase Release 13; *# depletion* - rank of synaptic depletion (Lugli et al., 2008; Siegel et al., 2009); *# DR* - rank of downregulation during development of forebrain cultures (of significantly downregulated in the development miRNAs, see text); × *FC* - fold downregulation during development of forebrain cultures; *adj. P downreg in cultures* - adjusted P-value of downregulation during development of forebrain cultures (Methods, section 2.7). Abbreviations: "*na*" - not applicable; "*not sign.*" - not significant.

Lastly, miRNAs with a known function and/or expression in the adult brain were found to be upregulated in development of primary cultures. For example, miR-132, a miRNA induced by neuronal activity and implicated in homeostatic regulation of neuronal function (Klein et al., 2007), was upregulated in development of cultures (differential expression adjusted $P < 0.01$). In addition, upregulation was detected for miRNAs transcribed from the distal end of mouse chromosome 12. Expression of these miRNAs was shown to be restricted to the brain in the adult mice (Seitz et al., 2004). The distal 12 region encodes 54 miRNA hairpins (i.e. pre-miRNAs) from which 80 distinct mature miRNAs are transcribed and processed (according to miRBase Release 13 (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006, 2008)). Of these mature miRNAs, 53 were profiled by microarrays in development of cultures (Methods, section 2.7) and 41 were found in the upregulated category (approximately 2.7 times more than expected by chance alone, $P < 2.12e-15$). At the same time, only three miRNAs of the distal 12 region were attributed to the downregulated category of miRNAs (significant depletion, $P < 2.4e-05$).

3.2.3 miRNAs were active in primary cultures: miR-124 and let-7 miRNAs shaped gene expression

The previous section demonstrated that expression of miRNAs in the development of cultures was similar to that previously reported in the brain and neurons. In addition, it was possible to obtain evidence of direct miRNA effects on gene expression in primary

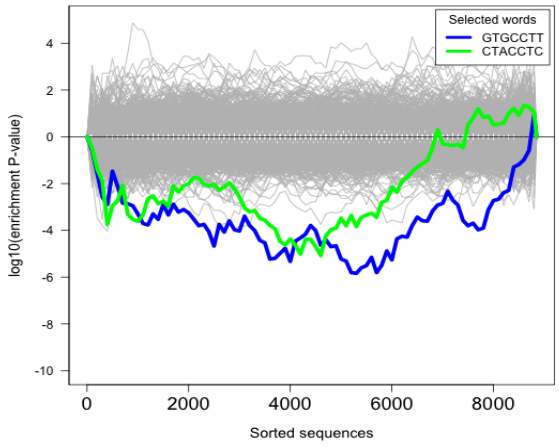
cultures. Sylamer, a method of word distribution analysis across sorted sequences[1] (van Dongen et al., 2008), was applied to estimate occurrence biases of miRNA seed matching sites in 3′UTRs of genes expressed in cultures (Methods, section 2.8). Results of the analysis were consistent with miR-124 and let-7 miRNAs (which were highly expressed in cultures, see Supplementary Data Table A.8) playing a direct role in shaping gene expression in the primary cultures.

Seed matching sites for miR-124 and let-7 miRNAs were significantly depleted from 3′UTRs of highly expressed genes in both forebrain and hippocampal cultures (Sylamer $P < 1e{-}04$ in all cases), which was consistent with the direct role of these miRNAs in regulation of mRNA levels in the primary cultures (Figure 3.9). Significant depletion of the seed matching sites (of the 7(2)-type, see Methods, section 1.2.1) was observed throughout the developmental timecourse: at 1DIV (Figures 3.9a and 3.9b), as well as at 8DIV (Figures 3.9c and 3.9d). The depletion in 3′UTRs of highly expressed genes suggested that miR-124 and let-7 miRNAs from the beginning of the developmental timecourse participated in shaping gene expression in primary cultures. Since miR-124 expression was shown to be specific to neurons (Christodoulou et al., 2010; Clark et al., 2010; Shkumatava et al., 2009), this result also meant that neuronal gene expression had a major contribution to the gene expression profile of the cultures.

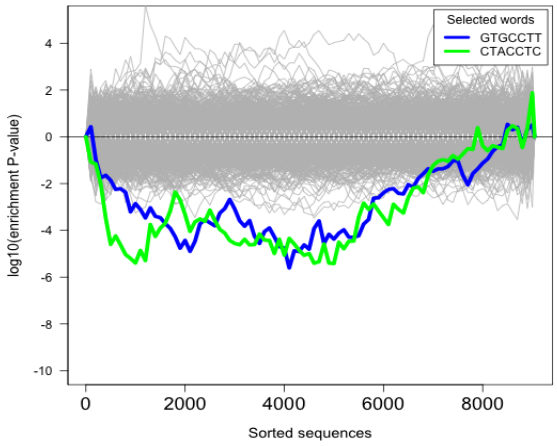
Apart from modulation of the level of gene expression, activity of miR-124 would likely to have a direct impact on dynamics of differential gene expression in cultures at the later stages in development of cultures. Sylamer analysis of 3′UTRs of genes expressed in the early stages of development, in transition from 1DIV to 2DIV, showed that the upregulated genes were significantly depleted from miR-124 seed matching sites (Figures 3.10a and 3.10b). This can be interpreted as miR-124 being permissive to upregulation in gene expression at this early stage, when the upregulated early genes faced little inhibition from miR-124. However, at the later stage, in transition from 4DIV to 8DIV, this depletion disappeared (forebrain primary cultures, Figure 3.10c) or even changed to enrichment (hippocampal primary cultures, Figure 3.10d). Therefore, it is conceivable that genes, which were upregulated later in the development, faced moderation by miR-124.

[1] A full description of Sylamer is in the Introduction (section 2.8).

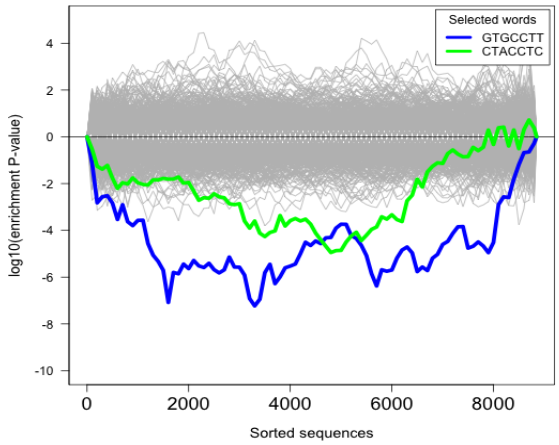
(a) miR-124 and let-7 at **1DIV**, FB culutres



(b) miR-124 and let-7 at **1DIV**, HP cultures



(c) miR-124 and let-7 **8DIV**, FB cultures



(d) miR-124 and let-7 **8DIV**, HP cultures

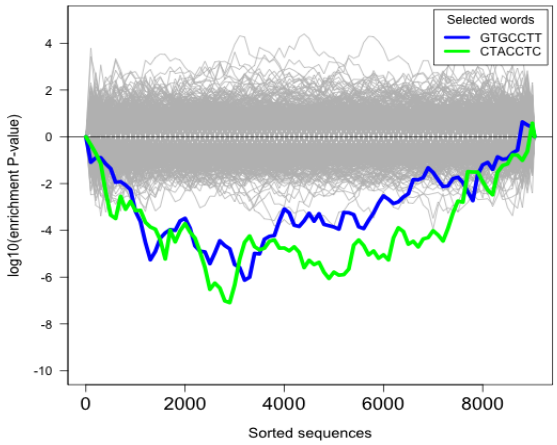
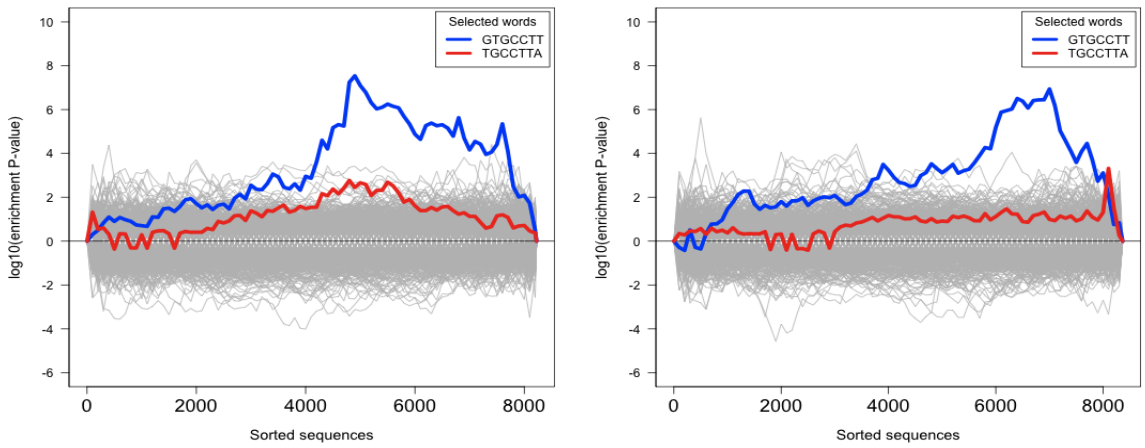


Figure 3.9: Signature of miR-124 and let-7 regulation of the level of gene expression in primary cultures.
The x-axes represent sorted 3′UTRs corresponding to detectably expressed genes (detected with the standard Illumina detection call $P < 0.01$, see Methods, section 2.7). These genes are ordered **from the most abundant to the least abundant** in replicates: 3.9a – of forebrain cultures at 1DIV; 3.9b – of hippocampal cultures at 1DIV; 3.9c – of forebrain cultures at 8DIV; 3.9d – of hippocampal cultures at 8DIV. The y-axes represent the hypergeometric P-values for occurrence biases of 876 nucleotide words complementary to the seed regions (7(2) and 7(1A)-types) of the complete set of 581 distinct mouse miRNAs, according to miRBase Release 14 (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006, 2008). Positive values on the y-axes correspond to an enrichment ($+|log_{10}$(P-value)|) and negative values to a depletion ($-|log_{10}$(P-value)|). The blue and the green lines show enrichment profiles of 7(2)-type seed matching sites complementary, respectively, to miR-124 and miRNAs of let-7 family. The grey lines show profiles of the rest of the distinct seed matching sites. The mapping of microarray probes to mRNA transcripts, and transcripts to genes, is described in Methods (section 2.7). The identification of the seed regions and parameters of Sylamer (van Dongen et al., 2008) is in Methods (section 2.8). The full description of the Sylamer method is in the Introduction (section 2.8). Abbreviations: *FB cultures* – primary forebrain cultures; *HP cultures* – primary hippocampal cultures.

(a) miR-124 in **1DIV** → **2DIV**, FB culutres **(b)** miR-124 in **1DIV** → **2DIV**, HP cultures



(c) miR-124 in **4DIV** → **8DIV**, FB cultures **(d)** miR-124 in **4DIV** → **8DIV**, HP cultures

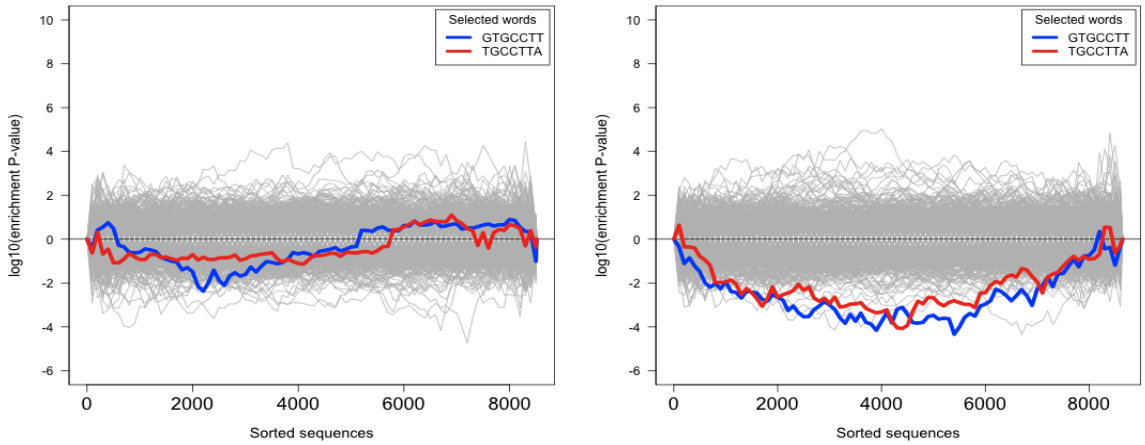


Figure 3.10: Signature of miR-124 regulation of differential gene expression in primary cultures.
The x-axes represent sorted 3′UTRs corresponding to detectably expressed genes (detected with the standard Illumina detection call $P < 0.01$, see Methods, section 2.7). These genes are ordered **from the most downregulated to the most upregulated** by fold change t-statistic for differential expression between replicates: 3.10a – of forebrain cultures at 1DIV compared to 2DIV; 3.10b – of hippocampal cultures at 1DIV compared to 2DIV; 3.10c – of forebrain cultures at 4DIV compared to 8DIV; 3.10d – of hippocampal cultures at 4DIV compared to 8DIV. The y-axes represent the hypergeometric P-values for occurrence biases of 876 nucleotide words complementary to the seed regions (7(2) and 7(1A)-types) of the complete set of 581 distinct mouse miRNAs, according to miRBase Release 14 (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006, 2008). Positive values on the y-axes correspond to an enrichment ($+|log_{10}$(P-value)|) and negative values to a depletion ($-|log_{10}$(P-value)|). The blue and the red lines show the enrichment profiles of 7(2) and 7(1A)-type seed matching sites for miR-124. The grey lines show the enrichment for the rest of the distinct seed matching sites. The mapping of microarray probes to mRNA transcripts, and transcripts to genes, is described in Methods (section 2.7). The identification of the seed regions and parameters of Sylamer (van Dongen et al., 2008) is in Methods (section 2.8). The full description of the Sylamer method is in the Introduction (section 2.8). Abbreviations: *FB cultures* – primary forebrain cultures; *HP cultures* – primary hippocampal cultures.

Summary of section 3.2

Profiling of miRNAs in developing E17.5 primary forebrain cultures (1DIV to 8DIV time-window) showed that expression of several miRNAs with a known role in differentiation of neuronal progenitors was high from the early stages in development of the cultures and remained so until the end of the time-window in question. This suggested that from 1DIV to 8DIV the cultures consisted predominantly of committed neurons. Composition of down- and upregulated categories of miRNAs was consistent with developing cultures being a model of neuronal growth: miRNAs previously reported as depleted from synapses were downregulated in the cultures, while miRNAs that were reported as enriched in the adult brain and neurons were upregulated. Additionally, it was also possible to establish, using the example of miR-124 and let-7 miRNAs, that miRNAs were likely to have been biologically active in cultures. Therefore I concluded that the forebrain cultures were a good model to study miRNA regulation of gene expression in neurons.

3.3 Selection of miRNAs for functional experiments

The goal of this work was to describe the role of miRNAs in the development and function of neurons (Introduction, section 1.2.2). Identification of three different modes of miRNA expression during development of primary forebrain cultures posed a question if miRNAs from different classes had similarly important roles in neurons.

Based on results of published works, miRNAs of the downregulated category were least likely to be functionally important for neurons. For example, miRNAs that were most downregulated in the cultures development were found to be most strongly depleted miRNAs from synapses in the adult mouse forebrain (Lugli et al., 2008) (Table 3.4). Additionally, a phenotype of a stable knock out mouse line lacking two of the four most downregulated miRNAs, miR-143 and miR-145 (Table 3.4), was published (Elia et al., 2009), and no significant abnormalities in brain development and function were reported. Therefore miRNAs of the downregulated category were assumed to be non-neuronal and non-functional in neurons under normal circumstances.

On the other hand, multiple miRNAs from the steady state highly expressed category were described as functionally important for neuronal development (Table 3.3). The upregulated in development miRNAs could also *a priori* be important for neuronal biology, as at least one of these miRNAs, miR-132, was shown to be induced by neuronal activity (Klein et al., 2007). Additionally, 41 miRNAs transcribed from the region in the distal end of chromosome 12 were among upregulated in cultures miRNAs (section 3.2.2). Previously, miRNAs from that region were shown to be highly expressed in the brain relative to other organs (Seitz et al., 2004), and misregulation of expression of that region was implicated in a mental disorder (Lewis and Redrup, 2005). Therefore miRNAs from the upregulated category were assumed to be functional and neuronal.

In total ten miRNAs were selected for functional experiments. Selection of miRNAs from down- and upregulated categories was based on two criteria: the level of expression during the developmental timecourse of primary cultures and the magnitude of change in expression between the first and the last developmental timepoints (i.e. between 1DIV and 8DIV). Table 3.5 summarises this information about the selected miRNAs from down- and upregulated categories.

Two miRNAs were selected from the steady state highly expressed category: **miR-124** and **miR-103**. Selection of miR-124 was due to its reported role in neuronal differentiation (summarised in Table 3.3) and also because of the indication of its direct role in development of primary cultures (Figures 3.9 and 3.10). Additionally, experiments

miRNA	# expression	# fold change
miR-143	*at 1DIV:* 16th	*Downregulated:* 1st
miR-145	*at 1DIV:* 12th	*Downregulated:* 2nd
miR-25	*at 1DIV:* 3rd	*Downregulated:* 10th
miR-551b	*at 8DIV:* 4th	*Upregulated:* 1st
miR-370	*at 8DIV:* 33rd	*Upregulated:* 2nd
miR-410	*at 8DIV:* 7th	*Upregulated:* 10th
miR-434-3p	*at 8DIV:* 2nd	*Upregulated:* 36th

Table 3.5: Selection of down- and upregulated miRNAs.
 # expression - rank of a miRNA by the level of expression at 1DIV or 8DIV among all miRNAs comprising a relevant category (either down- or upregualted); *# fold change* - rank of a miRNA by the level of fold change between 1DIV or 8DIV among all miRNAs comprising a relevant category (either down- or upregualted).

with this miRNA could serve as a positive control of methodology, because of the strong prior information arguing for the importance of miR-124: if the methods were suitable for studying neuronal function of miRNAs, such a function should be observable for miR-124. On the other hand, miR-103, although it was the second most highly expressed miRNAs at 8DIV (Supplementary Data, Table A.8), was not previously reported as having neuronal function. This miRNA was selected in order to establish if a high miRNA abundance was a good indication of importance and function of a miRNA.

To estimate the effects to be expected from an undoubtedly non-neuronal miRNA in neurons, a non-mouse miRNA, **cel-miR-67**, was selected for functional experiments. This miRNA was identified in *Caenorhabditis elegans*, and its seed region was different from any known mature mouse miRNA, as of miRBase Release 14 (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006, 2008). Despite cel-miR-67 not being naturally expressed in mouse primary cultures, the same methods that were used for overexpression of mouse miRNAs could be used for cel-miR-67 (Methods, section 2.5). Delivery of cel-miR-67 into mouse cells was expected to result in its loading and guiding mouse RISC to targets of cel-miR-67 in the same way as the endogenous miRNAs, because cases of functional activity of ectopically expressed miRNAs in a mammalian cell culture system were previously reported (Lim et al., 2005). It was assumed that cel-miR-67 could act similarly to mouse miRNAs and use its seed region for RISC guidance. Mouse neuronal mRNA transcripts had not evolved to avoid targeting by cel-miR-67. Therefore the effect of this miRNAs could serve as an estimate of effects expected from a "generic" miRNA, which targets a random sample of mouse transcripts susceptible to miRNA mediated destabilisation.

Chapter 4

A system to study microRNA function

As described in the Introduction (section 5.2) perturbation of miRNA expression in primary E17.5 forebrain cultures was the basis for experimental determination of miRNA function and targets in this thesis. Perturbation of miRNA expression levels was achieved through chemical transfection of miRNA mimics (for ectopic expression or overexpression of endogenous miRNAs) and inhibitors (for antisense mediated inhibition of endogenous miRNAs) in primary cultures (Methods, section 2.5). All transfection experiments in this work were conducted at one of the three developmental timepoints: **3DIV**, **4DIV** and **6DIV**. The 4DIV timepoint was selected because of its significance as a switch point in developmental gene expression program of primary E17.5 neuronal cultures, after which the ratio of neuritic and somatic transcripts resembles that in mature neurons (Chapter 3, section 3.1.3) (Valor et al., 2007; Manakov et al., 2009). The other two timepoints were selected on either side of the 4DIV timepoint. To transfect the mimics and the inhibitors, I used a protocol that was developed by a colleague in the laboratory, Dr. Erik MacLaren (Maclaren et al., 2011), for transfection of siRNAs into primary neuronal cultures. The first part of this chapter describes testing of the transfection protocol, and the second part describes adjustments to the protocol that improved detection of direct targets of miRNAs.

4.1 Efficient transfection of neurons

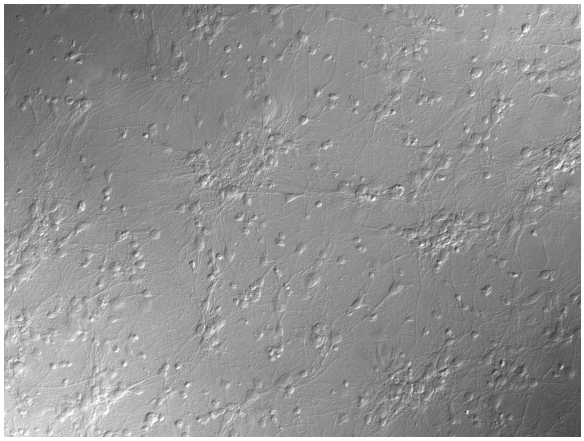
4.1.1 Microscopy confirmed efficient transfection of neurons in primary forebrain cultures

To confirm that transfection of neurons in primary E17.5 forebrain cultures was possible with the original siRNA transfection protocol, a plasmid expressing eGFP was transfected into the cultures at 6DIV. At 36 h after transfection, cultures were fixed and visualised (Methods, section 2.6). A strong fluorescence was detected in some neurons in the culture (Figure 4.1). This experiment unequivocally demonstrated that the protocol enabled transfection of neurons and it also showed that neurons were a predominant cell type among transfected cells (see images at a low magnification, Figure 4.1).

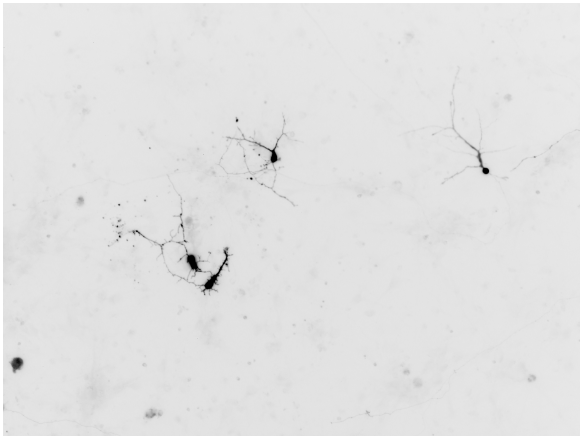
Transfection of a plasmid expressing eGFP demonstrated unambiguously that neurons were transfected, but it was not informative of the efficiencies to be expected in miRNA transfection experiments. The sizes of the plasmid and of miRNA mimics or inhibitors were different, thus, frequency of plasmid delivery was not a good estimate of the transfection efficiency for short polynucleotides. Additionally, the inefficient process of transfer to the nucleus is required for reporter gene expression (Zabner et al., 1995), while activity of miRNA mimics and inhibitors is thought to take place in the cytoplasm. Therefore, transfections of fluorescently labelled oligonucleotides were suggested by a manufacturer of miRNA mimics and inhibitors (Qiagen N.V.) to be a better estimate of the efficiency expected in transfections of miRNA mimics and inhibitors. The AlexaFlour 488 labelled oligonucleotide RNA was transfected into primary cultures at 3DIV and 6DIV (Methods, section 2.6). After 36 h of incubation, an abundant bright punctate fluorescence was detected (Figure 4.2). This was likely due to coagulation of the fluorescently labelled oligonucleotides outside the cells. At the same time, a fainter diffuse fluorescence was observed in a majority of the cells. This signal was unlikely to be due to the autofluorescence of the cells, as mock transfected cultures did not display the signal when viewed with the same settings (Figure 4.2e).

In summary, transfection of an eGFP expressing plasmid and AlexaFlour 488 labelled oligonucleotide demonstrated that the transfection protocol was capable of delivering various nucleic acids (e.g. a plasmid and a fluorescently labelled RNA oligomer) into primary neurons at different developmental timepoints. Additionally, these experiments showed that it was possible to transfect a majority of neurons with an RNA oligomer.

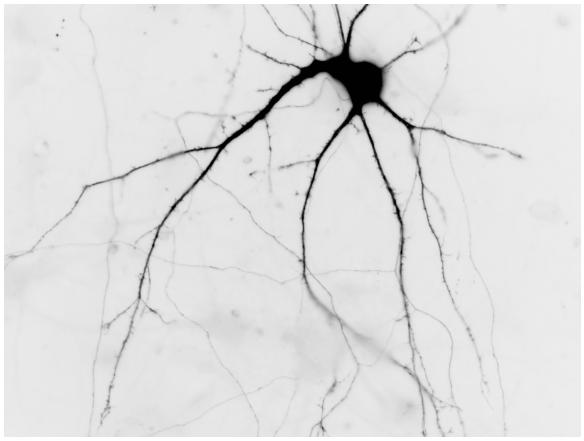
(a) transfections at 6DIV (DIC)



(b) transfections at 6DIV (eGFP)



(c) transfections at 6DIV (eGFP)



(d) transfections at 6DIV (eGFP)

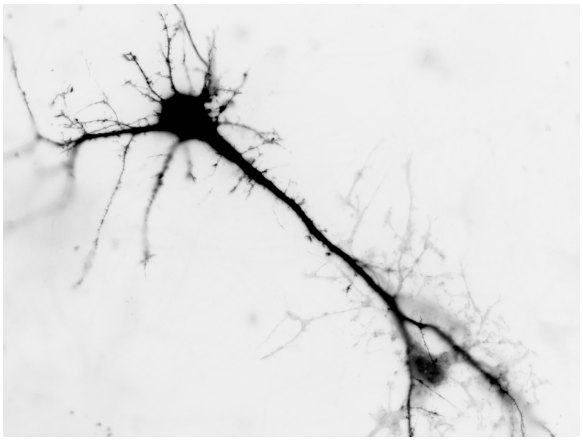
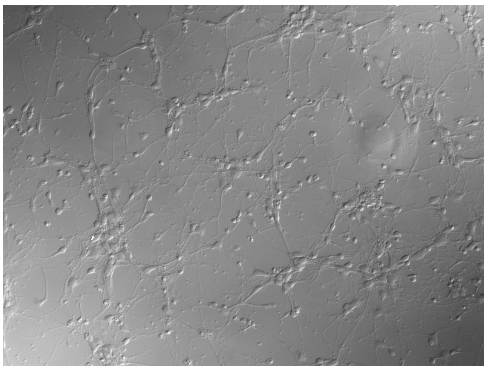
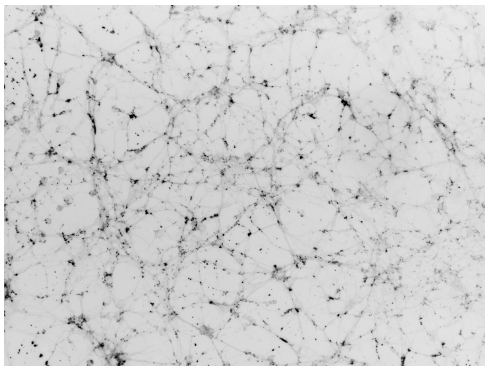


Figure 4.1: Transfection of primary cultures with eGFP-expressing plasmid.
Figures 4.1a and 4.1b show the same area of the culture viewed with differential interference contrast (DIC) or fluorescence microscopy (495 nm light for excitation of enhanced GFP (eGFP)) settings at a low magnification. Figures 4.1c and 4.1d show the fluorescence microscopy images (495 nm light for excitation of eGFP) at a high magnification. The slide preparation and the microscopy settings are described in Methods (section 2.6).

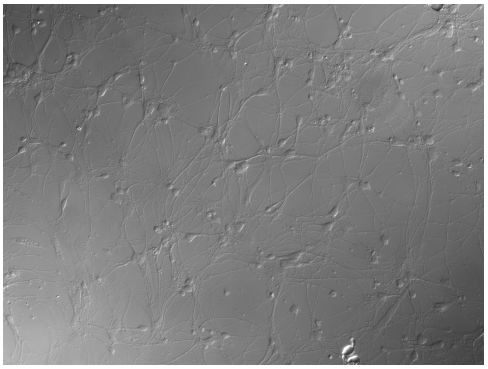
(a) transfections 3DIV (DIC)



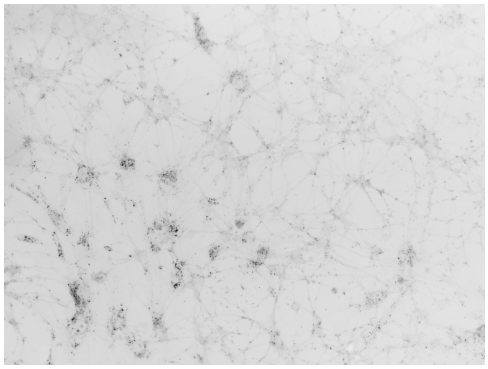
(b) transfections 3DIV (AlexaFlour 488)



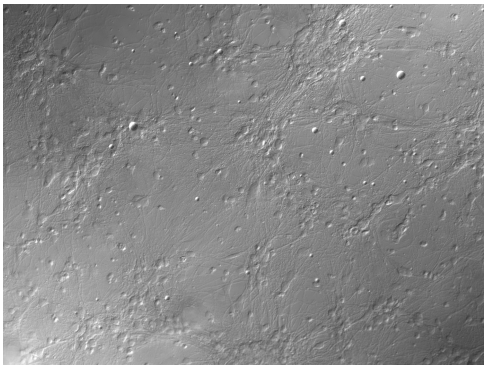
(c) transfections 6DIV (DIC)



(d) transfections 6DIV (AlexaFlour 488)



(e) mock 6DIV (DIC)



(f) mock 6DIV (AlexaFlour 488)

Figure 4.2: Transfection of primary cultures with AlexaFlour 488 labelled oligo at 3DIV and 6DIV.
Pairs of figures 4.2a and 4.2b, 4.2c and 4.2d, 4.2e and 4.2f, show the same areas of transfected cultures viewed with differential interference contrast (DIC) or fluorescence microscopy (495 nm light for excitation of AlexaFlour 488) settings at a low magnification. The exposure time for all AlexaFlour 488 images was fixed at 220ms. The slide preparation and the microscopy settings are described in Methods (section 2.6).

4.1.2 Transfection did not cause neuronal death

Four types of evidence showed that no significant neuronal loss was associated with transfections of primary cultures. These four sources of evidence came from visual examinations, measurement of differential gene expression, recordings of electrophysiological activity in transfected cultures and Trypan blue assay (Altman et al., 1993).

Visual inspections of transfected cultures did not reveal neuronal loss

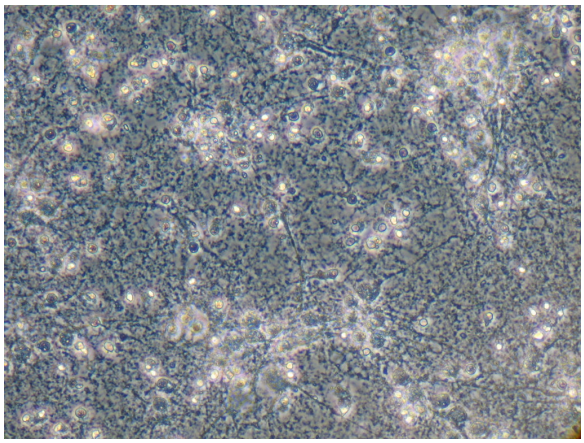
Concentration of mimics and inhibitors in transfection mixtures was either 0 nM (in case of mock transfection), 115 nM or 230 nM (Methods, section 2.5). No significant cell loss was visually detected in any of the experiments, neither in mock transfections nor in transfection of mimics and inhibitors. Moreover, addition of RNA in even higher concentration ($>$ 1,300 nM) to the mixture did not cause an observable neuronal loss (Figure 4.3). A complete degradation of nearly all neurites was associated with the death of neurons (Figure 4.3a), and this was not found to be the case in a mock transfected culture (Figure 4.3b) or cultures treated with a transfection mixture containing a mimic or an inhibitor in $>$ 1,300 nM concentration (Figure 4.3c and 4.3d).

Trends in differential gene expression were not compatible with a consistent significant neuronal loss

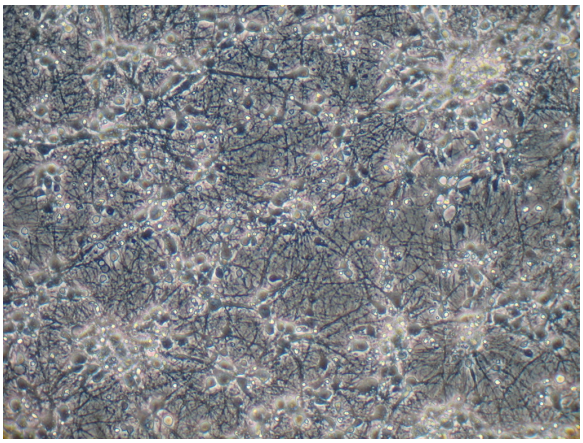
A significant loss of neurons would skew the results of gene expression analysis, because removal of neurons (i.e. neuronal mRNA) from transfected cultures would make neuron specific genes appear as downregulated. To test if this might have been the case, a list of putatively neuron specific genes was compiled (Methods, section 2.9) and their expression was compared to the rest of the genes.

Analysis of changes in expression of these neuron specific genes revealed that no consistent or significant neuronal loss was associated with transfection experiments. Although neuron-specific genes were found to be downregulated in some experiments (consistent with neuronal cell loss), they were upregulated in other experiments (inconsistent with significant neuronal loss). For example, overexpression of cel-miR-67 at 6DIV lead to significant downregulation of the neuron specific genes, while overexpression of miR-124, performed with the same protocol and at the same developmental timepoint, lead to upregulation of the neuron specific genes (Figure 4.4).

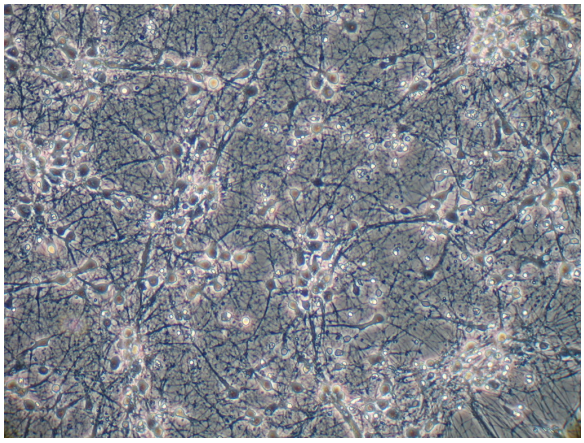
(a) killed by sodium azide



(b) mock transfection



(c) transfection with miR-103 mimic



(d) transfection with miR-103 inhibitor

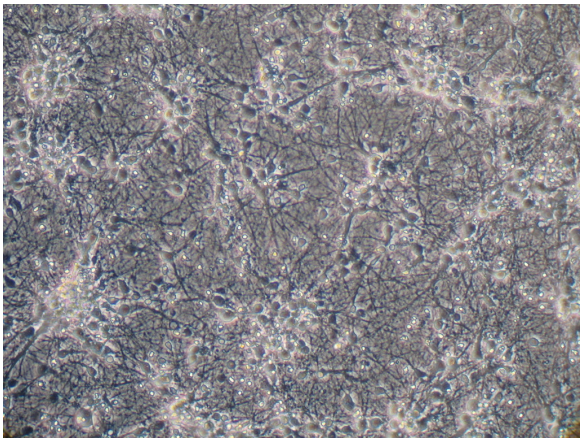


Figure 4.3: Visual inspection of transfected cultures.
The cultures were transfected at 6DIV, and the images were taken after 48h of incubation. 4.3a - killed with sodium azide at 8DIV (0.03%, 24h incubation); 4.3b - mock transfection; 4.3c - transfection with the mimic of miR-103 ($> 1,300$nM); 4.3d - transfection with the inhibitor of miR-103 ($> 1,300$nM).

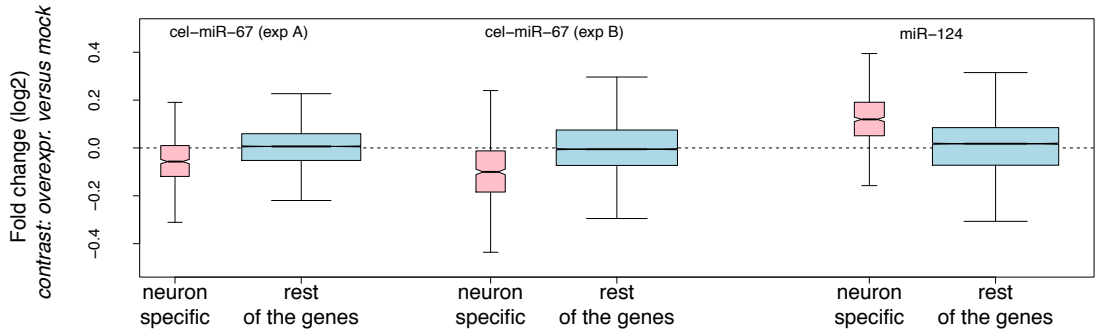


Figure 4.4: Differential regulation of neuron specific genes.
The y-axis shows the fold change (log_2) in gene expression of the *neuron specific* and the *rest of the genes* upon transfections of the cel-miR-67 and miR-124 mimics at 6DIV (the names of the over-expressed miRNAs are shown in the plot area, the types of the genelists are shown on the x-axis). The transfection of cel-miR-67 mimic at 6DIV was independently repeated twice (shown as *exp A* and *exp B*). The boxes correspond to the distribution of the fold changes (log_2) of the genes in the contrast of cultures transfected with the mimics to the matched mock transfected cultures. The width of the boxes corresponds to the number of genes within the two lists (the neuron specific and the rest of the genes) present among the genes detected in each of the experiments (using the standard Illumina detection call $P < 0.01$, see Methods, section 2.7). The notches of the boxes correspond to the median value of the distribution of the fold changes (log_2) of the genes in the genelists, the bottom and the top sides of the boxes correspond to the first and the third quartiles, the whiskers extend to no more than 1.5 times the interquartile range (IQR), or to the most extreme data-point, if it is closer to the median than 1.5 IQR. In all three experiments, the Wilcoxon test P-values for the differences between the medians of the neuron specific genes and the rest of the genes was beyond the precision limit of the test as implemented using the standard R libraries (RTeam, 2008) ($P < 1e - 320$).

Number of active synapses was not affected by transfections

While establishing the original protocol for siRNA transfection, it was shown that neither the mock transfection nor transfections with siRNAs (designed to target genes shown in Figure 4.5) reduced the number of neurons, or a number of active synapses in primary cultures (Maclaren et al., 2011). In order to demonstrate that, primary neuronal cultures were plated on microelectrode arrays (MEAs), where firing patterns (spikes) could be recorded for a number of days after transfections. Toxicity of the transfection reagent could be uncovered through these measurements, because it was previously shown that the total number of spikes correlated with the number of active synapses adjacent to the electrodes of MEAs (Wagenaar et al., 2006) and with synaptic density in the whole culture (Brewer et al., 2009). No significant differences ($P < 0.05$) were detected at any timepoint between mock transfected, untransfected and siRNA transfected cultures (Maclaren et al., 2011) (Figure 4.5).

Cell viability was not affected by transfections

Finally, I confirmed viability of transfected cells with Trypan assay (Altman et al., 1993). Cultures at 6DIV were transfected with the mimic of cel-miR-67 (220 nM concentration, see Methods, section 2.5) and stained with Trypan blue and DAPI at 48 h post transfection (Figures 4.6a and 4.6b, see Methods, section 2.6). Majority of both transfected and matched untransfected cells (Figures 4.6c and 4.6d) were viable (i.e. not stained with Trypan blue (Table 4.1), see for a comparison cells treated with sodium azide in Figures 4.6e and 4.6f). Additionally, the difference in percent of viable cells was not significant between transfected and untransfected cultures (t-test $P > 0.7$, based on three non-overlapping images per treatment).

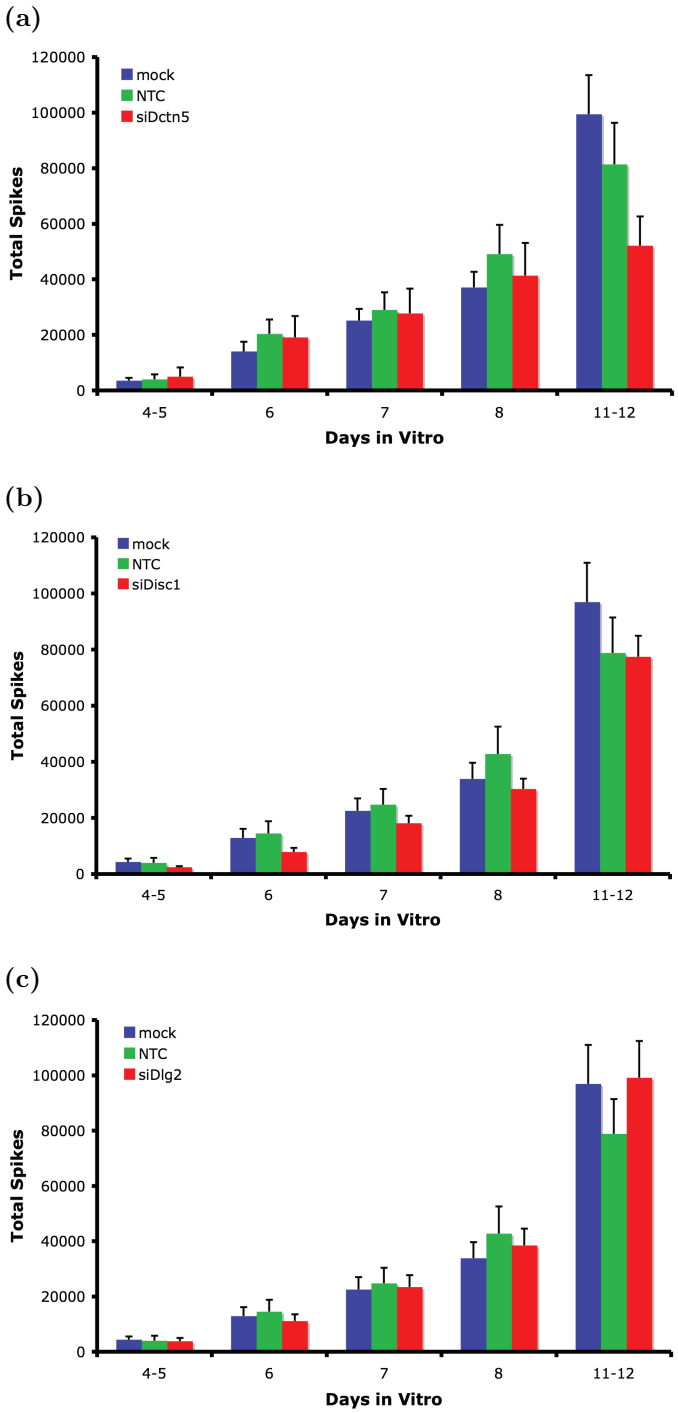
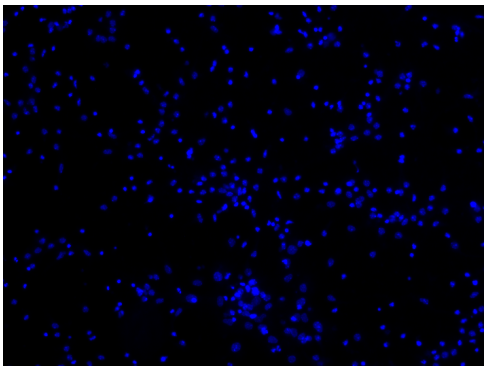
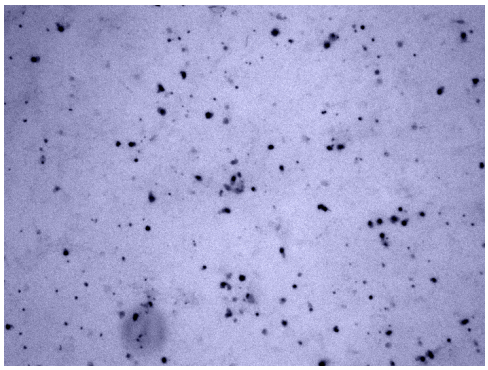


Figure 4.5: The total number of spikes recorded in cultures after transfections. The figure is reproduced from the manuscript by MacLaren *et al.*, with permission of Erik MacLaren. "*mock*" – mock transfected cultures; "*NTC*" – untransfected cultures; "*siDctn5*", "*siDisc1*" and "*siDlg2*" – cultures transfected with siRNAs designed to target Dctn5, Disc1 and Dlg2 transcripts (Maclaren et al., 2011).

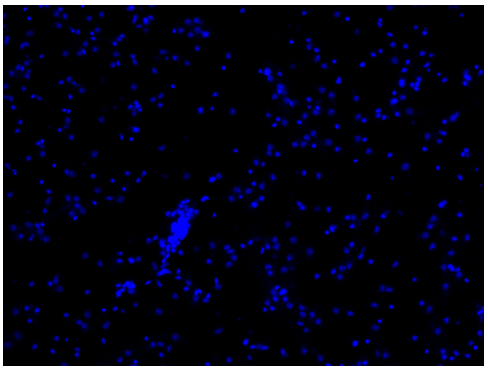
(a) transfected (DAPI)



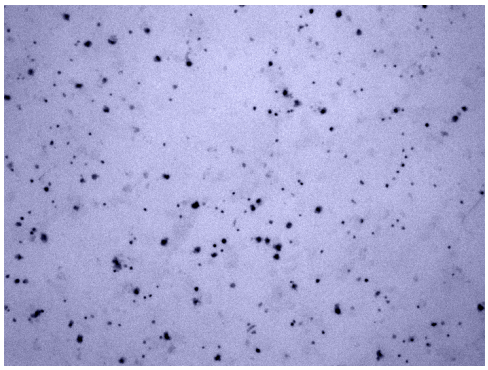
(b) transfected (Trypan)



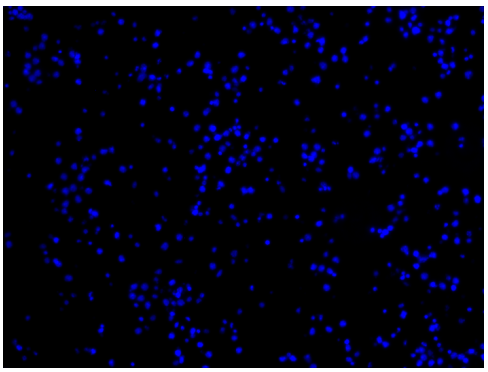
(c) untransfected (DAPI)



(d) untransfected (Trypan)



(e) killed by sodium azide (DAPI)



(f) killed by sodium azide (Trypan)

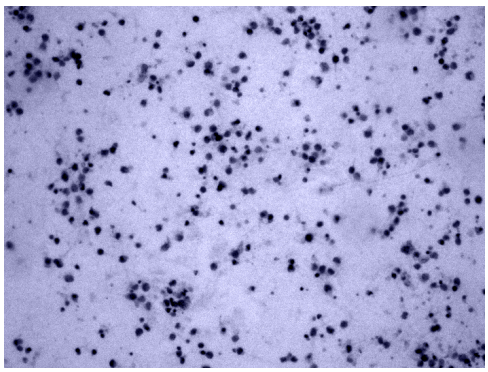


Figure 4.6: Viability of cells in transfected cultures.
Pairs of figures show the same areas of cultures stained with DAPI or Trypan blue (see titles of the subfigures). Figures 4.6a and 4.6b show a culture transfected at 6DIV with the mimic of cel-miR-67 (visualised at 8DIV); Figures 4.6c and 4.6d show a matched untransfected culture (visualised at 8DIV); Figures 4.6d and 4.6d) show a culture treated at 8DIV with sodium azide (0.03%, 24 h incubation). The assay is describe in Methods (section 2.6).

Days *in vitro*	**Viability**	**Standard deviation**
transfected	68.1%	± 2.85%
untransfected	66.4%	± 8.05%

Table 4.1: Viability of cells in transfected cultures.
Viability was estimated based on three non-overlapping 10× objective images (∼ 500 cells per image, see for example Figure 4.6). Total numbers of cells were estimated by counting DAPI stained nuclei, numbers of dead cells – counting Trypan stained cells (Methods, section 2.6).

4.1.3 Transfection of miRNA mimics consistently induced miRNA mediated changes in gene expression

Although transfections of a plasmid expressing eGFP and of a fluorescently labelled oligonucleotide (section 4.1.1) showed that neurons were efficiently transfected, it was not known if introduced miRNAs were active inside cells. It was also not known how well the changes in gene expression induced by transfections were reproducible between different transfection experiments.

Activity of introduced miRNAs and their widespread direct effect on gene expression in the cultures was made evident through analysis of the distribution of seed matching sites complementary to the transfected miRNAs. It was previously demonstrated that a widespread direct effect of an overexpressed miRNA on gene expression manifested itself as a significant enrichment of the sites complementary to the seed region of that miRNA in the 3′UTRs of downregulated genes (Lim et al., 2005; Giraldez et al., 2006). Such an effect was consistently observed upon transfection of various miRNA mimics into primary E17.5 forebrain neuronal cultures (for example see Chapter 5, Figure 5.1). Importantly, by using Sylamer (van Dongen et al., 2008), which simultaneously assess biases in occurrence of all nucleotide words (Methods, section 2.8), it was possible to show that such enrichment in many cases was exclusively specific to the seed matching sites of only the transfected miRNAs (for example see Chapter 5, Figure 5.2).

The concern that the variable nature of the primary culture system would dramatically compromise reproducibility of gene expression measurements was resolved by doing replicate transfections of the same miRNA mimic. By conducting replicate experiments of cel-miR-67 at 4DIV and 6DIV it was possible to show significant correlation in gene expression changes between the experiments (Figure 4.7). Additionally, correlation between genes that contained sites complementary to the seed region of cel-miR-67 was higher than for all genes. The latter observation was consistent with the changes induced directly by the miRNA to be among primary changes in transfected cells.

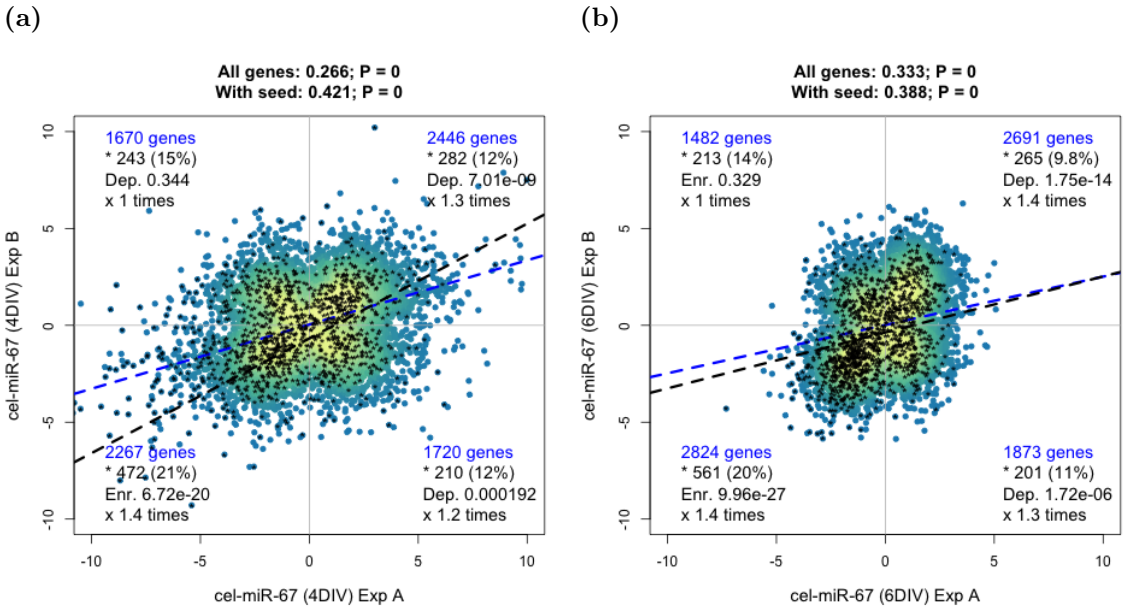


Figure 4.7: Correlation of cel-miR-67 experiments at 4DIV and 6DIV.
The points correspond to 8,103 and 8,870 genes detected in the replicate cel-miR-67 mimic transfection experiments at 4DIV (4.7a) and 6DIV (4.7b). The axes correspond to the moderated t-statistic for differential expression between the cultures transfected with the mimic of cel-miR-67 and the matched mock transfected cultures in each of the experiments. The differential expression analysis is described in Methods (section 2.7). The colors of the points depend on the density of the points in a given region of the plot (yellow – highest, blue – lowest). Black asterisks mark genes that encode transcripts with 3′UTRs harbouring one or more seed matching sites (7(2) or 7(1A)-types) for cel-miR-67. The text gives the following information: 1) The total number of genes in each of the quadrants of the plots; 2) The number (and percentage) of genes [encoding transcripts] with seed matching sites for cel-miR-67 in their 3′UTRs; 3) Hypergeometric P-value for the enrichment (*Enr.*) or depletion (*Dep.*) of genes with the seed matching sites in each of the quadrants; 4) Fold enrichment or depletion of genes with the seed matching sites for cel-miR-67 "$\times$ *times*" the number of genes that is expected by chance alone. Mapping of microarray probes to mRNA transcripts, and transcripts to genes is described in Methods (section 2.7). Pearson correlation between differential expression of all genes (*All genes* – all points) and genes with the seed sites (*With seed* – the points with asterisks), is given at the top of the plots, together with P-value of correlation (in each case it was beyond the precision of the correlation test as implemented via the standard R libraries (RTeam, 2008), which is equivalent to $P < 1e-320$). The blue dashed line is a linear model fitted through the all points (*All genes*), and the black dashed line - through points with asterisks (*With seed*).

Summary of section 4.1

Imaging cultures transfected with an eGFP expressing plasmid proved that the transfection protocol was efficient for delivery of nucleic acids into primary neurons. Transfection of the cultures with a fluorophore labelled oligonucleotide indicated that a majority of neurons was transfected in experiments at 3DIV and 6DIV. Visual inspection and directions of change in expression of neuron specific genes showed that transfections of primary cultures were unlikely to be associated with significant death of neurons. Additionally, measurement of electrophysiological parameters confirmed that transfections did not reduce the total number of active synapses in cultures (Maclaren et al., 2011), i.e. transfections had low toxicity to neurons. Significant miRNA mediated changes in gene expression were detected in transfection of cultures with mimics of several different miRNAs. A significant correlation was observed in differential gene expression that was triggered by transfections of cel-miR-67 mimics in replicate experiments. Therefore, I concluded that transfection protocol was capable of transfecting primary neurons, that transfections of miRNA mimics elicited miRNA mediated effects on gene expression in the cultures and that experimental results were reproducible.

4.2 Improving detection of miRNA targets

4.2.1 The use of miRNA inhibition instead of mock transfection improved detection of putative direct targets

The key modification of the design of the original siRNA transfection strategy (Maclaren et al., 2011), which improved detection of miRNA targets, was the use of cultures transfected with miRNA inhibitors instead of mock transfected cultures. Matched mock transfected cultures were previously used to contrast gene expression changes in transfection experiments (Lim et al., 2005; Selbach et al., 2008; Hendrickson et al., 2009). The contrast of transfected cultures with mock transfected cultures can be justified, because it compensates for the changes induced by technical manipulation and the transfection reagent itself. However, the use of cultures transfected with miRNA inhibitors has an important advantage for miRNA target identification, as it favours detection of direct miRNA targets, because the direct targets are expected to have an inverse response to miRNA overexpression and inhibition. For example, it was demonstrated that upon overexpression of miR-124, its targets were downregulated, and upon miR-124 inhibition – upregulated (Conaco et al., 2006). In the remainder of this thesis, transfection experiments where differential expression was identified through the contrast of cultures transfected with miRNA mimics to mock transfected cultures are referred to as **unidirectional** overexpression experiments. Experiments where cultures transfected with miRNA mimics were contrasted to the cultures transfected with miRNA inhibitors are called **bidirectional** perturbation experiments.

The first miRNA investigated with the bidirectional perturbation strategy was miR-124, as the inverse response of its targets was shown to take place upon overexpression of miR-124 versus its inhibition (Conaco et al., 2006). These bidirectional perturbation experiments were performed at two developmental timepoints (3DIV and 6DIV). The original transfection protocol was employed for this experiment (Methods, section 2.5), and both the mock transfection and the transfection with the miR-124 inhibitor were conducted for comparative purposes.

Transition from the contrast with mock transfected cultures to the contrast with the inhibition increased enrichment of transcripts with miR-124 seed matching sites among the downregulated transcripts (Figure 4.8). In transfections at 3DIV, the P-value of enrichment changed from $1.42e-14$ to $5.82e-29$, and at 6DIV from $6.67e-39$ to $3.11e-43$. Importantly, Sylamer analysis of the distribution of miRNA seed matching

sites also identified the increase of the enrichment (Figure 4.9). In 6DIV experiments, the peak of Sylamer enrichment increased by over 10 orders of magnitude in transition from the use of mock transfection to the miR-124 inhibitor (Figures 4.9c and 4.9d). I concluded that transition from unidirectional to bidirectional experimental strategy improved detection of putative miR-124 targets. This conclusion was based on previously reported observations, where, in miRNA overexpression experiments, downregulated transcripts with seed matching sites for the overexpressed miRNAs were enriched in validated direct targets of these miRNAs (Lim et al., 2005; Giraldez et al., 2006).

Data obtained from bidirectional miR-124 experiments was used to compile the list of miR-124 targets, which will be described in Chapter 5 (section 5.2.1). It was decided to use the bidirectional strategy for identification of targets of other mouse miRNAs, because it worked to improve detection of targets of miR-124. However, it was expected that of all selected miRNAs, transfection of miR-124 would likely have the strongest effect on gene expression in primary neuronal cultures (the selection is described in Chapter 3, section 3.3). Unlike other selected miRNAs, miR-124 was well known for its importance for neuronal biology (Chapter 3, Table 3.3), its expression was shown to neuron specific (Christodoulou et al., 2010; Clark et al., 2010) and it was shown to be one of the most abundant miRNAs in brain (Landgraf et al., 2007). Therefore, to maximise the chances of detection of targets of other miRNAs in bidirectional experiments, it was decided to use miR-124 transfections to optimise the original siRNA transfection protocol for miRNA target detection. The next section describes the use of miR-124 transfections for selection of an optimal post-transfection incubation time and cell plating density.

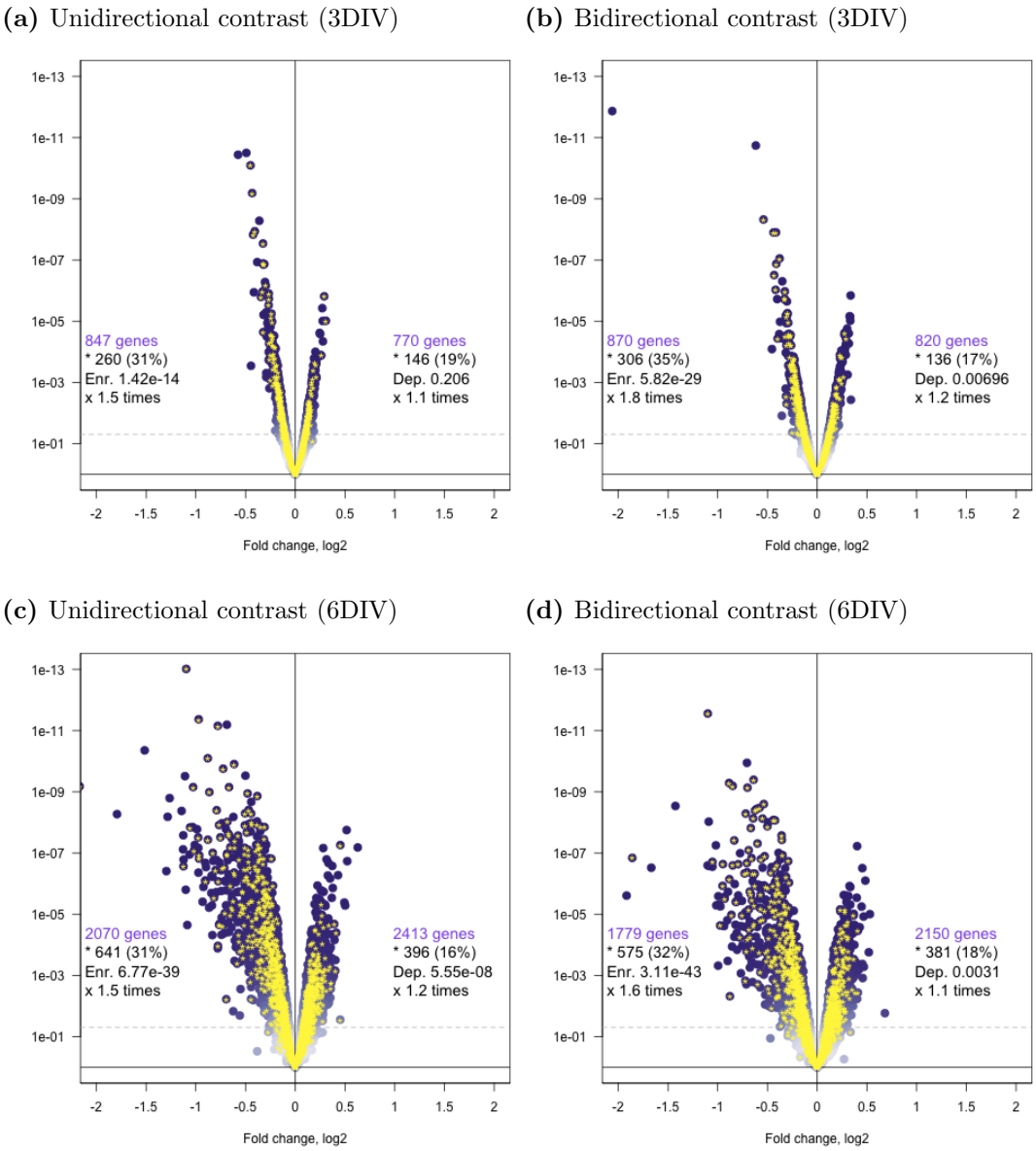


Figure 4.8: Differential gene expression and seed matching site enrichment in miR-124 transfection experiments at 3DIV and 6DIV.
Genes detected by microarrays (using the standard Illumina detection call $P < 0.01$) are shown as the purple dots (the analysis of microarray data is described in Methods, section 2.7). The x-axes represent log_2 of gene expression fold change between samples transfected: 4.8a – with the mimic of miR-124 in comparison to the matched mock transfection at 3DIV; 4.8b – with the mimic of miR-124 in comparison to the transfection with the inhibitor of miR-124 at 3DIV; 4.8c – with the mimic of miR-124 in comparison to the matched mock transfection at 6DIV; 4.8d – with the mimic of miR-124 in comparison to transfection with the inhibitor of miR-124 at 6DIV. The y-axes represent P-value of differential expression (log_{10} scale), and the horizontal dashed grey lines show P-value cutoff of 0.05. The yellow asterisks mark genes [encoding transcripts] with 3′UTRs harbouring one or more seed matching sites (7(2) or 7(1A)-types) for miR-124. The text in the two halves of the plot area provides the following information: 1) The total number of genes with differential expression P-value more significant than the cutoff (0.05); 2) The total number (and percentage) of genes with seed matching sites for miR-124; 3) The hypergeometric P-value of enrichment (*Enr.*) or depletion (*Dep.*); 4) Fold enrichment or depletion of genes with the seed matching sites "× *times*" the number that is expected by chance alone. The mapping of microarray probes to mRNA transcripts, and transcripts to genes, is described in Methods (section 2.7). The identification of the seed matching sites and the hypergeometric enrichment test is in Methods (section 2.8).

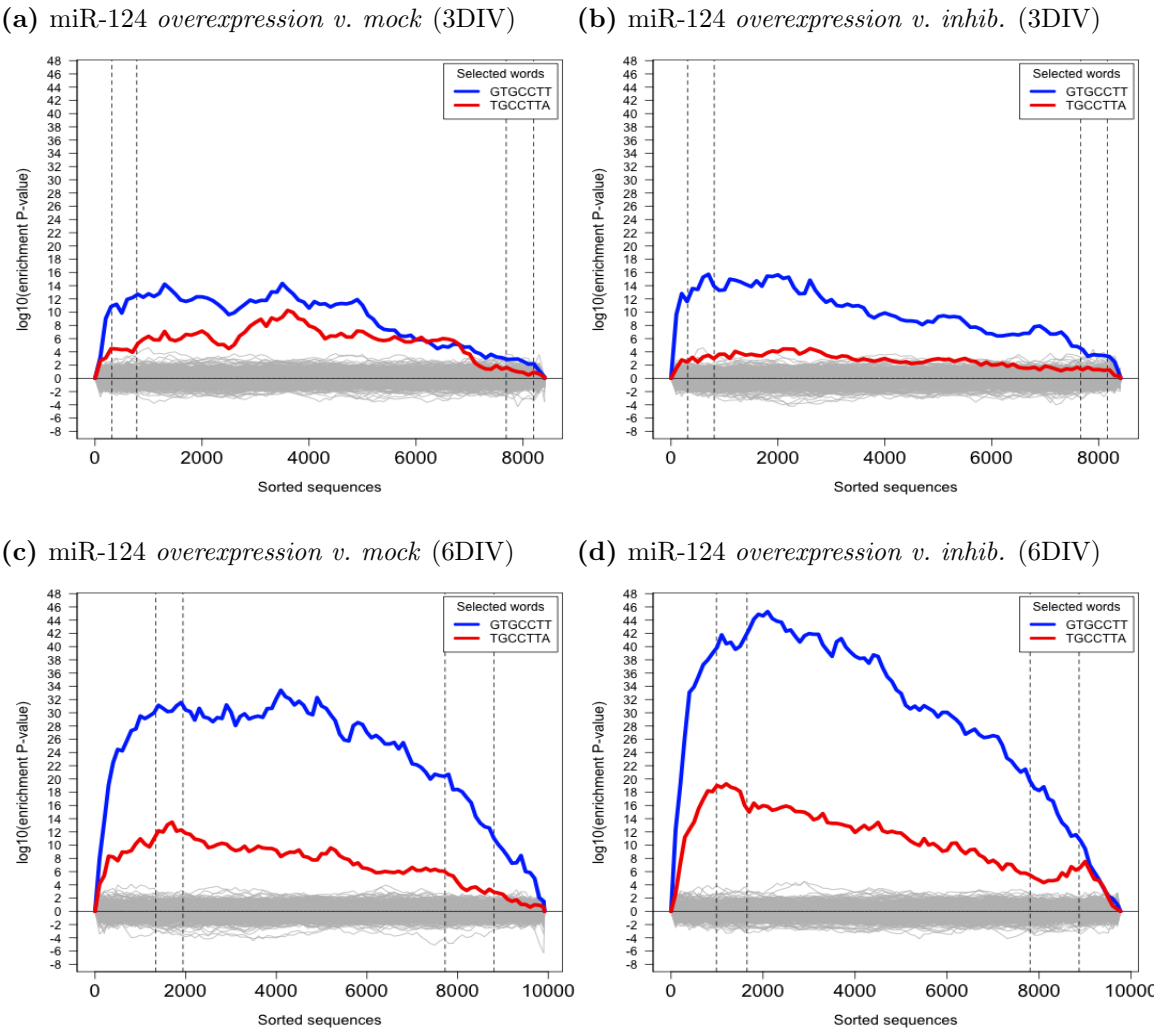


Figure 4.9: Sylamer analysis of biases in distributions of seed matching sites in miR-124 transfection experiments at 3DIV and 6DIV

The x-axes represent sorted 3′UTRs corresponding to detectably expressed genes (detected with the standard Illumina detection call $P < 0.01$, see Methods, section 2.7). These genes are ordered **from the most downregulated to the most upregulated** by fold change t-statistic for differential expression in the following transfections: 4.9a – with the mimic of miR-124 in comparison to a mock transfection at 3DIV; 4.9b – with the mimic of miR-124 in comparison to transfection with the inhibitor of miR-124 at 3DIV; 4.9c – with the mimic of miR-124 in comparison to a mock transfection at 6DIV; 4.9d – with the mimic of miR-124 in comparison to transfection with the inhibitor of miR-124 at 6DIV. The y-axes represent the hypergeometric P-values for occurrence biases of 876 nucleotide words complementary to the seed regions (7(2) and 7(1A)-types) of the complete set of 581 distinct mouse miRNAs, according to miRBase Release 14 (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006, 2008). Positive values on the y-axes correspond to an enrichment ($+|log_{10}$(P-value)|) and negative values to a depletion ($-|log_{10}$(P-value)|). The vertical dashed lines mark the P-value cutoffs (0.01 and 0.05) on both sides of the ranked gene lists. The blue and the red lines show the enrichment profiles of 7(2) and 7(1A)-type seed matching sites for miR-124. The grey lines show the enrichment for the rest of the distinct seed matching sites. The mapping of microarray probes to mRNA transcripts, and transcripts to genes, is described in Methods (section 2.7). The identification of the seed regions and parameters of Sylamer (van Dongen et al., 2008) is in Methods (section 2.8). The full description of the Sylamer method is in the Introduction (section 2.8).

4.2.2 Selection of an optimal incubation time and cell plating density

A long post-transfection incubation time was used in transfection experiments described up to now (36h or 48 h), because similarly long post-transfection times were shown to be efficient in siRNA experiments (Maclaren et al., 2011). However in mutant ES cell lines with deficient miRNA biogenesis, ectopic expression of miRNAs was shown to have the strongest effect on their targets at a short incubation time of $\approx$ 10 h (Matthew Davis, personal communication) and between 12 h to 16 h (Hanina et al., 2010)). Therefore, test experiments were carried out in order to assess the effect of miR-124 transfections on gene expression upon relatively short post-transfection incubation times (i.e. 24 h or less).

Based on miR-124 experiments described in section 4.2.1, two genes were selected as indicators of the impact of miR-124 transfection on gene expression. One of these was Lass2, a previously identified direct target of miR-124 (Conaco et al., 2006). As expected for a direct miRNA target, the changes in the expression level of Lass2 were subtle: the fold change difference between inhibition and over-expression in experiment at 3DIV was 1.074 fold ($P < 0.1$) and at 6DIV it was 1.61 fold ($P < 3.15e-05$). In addition to Lass2, another gene, Acta2, was chosen as an indicator of the effect of miR-124 transfection. Acta2 does not have a seed-matching site and may be an indirect target. However it showed bigger changes in expression (4.17 fold with $P < 1.36e-12$ at 3DIV, and 3.78 fold with $P < 2.48e-06$ at 6DIV), thus Acta2 was useful as an additional indicator of transfection efficiency.

To deduce if shorter incubation times were conducive to a bigger contrast in expression of targets in bidirectional miR-124 perturbation experiments, expression levels of Lass2 and Acta2 were assessed with qRT-PCR over a timecourse after transfection. For this, transfection of miR-124 was conducted at an intermediate developmental timepoint (4DIV) and total RNA was collected from cultures at 4 h, 12 h, 24 h and 36 h (see Methods, section 2.5). Subsequently, qRT-PCR with primers for Lass2 and Acta2 was performed as described in Methods (section 2.3). Differential expression of Lass2 and Acta2 was estimated using $\Delta\Delta$Ct method (Methods, section 2.3). Surprisingly, in contrast to results reported in the ES cells ((Hanina et al., 2010) and Matthew Davis, personal communication), the most significant differential expression of both Acta2 and Lass2 was observed at a relatively long incubation time of 36 h (Figure 4.10). This indicated that

a long post-transfection incubation time (36h) was likely to enable the most consistent detection of miRNA mediated effects.

In the next experiment, 36 h and 48 h incubations were compared. To achieve a better separation of differential expression values between 36 h and 48 h, transfection in this experiment was performed at 6DIV. The 6DIV timepoint was selected for this experiment, because results of the microarray profiling indicated a stronger miR-124 effect at 6DIV than at 3DIV (see above), and the qRT-PCR timecourse experiment at 4DIV confirmed that changes in expression of miR-124 targets were very subtle (Figure 4.10). Additionally, plating cells at a relatively lower cell plating density was tested in this experiment. For this experiment cells were plated at a density of 790 cells $\cdot$ mm^{-2}, which was over two times lower than the cell density used in experiments up to now (Methods, sections 2.5 and 2.3). The cell plating density was reduced, because in siRNA transfection experiments plating primary cultures at relatively low densities (400 to 800 cells $\cdot$ mm^{-2}) was found to increase the knock-down efficiency at the protein level (Esperanza Fernandez, personal communication).

In comparison between 36 h and 48 h post-transfection incubation, inhibition of the direct target, Lass2, was the strongest at 36 h: $\Delta\Delta$Ct was 0.924 and 0.584 at 36 h and 48 h respectively. Additionally, in agreement with the proposition of lower cell density to favour higher knock-down efficiency (Esperanza Fernandez, personal communication), lower cell plating density appeared to increase the differential expression contrast of Lass2 at mRNA level. If $\Delta\Delta$Ct at 36 h was converted into a fold change ($\Delta\Delta$Ct 1 $\approx$ 2 fold difference), then Lass2 inhibition in the lower density 6DIV experiment was bigger than that detected by microarrays in the 6DIV transfection experiment with the higher cell plating density (see above).

In summary, relatively short post-transfection incubation times (24h or less) were not identified to be significantly more efficient than longer incubation times at generating the expression contrast of miR-124 targets in bidirectional perturbation experiments. However, the results described in this section showed that several relatively minor changes to the transfection protocol could improve detection of direct miRNA targets. These changes were lowering the incubation time from 48 h to 36 h and reducing cell plating density. Therefore, these changes to the settings of the protocol were used for bidirectional per-

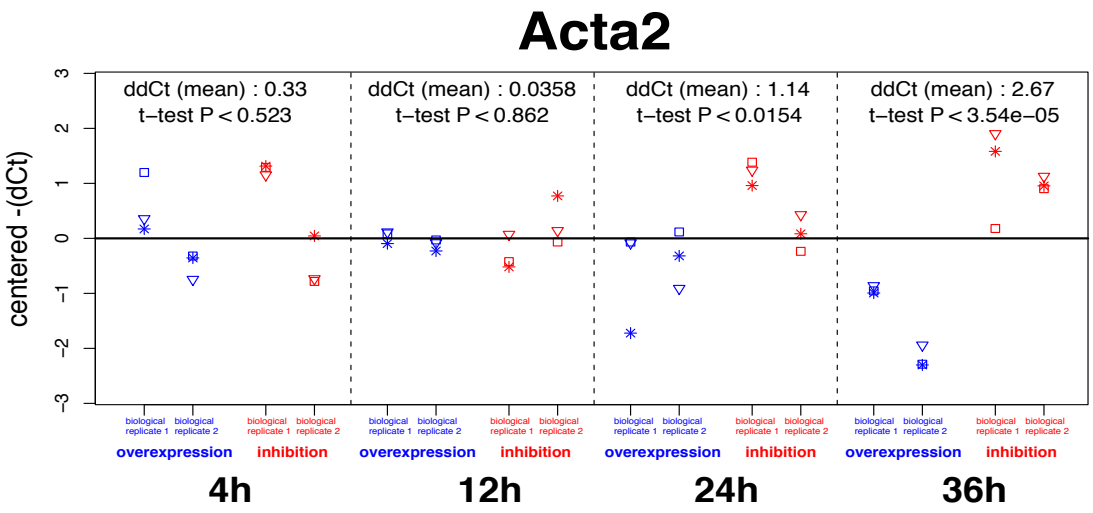
turbation experiments on the miRNAs selected for functional experiments[1] (the selection is described in Chapter 3, section 3.3).

Summary of section 4.2

The detection of miR-124 mediated inhibition of gene expression was more efficient when transfections with the mimic were compared to transfections with the inhibitor, rather than to mock transfections. Following a robust identification of miR-124 mediated effects in miR-124 transfection experiments, series of these experiments were used to improve upon the original transfection protocol. A drastic reduction of post-transfection incubation time did not significantly increase detection of miR-124 mediated inhibition of its targets, however an improvement was detected upon a 12 h reduction of post-transfection incubation time and reduction of cell plating density. Therefore the bidirectional strategy and the adjusted protocol were used to identify targets of all other selected miRNAs (with the exception of cel-miR-67, for which the inhibition was not available), which is described in the next chapter.

[1] Apart from miR-124, for which 48 h incubation and higher cell density worked satisfactorily (section 4.2.1), and cel-miR-67 that was not present in the mouse genome and its inhibition was not logistically possible.

(a)



(b)

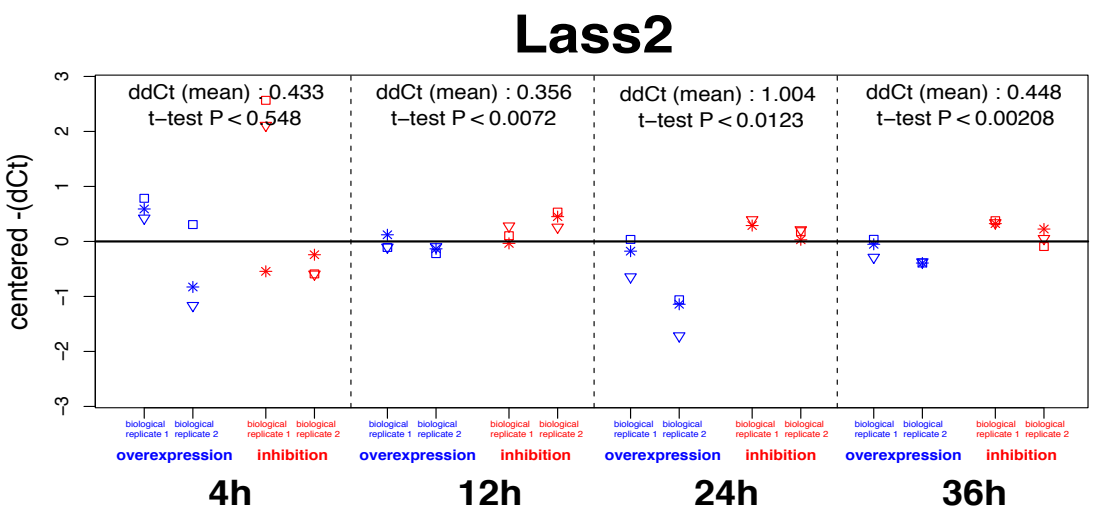
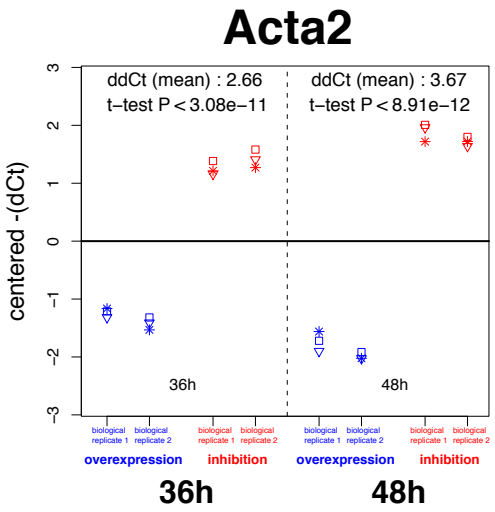


Figure 4.10: qRT-PCR profiling of the effect of miR-124 transfection in a high cell plating density timecourse at 4DIV.

The x-axes correspond to biological replicates of transfected samples (*overexpression* - with the miR-124 mimic, *inhibition* - with the miR-124 inhibitor. The y-axes correspond to inverse ΔCt values centered around the experimental medians. The styled points correspond to ΔCt values for technical replicates per one biological replicate. The text gives the mean of ΔΔCt values and the t-test P-value for the differential expression. The ΔΔCt method is described in the Methods (section 2.3).

(a)



(b)

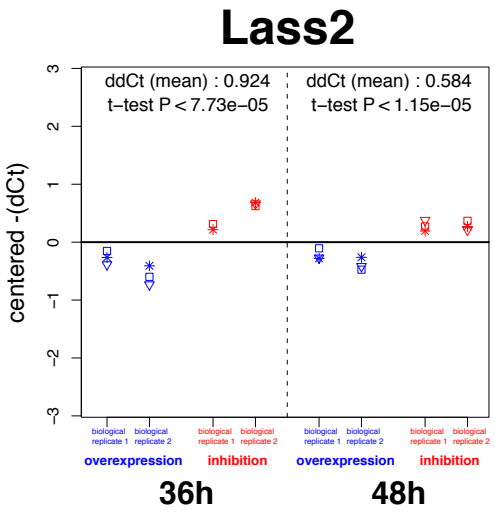


Figure 4.11: qRT-PCR profiling of the effect of miR-124 transfection in a low cell plating density timecourse at 6DIV.
See Figure 4.10 for description.

Chapter 5

Results of transfection experiments

The bidirectional perturbation experiments (i.e. the experiment in which cultures transfected with miRNA mimics were compared to cultures transfected with miRNA inhibitors) allowed to efficiently detect miR-124 mediated inhibition of gene expression (Chapter 4, section 4.2.1). Therefore, I decided to use this strategy to compile the list of putative direct targets of miR-124, and also to conduct bidirectional experiments to identify targets of other selected mouse miRNAs (the selection of miRNAs for these experiments is described in Chapter 3, section 3.3).

Before the start of these experiments, it was not known at what stage in development of cultures the bidirectional perturbation experiments would work most efficiently to identify miRNA targets. This uncertainty was due to differences between the selected miRNAs (Chapter 3). For example, miR-124 is highly expressed throughout the development of cultures, and miR-434-3p is upregulated (they are referred to as neuronal miRNAs), while miR-143, miR-145 and miR-25 are downregulated (they are referred to as non-neuronal miRNAs). It was not certain, whether the developmental timepoint optimal for identification of targets of neuronal miRNAs would be the same as for non-neuronal (and vice versa). Therefore, miRNA mediated effects of three non-neuronal miRNAs (miR-143, miR-145 and also cel-miR-67, which is a non-mouse miRNA and, hence, non-neuronal in the mouse) and of miR-124 were surveyed across the three developmental timepoints (**3DIV**, **4DIV** and **6DIV**[1]). This survey was conducted using the unidirectional overexpression strategy, where cultures transfected with miRNA mimics were compared to mock transfected cultures.

[1]The 4DIV timepoint was selected because of its importance as a switch point in gene expression of developing cultures: after 4DIV the ratio of abundances of neuritic transcripts and of somatic transcripts was above 1, i.e. similar to that in mature neurons (Chapter 3, section 3.1.3). The other two timepoints were picked around the 4DIV timepoint.

The first part of this chapter describes these unidirectional experiments, which resulted in the identification of the timepoints optimal for bidirectional perturbation experiments of neuronal and non-neuronal miRNAs. Additionally, results of these experiments suggested that the endogenous miRNA, miR-124, is a buffer of changes to the transcriptome of mature neurons, as discussed at the end of the first part of the chapter. The second part describes the bidirectional perturbation experiments, which resulted in lists of putative direct targets of several mouse miRNAs. Compilation of putative direct targets of cel-miR-67 from a unidirectional experiment will also be described. In the final part of the chapter, the methodology for identification of miRNA targets in this thesis is validated by comparing of miR-124 thesis targets to miR-124 targets previously reported in published literature.

5.1 Unidirectional overexpression experiments

5.1.1 The effect of cel-miR-67, miR-143 and miR-145 overexpression was maximal at 3DIV or 4DIV

Developmentally downregulated miRNAs were not expected to be highly expressed and be functional in mature neurons (Chapter 3, section 3.3). I selected two of these miRNAs, miR-143 and miR-145, for a series of unidirectional transfection experiments that are described in this section. In addition to these two mouse miRNAs, transfections of the mimic for a non-mouse miRNA, cel-miR-67, were also performed. This miRNA was identified in *Caenorhabditis elegans* and its seed region, *CACAACC*, was different from the seed region of any known mouse miRNA (as of miRBase release 14 (Griffiths-Jones et al., 2008, 2006; Griffiths-Jones, 2004)). Therefore, no specific category of mouse genes was expected to be under endogenous regulation of this miRNA.

Although these three miRNAs were unlikely to be involved in the regulation of normal neuronal function, their ectopic expression could still have an effect on neuronal gene expression. This is because the function of miRNAs is dependent on the generic RISC machinery (Introduction, section 1.1), and, in work by Lim and colleagues, the ectopic expression of a neuronal miR-124 and a muscular miR-1 in HeLa cells had a profound miRNA mediated effect on gene expression (Lim et al., 2005). Therefore it was not surprising that overexpression of cel-miR-67, miR-143 and miR-145 had an effect on gene expression in primary neuronal cultures, too (Figure 5.1).

Primary cultures were transfected with mimics of cel-miR-67 and miR-145 at 3DIV, 4DIV and 6DIV, and with mimics of miR-143 at 3DIV and 6DIV (Methods, section 2.5). Transfections of cel-miR-67 at 4DIV and 6DIV were independently repeated twice (specified by "A" and "B" indices). As these experiments were performed according to the unidirectional overexpression strategy, differential expression was estimated by comparison (contrast) of mRNA profiles of cultures transfected with miRNA mimics to those of mock transfected cultures. Analysis of the microarray data (Methods, section 2.7) revealed pronounced changes in mRNA profiles of cultures transfected with mimics of all three miRNAs (Figure 5.1). In each of the experiments downregulated genes ($P < 0.05$) were significantly enriched (hypergeometric test $P < 0.05$) in genes harbouring miRNA seed matching sites in their 3′UTRs (seed matching sites of 7(2) and 7(1A)-types, see Introduction, section 1.2.1 and Methods, section 2.7).

A trend toward decreasing impact of miRNA overexpression on gene expression was observed in these experiments. For all three miRNAs, the number of highly significantly downregulated genes decreased toward the 6DIV (e.g. numbers of genes downregulated beyond a strict cutoff $P < 1e-06$ are shown in Table 5.1). Additionally, the most significant enrichment of seed matching sites containing genes was achieved in experiments at 3DIV or 4DIV for all three miRNAs (Figure 5.1). The trend toward decreasing miRNA impact at 6DIV was especially clear when the seed matching site distribution was assessed using the Sylamer program (van Dongen et al., 2008), which accounted for length and composition biases (Figure 5.2, description of the Sylamer method is in Methods, section 2.8).

miRNA	**3DIV** $\boldsymbol{P < 1e-06}$	**4DIV** $\boldsymbol{P < 1e-06}$	**6DIV** $\boldsymbol{P < 1e-06}$
cel-miR-67	7	61 *(exp A)* 16 *(exp B)*	1 *(exp A)* 0 *(exp B)*
miR-143	40	*na*	0
miR-145	422	37	0

Table 5.1: Number of genes downregulated ($\boldsymbol{P < 1e-06}$) in miRNA overexpression experiments.

Another finding was the observation of significant biases in the distribution of miR-124 seed matching sites (7(2) and 7(1A)-types) in these experiments. Interestingly, the direction of these biases was reciprocal between 3DIV and 6DIV. At 3DIV transcripts that were upregulated in samples transfected with the mimics of miR-143, miR-145 and cel-miR-67 were depleted of seed matching sites for miR-124 (Figures 5.2a, 5.2b and 5.2c). In

transfections at 6DIV, on the other hand, the upregulated transcripts were enriched in the seed matching sites for miR-124 (Figures 5.2g, 5.2h, 5.2i and 5.2j). Interestingly, a similar trend was observed in gene expression profiles of the development of primary forebrain and hippocampal cultures (Chapter 3, section 3.2.3). There, transcripts that were upregulated early in development (in transition from 1DIV to 2DIV), were depleted of seed matching sites for miR-124 (Chapter 3, Figures 3.10a and 3.10b). On the other hand, transcripts that were upregulated later (in transition from 4DIV to 8DIV), were not depleted in the forebrain cultures (Chapter 3, Figure 3.10c), or enriched in the hippocampal cultures (Chapter 3, Figure 3.10d) for miR-124 seed matching sites.

In summary, the miRNA mediated effects of non-neuronal miRNAs, miR-143, miR-145 and cel-miR-67, were the strongest in transfection experiments at 3DIV or 4DIV. At 6DIV overexpression of these miRNAs had a lower direct effect on gene expression in the cultures. Unexpectedly, an inverse significant bias in the distribution of seed matching sites for miR-124 was observed.

Of the two timepoints (3DIV and 4DIV), the 4DIV timepoint was selected as optimal for the bidirectional experiments (section 5.2). This timepoint was preferred because the maximal seed matching site Sylamer enrichment was detected at 4DIV for miRNAs that were overexpressed at both 3DIV and 4DIV (Figures 5.2d and 5.2f).

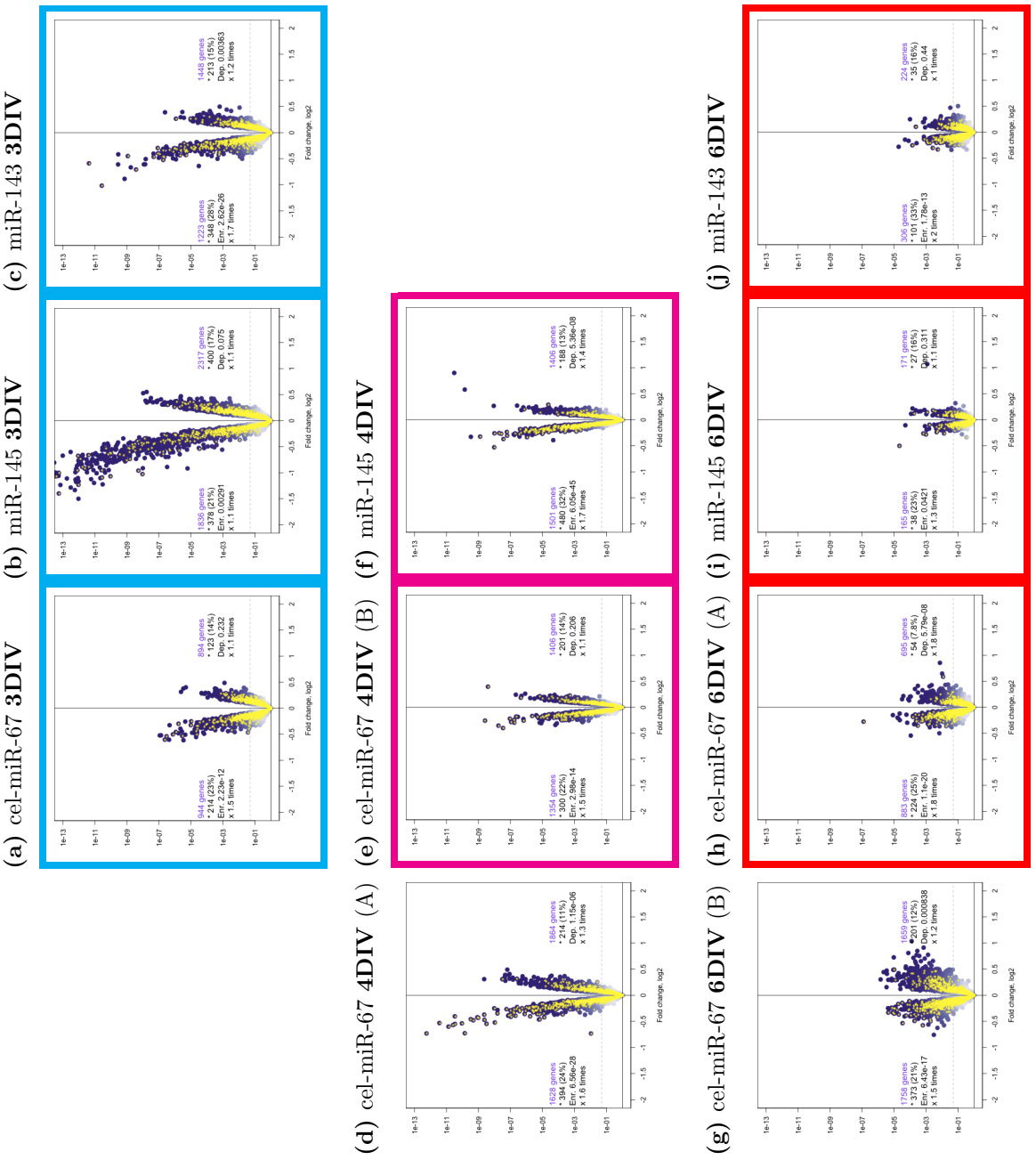


Figure 5.1: Differential gene expression upon unidirectional over-expression of miR-143, miR-145 and cel-miR-67. *[See the next page for caption]*

Figure 5.1: Differential gene expression and seed enrichment upon unidirectional overexpression of miR-143, miR-145 and cel-miR-67. *[The figure is on the previous page]*
Genes detected by microarrays (using the standard Illumina detection call $P < 0.01$) are shown as the purple dots (the analysis of microarray data is described in Methods, section 2.7). The x-axes represent log_2 of gene expression fold change between samples transfected with miRNA mimics (for the miRNAs named in the titles to the subfigures) in comparison to the matched mock transfected samples. The y-axes represent P-value of differential expression (log_{10} scale), and the horizontal dashed grey lines show P-value cutoff of 0.05. The yellow asterisks mark genes with 3′UTRs harbouring one or more seed matching sites (7(2) or (7(1A)-types) for the miRNAs named in the titles to the subfigures. The text in the two halves of the plot area provides the following information: 1) The total number of genes with differential expression P-value more significant than the cutoff (0.05); 2) The total number (and percentage) of genes with seed matching sites for miR-124; 3) The hypergeometric P-value of enrichment (*Enr.*) or depletion (*Dep.*); 4) Fold enrichment or depletion of genes with the seed matching sites "× *times*" the number that is expected by chance alone. The mapping of microarray probes to mRNA transcripts, and transcripts to genes, is described in Methods (section 2.7). The identification of the seed matching sites and the hypergeometric enrichment test is in Methods (section 2.8). The subfigures surrounded by the boxes of the same color describe experiments that were performed on the same batch of primary cultures (Methods, section 2.5).

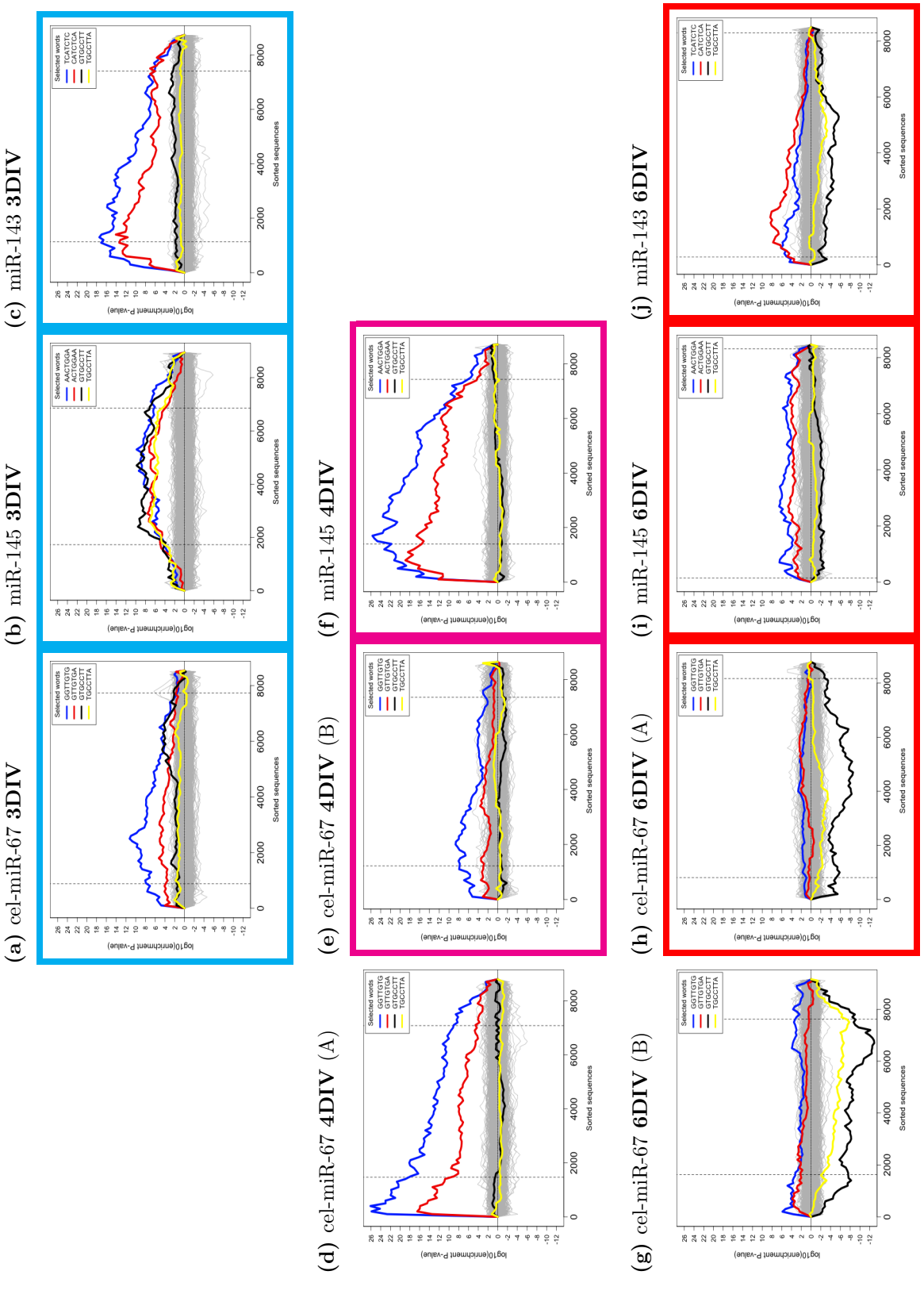


Figure 5.2: Sylamer analysis of biases in distribution of seed matching sites upon unidirectional over-expression of miR-143, miR-145 and cel-miR-67.
[See the next page for caption]

Figure 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. *[The figure is on the previous page]* The x-axes represent sorted 3′UTRs corresponding to detectably expressed genes (detected with the standard Illumina detection call $P < 0.01$, see Methods, section 2.7). These genes are ordered **from the most downregulated to the most upregulated** by fold change t-statistic for differential expression between samples transfected with miRNA mimics (for the miRNAs named in the titles to the subfigures) in comparison to the matched mock transfected samples. The y-axes represent the hypergeometric P-values for occurrence biases of 876 nucleotide words complementary to the seed regions (7(2) and 7(1A)-types) of the complete set of 581 distinct mouse miRNAs, according to miRBase Release 14 (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006, 2008). Positive values on the y-axes correspond to an enrichment ($+|log_{10}$(P-value)|) and negative values to a depletion ($-|log_{10}$(P-value)|). The vertical dashed lines mark the P-value cutoff (0.05) on both sides of the ranked gene lists. The blue and the red lines show enrichment profiles of 7(2) and 7(1A)-type seed matching sites for the miRNAs named in the titles to the subfigures, the black and the yellow lines – for miR-124, the grey lines – for the rest of the distinct seed matching sites. The mapping of microarray probes to mRNA transcripts, and transcripts to genes, is described in Methods (section 2.7). The identification of the seed regions and parameters of Sylamer (van Dongen et al., 2008) is in Methods (section 2.8). The full description of the Sylamer method is in the Introduction (section 2.8). The subfigures surrounded by the boxes of the same color describe experiments that were performed on the same batch of primary cultures (Methods, section 2.5).

5.1.2 The effect of miR-124 overexpression was maximal at 6DIV

Unlike the three miRNAs described in the previous section, miR-124 is known to have important functions in neuronal differentiation and development (Chapter 3, Table 3.3). Thus it was expected that its perturbation would have a significant impact on gene expression in primary cultures. It was not known, however, at which developmental timepoint perturbations of miR-124 would have the strongest impact.

In contrast to the outcome of the unidirectional overexpression of cel-miR-67, miR-143 and miR-145, the extent of gene expression changes increased dramatically between 3DIV and 6DIV in miR-124 overexpression experiments (Chapter 4, Figures 4.9a and 4.9c). At 3DIV, 847 genes were downregulated ($P < 0.05$) in contrast of cultures transfected with the mimic versus mock transfected cultures, while at 6DIV there were 2,070 downregulated ($P < 0.05$) genes. Additionally, the direct contribution of miR-124 to differential expression also appeared to have increased between 3DIV and 6DIV. Sylamer analysis of biases of miRNA seed matching site distributions in 3′UTRs showed that the peak of enrichment for miR-124 7(2)-type seed matching site increased between 3DIV and 6DIV from $P < 1e-14$ to $P < 1e-33$ (Chapter 4, Figures 4.8a and 4.8c). Therefore, the bidirectional perturbations at 6DIV were used to compile a list of putative miR-124 targets (see section 5.2.1). Additionally, based on the miR-124 effect at 6DIV as an example of an effect of a miRNA that is functional in neurons, the 6DIV timepoint was also selected for identification of targets of miRNAs upregulated in the development (section 5.2.3).

5.1.3 Endogenous miR-124 constrained gene expression in more mature primary neurons (6DIV)

As well as allowing identification of optimal timepoints for bidirectional experiments, the series of unidirectional perturbation experiments uncovered an interesting trend: the effect of overexpression of non-neuronal miRNAs decreased in more mature neurons (6DIV), while it increased toward 6DIV in the case of miR-124 overexpression experiments. Importantly, the increase in impact in miR-124 experiments demonstrated that the decrease in impact of non-neuronal miRNAs could not be explained by the decrease in transfection efficiency in more mature cultures.

Transfections of mimics of non-neuronal miRNAs (miR-143, miR-145 and cel-miR-67) had the biggest effect on gene expression in primary cultures at 3DIV and 4DIV, while it decreased at 6DIV. This effect is seen as shrinking of the plots of differential expression for all three non-neuronal miRNAs at 6DIV (Figure 5.1). Sylamer analysis

of seed matching site distributions in 3′UTRs of transcripts expressed in the cultures showed that enrichment of seed matching sites in 3′UTRs of downregulated transcripts also decreased at 6DIV (Figure 5.2). Therefore, both the overall effect of transfections and direct inhibition by the transfected non-neuronal miRNAs decreased in relatively more mature primary neuronal cultures (6DIV).

Interestingly, a reciprocal trend in the distribution of seed matching sites for miR-124 was observed in transfections of non-neuronal miRNAs at 3DIV and 6DIV. At 3DIV, transcripts that were upregulated upon the transfections of the mimics, were depleted of the seed matching sites for miR-124 (the black and yellow lines in Figures 5.2a, 5.2b and 5.2c). An inverse trend was observed at 6DIV, where upregulated transcripts were enriched in miR-124 seed matching sites (the black and yellow lines in Figures 5.2g, 5.2h, 5.2i and 5.2j). It is possible, that endogenous miR-124 imposes a limit to transcriptome changes in more mature neurons (6DIV) by moderating the extent of upregulation of the transcripts that harbour miR-124 seed matching sites. On the other hand, changes to the transcriptome in less mature neurons (3DIV and 4DIV), are relatively less restricted, because upregulated transcripts are neither enriched nor significantly depleted of the seed matching sites for miR-124.

Profiling of differential expression in developing primary neuronal cultures suggested an explanation for an increased capacity of miR-124 to moderate (or buffer) changes to the transcriptome in more mature neurons. Transcripts that were upregulated early in development of primary cultures (in transition from 1DIV to 2DIV) were seen to be depleted of miR-124 seed matching sites (Chapter 3, Figures 3.10a and 3.10b). Therefore, in very immature primary neurons transcripts that can in principle be inhibited by miR-124 (i.e. transcripts with seed matching sites for miR-124) may not yet be available. The relatively small scope for miR-124 activity in very immature neurons make sense biologically in the light of a rapid spurt of neurites and early synaptogenesis events that were observed in the early stages of development of primary neuronal cultures (Valor et al., 2007). Perhaps at these early stages it would be detrimental for neurons if limits to the changes in the transcriptome were imposed by highly expressed neuronal miRNAs, such as miR-124.

At later stages in development of cultures (the transition from 4DIV to 8DIV), 3′UTRs of developmentally upregulated transcripts were seen to be not depleted (primary forebrain cultures, Chapter 3, Figure 3.10c) or even enriched (primary hippocampal cultures, Chapter 3, Figure 3.10d) in seed matching sites for miR-124. Therefore, in more mature neurons there is scope for endogenous miR-124 to limit changes to the transcriptome,

which can explain the decrease in impact of transfections of non-neuronal miRNAs at 6DIV. Similarly, the increase in the impact of transfections of miR-124 itself at 6DIV can also be explained by the increased number of transcripts available for miR-124 mediated inhibition in more mature cultures.

It should be pointed out, that the hypothesis of miR-124 to buffer upregulated in mature neurons genes and by that restrict perturbations of the transcriptome, is speculative. However, this hypothesis is convenient as an effective theory at the moment, because it explains some of the observations that are described in Chapter 6. Experiments that can test this hypothesis are suggested in the Discussion (section 7.4).

Summary of section 5.1

Transfections of mimics of non-neuronal miRNAs (miR-143 and miR-145, cel-miR-67) in primary forebrain cultures at 3DIV, 4DIV and 6DIV elicited changes in gene expression, which were the most significant at 3DIV or 4DIV. Sylamer analysis suggested that the direct contribution of the transfected miRNAs to changes in differential gene expression was also greatest at 3DIV or 4DIV. In the end, the 4DIV timepoint was selected as the best timepoint to conduct bidirectional perturbation experiments aiming to identify putatively direct targets of non-neuronal miRNAs.

Transfections of mimics of the neuronal miR-124 at 3DIV and 6DIV elicited changes in gene expression at both of the timepoints. Sylamer analysis showed that the direct contribution of miR-124 mediated inhibition to differential gene expression increased from 3DIV to 6DIV. Therefore, the 6DIV timepoint was selected as the best timepoint for bidirectional experiments on neuronal miRNAs.

Additionally, significant biases in the distribution of miR-124 seed matching sites where observed in the experiments where non-neuronal miRNAs were exogenously added (transfected) into the primary neuronal cultures. The direction of these biases in the 6DIV experiments (an enrichment of miR-124 seed matching sites in the upregulated transcripts), in conjunction with the overall decrease in the effect of 6DIV transfections on gene expression, suggested that the endogenous miR-124 can act as a buffer for genes that are upregulated in more mature neurons (6DIV).

5.2 Bidirectional perturbation experiments

5.2.1 Identification of targets for a steady state expressed miRNA (miR-124)

Of miRNAs that were expressed at the steady state level in the development of primary neuronal cultures, I selected two for bidirectional transfection experiments[1]: miR-124 and miR-103 (Chapter 3, section 3.3). Bidirectional transfection experiments with miR-124 were conducted at all three experimental timepoints (3DIV, 4DIV and 6DIV), because of the special interest in this miRNA (Introduction, section 1.1.3). On the other hand, miR-103, although it was highly expressed, was not previously reported as a functionally important miRNA for neurons. Therefore, the miR-103 transfection experiment was performed at 4DIV, as this timepoint was conducive to detection of miRNA mediated effects in both miR-124 transfection (see below) and transfections of non-neuronal miRNAs (see section 5.2.2).

Both the bidirectional and unidirectional contrasts were available for analysis of miR-124 and miR-103 transfections (Chapter 4, section 4.2.1). In experiments on miR-124 at 3DIV and 6DIV, where matched mock transfected samples were available (Methods, section 2.5), the use of inhibition increased detection of enrichment of miR-124 seed matching sites (Chapter 4, Figure 4.8). In the miR-103 experiment both the mock transfection and the inhibition were conducted, and the bidirectional strategy increased enrichment of genes with miR-103 seed matching sites among downregulated genes (from approximately 1.1 times ($P < 0.039$) to 1.3 times ($P < 0.017$) more than expected by chance alone, Figure 5.3).

Despite the fact that the bidirectional strategy increased enrichment of genes with miRNA seed matching sites among downregulated genes for both miR-124 and miR-103, Sylamer analysis (Methods, section 2.8) showed that only in case of miR-124 the enrichment of seed matching sites was more significant than the background distribution (Figure 5.4). Therefore it was possible to compile putative direct targets for miR-124, but not for miR-103. Of the three miR-124 perturbation experiments, the biggest enrichment of miR-124 seed matching sites in 3′UTRs of downregulated genes was observed at 6DIV (Figure 5.4c). Therefore, putative direct targets of miR-124 were derived from the 6DIV experiments. The cutoff P-value of 0.01 in differential expression coincided well with the

[1] As introduced in Chapter 4 (section 4.2.1), the word **bidirectional** refers to comparison of transfections of miRNA mimics with transfections of miRNA inhibitors. On the other hand, the word **unidirectional** refers to comparison of transfections of miRNA mimics with mock transfections.

(a) miR-103 *overexpression v. mock* **(b)** miR-103 *overexpression v. inhib.*

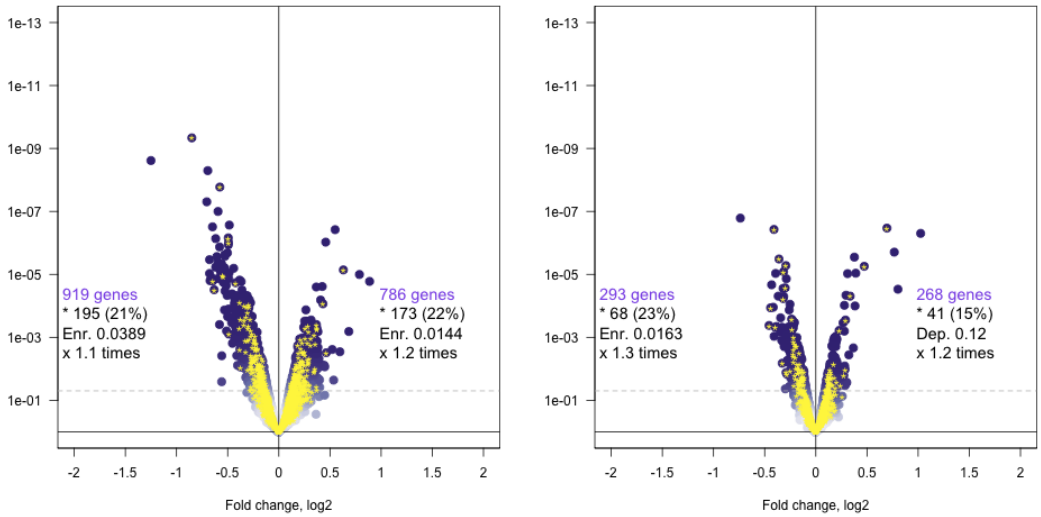


Figure 5.3: Differential gene expression and seed matching site enrichment in miR-103 transfection experiments at 4DIV.
Genes detected by microarrays (using the standard Illumina detection call $P < 0.01$) are shown as the purple dots (analysis of microarray data is described in Methods, section 2.7). The x-axes represent log_2 of gene expression fold change between samples transfected: 5.3a – with the mimic of miR-103 in comparison to the matched mock transfection at 4DIV; 5.3b – with the mimic of miR-103 in comparison to transfection with the inhibitor of miR-103 at 4DIV; The y-axes represent P-value of differential expression (log_{10} scale), and the horizontal dashed grey lines show P-value cutoff of 0.05. The yellow asterisks mark genes [encoding transcripts] with 3′UTRs harbouring one or more seed matching sites (7(2) or 7(1A)-types) for miR-103. The text in the two halves of the plot area provides the following information: 1) The total number of genes with differential expression P-value more significant than the cutoff (0.05); 2) The total number (and percentage) of genes with seed matching sites for miR-103; 3) The hypergeometric P-value of enrichment (*Enr.*) or depletion (*Dep.*); 4) Fold enrichment or depletion of genes with the seed matching sites "× *times*" the number that is expected by chance alone. The mapping of microarray probes to mRNA transcripts, and transcripts to genes, is described in Methods (section 2.7). The identification of the seed matching sites and the hypergeometric enrichment test is in Methods (section 2.8).

slope and the peak of the Sylamer distribution of miR-124 seed matching sites (both 7(2) and 7(1A)-types, see blue and red line and the first vertical dashed line in Figure 5.4c). Overall, there were 399 genes that contained one or more miR-124 seed matching sites in their 3′UTRs and were downregulated beyond the $P < 0.01$ cutoff. These putative direct miR-124 targets are listed in Supplementary Data, Table A.9.

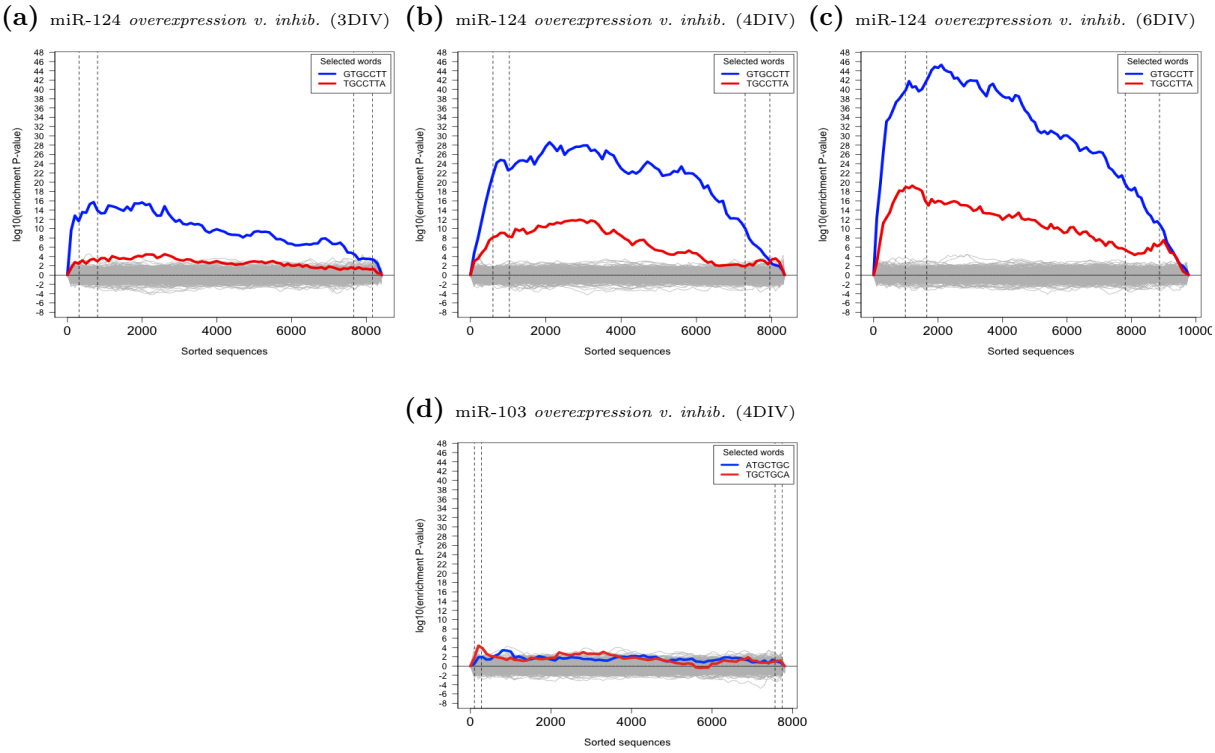


Figure 5.4: Sylamer analysis of biases in distributions of seed enrichment in miR-124 and miR-103 transfection experiments.
The x-axes represent sorted 3′UTRs corresponding to detectably expressed genes (detected with the standard Illumina detection call $P < 0.01$, see Methods, section 2.7). These genes are ordered **from the most downregulated to the most upregulated** by fold change t-statistic for differential expression in the following transfections: 5.4a – with the mimic of miR-124 in comparison to the transfection with the inhibitor of miR-124 at 3DIV; 5.4b – with the mimic of miR-124 in comparison to the transfection with the inhibitor of miR-124 at 4DIV; 5.4c – with the mimic of miR-124 in comparison to the transfection with the inhibitor of miR-124 at 6DIV; 5.4d – with the mimic of miR-103 in comparison to the transfection with the inhibitor of miR-103 at 4DIV. The y-axes represent the hypergeometric P-values for occurrence biases of 876 nucleotide words complementary to the seed regions (7(2) and 7(1A)-types) of the complete set of 581 distinct mouse miRNAs, according to miRBase Release 14 (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006, 2008). Positive values on the y-axes correspond to an enrichment ($+|log_{10}$(P-value)|) and negative values to a depletion ($-|log_{10}$(P-value)|). The vertical dashed lines mark the P-value cutoffs (0.01 and 0.05) on both sides of the ranked gene lists. The blue and the red lines show enrichment profiles of 7(2) and 7(1A)-type seed matching sites for miR-124 or miR-103 (see the titles for the subfigures), the grey lines – for the rest of the distinct seed matching sites. The mapping of microarray probes to mRNA transcripts, and transcripts to genes, is described in Methods (section 2.7). The identification of the seed regions and parameters of Sylamer (van Dongen et al., 2008) is in Methods (section 2.8). The full description of the Sylamer method is in the Introduction (section 2.8).

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)

Of miRNAs that were downregulated in the development of primary neuronal cultures, I selected three for bidirectional transfection experiments: miR-143, miR-145 and miR-25 (the selection is described in Chapter 3, section 3.3). These experiments were performed on a single batch of cultures (Methods, section 2.5) at 4DIV. The 4DIV timepoint was selected as results of unidirectional overexpression experiments (see section 5.1.1) suggested that detection of targets of putatively non-neuronal miRNAs was the most efficient at 4DIV (Figure 5.2). In addition to the compilation of lists of putative direct targets of miR-143, miR-145 and miR-25, identification of putative direct targets of cel-miR-67 (a non-mouse miRNA) is also described at the end of this section.

By conducting a mock transfection, in addition to overexpression and inhibition, it was possible to demonstrate that the use of bidirectional contrasts improved detection of miRNA-mediated effects for all three of the downregulated miRNAs (miR-143 (Figures 5.5a and 5.5d), miR-145 (Figures 5.5b and 5.5e) and miR-25 (Figures 5.5c and 5.5f)). Upon bidirectional perturbation of miR-143, the enrichment of the genes with the seed matching sites complementary to miR-143 increased among downregulated genes (differential expression $P < 0.05$) from approximately equal to that expected by chance alone to 1.1 times more than expected ($P < 0.05$). This increase was also true for miR-145 (1.5 fold enrichment in contrast with mock transfection and 1.6 times in contrast with inhibition), and for miR-25 (increasing from 2.4 times to 2.7 times).

Sylamer analysis (Methods, section 2.8) showed that bidirectional perturbation of miR-143, miR-145 and miR-25, leads to identification of significant miRNA-mediated effects in each of the experiments (Figure 5.6). In all experiments, the enrichment of seed matching sites (7(2) and 7(1A)-types) for the transfected miRNAs was more significant than that of all other tested nucleotide words (shown by the grey lines, Figure 5.6). However, in miR-143 and miR-145 experiments the peak of enrichment of the seed matching sites for several other nucleotide words approached that of the seed matching sites for miR-145 and miR-143. These nucleotide words were either of low complexity (e.g. *GCCCCGG* in the case of miR-143 experiment), were related to the polyadenylation site, or corresponded to miRNAs with low abundance and unknown function. Perhaps these distributions were indicative of interesting biological phenomena, however they were unlikely to be

directly related to activity of the transfected miRNAs[1], and therefore they were not studied further.

Interestingly, Sylamer enrichment of miR-25 seed matching sites was very significant (Figure 5.6c). The Sylamer enrichment peak for 7(2)-type seed matching site for miR-25 ($GTGCAAT$) corresponded to a hypergeometric enrichment P-value below $1e-52$, which was the most significant Sylamer enrichment P-value detected in this thesis project. Surprisingly, the effect of miR-25 on expression of transcripts harbouring seed matching sites for miR-25 was even more dramatic than that of miR-124 ($P < 1e-44$ in transfection at 6DIV, see Figure 5.4c). The extremely high enrichment of seed matching sites for miR-25 was likely a combination of high efficiency of miR-25 as a guide of the RISC to destabilise its targets (Introduction, section 1.1.2), and a relatively small repertoire of its potential targets[2]. It should be noted that in transfection experiments performed in this thesis, there were substantial fluctuations in Sylamer enrichment results[3]. Therefore the result of one miR-25 transfection experiment cannot serve as definitive evidence of its exceptional properties. However, transfection of miR-25 into mutant mouse embryonic stem cells was found to lead to a similarly significant destabilisation of seed matching site containing mRNAs (Matthew Davis, personal communication). Interestingly, expression of miR-25 was shown to be induced in tumors (Poliseno et al., 2010), which is a relatively unusual for miRNAs property (as miRNAs are generally reduced in tumors (Thomson et al., 2006; Lotterman et al., 2008)). Therefore, a strong inhibition of miR-25 targets in differentiated cell types, such as cells in primary neuronal cultures, may be related to the role of miRNAs in carcinogenesis. This makes the role of miR-25 in primary neuronal cultures a relevant subject for research of miRNAs in cancer (Discussion, section 7.4).

Identification of significant enrichment for miR-143, miR-145 and miR-25 seed matching sites in downregulated genes enabled the compilation of lists of their putative direct targets. As in the case of miR-124 (see section 5.2.1), the 0.01 P-value cutoff of differential expression corresponded well to the peaks in the enrichment of seed matching sites for

[1]It should be noted that these enrichments peaked in the right half of the plot (i.e. the dominant trend was depletion of these seed matching sites in the 3′UTRs of upregulated genes), whilst in case of direct miRNA mediated effects on gene expression the peak is expected to be in the left half (i.e. consistent with a enrichment of seed matching sites in the 3′UTRs of downregulated genes being a dominant trend). For interpretation of Sylamer plots see Introduction, section 1.2.3.

[2]Of genes represented on the microarray platform there were 2,083 that uniquely corresponded to transcripts harbouring one or more seed matching sites for miR-25, while this number is over 3,000 for miR-143 and miR-145.

[3]For example, the Sylamer enrichment P-values of enrichment of the seed matching sites for cel-miR-67 differed by more than 10 orders of magnitude in replicate transfections at 4DIV, see Figures 5.2d and 5.2e.

transfected miRNAs (the first dashed vertical lines in Figure 5.6). Selection of downregulated genes (differential expression $P < 0.01$), harbouring seed matching sites for the transfected miRNAs in their 3′UTRs, identified 272 putative direct targets of miR-143, 358 targets of miR-145 and 196 targets of miR-25. These putative direct miRNA targets are listed in Supplementary Data, Table A.10 (miR-143 targets), Table A.11 (miR-145 targets) and Table A.12 (miR-25 targets).

Transfections of cel-miR-67 (naturally expressed in *Caenorhabditis elegans*) were only possible using the unidirectional overexpression design as its inhibition was not available in mouse cells. The maximal enrichment of cel-miR-67 seed matching sites in 3′UTRs of downregulated in the transfections genes was detected in one of the experiments at 4DIV (the experiment marked by the index "A", see Figures 5.1d and 5.2d). Therefore this experiment was used to compile putative direct targets of cel-miR-67 in mouse primary neuronal cultures. A P-value cutoff of 0.05 was used[4] and it led to compilation of a list of 394 putative direct cel-miR-67 targets. These putative direct targets of cel-miR-67 in mouse primary neurons are listed in Supplementary Data, Table A.13.

[4]The 0.05 cutoff, rather than 0.01, was used for compilation of target of cel-miR-67, because the design of cel-miR-67 transfection experiments on cel-miR-67 was not optimised for detection of direct miRNA-mediated effects and it was performed according to the unidirectional strategy. For two out of three unidirectional transfections of the mimics of non-neuronal miRNAs at 4DIV the 0.05 cutoff approximately coincided with the Sylamer enrichment peak (Figure 5.2).

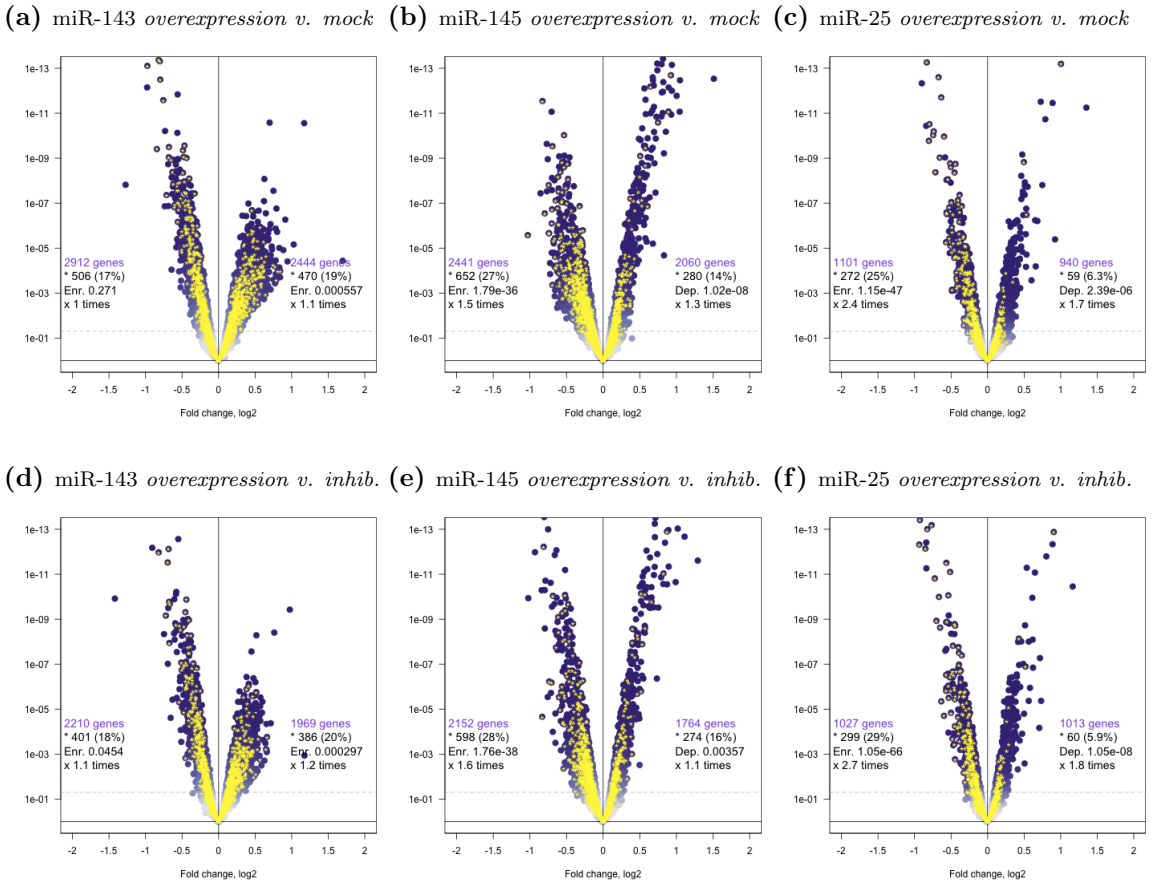


Figure 5.5: Differential gene expression and seed matching site enrichment in miR-143, miR-145 and miR-25 transfections experiments at 4DIV.
Genes detected by microarrays (using the standard Illumina detection call $P < 0.01$) are shown as the purple dots (analysis of microarray data is described in Methods, section 2.7). The x-axes represent log_2 of gene expression fold change between samples transfected: 5.5a – with the mimic of miR-143 in comparison to the matched mock transfection at 4DIV; 5.5b – with the mimic of miR-145 in comparison to the matched mock transfection at 4DIV; 5.5c – with the mimic of miR-25 in comparison to the matched mock transfection at 4DIV; 5.5d – with the mimic of miR-143 in comparison to the transfection with the inhibitor of miR-143 at 4DIV; 5.5e – with the mimic of miR-145 in comparison to the transfection with the inhibitor of miR-145 at 4DIV; 5.5f – with the mimic of miR-25 in comparison to the transfection with the inhibitor of miR-25 at 4DIV. The y-axes represent P-value of differential expression (log_{10} scale), and the horizontal dashed grey lines show P-value cutoff of 0.05. The yellow asterisks mark genes [encoding transcripts] with 3′UTRs harbouring one or more seed matching sites (7(2) or 7(1A)-types) for the miRNAs in the titles of the subfigures. The text in the two halves of the plot area provides the following information: 1) The total number of genes with differential expression P-value more significant than the cutoff (0.05); 2) The total number (and percentage) of genes with seed matching sites for the miRNAs in the titles to the subfigures; 3) The hypergeometric P-value of enrichment (*Enr.*) or depletion (*Dep.*); 4) Fold enrichment or depletion of genes with the seed matching sites "× *times*" the number that is expected by chance alone. The mapping of microarray probes to mRNA transcripts, and transcripts to genes, is described in Methods (section 2.7); the identification of seed matching sites and the hypergeometric enrichment test is in Methods (section 2.8).

(a) miR-143 *overexpression v. inhib.* **(b)** miR-145 *overexpression v. inhib.* **(c)** miR-25 *overexpression v. inhib.*

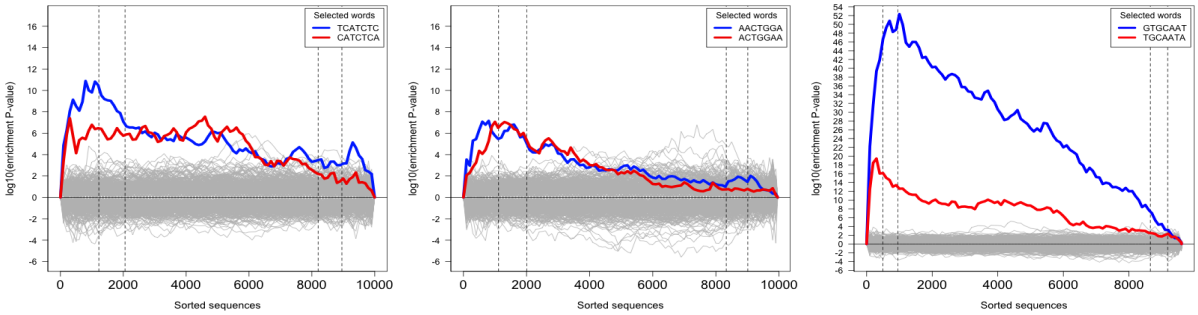


Figure 5.6: Sylamer analysis of biases in distributions of seed matching sites in miR-143, miR-145 and miR-25 transfection experiments at 4DIV.
The x-axes represent sorted 3′UTRs corresponding to detectably expressed genes (detected with the standard Illumina detection call $P < 0.01$, see Methods, section 2.7). These genes are ordered **from the most downregulated to the most upregulated** by fold change t-statistic for differential expression in the following transfections: 5.6a – with the mimic of miR-143 in comparison to the transfection with the inhibitor of miR-143 at 4DIV; 5.6b – with the mimic of miR-145 in comparison to the transfection with the inhibitor of miR-145 at 4DIV; 5.6c – with the mimic of miR-25 in comparison to the transfection with the inhibitor of miR-25 at 4DIV. The y-axes represent the hypergeometric P-values for occurrence biases of 876 nucleotide words complementary to the seed regions (7(2) and 7(1A)-types) of the complete set of 581 distinct mouse miRNAs, according to miRBase Release 14 (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006, 2008). Positive values on the y-axes correspond to an enrichment ($+|log_{10}$(P-value)|) and negative values to a depletion ($-|log_{10}$(P-value)|). The vertical dashed lines mark the P-value cutoffs (0.01 and 0.05) on both sides of the ranked gene lists. The blue and the red lines show enrichment profiles of 7(2) and 7(1A)-type seed matching sites for either of miR-143, miR-145 or miR-25 (see the titles of the subfigures), the grey lines – for the rest of the distinct seed matching sites. The mapping of microarray probes to mRNA transcripts, and transcripts to genes, is described in Methods (section 2.7). The identification of the seed regions and parameters of Sylamer (van Dongen et al., 2008) is in Methods (section 2.8). The full description of the Sylamer method is in the Introduction (section 2.8).

5.2.3 Identification of targets for an upregulated miRNA in development (miR-434-3p)

Of miRNAs that were upregulated in the development of primary cultures, I selected four for bidirectional transfection experiments. These miRNAs were miR-370, miR-410, miR-551b and miR-434-3p (the selection is described in Chapter 3, section 3.3). Unlike miRNAs of the downregulated category and the non-mouse cel-miR-67, the upregulated miRNAs were assumed to be functional in neurons (Chapter 3, section 3.3). Bidirectional perturbations at 6DIV worked best for induction of miRNA mediated effects in miR-124 transfection experiments (Figure 5.1.2), which has several known functions in neurons (Introduction, section 1.1.3). Therefore transfection experiments to identify targets of upregulated miRNAs were performed according to the bidirectional strategy at 6DIV. Additionally, levels of expression of the upregulated miRNAs were increasing during development (Chapter 3, section 3.2.1), therefore it was reasoned that inhibition of these miRNAs may have a greater effect at the later timepoint of 6DIV.

Bidirectional perturbation of three miRNAs (miR-370, miR-410 and miR-551b) was conducted in one batch of cultures, while the miR-434-3p experiment was performed separately (Methods, section 2.5). By carrying out the mock transfection as a part of the miR-434-3p experiment, it was possible to conclude that the bidirectional perturbation strategy had worked successfully to improve detection of direct miRNA mediated effects. Enrichment of transcripts with miR-434-3p seed matching sites among all downregulated transcripts ($P < 0.05$) was approximately 1.2 fold more than expected ($P < 1.48e-06$) in case of the unidirectional contrast, and 1.7 fold more than expected ($P < 3.1e-15$) in case of the bidirectional contrast (Figure 5.7).

Sylamer analysis (Methods, section 2.8), showed that bidirectional perturbation of only miR-434-3p produced a significant miRNA-mediated effect on gene expression (Figure 5.8). Significant enrichment for miR-434-3p seed matching sites in 3′UTRs of downregulated genes allowed the compilation of a list of its putative direct targets. The differential expression cutoff was again chosen to be $P < 0.01$, as it corresponded well to the Sylamer peak of the enrichment of the 7(2)-type seed matching site complementary to the seed region of miR-434-3p (the first dashed vertical line in Figure 5.8a). There were 112 genes downregulated beyond the cutoff ($P < 0.01$) and containing one or more seed matching sites (7(2) or 7(1A)-types) for miR-434-3p, which comprised the list of its putative direct targets. The putative direct targets of miR-434-3p are listed in Supplementary Data, Table A.14.

(a) miR-434-3p *overexpression v. mock* **(b)** miR-434-3p *overexpression v. inh*

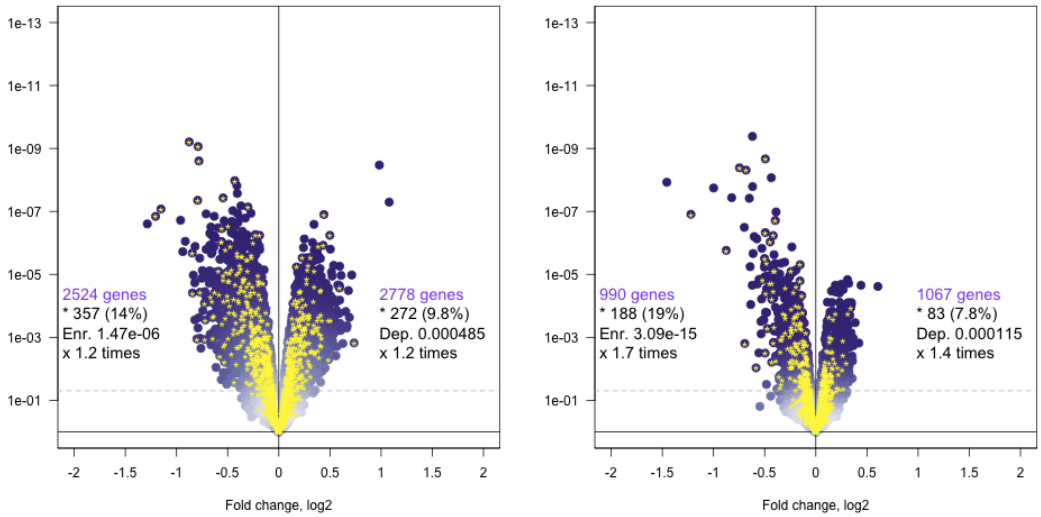
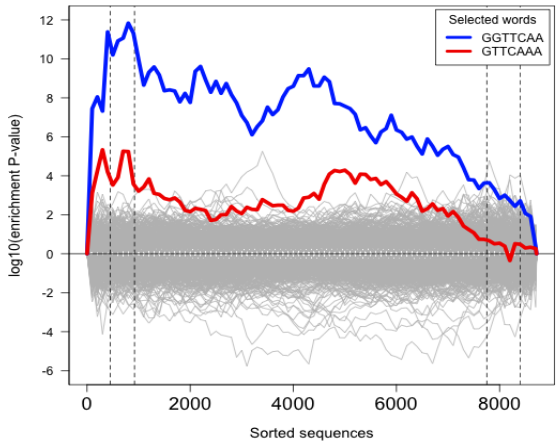


Figure 5.7: Differential gene expression and seed matching site enrichment in miR-434-3p transfection experiments at 6DIV.
Genes detected by microarrays (using the standard Illumina detection call $P < 0.01$) are shown as the purple dots (the analysis of microarray data is described in Methods, section 2.7). The x-axes represent log_2 of gene expression fold change between samples transfected: 5.7a – with the mimic of miR-434-3p in comparison to the matched mock transfection at 6DIV; 5.7b – with the mimic of miR-434-3p in comparison to the transfection with the inhibitor of miR-434-3p at 6DIV. The y-axes represent P-value of differential expression (log_{10} scale), and the horizontal dashed grey lines show P-value cutoff of 0.05. The yellow asterisks mark genes [encoding transcripts] with 3′UTRs harbouring one or more seed matching sites (7(2) or 7(1A)-types) for miR-434-3p. The text in the two halves of the plot area provides the following information: 1) The total number of genes with differential expression P-value more significant than the cutoff (0.05); 2) The total number (and percentage) of genes with seed matching sites for miR-434-3p; 3) The hypergeometric P-value of enrichment (*Enr.*) or depletion (*Dep.*); 4) Fold enrichment or depletion of genes with the seed matching sites "× *times*" the number that is expected by chance alone. The mapping of microarray probes to mRNA transcripts, and transcripts to genes, is described in Methods (section 2.7). The identification of the seed matching sites and the hypergeometric enrichment test is in Methods (section 2.8).

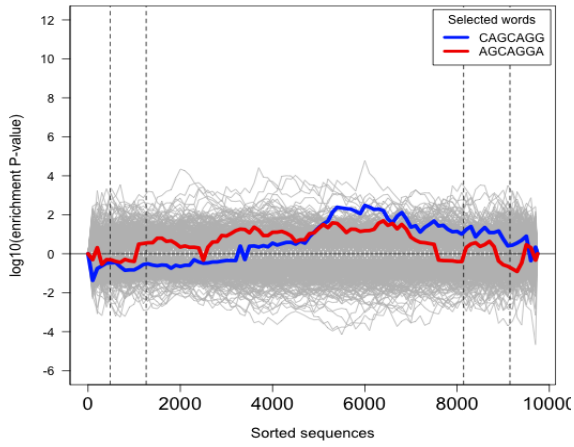
Summary of section 5.2

Using timepoints that were identified in unidirectional overexpression experiments (section 5.1) as best allowing identification of putative direct targets of miRNAs, the bidirectional transfection experiments were performed on the nine selected mouse miRNAs. As a result of these experiments, lists of putatively direct targets were compiled for miR-124, miR-434-3p, miR-143, miR-145 and miR-25. Additionally, using results of a unidirectional overexpression experiment, the list of putative direct targets was compiled for a non-mouse miRNA, cel-miR-67.

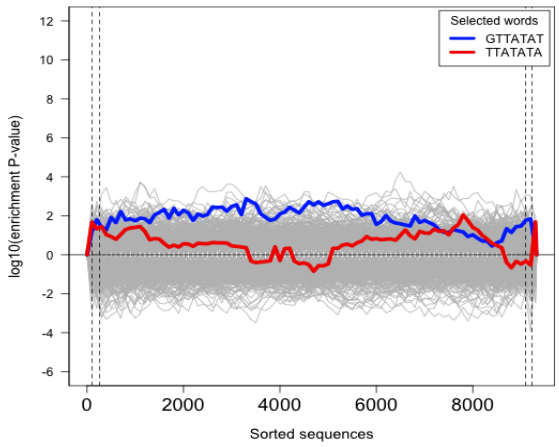
(a) miR-434-3p *overexpression v. inhib.*



(b) miR-370 *overexpression v. inhib.*



(c) miR-410 *overexpression v. inhib.*



(d) miR-551b *overexpression v. inhib.*

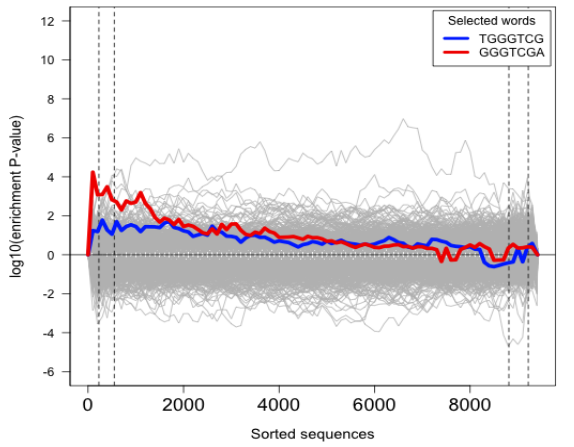


Figure 5.8: Sylamer analysis of distribution biases in distributions of seed matching sites in miR-434-3p, miR-370, miR-410 and miR-551b transfection experiments at 6DIV. The x-axes represent sorted 3′UTRs corresponding to detectably expressed genes (detected with the standard Illumina detection call $P < 0.01$, see Methods, section 2.7). These genes are ordered **from the most downregulated to the most upregulated** by fold change t-statistic for differential expression in the following transfections: 5.8a – with the mimic of miR-434-3p in comparison to the transfection with the inhibitor of miR-434-3p at 6DIV; 5.8b – with the mimic of miR-370 in comparison to the transfection with the inhibitor of miR-370 at 6DIV; 5.8c – with the mimic of miR-410 in comparison to the transfection with the inhibitor of miR-410 at 6DIV; 5.8d – with the mimic of miR-551b in comparison to the transfection with the inhibitor of miR-551b at 6DIV. The y-axes represent the hypergeometric P-values for occurrence biases of 876 nucleotide words complementary to the seed regions (7(2) and 7(1A)-types) of the complete set of 581 distinct mouse miRNAs, according to miRBase Release 14 (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006, 2008). Positive values on the y-axes correspond to an enrichment ($+|log_{10}$(P-value)|) and negative values to a depletion ($-|log_{10}$(P-value)|). The vertical dashed lines mark the P-value cutoffs (0.01 and 0.05) on both sides of the ranked gene lists. The blue and the red lines show enrichment profiles of 7(2) and 7(1A)-type seed matching sites for either of miR-434-3p, miR-370, miR-410 or miR-551b (see the titles of the subfigures), the grey lines – for the rest of the distinct seed matching sites. The mapping of microarray probes to mRNA transcripts, and transcripts to genes, is described in Methods (section 2.7). The identification of the seed regions and parameters of Sylamer (van Dongen et al., 2008) is in Methods (section 2.8). The full description of the Sylamer method is in the Introduction (section 2.8).

5.3 Validation of the methodology

The previous section described identification of putative direct targets for six miRNAs: miR-124, miR-434-3p, miR-25, miR-143, miR-145 and cel-miR-67. Targeting of miR-124 was previously studied and published before (Table 5.2). Therefore it was possible to validate methods to identify miRNA targets used in this thesis by comparison of miR-124 targets identified in the thesis to those previously published (or inferred from previously published data).

System	Detection	Reference
HeLa cell line	microarrays	(Lim et al., 2005)
HepG2 cell line	microarrays	(Wang and Wang, 2006)
293 cell line	microarrays	(Karginov et al., 2007)
CAD cell line	microarrays	(Makeyev et al., 2007)
HeLa cell line	protein mass-spectrometry	(Baek et al., 2008)
P13 mouse neocortex	RNA sequencing	(Chi et al., 2009)
HEK293T cell line	microarrays	(Hendrickson et al., 2009)

Table 5.2: Previously published genome-wide studies of miR-124 targeting

To date there are at least seven published experiments aiming at a genome wide description of miR-124 targeting at RNA and protein levels (Table 5.2). This section describes comparison of putative direct targets identified for miR-124 in this thesis (see Chapter 5, section 5.2.1) and targets identified in three of the previously published studies. These three studies were:

- Overexpression of miR-124 in HeLa cells published by **Lim *et al.*** in 2005 (Lim et al., 2005). This was the first study to characterize a global effect of miRNAs on the transcriptome, and multiple miR-124 targets from this study were subsequently experimentally validated elsewhere (Conaco et al., 2006).
- Overexpression of miR-124 in the mouse CAD cell line published by **Makeyev *et al.*** in 2007 (Makeyev et al., 2007). Comparison to the targets that could be derived from this work was more relevant to neurons than the HeLa targets, because CAD cells were originally obtained from a mouse brain tumor (Qi et al., 1997), while HeLa was derived from a human cervical carcinoma (Scherer et al., 1953).
- Identification of miR-124 targets with Argonaute HITS-CLIP method, published by **Chi *et al.*** in 2009 (Chi et al., 2009). Unlike the other two studies, in this study the cell cultures and transfections were not used. Instead, RNA was precipitated from the P13 neocortex and analyzed with new generation sequencing technology. Therefore miR-124 targets identified in this work provided an insight to what the

miR-124 targeting repertoire may be like in a relatively unperturbed and naïve brain (referred to as the *in vivo* targets).

5.3.1 Comparison to targets identified by Lim *et al.*

Lim et al. (Lim et al., 2005) described microarray analysis of gene expression changes upon delivery of miRNA mimics with cationic lipid transfection into HeLa cultures. The authors observed that two tissue specific miRNAs, miR-124 and miR-1, caused significant differential gene expression in HeLa cells. Downregulated genes were enriched for relevant miRNA seed matching sites in their 3′UTRs, which suggested direct miRNA-mediated inhibition of a fraction of these genes. Interestingly, genes that were downregulated by miR-124 in HeLa were found to be relatively lowly expressed in the brain, while genes downregulated by miR-1 were lowly expressed in the muscle. The authors made the suggestion that this observation indicated that the two miRNAs were involved in maintenance of cognate tissue identity, and that it was possible to identify functional targets of miR-124 and miR-1 in HeLa system. The list of miR-124 targets produced by Lim *et al.* was subsequently used as a bench-mark list of experimentally derived miR-124 direct targets (Conaco et al., 2006).

A significant overlap was detected between putative direct miR-124 targets identified in this thesis and the HeLa targets (Figure 5.9). Lim and colleagues identified 129 genes as significantly down-regulated upon miR-124 transfection and containing one or more putative miR-124 target sites. Of these putative direct targets in HeLa, 102 were homologous to mouse genes (according to HomoloGene, version 64 (Sayers et al., 2010)) and contained miR-124 heptamer seed matching sites. Exactly 50.0% of these genes ($P < 1.3e-22$) were identified as putative miR-124 targets in this work (Chapter 5, section 5.2.1).

5.3.2 Comparison to targets derived from Makeyev *et al.*

In a study of the role of miR-124 by Makeyev *et al.*, CAD cells were transfected with a plasmid expressing pre-miR-124-2, a precursor of miR-124 (Makeyev et al., 2007). The CAD cell line was derived from a mouse brain tumor, and it expressed a range of neuron-specific proteins and was capable of neuronal-like differentiation upon serum deprivation (Qi et al., 1997).

Makeyev and colleagues were able to purify a population of cells of which approximately 100% was transfected with a construct expressing miR-124. To achieve this, the

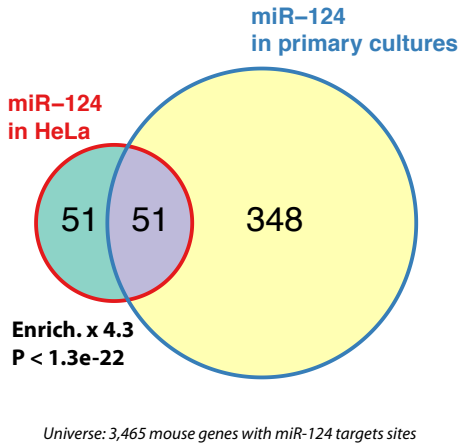


Figure 5.9: A significant intersection of miR-124 targets with those identified by Lim et al.
The Venn diagram shows counts of putative direct targets of miR-124 that were inferred from the transfection experiment of this thesis (*miR-124 in primary cultures*) and from the HeLa transfection experiment (Lim et al., 2005) (*miR-124 in HeLa*). The test universe was 3,465 mouse genes, with 3′UTRs containing one or more 7(2) or 7(1A)-type seed matching site for miR-124. The text shows fold enrichment above what is expected by chance alone and the hypergeometric P-value for the intersection.

pre-miR-124-2 sequence was inserted into an intron of a fluorescent reporter gene, which enabled the authors to FACS-sort transfected CAD cells. Subsequently, gene expression changes in cells transfected with pre-miR-124-2 were assessed using microarrays. Microarray results were deposited in the Gene Expression Omnibus (GEO) database (Sayers et al., 2010), GEO ID GSE8498.

Analysis of the microarray data obtained by Makeyev and colleagues (Methods, section 2.7) revealed that genes downregulated in CAD cells upon transfection with miR-124 expressing plasmid were significantly enriched in genes with miR-124 seed matching sites in their 3′UTRs (Figure 5.10a). Sylamer analysis of seed matching sites distribution showed that this enrichment was independent of length and composition biases (Figure 5.10b).

Analysis of the miR-124 seed matching site distribution enabled identification of putative direct miR-124 targets. The Sylamer enrichment peak approximately coincided with the 0.05 P-value cutoff for differential expression (Figure 5.10b). Of the genes that were downregulated with P-value < 0.05, 641 contained the 7(2) or 7(1A) seed matching sites for miR-124 and comprised a list of candidate direct targets of miR-124 in CAD cells.

A significant intersection (198 genes in common, which is approximately 2.7 more than expected by chance alone, $P < 1.6e-52$) was observed between putative miR-124 targets in CAD cell line (Makeyev et al., 2007) and miR-124 targets identified in this thesis (Figure 5.11).

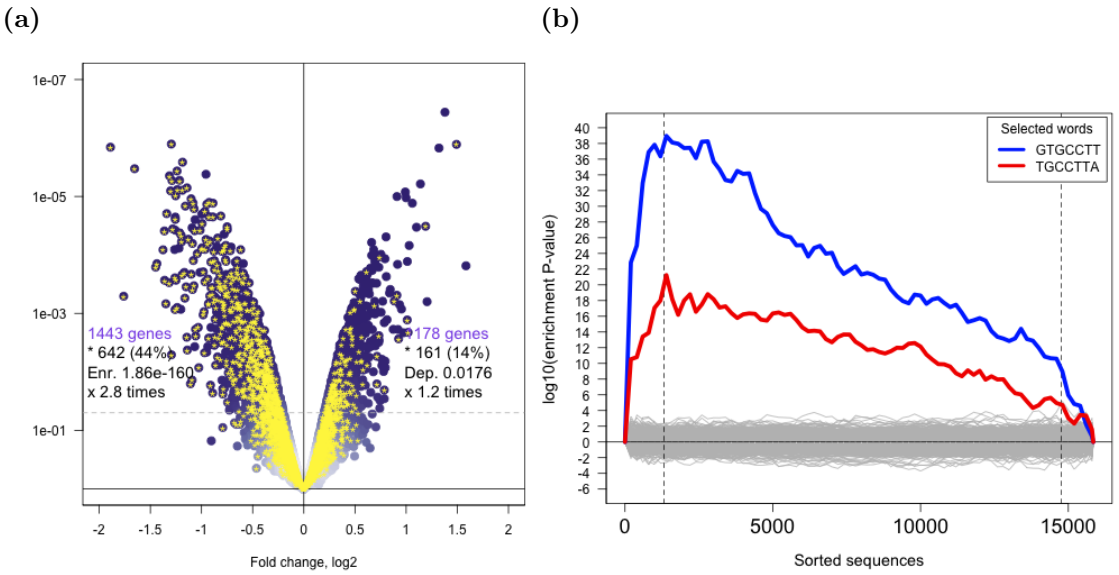


Figure 5.10: Differential expression and shifts in seed matching site distributions induced by miR-124 in CAD cells.
(Figure 5.10a) The x-axis represents log_2 of gene expression fold change of all genes represented on the Affymetrix platform (see Methods, section 2.7) between samples of the CAD cell line transfected with a miR-124 expressing plasmid in comparison the matched mock transfected samples. The y-axis represents P-value of differential expression (log_{10} scale), and the horizontal dashed grey lines show P-value cutoff of 0.05. The yellow asterisks mark genes [encoding transcripts] with 3′UTRs harbouring one or more seed matching sites (7(2) or 7(1A)-types) for miR-124. The text in the two halves of the plot area provides the following information: 1) The total number of genes with differential expression P-value more significant than the cutoff (0.05); 2) The total number (and percentage) of genes with seed matching sites for miR-124; 3) The hypergeometric P-value of enrichment (*Enr.*) or depletion (*Dep.*); 4) Fold enrichment or depletion of genes with the seed matching sites "× *times*" the number that is expected by chance alone. (Figure 5.10b) The x-axis represents sorted 3′UTRs corresponding to all genes represented on the Affymetrix platform (see Methods, section 2.7) ordered by fold change t-statistic for differential expression between samples of the CAD cell line transfected with a miR-124 expressing plasmid in comparison to the matched mock transfected samples. The sequences are **sorted from the most downregulated on the left to the most upregulated on the right**. The y-axis represents the hypergeometric P-values for occurrence biases of 876 nucleotide words complementary to the seed regions (7(2) and 7(1A)-types) of the complete set of 581 distinct mouse miRNAs, according to miRBase Release 14 (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006, 2008). Positive values on the y-axis corresponds to an enrichment ($+|log_{10}$(P-value)|) and negative values to a depletion ($-|log_{10}$(P-value)|). The vertical dashed lines mark the P-value cutoff (0.05) on both sides of the ranked gene lists. The blue and the red lines show the enrichment profiles of 7(2) and 7(1A)-type seed matching sites for miR-124. The grey lines show the enrichment for the rest of the distinct seed matching sites. The mapping of microarray probes to mRNA transcripts, and transcripts to genes, is described in Methods (section 2.7). The identification of the seed regions and parameters of Sylamer (van Dongen et al., 2008) is in Methods (section 2.8). The full description of the Sylamer method is in the Introduction (section 2.8).

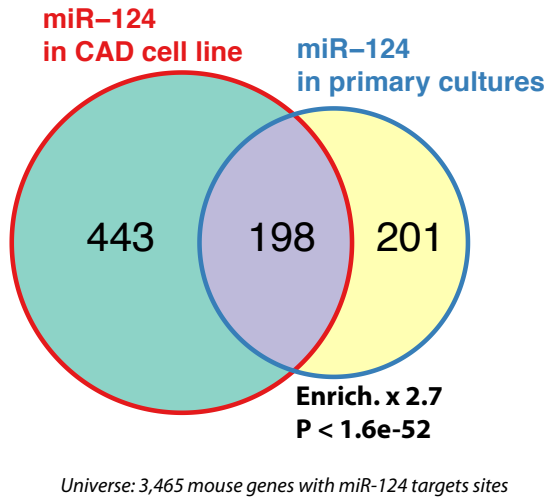


Figure 5.11: A significant intersection of miR-124 targets with those derived from Makeyev et al.
The Venn diagram shows counts of putative direct targets of miR-124 that were inferred from the transfection experiment of this thesis (*miR-124 in primary cultures*) and from the HeLa transfection experiment (Makeyev et al., 2007) (*miR-124 in CAD cell line*). The test universe was 3,465 mouse genes, with 3′UTRs containing one or more 7(2) or 7(1A)-type seed matching site for miR-124. The text shows fold enrichment above what is expected by chance alone and the hypergeometric P-value for the intersection.

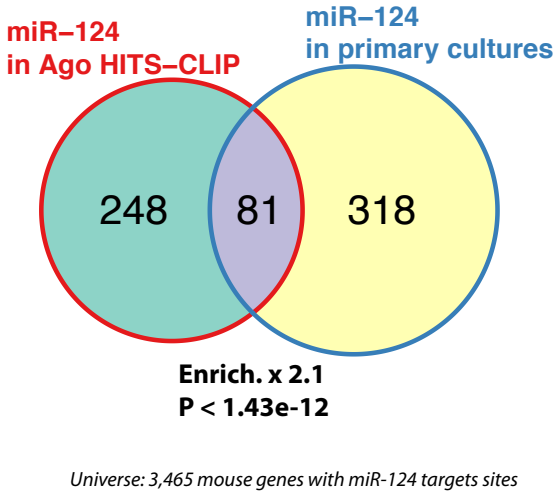


Figure 5.12: A significant intersection of miR-124 targets with those identified by Chi et al.
The Venn diagram shows counts of putative direct targets of miR-124 that were inferred from the transfection experiment of this thesis (*miR-124 in primary cultures*) and from the HITS-CLIP experiment (Chi et al., 2009) (*miR-123 in Ago HITS-CLIP*). The test universe was 3,465 mouse genes, with 3′UTRs containing one or more 7(2) or 7(1A)-type seed matching site for miR-124. The text shows fold enrichment above what is expected by chance alone and the hypergeometric P-value for the intersection.

5.3.3 Comparison to targets identified by Chi *et al.*

The basis of the two described above studies (Lim et al., 2005; Makeyev et al., 2007) was the overexpression of miR-124 through chemical transfection of nucleic acids, which led to an excess of miR-124. Therefore, even though targets identified in this thesis were in good agreement with these two experiments, the question remained whether the thesis targets were relevant to miR-124 activity outside cell cultures and transfection paradigms.

Lists of targets of miR-124 and 19 other highly expressed in brain miRNAs were recently identified using the HITS-CLIP method in an innate P13 mouse neocortex (Chi et al., 2009). The HITS-CLIP did not involve transfections, but was a combination of UV cross-linking of Ago proteins (a key component of miRNA silencing complex, see Introduction, section 1.1.2) and nucleic acids, and immunoprecipitation of bound RNA followed by a high-throughput RNA sequencing. The sequences, which also harboured the seed matching sites for one of the 20 miRNAs, corresponded to putative direct targets of these miRNAs[1] (the *in vivo* targets).

A significant overlap (approximately 2.1 times bigger than was expected by chance alone, $P < 1.44e-12$) was detected between HITS-CLIP miR-124 putative direct targets and miR-124 targets identified in this thesis (Figure 5.12). This suggested that a significant proportion of miR-124 targets identified in this thesis were direct miR-124 *in vivo* targets.

Summary of section 5.3

This chapter described the identification of putative direct miRNA targets in primary neuronal cultures of six different miRNAs: miR-124, miR-434-3p, miR-25, miR-143, miR-145 and of cel-miR-67. The targeting repertoire of miR-124 was previously studied, thus it was possible to compare published results to the list of miR-124 targets identified in this thesis. A good agreement was found in comparisons between the thesis targets and three published studies (Lim et al., 2005; Makeyev et al., 2007; Chi et al., 2009). This indicated that methods of this work to identify putative direct miRNA targets were valid and generated reproducible results.

[1]The targets identified with the HITS-CLIP method were available for download from the authors' website http://ago.rockefeller.edu/.

Chapter 6

Analysis of miRNA function in neurons

The goal of this thesis on a very basic level is to determine the function of miRNAs in neurons. Are neuronal miRNAs important? What is their repertoire of targets? Do these miRNAs share target pools? What can be learned about the overall function of these miRNAs from studying detected targets? Within this chapter I will attempt to answer these questions using the data derived during my research and previously published literature.

Previously, individual miRNAs were shown to regulate expression from dozens to hundreds of genes (Stark et al., 2003; Enright et al., 2003; Farh et al., 2005; Lim et al., 2005; Giraldez et al., 2006; Baek et al., 2008; Selbach et al., 2008), however functions of miRNAs are usually viewed through a prism of a handful of validated targets. Additionally, roles of miRNAs in the differentiation of neural progenitors are well established (Introduction, section 1.1.3). However after differentiation, the functions of a majority of miRNAs are poorly understood. Chapter 5 described the identification of hundreds of putative direct targets for six miRNAs (**miR-124**, **miR-434-3p**, **miR-143**, **miR-145**, **miR-25** and **cel-miR-67**) in committed primary neuronal cultures. Two of these miRNAs, miR-124 and miR-434-3p, were either highly expressed or upregulated in the development of primary forebrain cultures (Chapter 3, section 3.3). Therefore, they were expected to have endogenous functions in neurons. On the other hand, other miRNAs were either downregulated in the development of cultures, or absent from the mouse genome altogether (Chapter 3, section 3.3). Therefore, no specifically neuronal function was expected to be associated with these miRNAs in primary neuronal cultures. In this chapter I will compare and analyse lists of targets of these six miRNAs, and attempt to identify func-

tions associated with miRNA mediated regulation as a whole and functions that may be associated specifically with neuronal miRNAs.

The analysis of intersections between lists of targets, together with the analysis of enrichment of functionally annotated gene categories (GO and KEGG categories (terms) (Ashburner et al., 2000; Kanehisa et al., 2008)) in the lists of miRNA targets, is described in the first section (section 6.1) of this chapter. In the second section (section 6.2), I will explore the connection between functions of neuronal miRNAs and stress responses and highlight the importance of these processes for neuronal biology.

6.1 Characterisation of identified miRNA targets

6.1.1 Significant intersections between targets of different miRNAs

It was reasoned that between the lists of targets of six miRNAs intersections that were bigger than expected, would indicate related targeting repertoires, while smaller than expected intersections would suggest distinct repertoires. Assessment of the significance of an intersection between elements of two lists can be done using the hypergeometric test, where the lists are viewed as samples from a larger set of elements, which is referred to as the test universe. Selection of an appropriate test universe is critical for this type of analysis: an unsuitably large universe artificially enhances, while a restricted universe reduces the significance of intersections. As the targets of miR-124 were in good agreement with previously published results (Chapter 5, section 5.2.1), data derived from the miR-124 experiment (the 6DIV bidirectional transfection, see Chapter 5, section 5.2.1) was judged to be reliable, and 10,821 genes detected[1] in this experiment were used as the test universe.

Using the single gene universe from the miR-124 experiment, the hypergeometric test for enrichment showed significant intersections between the majority of the lists of miRNA targets (Figure 6.1). For example, targets of miR-124, miR-434-3p and miR-145 intersected significantly with the lists of targets of all of the other five miRNAs (enrichment P-values varied from 0.0383 to $6.41e-10$). Interestingly, the intersection of the targets of miR-124, miR-434-3p and miR-145 with the targets of cel-miR-67, which is not expressed in the mouse, was two or more times bigger than expected by chance alone (enrichment P-values were 0.00054, 0.0057 and $7.04e-07$). Overall, a cross comparison of the targets of

[1] Using the standard Illumina detection call threshold $P < 0.01$, see Methods (section 2.7)

the six miRNAs revealed 12 significant intersections with $P < 0.05$ (Figure 6.1). The observation of multiple significant intersections could not be explained by the increased false discovery rate associated with multiple testing, because nine results remained significant ($P < 0.05$) even with the strictest adjustment method (Bonferoni correction).

Significant intersections were also identified between the putative direct targets identified in this thesis, and miRNA targets identified elsewhere. For example, miR-124 targets derived from the experiment conducted by Makeyev and colleagues, which overexpressed miR-124 in the CAD cell line[2] (Makeyev et al., 2007), had a significant intersection with transfection targets of five out of six miRNAs from this thesis (Figure 6.1). Importantly, the intersection of the CAD miR-124 targets with targets of the non-mouse cel-miR-67 was also significant ($P < 0.0006$, see Figure 6.1). The latter indicated that the significant intersections where not only a feature of neuronal miRNAs, but also of miRNA mediated gene expression regulation as such.

6.1.2 Explaining the intersection: A hypothesis of a pool of transcripts primed for miRNA mediated regulation

One explanation for the multiple significant intersections could be artefacts in detection of differential expression. For example, due to technical biases, differential expression might have been detectable for only a small, common subset of transcripts in each of the transfection experiments. In such case, intersections between the lists of miRNA targets would be bigger than expected by chance alone, because identification of targets was itself reliant on detection of differential expression. An alternative explanation of significant intersections would be the existence of a pool of transcripts that is primed for miRNA mediated regulation. In other words, a higher than expected intersections could be explained by the fact that not all transcripts that were expressed and contained the seed matching sites were equally likely to be downregulated upon overexpression of miRNAs (I will sometimes refer to this proposition as a "pool of targets hypothesis"). To distinguish between these two explanations (i.e. artefacts versus the pool of targets hypothesis), targets of miRNAs obtained in this thesis were compared to targets identified with an experiment, which did not rely on microarrays or on detection of differential expression by other means.

[2]See Chapter 5 (section 5.3.2) for the description of derivation of putative direct miR-124 targets from the CAD cell experiment (Makeyev et al., 2007).

One such experiment is the Ago HITS-CLIP experiment conducted by Chi and colleagues (Chapter 5, section 5.3.3) (Chi et al., 2009). In this experiment Chi and colleagues identified putative direct targets of miR-124 and of 19 other miRNAs, which were all abundant in the P13 mouse neocortex (Chi et al., 2009). The putative miRNA targets were defined as transcripts that were bound by Ago, and which had a hexamer seed matching site near the Ago binding site for one of the 20 miRNAs. I will refer to the combined set of these putative targets as the Ago HITS-CLIP set[1].

There was a significant intersection between the Ago HITS-CLIP set and transfection targets of five out of six miRNAs identified in this thesis (Figure 6.1). This result was unexpected, because only one miRNA, miR-124, was investigated both in this thesis and in the HITS-CLIP experiment. Therefore the intersection could have been explained by the bias in the seed based target identification only for miR-124 (only transcripts with miR-124 seed matching sites can be defined as targets, which restricts possible selection of genes from the universe[2]). None of the other six miRNAs from this thesis shared the seed region (either hexamer or heptamer, see Introduction, section 1.2.1) with miRNAs from the HITS-CLIP study. Nevertheless, multiple significant intersections with the Ago HITS-CLIP were observed (Figure 6.1). Moreover, a significant intersection was identified between the Ago HITS-CLIP set and targets of cel-miR-67. This intersection was 87 genes (out of the 350 targets of cel-miR-67 present in the test universe), which was approximately 1.7 times more than expected by chance alone ($P < 7.35e-07$, see Figure 6.1). This was, perhaps, the strongest support for the pool of targets hypothesis, because targets of cel-miR-67 can be viewed as a sample of transcripts that could be regulated by miRNAs in primary neurons unbiased by previous evolutionary selection.

These observations implied that regardless of the method used for detection of putative direct miRNA targets, miRNAs in the neuronal systems (i.e. the P13 neocortex in the HITS-CLIP study, and primary forebrain cultures in this thesis) appeared to target a common subset of genes. Therefore, the pool of targets hypothesis was considered to be a likely explanation of the observed intersections between the lists of targets of different miRNAs.

[1]In experiments performed in this thesis, on average $\approx$ 1,500 genes encoding transcripts of the Ago HITS-CLIP set were detected as expressed (using the standard Illumina detection call threshold $P < 0.01$, see Methods, section 2.7).

[2]It should be noted that even for miR-124 the observed significant intersection was unlikely to had been explained by the bias in the seed based definition of targets. In Chapter 5 (section 5.3.3), I showed that the intersection between miR-124 targets identified in the thesis and in the HITS-CLIP experiment was significant ($P < 1.43e-12$) within the universe of only the genes that were encoding transcripts with the miR-124 seed matchings sites.

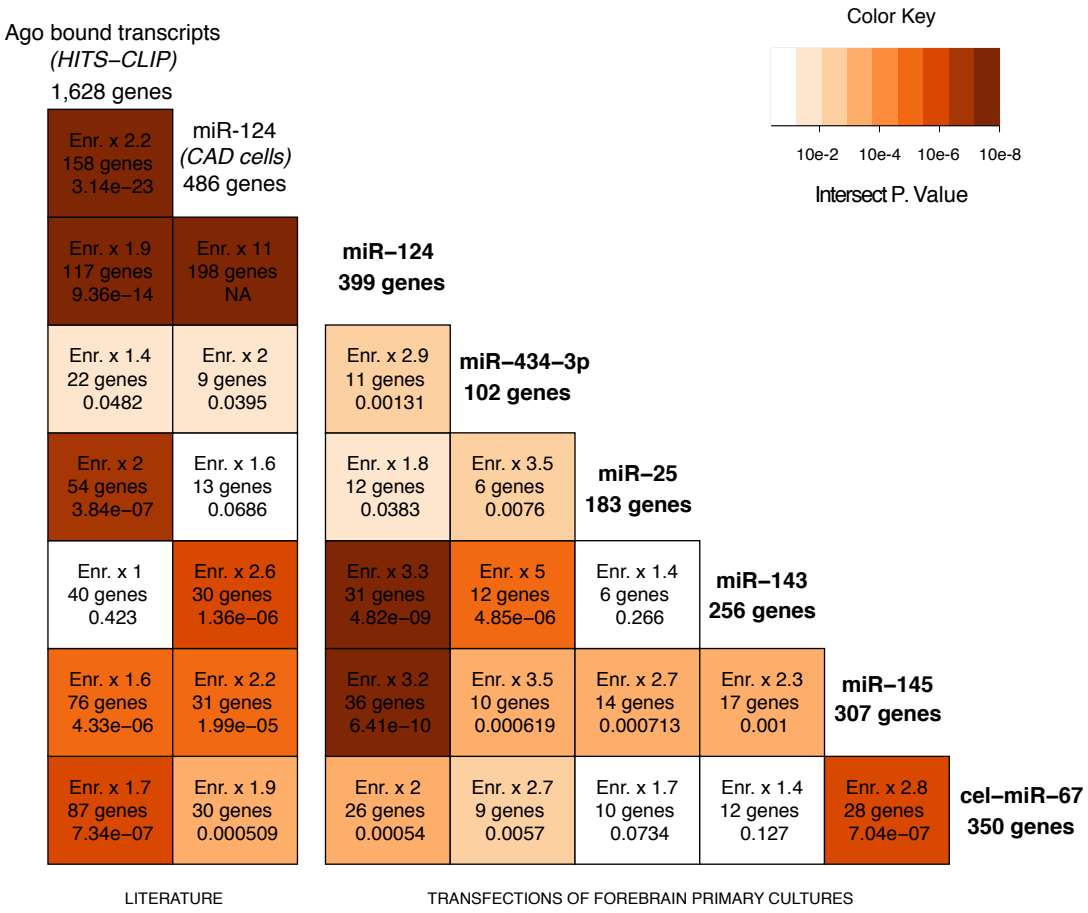


Figure 6.1: Intersections of the lists of putative direct miRNA targets.
Names of the genelists together with the total number of genes in the lists, belonging to the test universe is given on the sides of the boxes. The text inside of the boxes provides information about the intersection of the list that correspond to top and right sides of the boxes: 1) Fold enrichment (*Enr.*) the number of times (×) more than expected by chance alone; 2) The number of genes in the intersection between the two lists; 3) The hypergeometric P-value of the enrichment. The color of boxes corresponds to the hypergeometric P-value according to the color-scheme of the *Color Key*. The test universe (*Gene Universe*) for all tests was the complete set of 10,821 genes detected in the miR-124 bidirectional transfection experiment (6DIV, see Chapter 5, section 5.2.1).

6.1.3 Context-dependent nature of the pool: Over 20% of targets were induced by the transfection procedure itself

Having observed significant intersections between the lists of putative direct targets of different miRNAs, I proceeded to functional characterisation of these lists. If targets of more than one miRNA were associated with the same function, it would suggest existence of global functions of miRNA mediated regulation. To associate functions with the lists of miRNA targets, I evaluated enrichment of KEGG pathways (Kanehisa et al., 2008) in all lists (Methods, section 2.10).

To make KEGG enrichment analysis fully comprehensive, test universes were defined individually for each of the lists of targets. The universes were defined as all genes detectably expressed in the experiment that lead to generation of each list[1]. By pairing the universe with the corresponding transfection experiments ensured that all targets would be within the universe to be included in the analysis. Enrichment of KEGG pathways was assessed in the lists with the hypergeometric test (Methods, section 2.10), results of which depend on sizes of the lists tested (i.e. it is harder to obtain significant P-values for smaller lists of targets). Therefore, to make a comparison between the lists, for each of the lists an arbitrary cutoff was set at the 25 most enriched KEGG pathways, and all 25 pathways were considered irrespective of the enrichment P-value.

Analysis of KEGG term enrichment in the Ago HITS-CLIP target set (i.e. *in vivo* miRNA targets in P13 mouse neocortex), revealed that the set was enriched in the genes that were unlikely to be expressed constitutively. For example, three out of four most enriched KEGG pathways in the Ago HITS-CLIP set were "Long-term potentiation", "Regulation of actin cytoskeleton" and "Axon guidance" (Supplementary Data, Table A.22). It is possible that the 20 highly expressed in neurons miRNAs, which targets comprised the Ago HITS-CLIP set, acted as buffers of expression of the genes from the aforementioned pathways at the times when these pathways were induced (e.g. during neuronal plasticity). Such hypothesis agrees with previous reports of miR-124 (Rajasethupathy et al., 2009) and miR-134 (Gao et al., 2010) to reduce the plasticity, and a report of miR-134 to reduce the size of the synapse (Schratt et al., 2006) (functions of miRNAs as inhibitors of neuronal plasticity is described in the Introduction, section 1.1.3). This

[1]The exception to this was the Ago HITS-CLIP set, which was produced in the external experiment and where the full set of genes expressed was not known. Instead, the universe of genes from the miR-124 transfection experiment was used for testing KEGG enrichment in the Ago HITS-CLIP set. For all other experiments, all genes detected, using the standard Illumina detection call threshold $P < 0.01$ (Methods, section 2.7) were used as gene universes.

reasoning suggested that miRNAs exogenously added in the transfections could have also inhibited the genes that were induced during the experiments.

To test the hypothesis that the exogenously added miRNAs (i.e. transfected miRNAs) inhibited inducible pathways, I first identified genes that were induced during the transfection experiments. To achieve this, the expression profiles of mock transfected cultures (i.e. cultures treated with the transfection reagent, but without RNA added to it) were compared to the expression profiles of matched untransfected cultures (Methods, section 2.5). This comparison revealed that 1,293 genes were upregulated ($P < 0.05$) by the treatment with the transfection reagent (these genes will be referred to as the "induced by the transfection reagent" set). Analysis of KEGG pathway enrichment in this set[2] showed that the genes induced by the transfection reagent were enriched not in the pathways involved in normal neuronal function, but in pathways related to diseases and stresses. For example, the most highly enriched KEGG pathway was "p53 signalling pathway", while the pathway "Metabolism of xenobiotics by cytochrome P450" was also among the top 25 most highly enriched pathways. Additionally, of the 25 most enriched terms in the induced by the transfection reagent set, 12 were related to cancer or other diseases (Supplementary Data, Table A.21). Enrichment of stress and disease related pathways in the induced by the transfection set suggested that a significant fraction the induced genes was involved in offsetting the adverse effects of the transfection reagent.

Next, I tested if transfected miRNAs inhibited the induced by the transfection reagent genes. Using genes that were detected in the mock transfection experiment (using the standard Illumina detection call $P < 0.01$, see Methods, section 2.7) as the gene universe, I found that the intersections of the induced by the transfection reagent set and of the targets of five out six miRNAs was statistically significant (Figure 6.2). Interestingly, targets of the two neuronal miRNAs, miR-124 and miR-434-3p, were the most enriched in the genes induced by transfection reagent (3.3 times more than expected by chance alone for miR-124 ($P < 2.35e-41$) and 3.8 time more for miR-434-3p ($P < 3.56e-14$)). When the intersections were assessed outside of the gene universe, expression of 34.3% and 37.5% of miR-124 and miR-434-3p targets in total was found to have been induced by the transfection reagent. Also, the "p53 signalling pathway", which was the most highly enriched pathway in the genes induced by transfection (Supplementary Data, Table A.21) was among top 25 most enriched pathways in the targets of both miR-124 and miR-434-3p (Supplementary Data, Table A.43). If targets of all six miRNAs were combined,

[2]Genes detected using the standard Illumina detection call threshold $P < 0.01$ (Methods, section 2.7) in the mock transfection experiment were used as a gene universe for these tests.

producing a unique list of 1,512 genes, then 22.2% (337 genes) were identified as induced by the transfected reagent.

Furthermore, 14 out of the 25 most enriched KEGG pathways in the induced by the transfection reagent set were also in the top 25 most enriched pathways in the targets of one or more miRNAs (Supplementary Data, Table A.15 to A.20). Figure 6.3 shows this recurrent enrichment of KEGG pathways in targets of a selection of three miRNAs[3] (miR-124, miR-434-3p and cel-miR-67), and genes induced by the transfection reagent. Multiple disease related pathways (e.g. "Pathways in cancer") were found to be induced by the transfection reagent and targeted by several miRNAs. In fact, the KEGG term "Pathways in cancer" was among 25 most enriched terms in the targets of five miRNAs: miR-124, miR-434-3p, miR-25, miR-143 and cel-miR-67 (Figure 6.3 and Table 6.1). Identification of this and other cancer-related KEGG pathways as enriched in miRNA targets was in agreement with a large body of evidence that showed significance of miRNA mediated regulation in development of various types of tumors (Volinia et al., 2010).

In summary, results presented in this section showed that the significant intersection of lists of targets of different miRNAs (see section 6.1.2) could be due to different miRNAs having converged on inhibition of a common set of genes. According to the proposed pool of targets hypothesis, this set of genes encoded a pool of transcripts primed for miRNA mediated regulation. In transfection experiments conducted in this thesis, I found that genes induced by the transfection reagent contributed significantly to this pool of primed targets. Functional characterisation of miRNA targets and genes induced by transfection, showed that multiple transfected miRNAs inhibited pathways that were induced by the transfection reagent (i.e. many disease and stress associated pathways). Targets of the two neuronal miRNAs, miR-124 and miR-434-3p, were the most enriched in the genes induced by the transfection reagent. On the contrary, targets of an oncogenic miRNA, miR-25, and of a non-mouse miRNA, cel-miR-67, were least enriched in the genes induced by the transfection reagent. These observations suggested that targeting repertoire of neuronal miRNAs could have specifically evolved to buffer the expression of genes that can be induced by adverse treatments of neurons (such as the treatment by the transfection reagent).

These collected observations suggest that miRNA mediated regulation as a whole converges on inhibition of genes that are upregulated in the system, i.e. miRNAs can act

[3]These three miRNAs were selected to demonstrate that pathways induced by the transfection reagent contributed to targets of both mouse neuronal miRNAs (miR-124 and miR-434-3p), and a non-mouse miRNA (cel-miR-67). Recurrent enrichment of KEGG pathways in targets of these unrelated miRNAs supports the hypothesis of a pool of transcripts primed for miRNA mediated regulation.

as a buffer against deviation of the transcriptome from equilibrium of the differentiated state. In accordance with this proposition, in experiments that were performed in this thesis, transfected miRNAs acted to reduce the perturbation caused by the transfection reagent. In this function as a buffer of transcriptional changes, miRNAs can be particularly important for neurons, because these cells constantly receive a flux of stimuli, each of which, potentially, can alter gene expression and make the transcriptome to deviate from the *status quo*. The latter is in agreement with published reports of exogenous miRNAs reducing the plasticity of neurons (Gao et al., 2010; Rajasethupathy et al., 2009), while removal of the endogenous miRNA increases plasticity (Gao et al., 2010; Konopka et al., 2010).

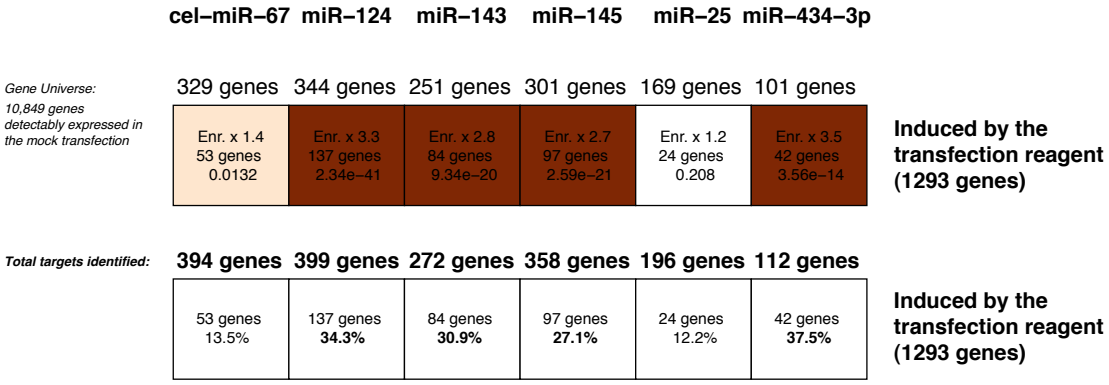


Figure 6.2: Intersection of genes induced by the transfection reagent and miRNA targets. Each of the boxes shows intersection between two gene lists: 1) A list of miRNA targets (the corresponding names of miRNAs are labelling the tops of the boxes); 2) The list of genes induced by the genes induced by the transfection reagent (labelling the right sides of the boxes). The upper row of boxes displays information about the intersections between the lists that were limited by the gene universe (10,849 genes detected with the standard Illumina detection call $P < 0.01$ (Methods, section 2.7) in the mock transfection experiment). Information about the intersections is presented in the same way (and coloring is according to the same color-scheme) as in Figure 6.1. The bottom row of clear boxes shows intersections of complete lists (i.e. not restricted by the universe of genes detected in the mock transfection experiment). The bottom row of boxes provides the following information: 1) The total number of genes in the intersections; 2) The percent of the intersections in the lists of miRNA targets.

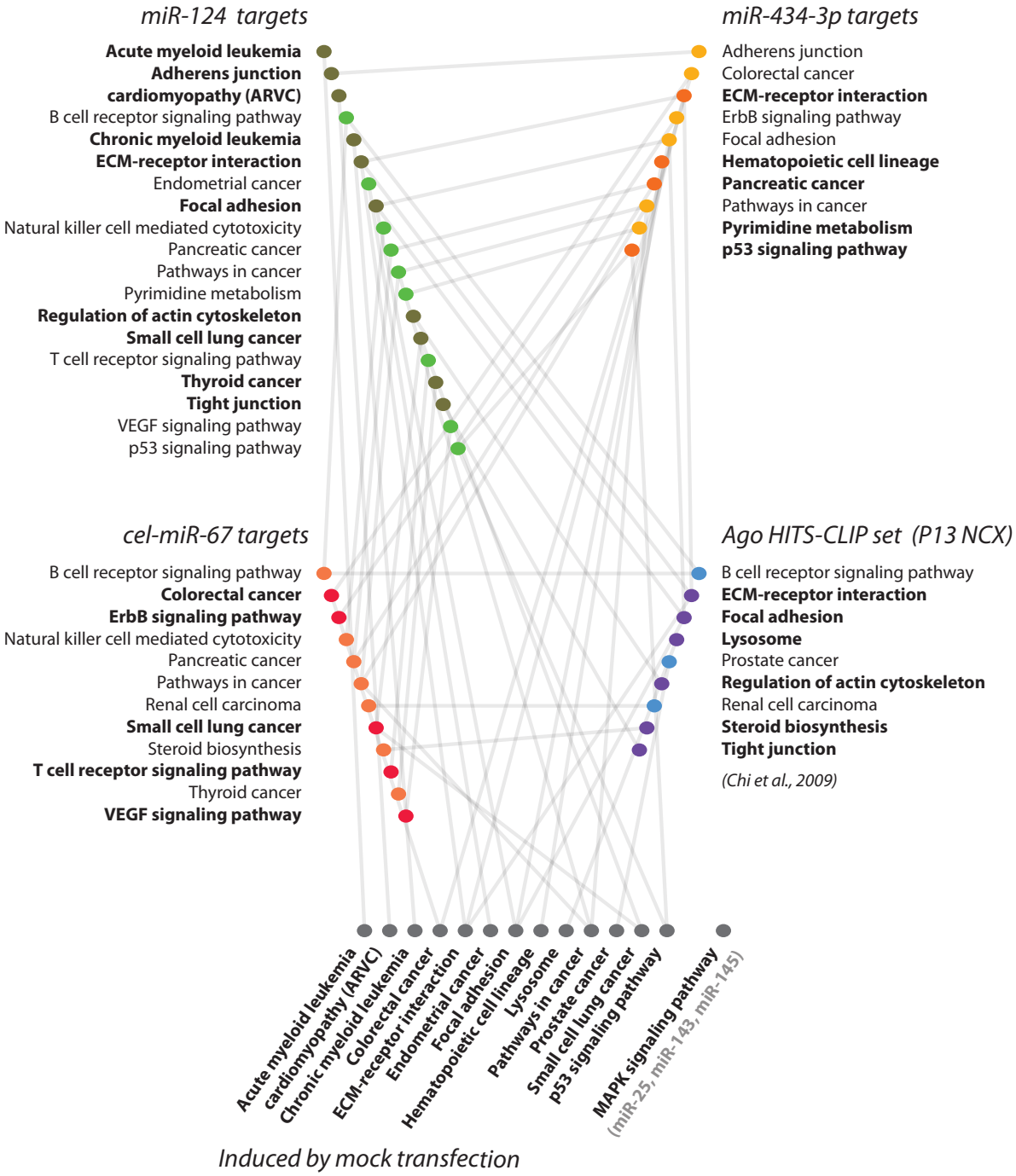


Figure 6.3: Recurrence in top 25 most enriched KEGG terms.
The text in italic is the gene lists in which KEGG enrichment was assessed. The text in regular font shows the lists of KEGG pathways that were among the top 25 most enriched pathways in more than one of these gene lists (with the exception of "MAPK signalling pathway", which was within the top 25 most enriched pathways in the targets of miR-143, miR-145 and miR-25). The pathways in bold, were enriched with the P-value < 0.05 (Methods, section 2.10).
The pathways that are enriched in multiple gene lists are connected with the grey lines. The full list of miRNA targets from within the following pathways is in the Supplementary Data: "ECM-receptor interaction" (Table A.39), "Gap junction" (Table A.40), "ErbB signaling pathway" (Table A.41), "Tight junction" (Table A.42), "p53 signaling pathway" (Table A.43), "Regulation of actin cytoskeleton" (Table A.44), "Focal adhesion" (Table A.45), "MAPK signaling pathway" (Table A.46), "VEGF signaling pathway" (Table A.47), "Toll-like receptor signaling pathway" (Table A.48).

Symbol	Description
cel-miR-67 (**Ranked 17, P ≈ 0.101**)	
Appl1	adaptor protein, phosphotyrosine interaction, PH domain and leucine zipper containing 1
Bcl2	B-cell leukemia/lymphoma 2
Cycs	cytochrome c, somatic
Fzd3	frizzled homolog 3 (Drosophila)
Kras	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
Pias3	protein inhibitor of activated STAT 3
Pik3cb	phosphatidylinositol 3-kinase, catalytic, beta polypeptide
Ralb	v-ral simian leukemia viral oncogene homolog B (ras related)
Rarb	retinoic acid receptor, beta
Rb1	retinoblastoma 1
Tpm3	tropomyosin 3, gamma
miR-124 (**Ranked 22, P ≈ 0.11**)	
Ccnd1	cyclin D1
Col4a1	collagen, type IV, alpha 1
Erbb2	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived [...]
Fadd	Fas (TNFRSF6)-associated via death domain
Itgb1	integrin beta 1 (fibronectin receptor beta)
Lamc1	laminin, gamma 1
Nras	neuroblastoma ras oncogene
Rela	v-rel reticuloendotheliosis viral oncogene homolog A (avian)
Skp2	S-phase kinase-associated protein 2 (p45)
Smad3	MAD homolog 3 (Drosophila)
Stat3	signal transducer and activator of transcription 3
Tcf7l1	transcription factor 7-like 1 (T-cell specific, HMG box)
Traf3	TNF receptor-associated factor 3
miR-143 (**Ranked 23, P ≈ 0.161**)	
Birc5	baculoviral IAP repeat-containing 5
Egfr	epidermal growth factor receptor
Fadd	Fas (TNFRSF6)-associated via death domain
Mmp2	matrix metallopeptidase 2
Pdgfb	platelet derived growth factor, B polypeptide
Pdgfra	platelet derived growth factor receptor, alpha polypeptide
Smad2	MAD homolog 2 (Drosophila)
Smo	smoothened homolog (Drosophila)
miR-145 (Ranked 26, P ≈ 0.222)	
Birc5	baculoviral IAP repeat-containing 5
Cycs	cytochrome c, somatic
Gli3	GLI-Kruppel family member GLI3
Ikbkg	inhibitor of kappaB kinase gamma
Nras	neuroblastoma ras oncogene
Pdgfra	platelet derived growth factor receptor, alpha polypeptide
Ptch1	patched homolog 1
Traf6	TNF receptor-associated factor 6
Wnt7b	wingless-related MMTV integration site 7B
miR-25 (**Ranked 24, P ≈ 0.111**)	
Fgf10	fibroblast growth factor 10
Fgf12	fibroblast growth factor 12
Igf1r	insulin-like growth factor I receptor
Mapk8	mitogen-activated protein kinase 8
Pik3cb	phosphatidylinositol 3-kinase, catalytic, beta polypeptide
Pik3r2	phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 2 (p85 beta)
Wnt5a	wingless-related MMTV integration site 5A
miR-434-3p (**Ranked 20, P ≈ 0.135**)	
Birc5	baculoviral IAP repeat-containing 5
Egfr	epidermal growth factor receptor
Fgf13	fibroblast growth factor 13
Stat3	signal transducer and activator of transcription 3
Tgfbr2	transforming growth factor, beta receptor II

Table 6.1: A list of the miRNA targets found within the "Pathways in cancer" KEGG pathway.
The text in parenthesis shows the rank of the enrichment of the "Pathways in cancer" among all the KEGG pathways and the P-value of that enrichment. The text is in bold if the enrichment was ranked within the top 25 most enriched pathways.

6.1.4 Context dependent nature of published miR-124 targets

The HITS-CLIP experiment (Chi et al., 2009) and the transfection experiments in this thesis were performed in related neuronal systems (P13 mouse neocortex and primary forebrain cultures). Therefore it was expected that similarities would be found between miRNA targets identified in both of these studies. Indeed, miR-124 targets identified in this thesis had a significant inetersection with the Ago HITS-CLIP as a whole (Figure 6.1), and specifically with the miR-124 HITS-CLIP targets (which were a part of the Ago HITS-CLIP set, see Chapter 5, section 5.3). Moreover, multiple KEGG pathways, which were overrepresented in the targets identified by miRNA transfection experiments and in genes induced by the transfection reagent alone, were also overrepresented in the Ago HITS-CLIP set (Figure 6.3).

Despite these similarities between the miRNA targets identified with transfections of primary cultures and the targets identified with the HITS-CLIP method, a line of evidence showed that transfection targets and the HITS-CLIP targets were focused around different aspects of miRNA function. In section 6.1.3, I described that in transfection experiments in this thesis project, a significant proportion of targets of five out of six miRNAs (including those of miR-124) was induced by the treatment with the transfection reagent (Figure 6.2). Interestingly, targets of miR-124 that were identified in published transfections of HeLa (Lim et al., 2005) and CAD cell lines (Makeyev et al., 2007) also had significant intersections with the set of genes induced by the transfection reagent (Figure 6.4). However, the HITS-CLIP miR-124 targets had approximately four orders of magnitude less significant P-value for the intersection, while the intersection of the whole Ago HITS-CLIP set and the genes induced by the transfectecion reagent was not not significant (Figure 6.4). This difference between the two types of targets (transfection versus HITS-CLIP) is important, because in this thesis the transfection targets of neuronal miRNAs (miR-124 and miR-434-3p) were comprised to a large and significant extent of the genes induced by the transfection reagent with an enrichment 3 times bigger than expected by chance alone, see Figure 6.2. Therefore, the transfection targets can be viewed as focusing on miRNAs function in disease and stress, because the "induced by the transfection reagent" set was enriched in disease and stress related pathways (Supplementary Data, Table A.21). At the same time, the HITS-CLIP targets are likely to focus on miRNA function in normal neurons, because the Ago HITS-CLIP set was highly enriched in several specifically neuronal pathways (Supplementary Data, Table A.22), and not enriched in the induced by the transfection reagent set itself (Figure 6.4).

These results support the proposition of a context dependent function of miRNAs as inhibitors of the inducible genes – in different transfection experiments miRNAs converge on inhibition of the genes induced in the transfection experiment, and it is not the case in the transfection-free experiment. This conclusion has direct implications for future studies aiming at identifying targets of miRNAs, as it is important to realise that miRNAs can act on different genes in different experimental contexts.

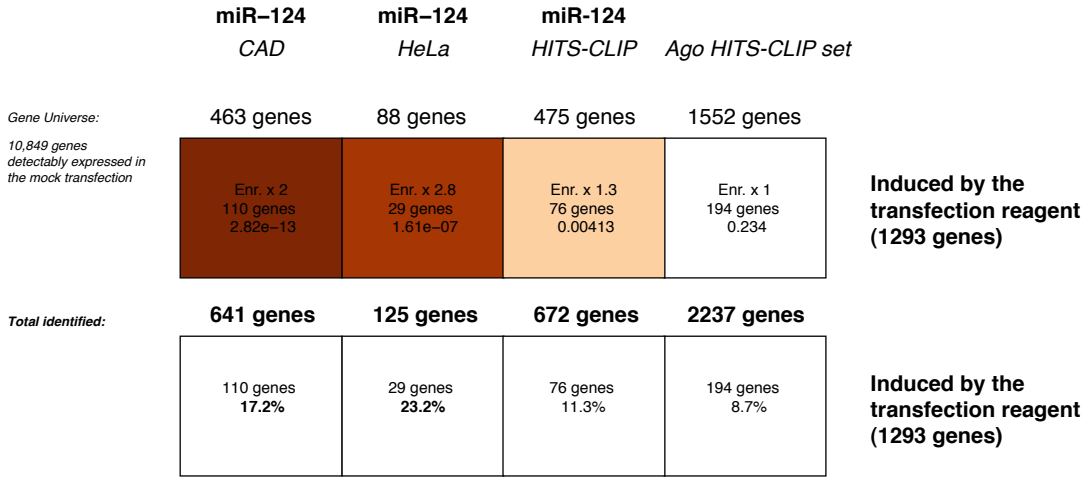


Figure 6.4: Intersection of published miR-124 targets and genes induced by the transfection reagent.
Each of the boxes shows the intersection between two gene lists: 1) A list of published miRNA targets (miR-124 targets or the complete list of the Ago HITS-CLIP targets of 20 most highly expressed miRNAs, see text). The corresponding names of miRNAs are labelling the tops of the boxes; 2) The list of genes induced by the mock transfection (labelling the right sides of the boxes). The upper row of boxes displays information about the intersections between the lists that were limited by the gene universe (10,849 genes detected with the standard Illumina detection call $P < 0.01$ (Methods, section 2.7) in the mock transfection experiment). Information about the intersections is presented in the same way (and colored according to the same color-scheme) as in Figure 6.1. The bottom row of clear boxes shows intersections of complete lists (i.e. not restricted by the universe of genes detected in the mock transfection experiment). The bottom row of boxes provides the following information: 1) The total number of genes in the intersections; 2) The percent of the intersections in the lists of published miRNA targets.

6.1.5 Recurrent enrichment of GO terms was in agreement with the pool of targets hypothesis

The hypothesis that a pool of transcripts was primed for miRNA mediated regulation was formulated in section 6.1.2, and it was based on the observation of significant intersections between transfection targets identified for unrelated miRNAs. In agreement with this hypothesis, several KEGG pathways (i.e. cell adhesion, cell signaling and stress related pathways, see sections 6.1.3) were recurrently enriched in the lists of targets of different miRNAs (Figure 6.3). In this section, the pool of targets hypothesis is further supported by the observation that several large Gene Ontology (GO) terms (Ashburner et al., 2000) were recurrently enriched in targets of unrelated miRNAs (Methods, section 2.10).

There are currently tens of thousands of annotated GO terms and the relationship between them is complicated (http://www.geneontology.org/). Therefore, to ease interpretation of GO enrichment across the lists of miRNA targets, the survey of GO enrichment was based on a selection of several relatively large representative GO terms (more that 100 genes in each, "Cellular compartment" and "Biological process" types of GO terms (Ashburner et al., 2000)). This selection of GO terms was based on 40 most enriched terms in the iduced by the transfection reagent and the Ago HITS-CLIP sets. Selection of the representative terms was made from these two sets, because the mock transfection and the HITS-CLIP experiments uncovered aspects of miRNA function in two different contexts: in neurons under the transfection stress and in neurons in the normal state (section 6.1.3). The selected terms are listed in the legend of Figure 6.5 and the complete list of the 40 most highly enriched GO terms in the induced by the transfection reagent, the Ago HITS-CLIP sets and also in the targets of the six miRNAs, are shown in Supplementary Data, Tables A.23 to A.38.

As expected based on the pool of targets hypothesis, most of the terms representative of GO enrichment in the induced by the transfection reagent and the Ago HITS-CLIP sets were found to be enriched in targets of one or more of the six miRNAs (Figure 6.5). For example, the most consistently enriched GO term was "signalling" (enriched in targets of four out of six miRNAs, and also in the mock transfection and the Ago HITS-CLIP sets). This was in agreement with results of KEGG enrichment analysis, where various signaling pathways were frequently observed among the 25 most enriched pathways in targets of different miRNAs (Figure 6.3). Several other selected GO terms were also enriched in targets of more than one miRNAs. Of "Cellular compartment" terms, these were "extracellular region", "cytoplasm", "endoplasmic reticulum", "plasma

membrane", "Golgi apparatus" and "cytoskeleton" (Figure 6.5a). Of the "Biological process" terms, these were "multicellular organism development", "cell adhesion", "cell cycle", "transport" and "cell communication" (Figure 6.5b).

Analyses presented in this section supports the pool of targets hypothesis, by showing that a representative selection of GO termes was recurrently enriched in targets of different miRNAs. This result was in agreement with observations of the recurrent enrichment of KEGG pathways (Figure 6.3), the enrichment of lists of targets in a set of genes induced by the transfection reagent (Figure 6.2), and multiple significant intersections between the target lists themselves (Figure 6.1). Additionally, the analysis of GO enrichment further described functions that were associated with miRNA targets in primary neurons. These functions were inhibition of genes from categories "signalling", "transport", "cytoskeleton" and etc. The nature of these categories suggests that a significant proportion of miRNA targets is likely not to be constitutively expressed, but to be inducible in certain contexts. Therefore, the results presented in this section are consistent with the proposition of miRNAs to buffer the expression of inducible genes.

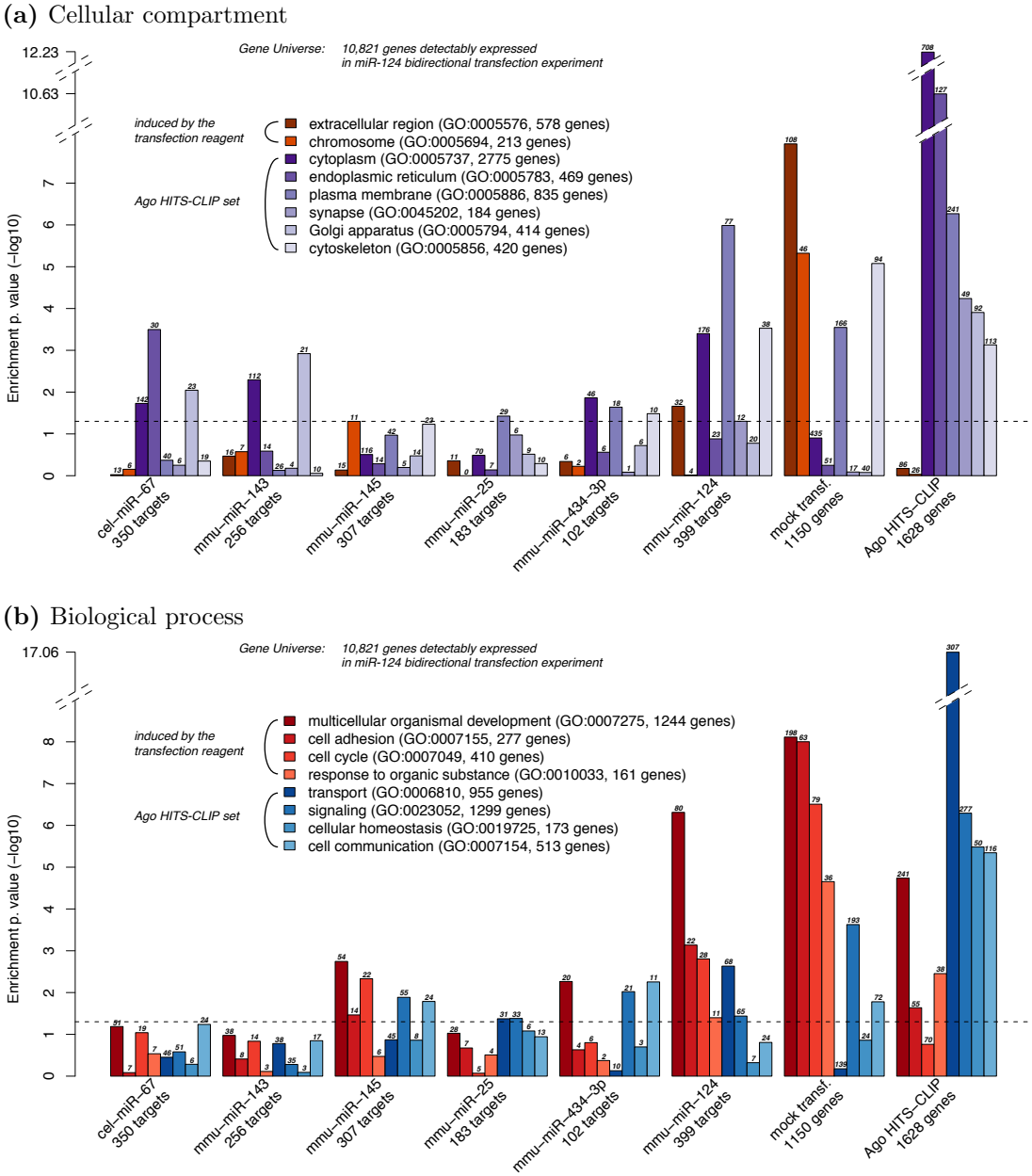


Figure 6.5: Enrichment of Gene Ontology (GO) terms in miRNA targets.
The y-axes show the enrichment P-value of the GO terms (log_{10}) that were selected from the top 40 most enriched GO terms in the genes induced by the transfection reagent of in the Ago HITS-CLIP set. The selected GO terms are listed in the plot areas. The colors that are attributed to the selected GO terms correspond to the colors of the bars for each of the gene lists (as specified on the x-axes). 6.5a – the enrichment of GO terms of the "Cellular compartment" type; 6.5b – the enrichment of GO terms of the "Biological process" type. The total number of genes from each of the selected GO terms that were identified in the corresponding list of targets is given at the top of the bars.

6.1.6 Enrichment of GO terms highlighted the importance of miR-124 and miR-434-3p

Significant intersections between the targets of unrelated miRNAs were described in previous sections. This overlap could be interpreted as an indication that all miRNAs were equally important functionally for neurons. In this section analysis of GO enrichment was used to explore the functional importance of the six transfected miRNAs, which leads to a conclusion that the two neuronal miRNAs (miR-124 and miR-434-3p) were more functionally important than other miRNAs investigated in this thesis.

In a study by Huang and colleagues on improving computational seed based miRNA target predictions, an assumption was made that functional miRNA targets "should have more consistent Gene Ontology annotations than random subsets of the sequence-based predictions" (Huang et al., 2007). A target prediction method, which was supported by the conclusion drawn from this assumption[1], was validated experimentally on the example of let-7 targets in human retinoblastoma (Huang et al., 2007). Therefore, the total number of GO enriched categories was assumed to be indicative of the functional significance of a gene list.

I have estimated the total number of enriched GO categories in targets of the six miRNAs from this thesis (as in Huang *et al.*, the type of "Biological process", size > 5 genes). This parameter was also estimated for the Ago HITS-CLIP set (section 6.1.2), where 267 GO categories were enriched at P-value threshold of < 0.05 (Figure 6.6a). This provided a benchmark for the scope of GO enrichment that may be expected in a broad set of functional miRNA targets. A similar number of GO categories (247 categories) was enriched in miR-124 transfection targets. At the same time, the number of GO categories enriched in targets of all other miRNAs (miR-434-3p, miR-143, miR-145, miR-25 and cel-miR-67) was more than two times smaller, i.e. around 100 categories and less. At a stricter P-value threshold of 0.001, there were still over twice as many GO categories enriched in miR-124 targets (32 categories) compare to miR-434-3p targets (14 categories), while targets of all other miRNAs had only one to three categories enriched (Figure 6.6b). It should be noted that this result could not be completely explained by differences in the number of identified targets for different miRNAs and resulting behaviour of the hypergeometric test (it is more difficult to obtain a significant results for smaller lists): at P-value threshold of 0.001 there were over three times as many GO categories enriched in

[1] with the focus on a "Biological process" type of GO categories of the size bigger than 5 genes (Ashburner et al., 2000)

miR-434-3p targets than in targets of any of the non-neuronal miRNAs (miR-25, miR-143, miR-145 and cel-miR-67), although the number of putative direct targets of miR-434-3p was the smallest of all miRNAs. The result was also not explained completely by the optimal timepoint to detect miR-124 and miR-434-3p targets being 6DIV, while for other miRNAs it was 4DIV (Chapter 5): analysis of 240 targets of miR-124 determined from a 4DIV experiment (a suboptimal timepoint for miR-124 target identification, see Chapter 5, section 5.2.1) showed enrichment of eight GO categories (P-value < 0.001), which was over twice as many enriched categories as was in targets of any of the non-neuronal miRNAs.

It should be pointed out that the results presented in this section are not, on their own, a definitive proof of greater functional significance of the two neuronal miRNAs for primary neurons, but they do provide additional support to the same conclusion being drawn from other results discussed in this chapter. For example, genes induced by the transfection reagent (i.e. the transfection stress) were the most significantly enriched in targets of miR-124 and miR-434-3p (Figure 6.2). Additionally, in the next section I will describe these two miRNAs as efficient inhibitors of genes that were induced in the brain by other two types of stresses (section 6.2.1). Together, these findings suggested that the neuronal miRNAs were more efficient than other miRNAs at inhibiting functionally related groups of genes (such as genes associated with GO terms, or genes co-expressed upon a treatment of the cultures).

Summary of section 6.1

Transfected miRNAs converged on inhibiting expression of a shared set of targets (which I referred to as a pool of targets primed for miRNA mediated regulation). Genes that were induced during the transfection experiments contributed significantly to the shared pool of targets: miRNA targets of five out of six miRNAs were enriched in genes induced by the transfection reagent. I also found that these genes were significantly enriched in published miR-124 transfection targets (Lim et al., 2005; Makeyev et al., 2007). These results suggest that miRNAs inhibit inducibly expressed genes, such as genes induced by the transfection reagent. Therefore targets of miRNAs (hence, their function) are defined, to a significant extent, by the experimental context.

In agreement with this proposition, targets identified by HITS-CLIP, which is a transfection-free method (Chi et al., 2009), were enriched in pathways related to a normal neuronal function rather than those affected by transfection. For example, KEGG path-

(a) Counts of the enriched terms ($P < 0.05$) **(b)** Counts of the enriched terms ($P < 0.001$)

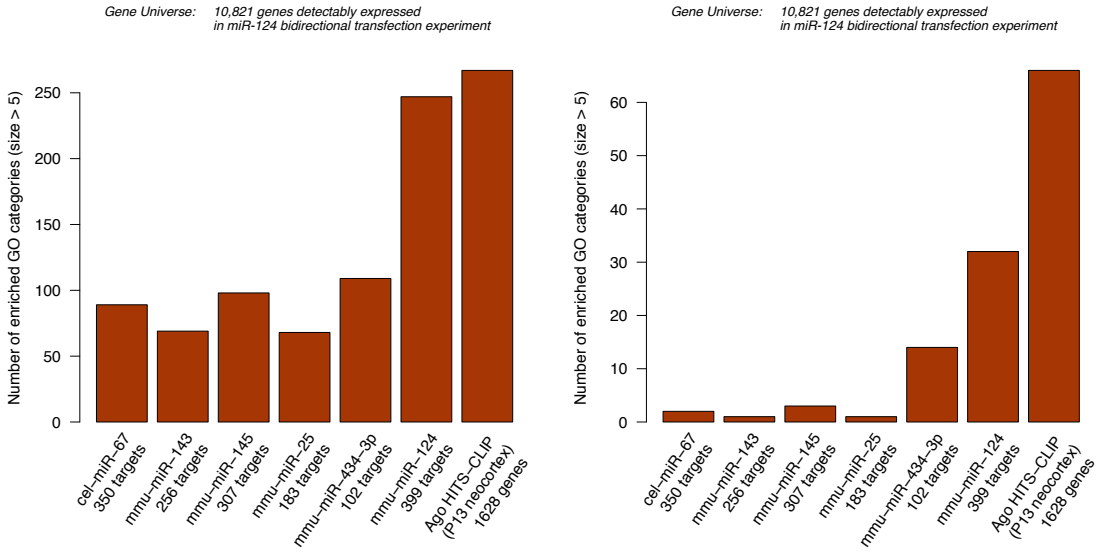


Figure 6.6: Counts of GO categories ("Biological process", size > 5 genes) enriched in miRNA targets.

ways "Long-term potentiation", "Regulation of actin cytoskeleton" and "Axon guidance" were three of the four most enriched pathways in the complete Ago HITS-CLIP set of targets. On the other hand, genes induced by the transfection reagent were enriched in disease and stress related pathways, and miRNA targets identified by transfection were also enriched in multiple disease and stress related pathways. Therefore, targets of miRNAs identified in transfection experiments may be informative of miRNA function in diseases and stresses, while HITS-CLIP targets will elucidate function in normal neurons.

Finally, neuronal miRNAs (miR-124 and miR-434-3p) were identified as having greater functional significance for primary neurons than non-neuronal miRNAs. Targets of neuronal miRNAs were more enriched in genes co-expressed in neurons than the targets of the non-neuronal miRNAs. For example, of all miRNAs, targets of neuronal miRNAs were the most significantly enriched in genes induced by the neuronal response to the transfection reagent, and also in genes associated with GO terms. Therefore, it is possible that neuronal miRNAs evolved to buffer genes that are inducible in neurons.

6.2 The function of miRNAs in neurons and the brain

6.2.1 In transfection experiments miRNAs downregulated stress inducible genes

Genes induced by transfection were enriched in KEGG pathways associated with diseases and stresses (section 6.1.3). For example, the KEGG pathways "p53 signaling pathway" (Supplementary Data, Table A.43), "Toll-like receptor signaling pathway" (Supplementary Data, Table A.48) and "pathways in cancer" (Table 6.1) were enriched among targets of several miRNAs. To test if miRNAs can inhibit a wide spectrum of genes induced by stress (i.e. not only the genes associated with KEGG terms), I obtained from published literature two lists of genes that were induced by the adverse treatments of the brain. These two published experiments were: an injection of kainate into the mouse hippocampus (Akahoshi et al., 2007) and ageing of the human brain (Lu et al., 2004).

Whole transcriptome microarray profiling data was available for the kainate injection (Akahoshi et al., 2007) and mock transfection experiments[1] (section 6.1.3), and I used Sylamer (Methods, section 2.8) to test if the innate biases in distribution of miRNA seed matching sites could be observed in these experiments. If stress upregulated genes encoded transcripts that were enriched in miRNA seed matching sites, that would suggest that there was scope for buffering of these genes by endogenous miRNAs. Such significant enrichment of miR-124 seed matching sites was observed in transcripts upregulated by both the mock transfection and kainate injection experiments, while in the latter a bias for miR-434-3p was also observed (Figure 6.7). One possible explanation to these biases can be that under normal conditions (i.e. before the kainate stress or the transfection) expression of the inducible transcripts is moderated by the endogenous miR-124 and miR-434-3p.

For the next step, I assessed whether lists of genes induced by the stress were significantly downregulated by transfected miRNAs. Genes induced by mock transfection and by the kainate stresses were derived from microarray expression profiling data[2] and I also obtained the list of genes induced by ageing of the human brain directly from a publisher's website[3]. The lists of genes induced by three types of stresses (the transfection, kainate

[1] Mock transfection experiment was performed and analysis as a part of this thesis project (section 6.1.3), and the raw microarray profiling data for the kainate injection experiment (Akahoshi et al., 2007) was available from Gene Expression Omnibus (GEO) database (Sayers et al., 2010), GEO ID GSE6388.

[2] as genes upregulated with differential expression $P < 0.05$ (Methods, section 2.7).

[3] http://www.nature.com/nature/journal/v429/n6994/suppinfo/nature02661.html. Mouse homologs of these human genes were obtained from HomoloGene Version 65 (Sayers et al., 2010).

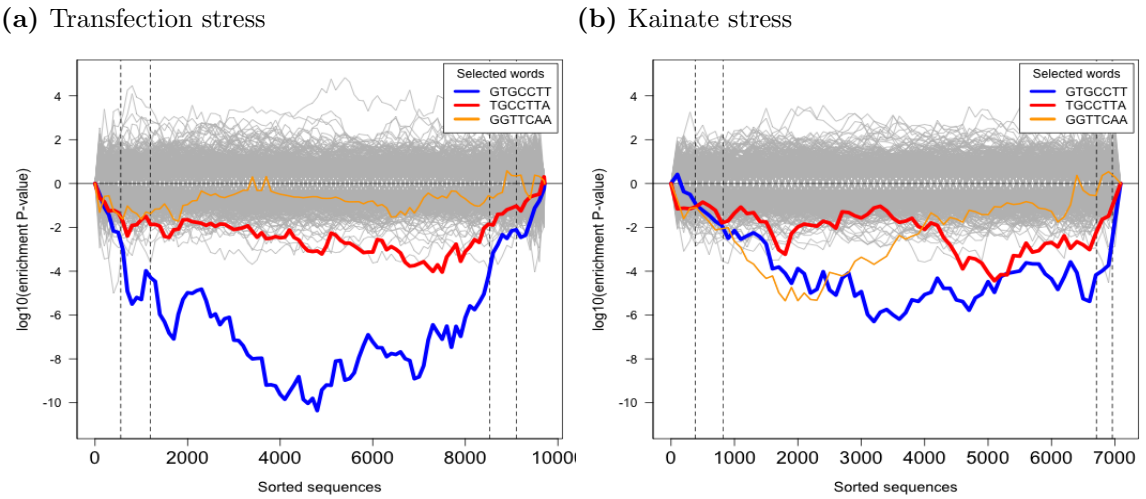


Figure 6.7: The innate miRNA seed matching site distribution biases in transfection and kainate stresses.
The x-axes represent 3′UTRs corresponding to genes sorted **from the most downregulated on the left to the most upregulated on the right** according to the fold change t-statistics in: 6.7a – the mock transfection experiment (the transfection stress); 6.7b – the kainate injection experiment (the kainate stress). The vertical dashed lines mark the P-value cutoffs (0.01 and 0.05) on both sides of the ranked gene lists. The y-axes represent the hypergeometric P-values ($-|log_{10}$(P-value)| if depletion, $+|log_{10}$(P-value)| if enrichment) for occurrence biases of 878 distinct seed matching sites (7(2) and 7(1A)-types) for a complete list of mature mouse miRNAs, which is 581 distinct miRNAs according to miRBase Release 14 (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006, 2008). Blue and red lines show enrichment profiles of 7(2) and 7(1A)-type seed matching sites for miR-124, the orange line – 7(2)-type seed matching site for miR-434-3p. The mapping of microarray probes to mRNA transcripts, and transcripts to genes, is described in Methods (section 2.7); the identification of the seed matching sites and parameters of Sylamer analysis (van Dongen et al., 2008) is in Methods (section 2.8).

and ageing stress) were frequently significantly downregulated in the miRNA transfection experiments of this thesis. The most significant relative downregulation of all three lists of the stress induced genes was achieved by neuronal miRNAs, miR-124 and miR-434-3p (Figure 6.8). Also, overexpression of miR-145 significantly downregulated genes induced in the ageing, and miR-143 – the kainate stress induced genes, while both miRNAs downregulated genes induced by the transfection stress (Supplementary Data, Figure A.7). Interestingly, neither miR-25 nor cel-miR-67 led to a significant downregulation of any of the three stress induced sets. This analysis suggested that although downregulation of stress induced genes may be a functional feature of miRNA mediated regulation as a whole, the neuronal miRNAs, such as miR-124 and miR-434-3p, may be specifically adapted for this function in the brain and neurons.

Another hypothesis generated by the analysis of genes induced by the three types of stresses, was that the stress response in neurons and in the brain may be canalised, or in

other words, the stress response is reproducible[4]. This conclusion was drawn from the observation of a significant similarity between changes in gene expression that was observed upon each of these stresses, despite their unrelated nature. The similarity manifested itself in significant intersections of the genes that were induced by the stresses[5] (Figure 6.9). The capability of several miRNAs to downregulate genes that were induced by the three unrelated stresses suggests that miRNAs may be involved in canalisation of the stress response. In agreement with this hypothesis, genes upregulated in all three stresses were targets of several miRNAs, including miR-124 (7 targets, which was ≈ 12 times more than expected by chance alone, $P < 7.6e-09$) and miR-434-3p (3 targets, ≈ 18 times more than expected by chance alone, $P < 0.00057$). The genes that were induced by all three stresses (unrestricted by the gene universe there were 19 of such genes) and their targeting by different miRNAs is listed in Table 6.2.

In summary, inhibition of genes induced by the stresses was identified as a feature of several miRNAs overexpressed in the transfection experiments. At the same time, the targeting repertoires of neuronal miRNAs (miR-124 and miR-434-3p) are likely to be more specifically adapted for inhibition of genes induced by stress of the brain and neurons, than random miRNA targets (such as the targets of cel-miR-67, for example). Moreover, a signature of the innate activity of miR-124 and miR-434-3p was discovered in the stress induction experiments, suggesting that both endogenous miR-124 and miR-434-3p may act as buffers of stress induced genes under normal circumstances. Genes that were shared between the three stresses were found to be highly enriched in targets of both miR-124 and miR-434-3p (≈ 12 and ≈ 18 times more than expected by chance alone), suggesting that miRNA mediated buffering of expression of these genes may be particularly important.

[4]Canalisation is the term that describes the reproducibility of the outcome of a developmental gene expression program, despite mutations of individual genes in the program (Siegal and Bergman, 2002; Hornstein and Shomron, 2006). I use this term to describe reproducibility of the outcome of the stress response gene expression program, despite the stress being triggered by unrelated factors.

[5]A strict test universe, consisting of 7,646 genes detectably expressed in development of both hippocampal and forebrain cultures (Chapter 3, section 3.1) was used to estimated the hypergeometric P-values for these intersections.

Symbol	Description	miR-124	miR-434-3′	miR-25	miR-143	miR-145	cel-miR-67
Anxa3	annexin A3		+		+		
Anxa5	annexin A5	+					
Prdx6	peroxiredoxin 6					+	
Ddr1	discoidin domain receptor family, member 1						
Cav1	caveolin 1, caveolae protein	+					
Cyp1b1	cytochrome P450, family 1, subfamily b, polypeptide 1						
Gfap	glial fibrillary acidic protein						
Gja1	gap junction protein, alpha 1		+				
Sdc2	syndecan 2		+	+			
Lamp2	lysosomal-associated membrane protein 2	+					
Myo10	myosin X						
Ntrk2	neurotrophic tyrosine kinase, receptor, type 2	+				+	+
Pmp22	peripheral myelin protein 22					+	
Tgif1	TGFB-induced factor homeobox 1						
Itpkb	inositol 1,4,5-trisphosphate 3-kinase B	+					
Pon2	paraoxonase 2				+	+	
Litaf	LPS-induced TN factor	+			+		
Tsc22d4	TSC22 domain family, member 4	+			+		
Wwtr1	WW domain containing transcription regulator 1						

Table 6.2: miRNA targeting of genes induced in three stress types.
The fist two columns of the table *Symbol* and *Description* give the official name symbols and descriptions of the 19 genes that were induced by all three stresses (the ageing, kainate and transfection stresses). The remaining columns indicate presence of these genes in the target lists of: miR-124 (Chapter 5, section 5.2.1); miR-434-3p (Chapter 5, section 5.2.3); miR-25 (Chapter 5, section 5.2.2); miR-143 (Chapter 5, section 5.2.2); miR-145 (Chapter 5, section 5.2.2); cel-miR-67 (Chapter 5, section 5.1.1).

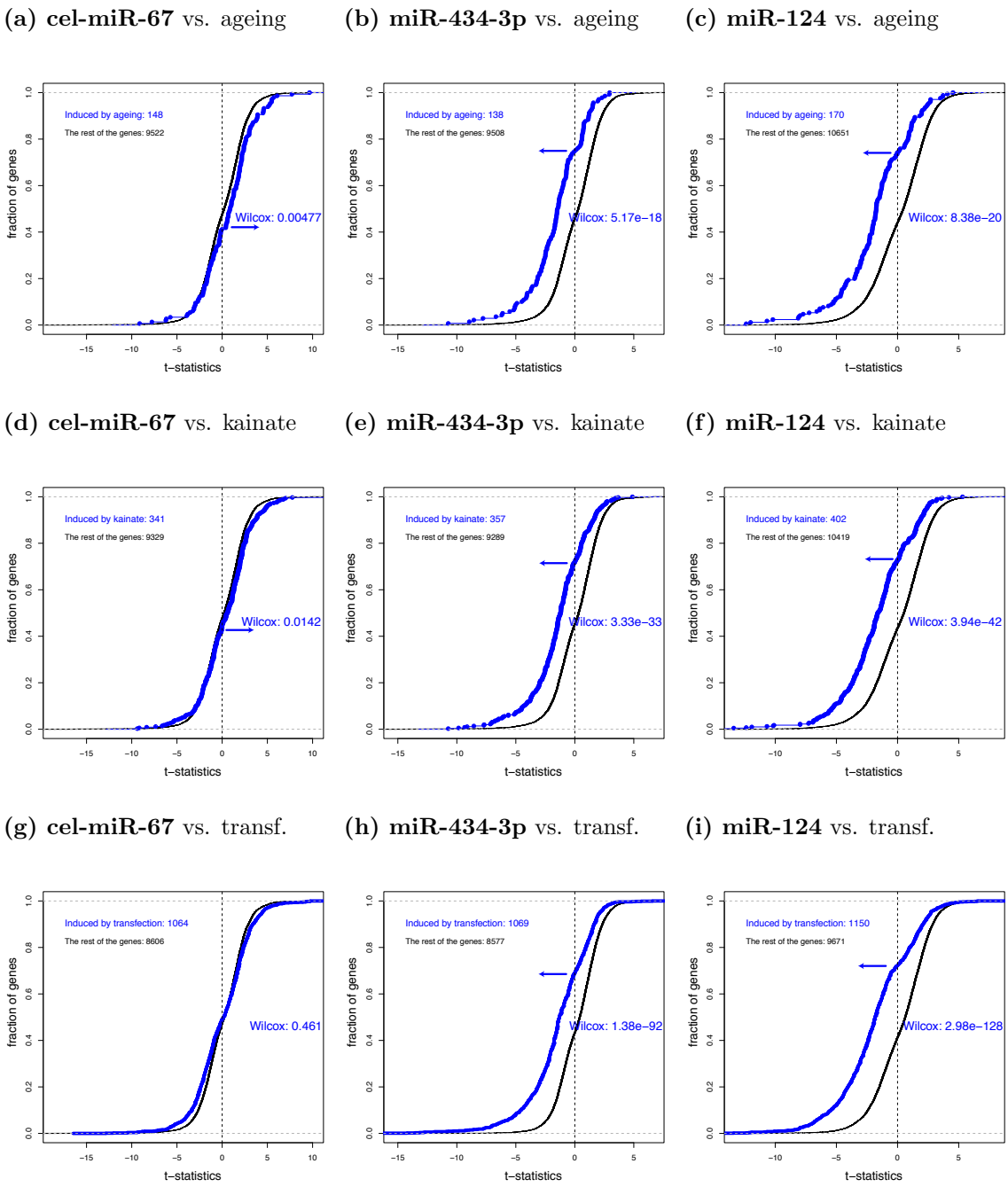


Figure 6.8: miR-124 and miR-434-3p downregulated stress induced genes.
The y-axes show the cumulative fraction of genes, the x-axes show the fold change t-statistics (Methods, section 2.7). Genes significantly induced (differential expression $P < 0.05$) by one of the three stresses (the ageing, kainate or transfection stresses) are shown as the blue line/points. The rest of the genes (except 0.01% most highly up- and downregulated genes, which were not plotted for the purpose of better scaling) is shown as the black lines. The text in the plot areas shows: 1) The number of genes *induced by* a stress that were expressed in the miRNA transfection experiments; 2) The number of other expressed genes (*The rest of the genes*); 3) The Wilcoxon test P-value for the difference in medians of the fold change t-statistics for the stress induced genes and the rest of the genes (*Wilcox*). The blue arrows show the direction of the shift in experiments where the Wilcoxon test P-value was significant ($P < 0.05$). The titles of the subfigures show: The names of the perturbed miRNAs (in bold) and the name of the stress experiment where the stress induced genes were identified.

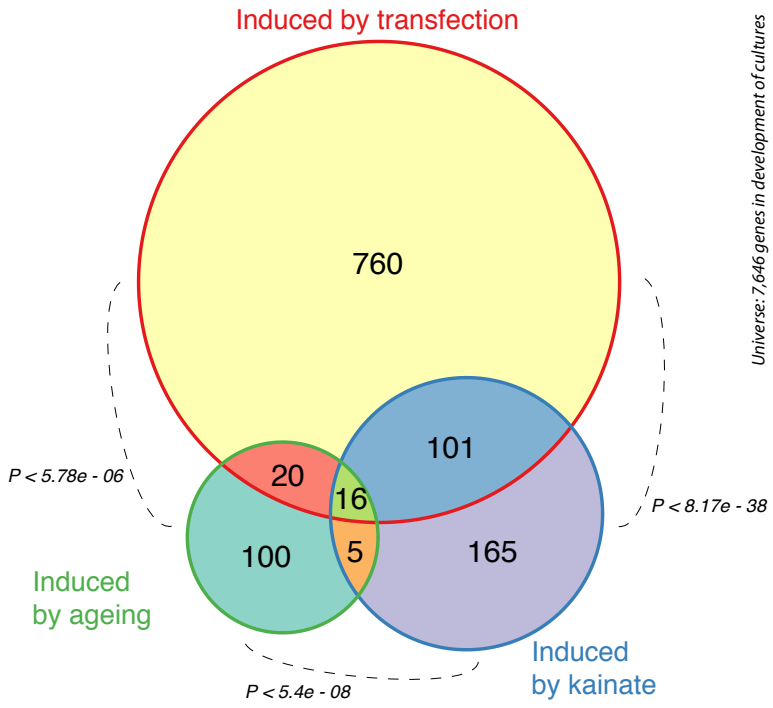


Figure 6.9: Significant intersections of genes induced in three stresses of neurons and the brain.
The Venn diagram shows the number of genes upregulated in three types of stress (the ageing, kainate and transfection and that are present in the test universe (*Gene Universe*). The test universe is 7,646 genes that were detectably expressed (using the standard Illumina detection call $P < 0.01$, see Methods, section 2.7) both in hippocampal and forebrain cultures development (Chapter 3, section 3.1). The hypergeometric test P-values are shown for each of the intersections between pairs of the lists.

6.2.2 Synaptic genes linked to neurological disease were enriched in miR-124 targets and in stress induced genes

The previous section showed that miRNAs can downregulate genes inducible by stresses in the brain and neurons. To gain an insight into how biologically important this miRNA mediated regulation of stress inducible genes may be, I evaluated the enrichment of neuronal genes linked to neurological diseases among genes that were induced by stresses and regulated by miRNAs. It has recently been demonstrated that 199 human genes encoding components of the synaptic proteome were genetically linked to neurological diseases (Bayés et al., 2011), of which 153 had mouse homologs expressed in primary forebrain cultures[1] (Chapter 3). This set of 153 genes (disease-linked genes) was assumed to be a list of neuronal genes, function of which was likely to be biologically important for the brain and neurons.

Next, I assessed whether the disease-linked genes were enriched in genes identified as stress induced. Indeed, enrichment was detected for all three lists of stress induced genes, which were discussed in the previous section (i.e. the kainate, the ageing and the transfection stresses, see section 6.2.1). Of the genes upregulated upon ageing of the brain, there were 11 disease-linked genes (approximately 4.4 times more than expected by chance alone[2], $P < 4.34e-05$), and among the genes upregulated by kainate injection there were 14 disease-linked genes induced (2.6 times more than expected, $P < 0.001$). The disease-linked genes were not significantly enriched in the genes induced by transfection (i.e. the set of 1,293 genes upregulated in mock transfected samples with P-value for differential expression < 0.05, see section 6.1.3). However, if genes induced by the transfection reagent were defined with a more stringent differential expression cutoff ($P < 0.01$), then the enrichment of disease linked genes became significant: 15 disease-linked genes were among the stricter set of genes induced by the transfection reagent (approximately 1.7 times more than expected by chance alone, $P < 0.0272$). In total, out of 153 disease-linked genes expressed in primary cultures, 31 were induced by one or more of the three stresses (Table 6.3), which was approximately 2.1 times more than expected by chance alone ($P < 3.67e-05$). These observations show that genes induced by different stresses

[1]The list of mouse homologs was retrieved directly from a publisher's website http://www.nature.com/neuro/journal/v14/n1/full/nn.2719.html#/supplementary-information.

[2]The gene universe that was used in to estimate and test the significance of enrichment of disease-linked genes in other gene lists was a complete list of genes expressed in primary forebrain cultures (9,826 genes).

were biologically important for neurons, because mutations in these genes were more frequently linked to neurological diseases than genes on average.

Similarly, the importance of miRNA regulation also manifested itself in the enrichment of disease-linked genes among miR-124 targets. Of the 153 disease-linked genes, twelve were identified as putative direct targets of miR-124 (Chapter 5, section 5.2.1), 2.4 times more than expected by chance alone ($P < 0.0041$). Interestingly, of the twelve targets of miR-124 that were linked to neurological diseases, seven were induced by one or more stresses (shown in bold in Table 6.3), which is approximately 2.9 times more than expected by chance alone, $P < 0.00281$. This observation indicated that genes, which were both biologically important for neurons (i.e. disease-linked genes) and at the same time stress inducible, were significantly more likely to be under miR-124 mediated regulation, than genes on average.

Table 6.3: Synaptic genes linked to neurological diseases and upregulated in stresses. The table provides a list of human synaptic genes that were induced by at least one of the three stresses (ageing, kainate or transfection) and linked to a neurological disease. *Symbol* – the Approved Gene Symbol (human); *Induced by* – description of a stress condition which induced the gene: *A* – the ageing stress; *K* – the kainate stress; *M* – the trasfection stress. Genes in **bold** were both induced in stresses and targeted by miR-124 (homologs of the targets identified in this thesis); *OMIM Disease Description* – the desease to to which the gene is linked in OMIM.

Symbol	Induced by	OMIM Disease Description
ALDH2	M	ALCOHOL SENSITIVITY, ACUTE
ALDH4A1	**A**	**HYPERPROLINEMIA, TYPE II; HPII**
APOE	K,M	ALZHEIMER DISEASE 2
APOE	K,M	LIPOPROTEIN GLOMERULOPATHY; LPG
APOE	K,M	MACULAR DEGENERATION, AGE-RELATED, 1; ARMD1
APOE	K,M	SEA-BLUE HISTIOCYTE DISEASE
C3	K,M	MACULAR DEGENERATION, AGE-RELATED, 9; ARMD9
CNTNAP2	M	CORTICAL DYSPLASIA-FOCAL EPILEPSY SYNDROME
CNTNAP2	M	AUTISM, SUSCEPTIBILITY TO, 15; AUTS15
CRYAB	A,M	ALPHA-B CRYSTALLINOPATHY
CST3	K	AMYLOIDOSIS VI
CST3	K	MACULAR DEGENERATION, AGE-RELATED, 11; ARMD11
DCX	K	LISSENCEPHALY, X-LINKED, 1; LISX1
DTNA	K	NONCOMPACTION OF LEFT VENTRICULAR MYOCARDIUM WITH CONGENITAL HEART
DTNA	K	NONCOMPACTION OF LEFT VENTRICULAR MYOCARDIUM, FAMILIAL ISOLATED, AUTOSOMAL
ENO3	K	GLYCOGEN STORAGE DISEASE XIII, GSD13

Continued on the next page

Synaptic genes linked to neurological diseases and upregulated in stresses.

Symbol	Induced by	OMIM Disease Description
ETFB	**M**	**MULTIPLE ACYL-CoA DEHYDROGENASE DEFICIENCY; MADD**
GFAP	A,K,M	ALEXANDER DISEASE
GJA1	A,K,M	ATRIOVENTRICULAR SEPTAL DEFECT; AVSD
GJA1	A,K,M	HYPOPLASTIC LEFT HEART SYNDROME
GJA1	A,K,M	SYNDACTYLY, TYPE III
GJA1	A,K,M	OCULODENTODIGITAL DYSPLASIA; ODDD
GNAI2	**K,M**	**VENTRICULAR TACHYCARDIA, FAMILIAL**
GPX1	M	GLUTATHIONE PEROXIDASE DEFICIENCY, HEMOLYTIC ANEMIA POSSIBLY DUE TO, INCLUDED
GRIA3	**M**	**MENTAL RETARDATION, X-LINKED 94; MRX94**
HSPB1	M	CHARCOT-MARIE-TOOTH DISEASE, AXONAL, TYPE 2F
HSPB1	M	NEURONOPATHY, DISTAL HEREDITARY MOTOR, TYPE IIB; HMN2B
HSPB8	**M**	**CHARCOT-MARIE-TOOTH DISEASE, AXONAL, TYPE 2L**
HSPB8	**M**	**NEURONOPATHY, DISTAL HEREDITARY MOTOR, TYPE IIA; HMN2A**
MYO6	K	DEAFNESS, AUTOSOMAL DOMINANT NONSYNDROMIC SENSORINEURAL 22; DFNA22
MYO6	K	DEAFNESS, CONGENITAL NEUROSENSORY, AUTOSOMAL RECESSIVE 37; DFNB37
NDRG1	A	CHARCOT-MARIE-TOOTH DISEASE, TYPE 4D; CMT4D
NDUFA2	M	LEIGH SYNDROME; LS
PC	K	PYRUVATE CARBOXYLASE DEFICIENCY
PLEC1	**K**	**EPIDERMOLYSIS BULLOSA SIMPLEX WITH MUSCULAR DYSTROPHY**
PLEC1	**K**	**EPIDERMOLYSIS BULLOSA SIMPLEX WITH PYLORIC ATRESIA**
PLEC1	**K**	**EPIDERMOLYSIS BULLOSA SIMPLEX, OGNA TYPE**
PLP1	A	SPASTIC PARAPLEGIA 2, X-LINKED; SPG2
PLP1	A	PELIZAEUS-MERZBACHER DISEASE; PMD
PTPN11	**K**	**NOONAN SYNDROME 1; NS1**
PTPN11	**K**	**LEOPARD SYNDROME 1**
RDX	A	DEAFNESS, AUTOSOMAL RECESSIVE, 24; DFNB24
SLC4A4	A	RENAL TUBULAR ACIDOSIS, PROXIMAL, WITH OCULAR ABNORMALITIES AND MENTAL
TPP1	K	CEROID LIPOFUSCINOSIS, NEURONAL, 2; CLN2
VCAN	A	WAGNER SYNDROME 1; WGN1

Continued on the next page

Synaptic genes linked to neurological diseases and upregulated in stresses.

Symbol	Induced by	OMIM Disease Description
WFS1	A	DEAFNESS, AUTOSOMAL DOMINANT NONSYNDROMIC SENSORINEURAL 6; DFNA6
WFS1	A	WOLFRAM SYNDROME 1; WFS1
WNK1	A	PSEUDOHYPOALDOSTERONISM, TYPE II; PHA2

6.2.3 miR-124 in development: Reduction of variability in gene expression

In this thesis I have identified several aspects of miR-124 mediated regulation that are changing with progression of development of the cultures. For example, transcripts with 3′UTRs that are not depleted of miR-124 seed matching sites, were found to be upregulated relatively late in development of cultures (from 4DIV to 8DIV, see Chapter 3, Figure 3.10). At the same time, the effect of miR-124 overexpression on gene expression in primary neuronal cultures consistently increased as development progressed, and was maximal in the experiment at the 6DIV timepoint (Chapter 5, section 5.2.1). Together with these developmental changes, genes that were induced by three types of stresses were significantly downregulated by miR-124 (Figure 6.8), therefore I investigated expression of stress induced genes in development of cultures.

Previously, miRNAs were suggested to be involved in canalisation of developmental gene expression programs (Hornstein and Shomron, 2006), therefore I investigated variability of stress induced genes in development of primary cultures. The standard deviation of intensities of microarray probes corresponding to these genes (the mapping of the probes is described in Methods, section 2.7) between the replicates within one developmental timepoint was used as a measure of the variability of gene expression.

Standard deviation in intensities of probes for stress induced genes were compared to that of the rest of the genes[1]. Additionally, as an internal control, I estimated standard deviations in intensities of probes for genes, expression of which was reduced in the stress[2]. A progressive increase in standard deviation (i.e. variability) of stress induced

[1]To avoid a bias from correlation of the standard deviation and the level of expression, the analysis was confined to top 25% most highly expressed genes, where no such correlation was observed (Supplementary Data, Figures A.8 and A.9).

[2]Genes that were reduced in the mock transfection and kainate experiments were defined as genes that were downregulated with differential expression $P < 0.05$. Genes that were reduced by ageing were obtained directly from a publisher's website (http://www.nature.com/nature/journal/v429/n6994/suppinfo/nature02661.html). Mouse homologs of these human genes were obtained from HomoloGene Version 65 (Sayers et al., 2010).

genes was observed in development of both hippocampal and forebrain cultures (Figure 6.10). Importantly, variability of the stress reduced genes did not increase in all cases but one[3] (Figure 6.10).

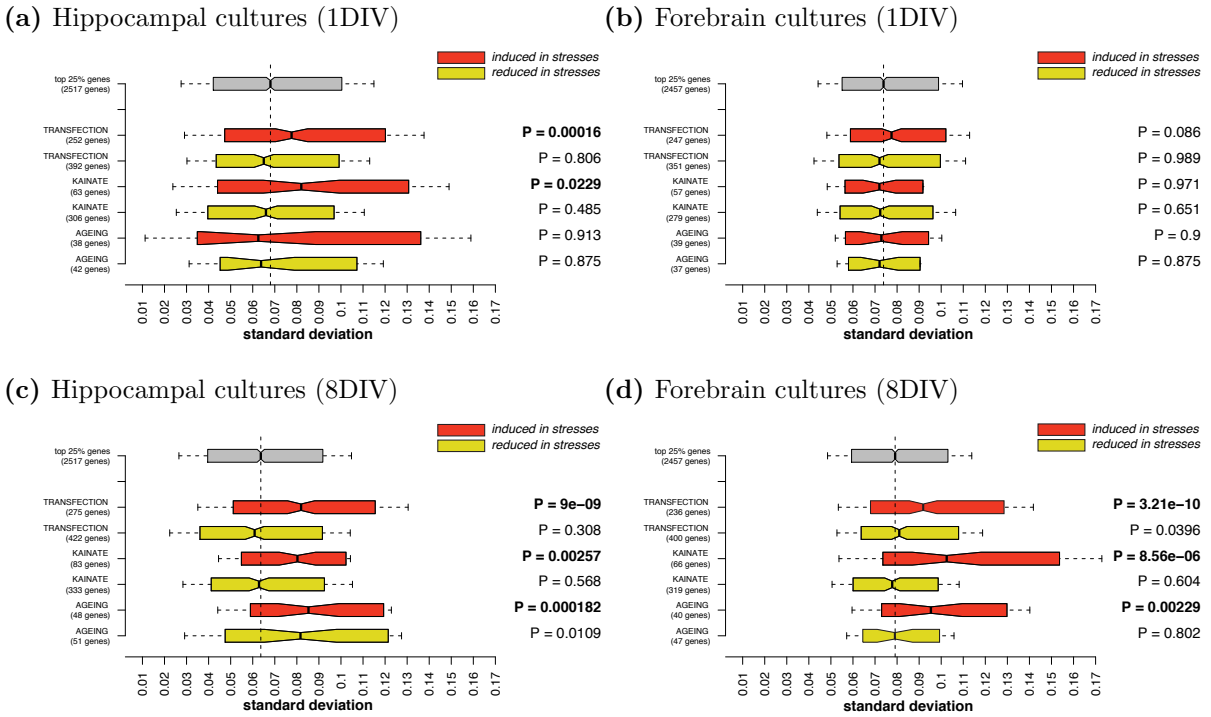


Figure 6.10: Variability in expression of stress induced genes increased with progression of development.
The x-axes show the standard deviation of gene expression between biological replicates within one developmental timepoint, the y-axes show the names of the gene lists and the number of the genes from the lists that were selected from the top 25% most highly expressed genes in the experiment (the numbers are in parentheses). The grey boxes correspond to the distribution of standard deviations of all 25% most highly expressed genes, all other boxes correspond to the genes belonging to the stress genelists (as indicated on the y-axes and according to the color-scheme). The notches in the boxes correspond to the median value of the distribution of standard deviations of the genes in the genelists. The left and the right sides of the boxes correspond to the first and the third quartiles. The whiskers extend to no more than 0.25 times the interquartile range (IQR), or to the most extreme data-point, if it is closer to the median than 0.25 IQR. The Wilcoxon test P-values are for the difference between the medians of the distributions of the genes within the gene lists (confined to the 25% most highly expressed genes) and the rest of the genes in the top 25%.

Next, I investigated whether miR-124 could play a role in controlling this variability of the stress induced genes. In order to do this, the standard deviation of probe intensities was estimated between replicates within the following treatments: transfection with the

[3]In hippocampal cultures at the 8DIV timepoint, variability in expression of genes reduced by ageing was significantly higher than for the rest of the genes. This difference, however, was smaller than that of the genes induced by ageing.

mimics of miR-124, transfection with the inhibitor of miR-124, transfection with the mimic of cel-miR-67 and a mock transfection. In samples transfected with a mimic of miR-124 the variability of the stress induced genes was not significantly different from the rest of the genes[4] (Figure 6.11a). On the other hand, in mock transfected samples or samples transfected with an inhibitor of miR-124 the variability of stress induced genes remained significantly higher than of the rest of the genes (Figures 6.11b and 6.11c). Transfection of a mimic of a non-mouse miRNA, cel-miR-67[5], at the same developmental timepoint and with the same protocol (Methods, section 2.5), did not reduce the variability in expression of the stress induced genes (Figure 6.11d).

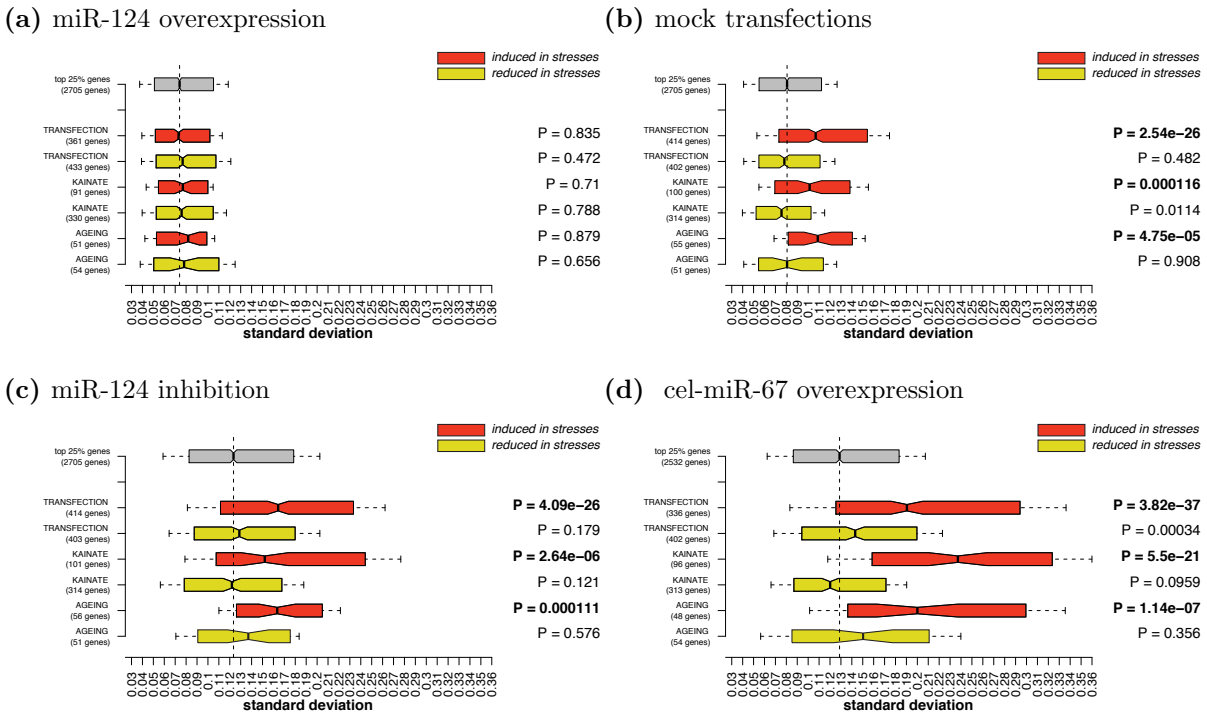


Figure 6.11: Variability in expression of stress induced genes was reduced by miR-124 overexpression.
The x-axes show the standard deviation of gene expression between biological replicates within one treatment (the treatments are named in the titles to the subfigures). See the legend to Figure 6.10 for the description.

[4]As in the case of the developmental profiling, the analysis was confined to the top 25% most highly expressed genes, where no significant correlation between the level and standard deviation of gene expression was observed in miR-124 and cel-miR-67 transfection experiments (Supplementary Data, Figures A.10 and A.11).

[5]Of the two cel-miR-67 overexpression experiments at 6DIV (Methods, section 2.5), the data from the experiment "B" was used for analysis in this section, as in the experiment "B" bigger effects on differential gene expression were observed than in the experiment "A" (compare Figure 5.1d and Figure 5.1e).

In summary, in relatively mature primary neurons (8DIV) variability in expression of genes induced by three types of stress (transfection, kainate and ageing stresses) was found to be higher than for other genes. Overexpression of miR-124 was found to reduce this variability (Figure 6.11). Therefore, it is possible that endogenous miR-124 can also act to limit variability in expression of stress induced genes in mature neurons.

Summary of section 6.2

The analyses described in this section showed that exogenously introduced miRNAs inhibited genes that were induced in three types of stress to both neurons and the brain (the transfection, kainate and ageing stress). This inhibition was most significant in the case of the two neuronal miRNAs (miR-124 and miR-434-3p), while it was not significant upon transfection of the non-mouse miRNA, cel-miR-67, or the oncogenic miRNA, miR-25. Therefore, targeting repertoires of endogenous miRNAs may have specifically evolved to inhibit genes that can be induced by the stresses of their cognate tissues and cell types.

Sylamer analysis of occurrence biases of miRNA seed matching sites in 3′UTRs of transcripts in the mock transfection experiment (transfection stress) revealed the innate signature of miR-124 mediated regulation (seed matching sites for miR-124 were enriched in 3′UTRs of transcripts upregulated by the treatment with the transfection reagent). A very similar distribution of miR-124 seed matching sites was observed upon the kainate stress. Additionally, enrichment of seed matching sites for miR-434-3p was also revealed. One possible explanation of these observations, is that in normal circumstances endogenous miR-124 and miR-434-3p can act as buffers of genes inducible by stress.

The stress induced genes targeted by miR-124 were enriched in synaptic genes and linked to neurological diseases. Therefore, the role of miR-124 in regulation of expression of stress induced genes, is likely to be of importance for neuronal biology.

In more mature primary neurons (8DIV), variability in expression of genes induced by the stresses was higher than that of the rest of the genes. Interestingly, overexpression of miR-124 suppressed this variability, while overexpression of cel-miR-67 did not. I suggest that endogenous neuronal miRNAs also regulate variability of stress inducible genes. This proposition is speculative, however it can be tested experimentally (Discussion, section 7.4).

Returning to the questions posed at the beginning of this chapter, I believe my results show that neuronal miRNAs have an important function in maintaining equilibrium in neuronal gene expression. I analysed hundreds of targets for miRNAs in neurons, and

demonstrated that targets of different miRNAs converged on the pool of transcripts that was, to a signficant extent, induced by the transfection procedure itself. At the same time, I have found that targeting repertoires of two neuronal miRNAs, miR-124 and miR-434-3p, seem to be adapted for inhibition of incucible in neuorns and the brain genes. Also, miR-124 and miR-434-3p appear to be endogenous buffers of neuronal transcription. This is important for regulation of stress response and has implications for disease. In summary, my data leads me to conclude that neuronal miRNAs act to maintain gene expression of their targets within steady state boundaries.

Chapter 7

Discussion

The aim of this thesis was to identify the roles of miRNAs at the level of the whole transcriptome in neurons (Introduction, section 1.2). Through the profiling of mRNA and miRNA expression in primary neuronal cultures I established that the cultures were a suitable model system to study miRNAs in neurons (Chapter 3). Secondly, profiling miRNA abundance in the cultures allowed me to characterise the dynamic expression of miRNAs during neuronal growth and to contrast this with patterns of mRNA expression (Chapter 3, section 3.2). Additionally, it allowed me to select miRNAs for further study in primary neurons (Chapter 3, section 3.3).

Based on previously published works, it was known that transcripts destabilised upon miRNA overexpression and which contained seed matching sites in their 3′UTRs, were enriched in direct targets (Lim et al., 2005; Giraldez et al., 2006). I used this approach to identify putative direct targets of the selected miRNAs. Primary neuronal cultures were transfected with mimics and inhibitors of nine mouse miRNAs and with a mimic of one non-mouse miRNA (Chapter 5). Subsequent differential expression was detected using mRNA microarrays, and putative direct miRNA targets were derived using seed matching site enrichment as an indication of miRNA mediated regulation. With this strategy I identified hundreds of putative direct targets for six of the selected miRNAs (Chapter 5, section 5.2). These miRNAs were: neuronal miR-124 and miR-434-3p, non-neuronal miR-143, miR-145 and miR-25, and cel-miR-67 (a miRNA not present in the mouse). Functions associated with the lists of targets of these miRNAs were characterised through KEGG and GO enrichment analyses. This identified several biological processes associated with miRNA targets, including processes related to cell signalling, transport and cytoskeleton remodelling (Chapter 6, sections 6.1.3 and 6.1.5).

Further analysis of these targets lead to the conclusion that transfected miRNAs and most significantly the two neuronal miRNAs (miR-124 and miR-434-3p), act to inhibit genes induced during the transfection experiments (Chapter 6, Figure 6.2). Therefore, the role of exogenously added (transfected) miRNAs is context dependent, i.e. their targeting repertoires were to a significant extent defined by the context of genes that were induced in the cultures by the transfection procedure (Chapter 6, section 6.1.3).

The identified targets were enriched in multiple disease and stress related KEGG pathways (Chapter 6, section 6.1.3). Therefore, I tested whether transfected miRNAs inhibited genes that were induced by stresses other than transfection, but not necessarily associated with KEGG pathways. Indeed, genes that were induced in the brain by two types of stress (the kainate (Akahoshi et al., 2007) and ageing (Lu et al., 2004) stress), were inhibited by some miRNAs that were transfected in the primary cultures (Chapter 6, section 6.2.1). Importantly, this inhibition was most significant for the two neuronal miRNAs (Chapter 6, Figure 6.8). miRNA mediated regulation of stress inducible genes is likely to be of biological importance, because a significant fraction of these genes was previously genetically linked to neurological disease (Chapter 6, section 6.2.2).

Although the function of miRNAs as inhibitors of stress inducible genes was identified for the transfected miRNAs, several observations suggested that endogenous miRNAs can also inhibit (buffer) inducibly expressed genes. For example, if miR-124 were to act as a buffer, constraints to differential gene expression imposed by endogenous miR-124 can explain the observation that transfected non-neuronal miRNAs have a less significant impact on gene expression in relatively mature cultures (Chapter 5, section 5.1.3). In addition, upon transfection induced stress, the induced genes are specifically enriched in miR-124 seed matching sites; Furthermore, genes induced by kainate stress are specifically enriched in miR-124 and miR-434-3p seed matching sites (Chapter 6, section 6.2.1). The latter observation is consistent with inhibition of stress inducible genes by the endogenous miR-124 and miR-434-3p in normal conditions. This observation leads me to believe that endogenous miRNAs function to reduce variability in expression of stress inducible genes. Such a reduction was indeed detected upon transfection of the miR-124 mimic (Chapter 6, section 6.2.3). Analysis of variability in the expression of stress induced genes gives biological credence to miRNAs acting to buffer stress inducible genes (Chapter 6, section 6.2.2). Expression of these stress inducible genes was found to be more variable between biological replicates than that of other genes, and exogenous miR-124 reduced this variability. This observation suggests that stress inducible genes may require an additional level of post-transcriptional control, because their expression is inherently “noisy”.

The precise control of expression of these genes is likely to be of importance, because a significant proportion of the stress inducible genes were previously linked genetically to neurological diseases (Chapter 6, section 6.2.2).

Below I will discuss these results in more detail. Additionally I will outline further work to test hypotheses proposed in the course of this study.

7.1 Characterising the experimental system

Before proceeding to the identification of miRNA targets in primary neuronal cultures, I first established that E17.5 primary forebrain cultures were a suitable model to study functions of miRNAs in neurons[1] (Chapter 3).

First, I described trends in gene expression during development of the cultures at four timepoints (at 1DIV, 2DIV, 4DIV and 8DIV). I profiled abundance of mRNAs using microarrays and characterised the associated function of differentially expressed genes through GO term and KEGG pathway enrichment (Chapter 3, section 3.1.3) (Manakov et al., 2009). Among terms and pathways which were found to be significantly upregulated during the development of cultures were "synapse", "neurological system process" and "Long-term potentiation". The nature of these upregulated categories suggested that they were enriched in neuritic genes. At the same time, pathways "DNA replication" and "Cell cycle" were significantly downregulated (Chapter 3, section 3.1.3). These downregulated categories suggested that they were enriched in somatic[2] genes, as the associated biological process are taking place in the nucleus. Together these observations show that the development of cultures was dominated by neuritic expansion, and it was unlikely that proliferating secondary cell types (e.g. fibroblasts, endothelial cells and etc.) contributed significantly to the mRNA profiles of the cultures. When expression trends of downregulated genes were overlayed with upregulated genes over the course of development, the intersection was at the 4DIV timepoint (Chapter 3, Figure 3.7). In other words, after 4DIV, the overall abundance of neuritic transcripts was higher than the abundance of somatic transcripts. A ratio of neurtic to somatic transcript abundance above one, is likely to be similar to that of mature forebrain neurons, as they are characterised by extensive arborisation. Therefore, 4DIV can be viewed as a timepoint at which a developmental

[1] Before profiling gene expression there were good reasons to consider E17.5 as a predominantly neuronal culture, rather than glial. Performing brain dissections at 17.5 days of mouse prenatal development and using a specifically optimised cell culture protocol already favored survival of neurons over that of glial cells (Introduction, section 1.2.4).

[2] By somatic, I refer to the soma of a neuron.

switch occurs in the transition from immature to mature gene expression of primary cultures[3]. miRNAs were previously demonstrated to be of importance for developmental switch timepoints (Giraldez et al., 2005). I previously proposed that miRNAs play a similar role during neuronal development (Manakov et al., 2009).

I demonstrated that miRNA expression trends in the development of primary forebrain cultures were in agreement with previously published reports of miRNA activity in differentiated neurons (Chapter 3, section 3.2). Abundances of 362 mature miRNAs were profiled (at 1DIV, 2DIV, 4DIV and 8DIV) using a miRNA microarray platform, and three major expression trends were identified (Manakov et al., 2009): 1) The upregulated miRNAs (105 sequences); 2) The downregulated miRNAs (99 sequences); 3) The steady state highly expressed miRNAs (26 sequences) (Chapter 3, section 3.2.1). The steady state highly expressed category included several miRNAs that were previously shown to be induced upon neural differentiation (these included miRNAs of the let-7 family, miR-124 and miR-125) (Chapter 3, Table 3.3). The observation that these miRNAs are highly expressed as early as 1DIV was in agreement with the plating material being comprised, at least to a large extent, of committed neural cell types. Moreover, some of these miRNAs, for example miR-124, were previously shown to be neuron specific (Shkumatava et al., 2009; Hanina et al., 2010). Therefore their high steady state expression throughout the developmental timecourse was in agreement with these cultures consisting predominantly of differentiated neurons throughout the 1DIV to 8DIV developmental time-window. Analysis of differentially expressed miRNAs further supported the use of primary forebrain cultures as an accurate model system to study forebrain neurons (Chapter 3, section 3.2.2). For example, the most strongly downregulated miRNAs in culture (e.g. miR-143, miR-145, and etc.) were previously shown to be lowly expressed in forebrain synapses in comparison to the whole brain homogenate (Lugli et al., 2008; Siegel et al., 2009). On the other hand, some of the upregulated miRNAs were previously demonstrated to be significantly enriched in the adult brain (e.g. miRNAs of the mouse distal 12 cluster (Seitz et al., 2004)) or to be induced by neuronal activity (e.g. miR-132 (Klein et al., 2007)).

Profiling miRNA and mRNA expression in the development of cultures suggested that miRNAs can directly shape gene expression in the cultures (Chapter 3, section 3.2.3). Significant depletion of seed matching sites for miRNAs highly expressed in the cultures (miR-124 and let-7 family miRNAs) was observed in 3′UTRs of highly abundant tran-

[3]Additionally, by characterising development of hippocampal and forebrain cultures in parallel, the development of forebrain cultures was shown to be highly similar to that of more commonly used hippocampal cultures (Chapter 3, sections 3.1.1 and 3.1.2). This observations supported the use of primary forebrain cultures as a model of growing neurons.

scripts (Chapter 3, Figure 3.9) (Manakov et al., 2009). This finding was in agreement with the reported role of miRNAs as major modulators of tissue and cell-type specific gene expression profiles (Farh et al., 2005; Sood et al., 2006). Significant biases in the distribution of the seed matching sites for miR-124 were also observed in differentially expressed genes. Transcripts that were upregulated early in development (between 1DIV and 2DIV) were depleted in miR-124 seed matching sites (Chapter 3, Figures 3.10c and 3.10d). Therefore, it was unlikely that endogenous miR-124 would inhibit genes upregulated in early stages of the development, at timepoints associated with the initial spurt of neurite growth and early synaptogenesis events (Valor et al., 2007). However, as cultures matured, opportunities appeared for endogenous miR-124 to constrain the expression of the upregulated genes: 3′UTRs of the transcripts that were upregulated in transition between 4DIV to 8DIV were either not depleted in miR-124 seed matching sites (forebrain cultures) or were enriched (hippocampal cultures) in these sites (Chapter 3, Figures 3.10a and 3.10b).

Lastly, profiling of trends in miRNA expression during the development of cultures enabled me to make a selection of miRNAs with distinct expression modes for the identification of miRNA targets. Two miRNAs from the steady state highly expressed category were selected (miR-124 and miR-103), three – from the downregulated category (miR-143, miR-145 and miR-25) and four – from the upregulated category (miR-434-3p, miR-370, miR-551b and miR-410). The selection procedure is described in Chapter 3, section 3.3. Additionally, one miRNA that was not related to any of the known mouse miRNAs (a *Caenorhabditis elegans* miRNA, cel-miR-67) was also selected. The non-mouse miRNA was selected in order to identify targets that were equivalent to a random sample of the transcripts that were susceptible to miRNA mediated regulation (without constraints imposed by the evolutionary selection). Profiling of mRNA expression in development of cultures identified timepoints at which to conduct experiments for the identification of miRNA targets: the 4DIV timepoint was selected because of its importance as a switch timepoint in the developmental gene expression program (see above), and two timepoints were picked around the 4DIV timepoint (3DIV and 6DIV) in order define a timepoint at which to derive optimal results (see below).

7.2 Identification of miRNA targets

It has previously been demonstrated that the introduction of exogenous miRNAs directly target transcripts whose 3′UTRs contain seed matching sites for that miRNA (Lim et al.,

2005; Giraldez et al., 2006). Conversely, inhibition of a miRNA through transfection of an inhibitor, causes upregulation of its targets (Conaco et al., 2006). Hence, I decided to transfect primary neuronal cultures with miRNA mimics and inhibitors and attempt to derive lists of putative direct targets from these experiments. miRNA targets are identified by selecting transcripts that are differentially expressed upon miRNA perturbation and also contain seed matching sites for the perturbed miRNAs.

Before conducting experiments to identify miRNA targets, it was necessary to identify the developmental timepoints which would enable the most efficient identification of the targets. To do this, I conducted a series of transfection experiments[1] with the mimics of miR-124, miR-143, miR-145, cel-miR-67 and also with the inhibitor of miR-124 (Chapter 5, sections 5.1.1 and 5.1.2). These transfection experiments were performed at either 3DIV, 4DIV or 6DIV (see above). The transfection of miR-124 elicited differential gene expression, characterised by significant enrichment of miR-124 seed matching sites in 3′UTRs of downregulated transcripts (Chapter 5, sections 5.1.2 and 5.2.1). This was an indication that the exogenously added miR-124 directly caused a significant proportion of the observed changes in gene expression.

In the case of miR-124 transfection experiments, the enrichment of miR-124 seed matching sites was most significant at 6DIV (Chapter 3, section 5.1.2). This observation suggested that the 6DIV timepoint was the stage at which the direct contribution of miR-124 to differential gene expression was most significant. Therefore 6DIV was selected to be the optimal timepoint for the identification of putative direct targets of miR-124 and of other neuronal miRNAs (see below). Transcripts, which were downregulated by the transfection of the mimics of non-neuronal miRNAs (miR-143, miR-145 and cel-miR-67), were significantly enriched in seed matching sites for the transfected miRNA in a majority of the experiments (Chapter 5, section 5.1.1). However, these enrichments were more significant at 3DIV and 4DIV, rather than at 6DIV (Chapter 5, Figure 5.2). The maximal enrichment of seed matching sites was detected at 4DIV, therefore this was selected as the best timepoint at which to identify targets of non-neuronal miRNAs.

The enrichment of the seed matching sites for miR-124 was more significant in bidirectional perturbation experiments (Chapter 4, section 4.2.1). These experiments directly contrast overexpression and inhibition of the same miRNA. This method was favoured over the alternative approach involving contrasting mimic with mock transfected cultures

[1]These transfections were performed with an siRNA transfection protocol (Maclaren et al., 2011). I confirmed that this protocol was efficient for transfection of neurons in primary forebrain cultures by imaging cultures transfected with eGFP expressing plasmid and a fluorophore labelled oligonucleotide (Chapter 4, section 4.1.1).

(the unidirectional contrast). From the bidirectional experiment at 6DIV, I compiled a list of 399 putative direct miR-124 targets in primary neurons (Supplementary Data, Table A.9). The targets were defined as significantly downregulated genes (differential expression $P < 0.01$) that encode transcripts harbouring miR-124 seed matching sites in 3′UTRs. This approach is further validated by significant intersections of identified miR-124 targets and targets from previously published works (Chapter 5, section 5.3). Therefore, I believe that both the bidirectional transfection strategy and the identified list of 399 miR-124 targets are useful beyond the scope of this work, and publication of these data will be of use to scientific community.

Using data from the miR-124 bidirectional transfection experiments, I optimised the original siRNA transfection protocol (Maclaren et al., 2011) by adjusting the posttransfection incubation time and the cell plating density. These adjustments improved detection of differential expression of seed matching site containing transcripts in the bidirectional contrast (Chapter 4, section 4.2.2). With the adjusted protocol I performed the bidirectional transfection experiments on the remaining selected mouse miRNAs[4]. As a result of these bidirectional experiments, lists of putative targets were compiled for four miRNAs: 251 targets of miR-143 (Supplementary Data, Table A.10), 301 targets of miR-145 (Supplementary Data, Table A.11), 169 targets of miR-25 (Supplementary Data, Table A.12) and 101 targets of miR-434-3p (Supplementary Data, Table A.14). To my knowledge, this is the first report of direct targets of miR-434-3p, despite the fact that this miRNA is transcribed from the mouse chromosome 12 distal region (Davis et al., 2005), a region highly expressed in the adult brain relative to other organs (Seitz et al., 2004), and which has previously been implicated in cognitive disfunction (Lewis and Redrup, 2005).

Additionally, identification of targets of a non-mouse, non-neuronal miRNA have not previously been reported in mouse primary neuronal cultures. Therefore, targets of cel-miR-67 will be useful for researchers of neuronal miRNAs as a control of specificity for neuronal miRNA targets identified in the future. Indeed, I have been approached by many researchers already interested in these datasets. Although these targets are described within this thesis, I also intend to publish these data in the near future to be made available as a resource to the community.

[4]Targets of cel-miR-67 were derived from a unidirectional overexpression experiment (i.e. a contrast of cultures transfected with the mimic of cel-miR-67 with mock transfected cultures), because it was not represented in the mouse genome and so its inhibition was not possible. The experiment at 4DIV (marked with the index “A”, Figure 5.2d), which resulted in the highest enrichment of the seed matching sites for cel-miR-67, was used to derive the list of 394 putatively direct targets of cel-miR-67 in mouse primary neurons (Chapter 5, section 5.2.2 and Supplementary Data, Table A.13).

7.3 Context dependent function of miRNAs

Analysis of miRNA targets identified in this thesis showed that transfected miRNAs inhibited transfection induced genes, and that the identification of targets was dependent on the context of genes induced during the transfection experiments. The targeting repertoires of neuronal miRNAs were found to be better adapted for the inhibition of genes that were induced by stress in primary cultures and the brain, than targeting repertoires of random miRNAs. In the first part of this section I will discuss this context dependent inhibition of inducible genes by transfected miRNAs. In the second part, I will describe a line of evidence that supports a similar role for endogenous neuronal miR-124 and miR-434-3p in neurons and the brain.

Exogenously added (transfected) miRNAs are context dependent inhibitors of inducibly expressed genes

The hypothesis of context dependent miRNA-mediated regulation was prompted by significant intersections between targets of six unrelated miRNAs and genes induced by the transfection reagent (Chapter 6, sections 6.1.1, 6.1.2 and 6.1.3). Additionally, various GO terms and KEGG pathways were identified to be frequently enriched in both transfection induced genes and in targets of multiple miRNAs. Terms related to cell signaling, molecular transport and cytoskeleton remodelling were frequently enriched in targets and in the transfection induced genes (Chapter 6, sections 6.1.3 and 6.1.5).

The genes induced by the transfection reagent were also enriched in published miR-124 targets derived similarly (Chapter 3, section 6.1.4). At the same time, these genes were not enriched in a large independent set of miRNA targets (the Ago HITS-CLIP set) identified in P13 neocortex with a transfection-free method (Chapter 6, Figure 6.4). This observation shows that the repertoire of miRNA targets identified in transfection experiments was dependent on the context of genes induced by the experimental procedure.

The targets for miR-124 and miR-434-3p were the most significantly enriched in the genes induced by the transfection reagent. In total, 34.3% (3.3 times more than expected by chance alone, if the intersections were determined within the experimental test universe) and 37.5% (3.5 times more than expected) of miR-124 and miR-434-3p targets were induced by the transfections (enrichment P-values were $2.34e-41$ and $3.56e-14$, respectively). At the same time, enrichment was weakest in targets of a non-mouse miRNA, cel-miR-67, and an oncogenic miRNA, miR-25 (Poliseno et al., 2010) (Chapter 6, Figure 6.2).

To test if miRNAs inhibited genes that were induced by stresses other than transfection, I compiled lists of genes that were induced by two additional types of brain stress: the injection of a kainate into the mouse hippocampus (kainate stress) (Akahoshi et al., 2007), and ageing of the human brain (ageing stress) (Lu et al., 2004). Transfection of primary cultures with neuronal miRNAs, miR-124 and miR-434-3p, was found to significantly downregulate genes that were induced in kainate and ageing stresses (as well as the genes induced by the transfection reagent). At the same time, transfection of the non-mouse miRNA, cel-miR-67, did not significantly downregulate any of the sets. This observation suggested that the targeting repertoire of neuronal miRNAs was better adapted to inhibit genes inducible by stress in the brain than targets of a random miRNA.

In summary, the observations discussed above showed that a significant proportion of the targets of different transfected miRNAs were shared, and that almost all transfected miRNAs converged on the inhibition of genes that were induced by the transfection reagent. Therefore, inhibition of inducible genes may be a common feature of miRNA mediated regulation as a whole. At the same time, the neuronal miRNAs, miR-124 and miR-434-3p, were most efficient in causing the widespread inhibition of genes induced by the transfection reagent in primary cultures, as well as inhibition of the genes induced by the other two types of stress. Therefore, targeting repertoires of neuronal miRNAs may be specifically adapted to inhibit genes that can be induced in neurons and the brain.

Endogenous miR-124 and miR-434-3p as buffers of inducible genes in neurons and the brain.

Although the experiments in this thesis directly studied the activity of only transfected miRNAs, several observations indirectly provide insights into the function of endogenous miRNAs in neurons and the brain (see above). Endogenous neuronal miRNAs appear to be buffers of perturbations to the equilibrium in the neuronal transcriptome. I discuss these propositions below.

1. Endogenous miR-124 buffers differential gene expression in mature neurons. In the development of untransfected cultures, transcripts with 3′UTRs not depleted in miR-124 seed matching sites are upregulated relatively late in development (in transition from 4DIV to 8DIV), while transcripts upregulated early (1DIV to 2DIV) are depleted in miR-124 sites (Chapter 3, section 3.2.3). This means that the scope for endogenous miR-124 mediated inhibition of developmentally upregulated transcripts normally appears only in more mature neurons.

Transfections of non-neuronal miRNAs had a pronounced effect on gene expression in primary cultures at 3DIV and 4DIV, but the effect diminished at 6DIV (Chapter 5, sections 5.1.1). At the same time, a significant enrichment of seed matching sites for miR-124 was observed in transcripts upregulated upon transfection of non-neuronal miRNAs at 6DIV, while this was not the case in experiments at 3DIV or 4DIV (Chapter 5, Figure 5.2). The increasing scope for miR-124 mediated inhibition can account for differences between transfection experiments at later developmental timepoints. It is possible that in more mature cultures (6DIV) endogenous miR-124 buffers the induced transcripts and reduces the extent of their upregulation. This potentially leads to an overall decrease in differential expression (Chapter 5, section 5.1.3). In agreement with the bigger scope for miR-124 action at 6DIV, transfections of the mimic of miR-124 had the greatest effect on the transcriptome at the 6DIV timepoint (Chapter 5, section 5.2.1).

The proposition of endogenous miR-124 imposing constraints on differential gene expression in mature neurons is speculative, however it can be experimentally tested. These experiments will be suggested later.

2. Endogenous miR-124 and miR-434-3p buffer genes inducible by stresses

Buffering of changes to the transcriptome by miR-124 may be a general phenomenon. In order to study this, I assessed the distribution of seed matching sites for miRNAs in the 3′UTRs of transcripts after the injection of a kainate into the mouse hippocampus (Akahoshi et al., 2007). Kainate stress was found to upregulate transcripts with 3′UTRs enriched in seed matching sites for miR-124 and miR-434-3p, but not for other miRNAs (Chapter 6, section 6.7b). This enrichment, together with the observation that exogenous miR-124 and miR-434-3p can significantly inhibit genes induced by the kainate stress (Chapter 6, Figure 6.8), suggests that genes induced by kainate stress are enriched in targets of miR-124 and miR-434-3p. This observation, in conjunction with recent reports that targets of miR-124 are normally co-expressed with miR-124 in the same cells (Shkumatava et al., 2009; Clark et al., 2010), suggests that normally (before stress) miR-124 and miR-434-3p buffer expression of genes that can be induced by the stressful condition.

In addition to genes induced by the kainate stress, I obtained a list of the mouse homologs of genes induced by ageing of the human brain (Lu et al., 2004). These genes were inhibited significantly in primary forebrain cultures by both transfected miR-124 and miR-434-3p (Chapter 6, Figure 6.8). This suggests that endogenous miR-124 and miR-434-3p can buffer expression of genes induced by ageing.

Results of the transfection experiments in this thesis show that miRNAs are targeting genes induced by stress. This raises the question: what is the biological purpose for buffering stress inducible genes under normal conditions? One possible explanation comes from assessing variability in the expression of genes inducible by stresses in normal (untransfected) primary neuronal cultures. Genes inducible by transfection, kainate and ageing stresses were found to have greater variability in their expression between replicates of cultures at 8DIV than other genes (Chapter 6, section 6.10). Interestingly, transfection of cultures with miR-124 reduced this variability, while transfection with a non-mouse miRNA, cel-miR-67, did not (Chapter 6, Figure 6.11). Therefore, endogenous miR-124 may also normally act to reduce variability in expression of genes inducible by stress. One would expect precise control of gene expression to be important for neurons. My data indicates that miRNA mediated regulation of stress inducible genes is necessary, because precise control of expression of these genes is of particularly importance for neurons: the mutations in stress inducible genes are significantly more frequently linked to neurological disorders, than in genes on average ($P < 3.67e-05$, see Chapter 6, section 6.2.2). I believe the results presented here are a starting point for establishing a more general model of miRNA buffering of gene expression in neurons. The next section will detail some proposed experiments that can further test this hypothesis.

7.4 Directions for future work

In this thesis I directly demonstrated that transfected miRNAs inhibited genes that were induced by the transfection reagent (i.e. the transfection stress). The proposition of that endogenous miR-124 and miR-434-3p may also inhibit genes inducible by stress merits further investigation. This could be tested by studying mutant neuronal cultures that do not express miR-124 and/or miR-434-3p. The hypothesis that these miRNAs act as buffers of genes inducible by stress leads to the following predictions that could be tested in these cultures:

- The inter-replicate variability in the expression of genes inducible by stress is expected to be higher between cultures of mutant neurons. This prediction is based on the observation that exogenous (transfected) miR-124 reduces variability in expression of these genes in wild type cultures (Chapter 6, section 6.2.3).
- Transfection of mutant cultures at 6DIV with non-neuronal miRNAs is predicted to lead to similarly significant miRNA mediated changes in differential gene expression, as in 3DIV and 4DIV transfections. This prediction is based on the observation that

the decrease in the effect of non-neuronal miRNAs at 6DIV was associated with simultaneous enrichment of miR-124 seed matching sites in upregulated transcripts (Chapter 5, section 5.1.2).

Once mutant mouse lines that lack miR-124 and miR-434-3p become available, it will allow to test other predictions of the hypothesis of endogenous miRNAs as buffers of perturbations in the transcriptome. For example, in Chapter 6 I described the innate bias in the distribution of miR-124 and miR-434-3p seed matching sites upon kainate injection into mouse hippocampus. Therefore, it seems likely that these two miRNAs are particularly important in mouse neurons for buffering changes in the transcriptome that are associated with neuronal activity. Based on the results of a recent study by Konopka and colleagues (Konopka et al., 2010), in which mice with reduced abundance of nearly all miRNAs displayed enhanced learning, it is possible that mice lacking just miR-124 and/or miR-434-3p will display a similar enhancement. On the other hand, genes induced by ageing of the brain were efficiently inhibited by transfections of both miR-124 and miR-434-3p (Chapter 6, Figure 6.8). Therefore, it is possible that earlier than normal induction of genes associated with ageing will take place in the mutant mice.

The observation that miRNA targets are upregulated in stressful conditions (Chapter 6, section 2.7) raises a question concerning the mechanism behind this upregulation. One possibility is that upregulation was caused by miRNA independent mechanisms (for example, activation of transcription). However, an alternative explanation is a relief of miRNA mediated regulation in stress, or even a switch of miRNA mediated regulation to an activatory mode. Intriguingly, both relief of miRNA mediated regulation (Bhattacharyya et al., 2006) and a switch to activation (Vasudevan et al., 2007) were previously reported in stress, however this subject has not been studied extensively. My observations suggest a similar effect to possibly take place in neuronal cultures which makes them a suitable model system to study this enigmatic phenomenon. According to reports from the laboratory of Philip Sharp, localisation of components of RNA silencing machinery may be important for miRNA function (Leung and Sharp, 2006; Leung et al., 2006). Neuronal cultures, like other cell culture systems, allow direct access to cells which makes the study of subcellular localisation of miRNAs and RNA silencing machinery possible. Additionally, it was shown that the activity of several miRNAs, most notably of miR-124, is likely to play a major role in shaping gene expression in the brain (Farh et al., 2005; Sood et al., 2006) and in primary neuronal cultures (Manakov et al., 2009). Therefore, primary neuronal cultures allow us to make use of global changes in the transcriptome as a robust marker of changes in neuronal miRNA mediated activity.

Inhibition of targets by miR-25 in the transfection experiment described in Chapter 5 (section 5.2.2) was extremely efficient. In fact, enrichment of miR-25 seed matching sites in 3′UTRs of downregulated transcripts was more significant than in miR-124 experiments (Chapter 5, section 5.2.1). Interestingly, expression of miR-25 was found to be induced in tumours (Poliseno et al., 2010), while expression of a majority of miRNAs is downregulated in tumours (Thomson et al., 2006; Lotterman et al., 2008). Therefore, the high impact of miR-25 on transcriptome of neurons may be related to its potential role in reprogramming cells during carcinogenesis. I propose that transfection of mimics of other miRNAs that are co-expressed with miR-25 in tumours will have a similarly strong effect on the transcriptome of differentiated cell types, such as the cells of primary neuronal cultures. Further research of miR-25 and other oncogenic miRNAs in neuronal cultures may help to understand mechanisms of carcinogenic reprogramming of differentiated cell types.

7.5 Conclusion

In this thesis I identified hundreds of putative direct miRNA targets for six different miRNAs (both neuronal and non-neuronal) in primary neuronal cultures. This large resource of novel miRNA targets allows an in-depth analysis of the roles of miRNAs in neurons. Analysis of these target lists indicates that the major function of miRNA mediated regulation is buffering of gene expression. This effect is context dependent. The targets in this thesis were identified using chemical transfections of miRNA mimics and inhibitors, therefore the functions of miRNAs were elucidated in the context of stress associated with transfection. In this context, I identified that the targeting repertoire of neuronal miRNAs is adapted for the inhibition of genes induced by different stresses in neurons and the brain. I have identified lists of putative direct miRNA targets (see Supplementary Data), which I hope will be a useful resource for future research into both the function of miRNAs and their role during stress.

Context dependent inhibition of inducibly expressed genes at the level of the whole genome is a novel concept. However, taking the experimental context into account is necessary for understating the results of previously published miRNA transfection experiments and for the design of the future experiments into miRNA function. If context dependent inhibition of inducibly expressed genes is confirmed for endogenous miRNAs, it will make miRNAs guardians of transcriptional equilibrium. This would contribute to our understanding of the role of miRNAs in general and of neuronal miRNAs in particular. It would also provide an explanation for the mechanism of inhibition of neuronal plasticity and learning by neuronal miRNAs (Rajasethupathy et al., 2009; Gao et al., 2010; Konopka et al., 2010).

Bibliography

Ahlemeyer, B. and Baumgart-Vogt, E. (2005). Optimized protocols for the simultaneous preparation of primary neuronal cultures of the neocortex, hippocampus and cerebellum from individual newborn (P0.5) C57Bl/6J mice. *J Neurosci Methods* **149**, 110–20.

Akahoshi, N., Murashima, Y. L., Himi, T., Ishizaki, Y. and Ishii, I. (2007). Increased expression of the lysosomal protease cathepsin S in hippocampal microglia following kainate-induced seizures. *Neurosci Lett* **429**, 136–41.

Alexopoulou, A. N., Couchman, J. R. and Whiteford, J. R. (2008). The CMV early enhancer/chicken beta actin (CAG) promoter can be used to drive transgene expression during the differentiation of murine embryonic stem cells into vascular progenitors. *BMC Cell Biol* **9**, 2.

Altman, S. A., Randers, L. and Rao, G. (1993). Comparison of trypan blue dye exclusion and fluorometric assays for mammalian cell viability determinations. *Biotechnol Prog* **9**, 671–4.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–10.

Alvarez-Saavedra, E. and Horvitz, H. R. (2010). Many families of C. elegans microRNAs are not essential for development or viability. *Curr Biol* **20**, 367–73.

Andersson, T., Rahman, S., Sansom, S. N., Alsiö, J. M., Kaneda, M., Smith, J., O'Carroll, D., Tarakhovsky, A. and Livesey, F. J. (2010). Reversible block of mouse neural stem cell differentiation in the absence of dicer and microRNAs. *PLoS ONE* **5**, e13453.

Aravin, A. A., Hannon, G. J. and Brennecke, J. (2007). The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* **318**, 761–4.

Arnold, F. J. L., Hofmann, F., Bengtson, C. P., Wittmann, M., Vanhoutte, P. and Bading, H. (2005). Microelectrode array recordings of cultured hippocampal networks reveal a simple model for transcription and protein synthesis-dependent plasticity. *J Physiol (Lond)* **564**, 3–19.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M. and Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25–9.

Azuma-Mukai, A., Oguri, H., Mituyama, T., Qian, Z. R., Asai, K., Siomi, H. and Siomi, M. C. (2008). Characterization of endogenous human Argonautes and their miRNA partners in RNA silencing. *Proc Natl Acad Sci USA* **105**, 7964–9.

Babiarz, J. E., Ruby, J. G., Wang, Y., Bartel, D. P. and Blelloch, R. (2008). Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes & Development* **22**, 2773–85.

Bading, H., Segal, M. M., Sucher, N. J., Dudek, H., Lipton, S. A. and Greenberg, M. E. (1995). N-methyl-D-aspartate receptors are critical for mediating the effects of glutamate on intracellular calcium concentration and immediate early gene expression in cultured hippocampal neurons. *Neuroscience* **64**, 653–64.

Baek, D., Villén, J., Shin, C., Camargo, F. D., Gygi, S. P. and Bartel, D. P. (2008). The impact of microRNAs on protein output. *Nature* **455**, 64–71.

Baillat, D. and Shiekhattar, R. (2009). Functional dissection of the human TNRC6 (GW182-related) family of proteins. *Mol Cell Biol* **29**, 4144–55.

Bak, M., Silahtaroglu, A., Moller, M., Christensen, M., Rath, M. F., Skryabin, B., Tommerup, N. and Kauppinen, S. (2008). MicroRNA expression in the adult mouse central nervous system. *RNA* **14**, 432–444.

Ballas, N., Grunseich, C., Lu, D. D., Speh, J. C. and Mandel, G. (2005). REST and its corepressors mediate plasticity of neuronal gene chromatin throughout neurogenesis. *Cell* **121**, 645–57.

Ballas, N. and Mandel, G. (2005). The many faces of REST oversee epigenetic programming of neuronal genes. *Curr Opin Neurobiol* **15**, 500–6.

Barrett, T., Troup, D. B., Wilhite, S. E., Ledoux, P., Rudnev, D., Evangelista, C., Kim, I. F., Soboleva, A., Tomashevsky, M. and Edgar, R. (2007). NCBI GEO: mining tens of millions of expression profiles–database and tools update. *Nucleic Acids Research* **35**, D760–5.

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–97.

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215–33.

Basyuk, E., Suavet, F., Doglio, A., Bordonné, R. and Bertrand, E. (2003). Human let-7 stem-loop precursors harbor features of RNase III cleavage products. *Nucleic acids research* **31**, 6593–7.

Bayés, A., van de Lagemaat, L. N., Collins, M. O., Croning, M. D. R., Whittle, I. R., Choudhary, J. S. and Grant, S. G. N. (2011). Characterization of the proteome, diseases and evolution of the human postsynaptic density. *Nat Neurosci* **14**, 19–21.

Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P. and Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes & Development* **20**, 1885–98.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289–300.

Bernstein, E., Caudy, A. A., Hammond, S. M. and Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363–6.

Bernstein, E., Kim, S. Y., Carmell, M. A., Murchison, E. P., Alcorn, H., Li, M. Z., Mills, A. A., Elledge, S. J., Anderson, K. V. and Hannon, G. J. (2003). Dicer is essential for mouse development. *Nat Genet* **35**, 215–7.

Bhattacharyya, S. N., Habermacher, R., Martine, U., Closs, E. I. and Filipowicz, W. (2006). Relief of microRNA-mediated translational repression in human cells subjected to stress. *Cell* **125**, 1111–24.

Bohnsack, M. T., Czaplinski, K. and Gorlich, D. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA (New York, NY)* **10**, 185–91.

Borchert, G. M., Lanier, W. and Davidson, B. L. (2006). RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol* **13**, 1097–101.

Brennecke, J., Stark, A., Russell, R. B. and Cohen, S. M. (2005). Principles of microRNA-target recognition. *PLoS Biol* **3**, e85.

Brenner, J. L., Jasiewicz, K. L., Fahley, A. F., Kemp, B. J. and Abbott, A. L. (2010). Loss of individual microRNAs causes mutant phenotypes in sensitized genetic backgrounds in C. elegans. *Curr Biol* **20**, 1321–5.

Brewer, G. J. (1997). Isolation and culture of adult rat hippocampal neurons. *J Neurosci Methods* **71**, 143–55.

Brewer, G. J., Boehler, M. D., Pearson, R. A., Demaris, A. A., Ide, A. N. and Wheeler, B. C. (2009). Neuron network activity scales exponentially with synapse density. *J. Neural Eng.* **6**, 014001.

Brewer, G. J. and Torricelli, J. R. (2007). Isolation and culture of adult neurons and neurospheres. *Nat Protoc* **2**, 1490–8.

Brewer, G. J., Torricelli, J. R., Evege, E. K. and Price, P. J. (1993). Optimized survival of hippocampal neurons in B27-supplemented Neurobasal, a new serum-free medium combination. *J Neurosci Res* **35**, 567–76.

Cai, X., Hagedorn, C. H. and Cullen, B. R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA (New York, NY)* **10**, 1957–66.

Cao, X., Pfaff, S. L. and Gage, F. H. (2007). A functional study of miR-124 in the developing neural tube. *Genes & Development* **21**, 531–6.

Carthew, R. W. and Sontheimer, E. J. (2009). Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136**, 642–55.

Chekulaeva, M., Filipowicz, W. and Parker, R. (2009). Multiple independent domains of dGW182 function in miRNA-mediated repression in Drosophila. *RNA (New York, NY)* **15**, 794–803.

Chendrimada, T. P., Finn, K. J., Ji, X., Baillat, D., Gregory, R. I., Liebhaber, S. A., Pasquinelli, A. E. and Shiekhattar, R. (2007). MicroRNA silencing through RISC recruitment of eIF6. *Nature* **447**, 823–8.

Chendrimada, T. P., Gregory, R. I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K. and Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **436**, 740–4.

Cheng, L.-C., Pastrana, E., Tavazoie, M. and Doetsch, F. (2009). miR-124 regulates adult neurogenesis in the subventricular zone stem cell niche. *Nat Neurosci* **12**, 399–408.

Chi, S., Zang, J., Mele, A. and Darnell, R. (2009). Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* **460**, 479–86.

Christodoulou, F., Raible, F., Tomer, R., Simakov, O., Trachana, K., Klaus, S., Snyman, H., Hannon, G. J., Bork, P. and Arendt, D. (2010). Ancient animal microRNAs and the evolution of tissue identity. *Nature* **463**, 1084–8.

Cifuentes, D., Xue, H., Taylor, D. W., Patnode, H., Mishima, Y., Cheloufi, S., Ma, E., Mane, S., Hannon, G. J., Lawson, N. D., Wolfe, S. A. and Giraldez, A. J. (2010). A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* **328**, 1694–8.

Clark, A. M., Goldstein, L. D., Tevlin, M., Tavaré, S., Shaham, S. and Miska, E. A. (2010). The microRNA miR-124 controls gene expression in the sensory nervous system of Caenorhabditis elegans. *Nucleic Acids Research* **38**, 3780–93.

Cogoni, C., Irelan, J. T., Schumacher, M., Schmidhauser, T. J., Selker, E. U. and Macino, G. (1996). Transgene silencing of the al-1 gene in vegetative cells of Neurospora is mediated by a cytoplasmic effector and does not depend on DNA-DNA interactions or DNA methylation. *EMBO J* **15**, 3153–63.

Conaco, C., Otto, S., Han, J.-J. and Mandel, G. (2006). Reciprocal actions of REST and a microRNA promote neuronal identity. *Proc Natl Acad Sci USA* **103**, 2422–7.

Connelly, S. and Manley, J. L. (1988). A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes & Development* **2**, 440–52.

Davis, E., Caiment, F., Tordoir, X., Cavaillé, J., Ferguson-Smith, A., Cockett, N., Georges, M. and Charlier, C. (2005). RNAi-mediated allelic trans-interaction at the imprinted Rtl1/Peg11 locus. *Curr Biol* **15**, 743–9.

Davis, T. H., Cuellar, T. L., Koch, S. M., Barker, A. J., Harfe, B. D., Mcmanus, M. T. and Ullian, E. M. (2008). Conditional loss of Dicer disrupts cellular and tissue morphogenesis in the cortex and hippocampus. *J Neurosci* **28**, 4322–30.

de Carvalho, F., Gheysen, G., Kushnir, S., Montagu, M. V., Inzé, D. and Castresana, C. (1992). Suppression of beta-1,3-glucanase transgene expression in homozygous plants. *EMBO J* **11**, 2595–602.

Dichter, M. A. (1978). Rat cortical neurons in cell culture: culture methods, cell morphology, electrophysiology, and synapse formation. *Brain Res* **149**, 279–93.

Doench, J. G. and Sharp, P. A. (2004). Specificity of microRNA target selection in translational repression. *Genes & Development* **18**, 504–11.

Du, P., Kibbe, W. A. and Lin, S. M. (2008). lumi: a pipeline for processing Illumina microarray. *Bioinformatics* **24**, 1547–8.

Elbashir, S. M., Lendeckel, W. and Tuschl, T. (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes & Development* **15**, 188–200.

Elia, L., Quintavalle, M., Zhang, J., Contu, R., Cossu, L., Latronico, M. V. G., Peterson, K. L., Indolfi, C., Catalucci, D., Chen, J., Courtneidge, S. A. and Condorelli, G. (2009). The knockout of miR-143 and -145 alters smooth muscle cell maintenance and vascular homeostasis in mice: correlates with human disease. *Cell Death Differ* **16**, 1590–8.

Enright, A. J., John, B., Gaul, U., Tuschl, T., Sander, C. and Marks, D. S. (2003). MicroRNA targets in Drosophila. *Genome Biol* **5**, R1.

Eulalio, A., Huntzinger, E. and Izaurralde, E. (2008). GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay. *Nat Struct Mol Biol* **15**, 346–53.

Eulalio, A., Huntzinger, E., Nishihara, T., Rehwinkel, J., Fauser, M. and Izaurralde, E. (2009). Deadenylation is a widespread effect of miRNA regulation. *RNA (New York, NY)* **15**, 21–32.

Fabian, M. R., Sonenberg, N. and Filipowicz, W. (2010). Regulation of mRNA translation and stability by microRNAs. *Annu Rev Biochem* **79**, 351–79.

Falcon, S. and Gentleman, R. (2007). Using GOstats to test gene lists for GO term association. *Bioinformatics* **23**, 257–8.

Farh, K. K.-H., Grimson, A., Jan, C., Lewis, B. P., Johnston, W. K., Lim, L. P., Burge, C. B. and Bartel, D. P. (2005). The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science* **310**, 1817–21.

Felgner, P. L., Gadek, T. R., Holm, M., Roman, R., Chan, H. W., Wenz, M., Northrop, J. P., Ringold, G. M. and Danielsen, M. (1987). Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proc Natl Acad Sci USA* **84**, 7413–7.

Flicek, P., Aken, B. L., Beal, K., Ballester, B., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cunningham, F., Cutts, T., Down, T., Dyer, S. C., Eyre, T., Fitzgerald, S., Fernandez-Banet, J., Gräf, S., Haider, S., Hammond, M., Holland, R., Howe, K. L., Howe, K., Johnson, N., Jenkinson, A., Kähäri, A., Keefe, D., Kokocinski, F., Kulesha, E., Lawson, D., Longden, I., Megy, K., Meidl, P., Overduin, B., Parker, A., Pritchard, B., Prlic, A., Rice, S., Rios, D., Schuster, M., Sealy, I., Slater, G., Smedley, D., Spudich, G., Trevanion, S., Vilella, A. J., Vogel, J., White, S., Wood, M., Birney, E., Cox, T., Curwen, V., Durbin, R., Fernandez-Suarez, X. M., Herrero, J., Hubbard, T. J. P., Kasprzyk, A., Proctor, G., Smith, J., Ureta-Vidal, A. and Searle, S. (2008). Ensembl 2008. *Nucleic Acids Res* **36**, D707–14.

Flynt, A. S., Thatcher, E. J., Burkewitz, K., Li, N., Liu, Y. and Patton, J. G. (2009). miR-8 microRNAs regulate the response to osmotic stress in zebrafish embryos. *J Cell Biol* **185**, 115–27.

Freeman, M. R. (2010). Specification and morphogenesis of astrocytes. *Science* **330**, 774–8.

Freeman, T. C., Goldovsky, L., Brosch, M., van Dongen, S., Mazière, P., Grocock, R. J., Freilich, S., Thornton, J. and Enright, A. J. (2007). Construction, visualisation, and clustering of transcription networks from microarray expression data. *PLoS Comput Biol* **3**, 2032–42.

Gao, J., Wang, W.-Y., Mao, Y.-W., Gräff, J., Guan, J.-S., Pan, L., Mak, G., Kim, D., Su, S. C. and Tsai, L.-H. (2010). A novel pathway regulates memory and plasticity via SIRT1 and miR-134. *Nature* **466**, 1105–9.

Gautier, L., Cope, L., Bolstad, B. M. and Irizarry, R. A. (2004). affy–analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307–15.

Gelder, R. N. V., von Zastrow, M. E., Yool, A., Dement, W. C., Barchas, J. D. and Eberwine, J. H. (1990). Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc Natl Acad Sci USA* **87**, 1663–7.

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A. J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J. Y. H. and Zhang, J. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* **5**, R80.

Giraldez, A. J., Cinalli, R. M., Glasner, M. E., Enright, A. J., Thomson, J. M., Baskerville, S., Hammond, S. M., Bartel, D. P. and Schier, A. F. (2005). MicroRNAs regulate brain morphogenesis in zebrafish. *Science* **308**, 833–8.

Giraldez, A. J., Mishima, Y., Rihel, J., Grocock, R. J., van Dongen, S., Inoue, K., Enright, A. J. and Schier, A. F. (2006). Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* **312**, 75–9.

Götz, M. and Huttner, W. (2005). The cell biology of neurogenesis. *Nat Rev Mol Cell Biol* **6**, 777–788.

Griffiths-Jones, S. (2004). The microRNA Registry. *Nucleic Acids Research* **32**, D109–11.

Griffiths-Jones, S., Grocock, R. J., van Dongen, S., Bateman, A. and Enright, A. J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Research* **34**, D140–4.

Griffiths-Jones, S., Saini, H. K., van Dongen, S. and Enright, A. J. (2008). miRBase: tools for microRNA genomics. *Nucleic Acids Res* **36**, D154–8.

Grimson, A., Farh, K. K.-H., Johnston, W. K., Garrett-Engele, P., Lim, L. P. and Bartel, D. P. (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Molecular Cell* **27**, 91–105.

Grimson, A., Srivastava, M., Fahey, B., Woodcroft, B. J., Chiang, H. R., King, N., Degnan, B. M., Rokhsar, D. S. and Bartel, D. P. (2008). Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature* **455**, 1193–7.

Grishok, A., Pasquinelli, A. E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D. L., Fire, A., Ruvkun, G. and Mello, C. C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. *Cell* **106**, 23–34.

Grün, D., Wang, Y.-L., Langenberger, D., Gunsalus, K. C. and Rajewsky, N. (2005). microRNA target predictions across seven Drosophila species and comparison to mammalian targets. *PLoS Comput Biol* **1**, e13.

Gu, S., Jin, L., Zhang, F., Sarnow, P. and Kay, M. A. (2009). Biological basis for restriction of microRNA targets to the 3' untranslated region in mammalian mRNAs. *Nat Struct Mol Biol* **16**, 144–50.

Guo, H., Ingolia, N. T., Weissman, J. S. and Bartel, D. P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* **466**, 835–40.

Guo, S. and Kemphues, K. J. (1995). par-1, a gene required for establishing polarity in C. elegans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. *Cell* **81**, 611–20.

Hamilton, A. J. and Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **286**, 950–2.

Hammond, S. M., Bernstein, E., Beach, D. and Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* **404**, 293–6.

Han, J., Lee, Y., Yeom, K.-H., Kim, Y.-K., Jin, H. and Kim, V. N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. *Genes & Development* **18**, 3016–27.

Hanina, S. A., Mifsud, W., Down, T. A., Hayashi, K., O'Carroll, D., Lao, K., Miska, E. A. and Surani, M. A. (2010). Genome-Wide Identification of Targets and Function of Individual MicroRNAs in Mouse Embryonic Stem Cells. *PLoS Genetics* **6**, e1001163.

Hardingham, G. E., Arnold, F. J. and Bading, H. (2001). Nuclear calcium signaling controls CREB-mediated gene expression triggered by synaptic activity. *Nat Neurosci* **4**, 261–7.

Hébert, S. S., Papadopoulou, A. S., Smith, P., Galas, M.-C., Planel, E., Silahtaroglu, A. N., Sergeant, N., Buée, L. and Strooper, B. D. (2010). Genetic ablation of Dicer in adult forebrain neurons results in abnormal tau hyperphosphorylation and neurodegeneration. *Hum Mol Genet* **19**, 3959–69.

Hendrickson, D. G., Hogan, D. J., McCullough, H. L., Myers, J. W., Herschlag, D., Ferrell, J. E. and Brown, P. O. (2009). Concordant regulation of translation and mRNA abundance for hundreds of targets of a human microRNA. *PLoS Biol* **7**, e1000238.

Henke, J. I., Goergen, D., Zheng, J., Song, Y., Schüttler, C. G., Fehr, C., Jünemann, C. and Niepmann, M. (2008). microRNA-122 stimulates translation of hepatitis C virus RNA. *EMBO J* **27**, 3300–10.

Herranz, H. and Cohen, S. M. (2010). MicroRNAs and gene regulatory networks: managing the impact of noise in biological systems. *Genes & Development* **24**, 1339–44.

Höck, J. and Meister, G. (2008). The Argonaute protein family. *Genome Biol* **9**, 210.

Hornstein, E. and Shomron, N. (2006). Canalization of development by microRNAs. *Nat Genet* **38 Suppl**, S20–4.

Huang, J. C., Babak, T., Corson, T. W., Chua, G., Khan, S., Gallie, B. L., Hughes, T. R., Blencowe, B. J., Frey, B. J. and Morris, Q. D. (2007). Using expression profiling data to identify human microRNA targets. *Nat Meth* **4**, 1045–9.

Hubbard, T. J. P., Aken, B. L., Ayling, S., Ballester, B., Beal, K., Bragin, E., Brent, S., Chen, Y., Clapham, P., Clarke, L., Coates, G., Fairley, S., Fitzgerald, S., Fernandez-Banet, J., Gordon, L., Graf, S., Haider, S., Hammond, M., Holland, R., Howe, K., Jenkinson, A., Johnson, N., Kahari, A.,

Keefe, D., Keenan, S., Kinsella, R., Kokocinski, F., Kulesha, E., Lawson, D., Longden, I., Megy, K., Meidl, P., Overduin, B., Parker, A., Pritchard, B., Rios, D., Schuster, M., Slater, G., Smedley, D., Spooner, W., Spudich, G., Trevanion, S., Vilella, A., Vogel, J., White, S., Wilder, S., Zadissa, A., Birney, E., Cunningham, F., Curwen, V., Durbin, R., Fernandez-Suarez, X. M., Herrero, J., Kasprzyk, A., Proctor, G., Smith, J., Searle, S. and Flicek, P. (2009). Ensembl 2009. *Nucleic acids research* **37**, D690–7.

Humphreys, D. T., Westman, B. J., Martin, D. I. K. and Preiss, T. (2005). MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. *Proc Natl Acad Sci USA* **102**, 16961–6.

Hutvágner, G., McLachlan, J., Pasquinelli, A. E., Bálint, E., Tuschl, T. and Zamore, P. D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293**, 834–8.

Hutvágner, G. and Zamore, P. D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. *Science* **297**, 2056–60.

Inui, M., Martello, G. and Piccolo, S. (2010). MicroRNA control of signal transduction. *Nat Rev Mol Cell Biol* **11**, 252–63.

Irizarry, R. A., Bolstad, B. M., Collin, F., Cope, L. M., Hobbs, B. and Speed, T. P. (2003a). Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Research* **31**, e15.

Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U. and Speed, T. P. (2003b). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249–64.

Joshua-Tor, L. and Hannon, G. J. (2010). Ancestral Roles of Small RNAs: An Ago-Centric Perspective. *Cold Spring Harbor perspectives in biology* **10**, a003772.

Judson, R. L., Babiarz, J. E., Venere, M. and Blelloch, R. (2009). Embryonic stem cell-specific microRNAs promote induced pluripotency. *Nat Biotechnol* **27**, 459–61.

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T. and Yamanashi, Y. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic acids research* **28**, 27–30.

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T. and Yamanashi, Y. (2008). KEGG for linking genomes to life and the environment. *Nucleic acids research* **36**, D480–4.

Kanellopoulou, C., Muljo, S. A., Kung, A. L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D. M. and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes & Development* **19**, 489–501.

Karginov, F. V., Conaco, C., Xuan, Z., Schmidt, B. H., Parker, J. S., Mandel, G. and Hannon, G. J. (2007). A biochemical approach to identifying microRNA targets. *Proc Natl Acad Sci USA* **104**, 19291–6.

Kennerdell, J. R. and Carthew, R. W. (1998). Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in the wingless pathway. *Cell* **95**, 1017–26.

Ketting, R. F., Fischer, S. E., Bernstein, E., Sijen, T., Hannon, G. J. and Plasterk, R. H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. *Genes & Development* **15**, 2654–9.

Khvorova, A., Reynolds, A. and Jayasena, S. D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**, 209–16.

Kim, J., Inoue, K., Ishii, J., Vanti, W. B., Voronov, S. V., Murchison, E., Hannon, G. and Abeliovich, A. (2007). A MicroRNA feedback circuit in midbrain dopamine neurons. *Science* **317**, 1220–4.

Kim, V. N., Han, J. and Siomi, M. C. (2009). Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* **10**, 126–39.

Klein, M. E., Lioy, D. T., Ma, L., Impey, S., Mandel, G. and Goodman, R. H. (2007). Homeostatic regulation of MeCP2 expression by a CREB-induced microRNA. *Nat Neurosci* **10**, 1513–4.

Knight, S. W. and Bass, B. L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. *Science* **293**, 2269–71.

Konopka, W., Kiryk, A., Novak, M., Herwerth, M., Parkitna, J. R., Wawrzyniak, M., Kowarsch, A., Michaluk, P., Dzwonek, J., Arnsperger, T., Wilczynski, G., Merkenschlager, M., Theis, F. J., Köhr, G., Kaczmarek, L. and Schütz, G. (2010). MicroRNA Loss Enhances Learning and Memory in Mice. *J Neurosci* **30**, 14835–14842.

Kriegstein, A. R. and Dichter, M. A. (1983). Morphological classification of rat cortical neurons in cell culture. *J Neurosci* **3**, 1634–47.

Krützfeldt, J., Rajewsky, N., Braich, R., Rajeev, K. G., Tuschl, T., Manoharan, M. and Stoffel, M. (2005). Silencing of microRNAs in vivo with 'antagomirs'. *Nature* **438**, 685–689.

Lagos-Quintana, M., Rauhut, R., Lendeckel, W. and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853–8.

Lagos-Quintana, M., Rauhut, R., Yalcin, A., Meyer, J., Lendeckel, W. and Tuschl, T. (2002). Identification of tissue-specific microRNAs from mouse. *Curr Biol* **12**, 735–9.

Lai, E. C. (2002). Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. *Nat Genet* **30**, 363–4.

Lai, E. C., Tam, B. and Rubin, G. M. (2005). Pervasive regulation of Drosophila Notch target genes by GY-box-, Brd-box-, and K-box-class microRNAs. *Genes & Development* **19**, 1067–80.

Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., Pfeffer, S., Rice, A., Kamphorst, A. O., Landthaler, M., Lin, C., Socci, N. D., Hermida, L., Fulci, V., Chiaretti, S., Foà, R., Schliwka, J., Fuchs, U., Novosel, A., Müller, R.-U., Schermer, B., Bissels, U., Inman, J., Phan, Q., Chien, M., Weir, D. B., Choksi, R., Vita, G. D., Frezzetti, D., Trompeter, H.-I., Hornung, V., Teng, G., Hartmann, G., Palkovits, M., Lauro, R. D., Wernet, P., Macino, G., Rogler, C. E., Nagle, J. W., Ju, J., Papavasiliou, F. N., Benzing, T., Lichter, P., Tam, W., Brownstein, M. J., Bosio, A., Borkhardt, A., Russo, J. J., Sander, C., Zavolan, M. and Tuschl, T. (2007). A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**, 1401–14.

Landthaler, M., Gaidatzis, D., Rothballer, A., Chen, P. Y., Soll, S. J., Dinic, L., Ojo, T., Hafner, M., Zavolan, M. and Tuschl, T. (2008). Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. *RNA (New York, NY)* **14**, 2580–96.

Lattin, J. E., Schroder, K., Su, A. I., Walker, J. R., Zhang, J., Wiltshire, T., Saijo, K., Glass, C. K., Hume, D. A., Kellie, S. and Sweet, M. J. (2008). Expression analysis of G Protein-Coupled Receptors in mouse macrophages. *Immunome Res* **4**, 5.

Le, M. T. N., Xie, H., Zhou, B., Chia, P. H., Rizk, P., Um, M., Udolph, G., Yang, H., Lim, B. and Lodish, H. F. (2009). MicroRNA-125b promotes neuronal differentiation in human cells by repressing multiple targets. *Mol Cell Biol* **29**, 5290–305.

Lee, M. K., Tuttle, J. B., Rebhun, L. I., Cleveland, D. W. and Frankfurter, A. (1990). The expression and posttranslational modification of a neuron-specific beta-tubulin isotype during chick embryogenesis. *Cell Motil Cytoskeleton* **17**, 118–32.

Lee, R. C., Feinbaum, R. L. and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843–54.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., Kim, S. and Kim, V. N. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415–9.

Lee, Y., Jeon, K., Lee, J.-T., Kim, S. and Kim, V. N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J* **21**, 4663–70.

Lee, Y., Kim, M., Han, J., Yeom, K.-H., Lee, S., Baek, S. H. and Kim, V. N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **23**, 4051–60.

Leucht, C., Stigloher, C., Wizenmann, A., Klafke, R., Folchert, A. and Bally-Cuif, L. (2008). MicroRNA-9 directs late organizer activity of the midbrain-hindbrain boundary. *Nat Neurosci* **11**, 641–8.

Leung, A. K. L., Calabrese, J. M. and Sharp, P. A. (2006). Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules. *Proc Natl Acad Sci USA* **103**, 18125–30.

Leung, A. K. L. and Sharp, P. A. (2006). Function and localization of microRNAs in mammalian cells. *Cold Spring Harb Symp Quant Biol* **71**, 29–38.

Leung, A. K. L. and Sharp, P. A. (2010). MicroRNA Functions in Stress Responses. *Molecular Cell* **40**, 205–15.

Leung, A. K. L., Young, A. G., Bhutkar, A., Zheng, G. X., Bosson, A. D., Nielsen, C. B. and Sharp, P. A. (2011). Genome-wide identification of Ago2 binding sites from mouse embryonic stem cells with and without mature microRNAs. *Nat Struct Mol Biol* **18**, 237–44.

Leuschner, P. J. F., Ameres, S. L., Kueng, S. and Martinez, J. (2006). Cleavage of the siRNA passenger strand during RISC assembly in human cells. *EMBO Rep* **7**, 314–20.

Lewis, A. and Redrup, L. (2005). Genetic imprinting: conflict at the Callipyge locus. *Curr Biol* **15**, R291–4.

Lewis, B. P., Burge, C. B. and Bartel, D. P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15–20.

Lewis, B. P., hung Shih, I., Jones-Rhoades, M. W., Bartel, D. P. and Burge, C. B. (2003). Prediction of mammalian microRNA targets. *Cell* **115**, 787–98.

Lewis, M. A., Quint, E., Glazier, A. M., Fuchs, H., Angelis, M. H. D., Langford, C., van Dongen, S., Abreu-Goodger, C., Piipari, M., Redshaw, N., Dalmay, T., Moreno-Pelayo, M. A., Enright, A. J. and Steel, K. P. (2009). An ENU-induced mutation of miR-96 associated with progressive hearing loss in mice. *Nat Genet* **41**, 614–8.

Li, X., Cassidy, J. J., Reinke, C. A., Fischboeck, S. and Carthew, R. W. (2009). A microRNA imparts robustness against environmental fluctuation during development. *Cell* **137**, 273–82.

Li, X. and Jin, P. (2010). Roles of small regulatory RNAs in determining neuronal identity. *Nat Rev Neurosci* **11**, 329–38.

Licatalosi, D. D., Mele, A., Fak, J. J., Ule, J., Kayikci, M., Chi, S. W., Clark, T. A., Schweitzer, A. C., Blume, J. E., Wang, X., Darnell, J. C. and

Darnell, R. B. (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**, 464–9.

Lim, L. P., Lau, N. C., Garrett-Engele, P., Grimson, A., Schelter, J. M., Castle, J., Bartel, D. P., Linsley, P. S. and Johnson, J. M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433**, 769–73.

Lim, L. P., Lau, N. C., Weinstein, E. G., Abdelhakim, A., Yekta, S., Rhoades, M. W., Burge, C. B. and Bartel, D. P. (2003). The microRNAs of Caenorhabditis elegans. *Genes & Development* **17**, 991–1008.

Lin, S. M., Du, P., Huber, W. and Kibbe, W. A. (2007). Model-based variance-stabilizing transformation for Illumina microarray data. *Nucleic Acids Research* **36**, e11–e11.

Liu, J., Carmell, M. A., Rivas, F. V., Marsden, C. G., Thomson, J. M., Song, J.-J., Hammond, S. M., Joshua-Tor, L. and Hannon, G. J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**, 1437–41.

Liu, J., Valencia-Sanchez, M. A., Hannon, G. J. and Parker, R. (2005). MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol* **7**, 719–23.

Lotterman, C. D., Kent, O. A. and Mendell, J. T. (2008). Functional integration of microRNAs into oncogenic and tumor suppressor pathways. *Cell Cycle* **7**, 2493–9.

Lu, T., Pan, Y., Kao, S.-Y., Li, C., Kohane, I., Chan, J. and Yankner, B. A. (2004). Gene regulation and DNA damage in the ageing human brain. *Nature* **429**, 883–91.

Lugli, G., Torvik, V. I., Larson, J. and Smalheiser, N. R. (2008). Expression of microRNAs and their precursors in synaptic fractions of adult mouse forebrain. *Journal of Neurochemistry* **106**, 650–61.

Maclaren, E. J., Charlesworth, P., Coba, M. P. and Grant, S. G. N. (2011). Knockdown of mental disorder susceptibility genes disrupts neuronal network physiology in vitro. *Molecular and cellular neurosciences* .

Makeyev, E. V., Zhang, J., Carrasco, M. A. and Maniatis, T. (2007). The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Molecular Cell* **27**, 435–48.

Malone, C. D. and Hannon, G. J. (2009). Small RNAs as guardians of the genome. *Cell* **136**, 656–68.

Manakov, S., Grant, S. and Enright, A. (2009). Reciprocal regulation of microRNA and mRNA profiles in neuronal development and synapse formation. *BMC Genomics* **10**, 419.

Maniataki, E. and Mourelatos, Z. (2005). A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes & Development* **19**, 2979–90.

Maroney, P. A., Yu, Y., Fisher, J. and Nilsen, T. W. (2006). Evidence that microRNAs are associated with translating messenger RNAs in human cells. *Nat Struct Mol Biol* **13**, 1102–7.

Mathonnet, G., Fabian, M. R., Svitkin, Y. V., Parsyan, A., Huck, L., Murata, T., Biffo, S., Merrick, W. C., Darzynkiewicz, E., Pillai, R. S., Filipowicz, W., Duchaine, T. F. and Sonenberg, N. (2007). MicroRNA inhibition of translation initiation in vitro by targeting the cap-binding complex eIF4F. *Science* **317**, 1764–7.

Matranga, C., Tomari, Y., Shin, C., Bartel, D. P. and Zamore, P. D. (2005). Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* **123**, 607–20.

Miska, E. A., Alvarez-Saavedra, E., Abbott, A. L., Lau, N. C., Hellman, A. B., McGonagle, S. M., Bartel, D. P., Ambros, V. R. and Horvitz, H. R. (2007). Most Caenorhabditis elegans microRNAs are individually not essential for development or viability. *PLoS Genet* **3**, e215.

Napoli, C., Lemieux, C. and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell* **2**, 279–289.

Nelson, P. G. (1975). Nerve and muscle cells in culture. *Physiol Rev* **55**, 1–61.

Ning, Z., Cox, A. J. and Mullikin, J. C. (2001). SSAHA: a fast search method for large DNA databases. *Genome Res* **11**, 1725–9.

Nottrott, S., Simard, M. J. and Richter, J. D. (2006). Human let-7a miRNA blocks protein production on actively translating polyribosomes. *Nat Struct Mol Biol* **13**, 1108–14.

Oelgeschläger, M., Larraín, J., Geissert, D. and Robertis, E. M. D. (2000). The evolutionarily conserved BMP-binding protein Twisted gastrulation promotes BMP signalling. *Nature* **405**, 757–63.

Olsen, P. H. and Ambros, V. (1999). The lin-4 regulatory RNA controls developmental timing in Caenorhabditis elegans by blocking LIN-14 protein synthesis after the initiation of translation. *Dev Biol* **216**, 671–80.

Ørom, U. A., Nielsen, F. C. and Lund, A. H. (2008). MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Mol Cell* **30**, 460–71.

Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., Hayward, D. C., Ball, E. E., Degnan, B., Müller, P., Spring, J., Srinivasan, A., Fishman, M., Finnerty, J., Corbo, J., Levine, M., Leahy, P., Davidson, E. and Ruvkun, G. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* **408**, 86–9.

Peters, L. and Meister, G. (2007). Argonaute proteins: mediators of RNA silencing. *Molecular Cell* **26**, 611–23.

Petersen, C. P., Bordeleau, M.-E., Pelletier, J. and Sharp, P. A. (2006). Short RNAs repress translation after initiation in mammalian cells. *Molecular Cell* **21**, 533–42.

Pillai, R. S., Artus, C. G. and Filipowicz, W. (2004). Tethering of human Ago proteins to mRNA mimics the miRNA-mediated repression of protein synthesis. *RNA (New York, NY)* **10**, 1518–25.

Pillai, R. S., Bhattacharyya, S. N., Artus, C. G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E. and Filipowicz, W. (2005). Inhibition of translational initiation by Let-7 MicroRNA in human cells. *Science* **309**, 1573–6.

Poliseno, L., Salmena, L., Riccardi, L., Fornari, A., Song, M. S., Hobbs, R. M., Sportoletti, P., Varmeh, S., Egia, A., Fedele, G., Rameh, L., Loda, M. and

Pandolfi, P. P. (2010). Identification of the miR-106b 25 microRNA cluster as a proto-oncogenic PTEN-targeting intron that cooperates with its host gene MCM7 in transformation. *Sci Signal* **3**, ra29.

Pruitt, K. D., Tatusova, T., Klimke, W. and Maglott, D. R. (2009). NCBI Reference Sequences: current status, policy and new initiatives. *Nucleic acids research* **37**, D32–6.

Qi, Y., Wang, J. K., McMillian, M. and Chikaraishi, D. M. (1997). Characterization of a CNS cell line, CAD, in which morphological differentiation is initiated by serum deprivation. *J Neurosci* **17**, 1217–25.

Quintavalle, M., Elia, L., Condorelli, G. and Courtneidge, S. A. (2010). MicroRNA control of podosome formation in vascular smooth muscle cells in vivo and in vitro. *J Cell Biol* **189**, 13–22.

Rajasethupathy, P., Fiumara, F., Sheridan, R., Betel, D., Puthanveettil, S. V., Russo, J. J., Sander, C., Tuschl, T. and Kandel, E. (2009). Characterization of small RNAs in aplysia reveals a role for miR-124 in constraining synaptic plasticity through CREB. *Neuron* **63**, 803–17.

Rao, Y., Lee, Y., Jarjoura, D., Ruppert, A. S., Liu, C.-G., Hsu, J. C. and Hagan, J. P. (2008). A comparison of normalization techniques for microRNA microarray data. *Statistical Applications in Genetics and Molecular Biology* **7**, Article22.

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R. and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* **403**, 901–6.

Rodriguez, A., Vigorito, E., Clare, S., Warren, M. V., Couttet, P., Soond, D. R., Dongen, S. V., Grocock, R. J., Das, P. P., Miska, E. A., Vetrie, D., Okkenhaug, K., Enright, A. J., Dougan, G., Turner, M. and Bradley, A. (2007). Requirement of bic/microRNA-155 for Normal Immune Function. *Science* **316**, 608–611.

Romano, N. and Macino, G. (1992). Quelling: transient inactivation of gene expression in Neurospora crassa by transformation with homologous sequences. *Mol Microbiol* **6**, 3343–53.

RTeam (2008). R: A language and environment for statistical computing. *R Foundation for Statistical Computing* **Vienna Austria**.

Rybak, A., Fuchs, H., Smirnova, L., Brandt, C., Pohl, E. E., Nitsch, R. and Wulczyn, F. G. (2008). A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. *Nat Cell Biol* **10**, 987–93.

Sayers, E. W., Barrett, T., Benson, D. A., Bolton, E., Bryant, S. H., Canese, K., Chetvernin, V., Church, D. M., Dicuccio, M., Federhen, S., Feolo, M., Fingerman, I. M., Geer, L. Y., Helmberg, W., Kapustin, Y., Landsman, D., Lipman, D. J., Lu, Z., Madden, T. L., Madej, T., Maglott, D. R., Marchler-Bauer, A., Miller, V., Mizrachi, I., Ostell, J., Panchenko, A., Phan, L., Pruitt, K. D., Schuler, G. D., Sequeira, E., Sherry, S. T., Shumway, M., Sirotkin, K., Slotta, D., Souvorov, A., Starchenko, G., Tatusova, T. A., Wagner, L., Wang, Y., Wilbur, W. J., Yaschenko, E. and Ye, J. (2010). Database resources of the National Center for Biotechnology Information. *Nucleic acids research* **39**, D38–51.

Schaefer, A., O'Carroll, D., Tan, C. L., Hillman, D., Sugimori, M., Llinas, R. and Greengard, P. (2007). Cerebellar neurodegeneration in the absence of microRNAs. *J Exp Med* **204**, 1553–8.

Scherer, W., Syverton, J. and Gey, G. (1953). Studies on the propagation in vitro of poliomyelitis viruses. IV. Viral multiplication in a stable strain of human malignant epithelial cells (strain HeLa) derived from an epidermoid carcinoma of the cervix. *J Exp Med* **97**, 695–710.

Schratt, G. (2009). microRNAs at the synapse. *Nat Rev Neurosci* **10**, 842–9.

Schratt, G. M., Tuebing, F., Nigh, E. A., Kane, C. G., Sabatini, M. E., Kiebler, M. and Greenberg, M. E. (2006). A brain-specific microRNA regulates dendritic spine development. *Nature* **439**, 283–9.

Schroeder, A., Levins, C. G., Cortez, C., Langer, R. and Anderson, D. G. (2010). Lipid-based nanotherapeutics for siRNA delivery. *J Intern Med* **267**, 9–21.

Schwarz, D. S., Hutvágner, G., Du, T., Xu, Z., Aronin, N. and Zamore, P. D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**, 199–208.

Seitz, H., Royo, H., Bortolin, M.-L., Lin, S.-P., Ferguson-Smith, A. C. and Cavaillé, J. (2004). A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. *Genome Res* **14**, 1741–8.

Selbach, M., Schwanhäusser, B., Thierfelder, N., Fang, Z., Khanin, R. and Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. *Nature* **455**, 58–63.

Sempere, L. F., Cole, C. N., McPeek, M. A. and Peterson, K. J. (2006). The phylogenetic distribution of metazoan microRNAs: insights into evolutionary complexity and constraint. *J Exp Zool B Mol Dev Evol* **306**, 575–88.

Sharp, P. A. (2009). The centrality of RNA. *Cell* **136**, 577–80.

Shibata, M., Kurokawa, D., Nakao, H., Ohmura, T. and Aizawa, S. (2008). MicroRNA-9 modulates Cajal-Retzius cell differentiation by suppressing Foxg1 expression in mouse medial pallium. *J Neurosci* **28**, 10415–21.

Shkumatava, A., Stark, A., Sive, H. and Bartel, D. P. (2009). Coherent but overlapping expression of microRNAs and their targets during vertebrate development. *Genes & Development* **23**, 466–81.

Siegal, M. L. and Bergman, A. (2002). Waddington's canalization revisited: developmental stability and evolution. *Proc Natl Acad Sci USA* **99**, 10528–32.

Siegel, G., Obernosterer, G., Fiore, R., Oehmen, M., Bicker, S., Christensen, M., Khudayberdiev, S., Leuschner, P. F., Busch, C. J. L., Kane, C., Hübel, K., Dekker, F., Hedberg, C., Rengarajan, B., Drepper, C., Waldmann, H., Kauppinen, S., Greenberg, M. E., Draguhn, A., Rehmsmeier, M., Martinez, J. and Schratt, G. M. (2009). A functional screen implicates microRNA-138-dependent regulation of the depalmitoylation enzyme APT1 in dendritic spine morphogenesis. *Nature* **11**, 705–16.

Silber, J., Lim, D. A., Petritsch, C., Persson, A. I., Maunakea, A. K., Yu, M., Vandenberg, S. R., Ginzinger, D. G., James, C. D., Costello, J. F., Bergers, G., Weiss, W. A., Alvarez-Buylla, A. and Hodgson, J. G. (2008). miR-124 and miR-137 inhibit proliferation of glioblastoma multiforme cells and induce differentiation of brain tumor stem cells. *BMC medicine* **6**, 14.

Smith, C. J., Watson, C. F., Bird, C. R., Ray, J., Schuch, W. and Grierson, D. (1990). Expression of a truncated tomato polygalacturonase gene inhibits expression of the endogenous gene in transgenic plants. *Mol Gen Genet* **224**, 477–81.

Smyth, G. (2004). Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. *Statistical Applications in Genetics and Molecular Biology* **3**, Article 3.

Sood, P., Krek, A., Zavolan, M., Macino, G. and Rajewsky, N. (2006). Cell-type-specific signatures of microRNAs on target mRNA expression. *Proc Natl Acad Sci USA* **103**, 2746–51. Notes.

Stark, A., Brennecke, J., Bushati, N., Russell, R. B. and Cohen, S. M. (2005). Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell* **123**, 1133–46.

Stark, A., Brennecke, J., Russell, R. B. and Cohen, S. M. (2003). Identification of Drosophila MicroRNA targets. *PLoS Biol* **1**, E60.

Svoboda, P., Stein, P., Hayashi, H. and Schultz, R. M. (2000). Selective reduction of dormant maternal mRNAs in mouse oocytes by RNA interference. *Development* **127**, 4147–56.

Taft, R. J., Glazov, E. A., Cloonan, N., Simons, C., Stephen, S., Faulkner, G. J., Lassmann, T., Forrest, A. R. R., Grimmond, S. M., Schroder, K., Irvine, K., Arakawa, T., Nakamura, M., Kubosaki, A., Hayashida, K., Kawazu, C., Murata, M., Nishiyori, H., Fukuda, S., Kawai, J., Daub, C. O., Hume, D. A., Suzuki, H., Orlando, V., Carninci, P., Hayashizaki, Y. and Mattick, J. S. (2009). Tiny RNAs associated with transcription start sites in animals. *Nat Genet* **41**, 572–8.

Taft, R. J., Simons, C., Nahkuri, S., Oey, H., Korbie, D. J., Mercer, T. R., Holst, J., Ritchie, W., Wong, J. J.-L., Rasko, J. E. J., Rokhsar, D. S., Degnan, B. M. and Mattick, J. S. (2010). Nuclear-localized tiny RNAs are associated with transcription initiation and splice sites in metazoans. *Nat Struct Mol Biol* **17**, 1030–4.

Thermann, R. and Hentze, M. W. (2007). Drosophila miR2 induces pseudo-polysomes and inhibits translation initiation. *Nature* **447**, 875–8.

Thomas-Chollier, M., Sand, O., Turatsinze, J.-V., Janky, R., Defrance, M., Vervisch, E., Brohée, S. and van Helden, J. (2008). RSAT: regulatory sequence analysis tools. *Nucleic acids research* **36**, W119–27.

Thomson, J. M., Newman, M., Parker, J. S., Morin-Kensicki, E. M., Wright, T. and Hammond, S. M. (2006). Extensive post-transcriptional regulation of microRNAs and its implications for cancer. *Genes & Development* **20**, 2202–7.

Valor, L. M., Charlesworth, P., Humphreys, L., Anderson, C. N. G. and Grant, S. G. N. (2007). Network activity-independent coordinated gene expression program for synapse assembly. *Proc Natl Acad Sci USA* **104**, 4658–63.

van der Krol, A. R., Mur, L. A., Beld, M., Mol, J. N. and Stuitje, A. R. (1990). Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. *Plant Cell* **2**, 291–9.

van Dongen, S. (2000). Graph clustering by flow simulation. *PhD thesis* **University of Utrecht**.

van Dongen, S., Abreu-Goodger, C. and Enright, A. J. (2008). Detecting microRNA binding and siRNA off-target effects from expression data. *Nat Meth* **5**, 1023–5.

van Rooij, E., Sutherland, L. B., Qi, X., Richardson, J. A., Hill, J. and Olson, E. N. (2007). Control of stress-dependent cardiac growth and gene expression by a microRNA. *Science* **316**, 575–9.

Vasudevan, S., Tong, Y. and Steitz, J. A. (2007). Switching from repression to activation: microRNAs can up-regulate translation. *Science* **318**, 1931–4.

Visvanathan, J., Lee, S., Lee, B., Lee, J. W. and Lee, S.-K. (2007). The microRNA miR-124 antagonizes the anti-neural REST/SCP1 pathway during embryonic CNS development. *Genes & Development* **21**, 744–9.

Voinnet, O. (2009). Origin, biogenesis, and activity of plant microRNAs. *Cell* **136**, 669–87.

Volinia, S., Galasso, M., Costinean, S., Tagliavini, L., Gamberoni, G., Drusco, A., Marchesini, J., Mascellani, N., Sana, M. E., Jarour, R. A., Desponts, C., Teitell, M., Baffa, R., Aqeilan, R., Iorio, M. V., Taccioli, C., Garzon,

R., Leva, G. D., Fabbri, M., Catozzi, M., Previati, M., Ambs, S., Palumbo, T., Garofalo, M., Veronese, A., Bottoni, A., Gasparini, P., Harris, C. C., Visone, R., Pekarsky, Y., de la Chapelle, A., Bloomston, M., Dillhoff, M., Rassenti, L. Z., Kipps, T. J., Huebner, K., Pichiorri, F., Lenze, D., Cairo, S., Buendia, M.-A., Pineau, P., Dejean, A., Zanesi, N., Rossi, S., Calin, G. A., Liu, C.-G., Palatini, J., Negrini, M., Vecchione, A., Rosenberg, A. and Croce, C. M. (2010). Reprogramming of miRNA networks in cancer and leukemia. *Genome Res* **20**, 589–99.

Wagenaar, D. A., Pine, J. and Potter, S. M. (2006). Searching for plasticity in dissociated cortical cultures on multi-electrode arrays. *J Negat Results BioMed* **5**, 16.

Wakiyama, M., Takimoto, K., Ohara, O. and Yokoyama, S. (2007). Let-7 microRNA-mediated mRNA deadenylation and translational repression in a mammalian cell-free system. *Genes & Development* **21**, 1857–62.

Wang, B., Love, T. M., Call, M. E., Doench, J. G. and Novina, C. D. (2006). Recapitulation of short RNA-directed translational gene silencing in vitro. *Molecular Cell* **22**, 553–60.

Wang, B., Yanez, A. and Novina, C. D. (2008a). MicroRNA-repressed mRNAs contain 40S but not 60S components. *Proc Natl Acad Sci USA* **105**, 5343–8.

Wang, X. and Wang, X. (2006). Systematic identification of microRNA functions by combining target prediction and expression profiling. *Nucleic Acids Res* **34**, 1646–52.

Wang, Y., Juranek, S., Li, H., Sheng, G., Tuschl, T. and Patel, D. J. (2008b). Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature* **456**, 921–6.

Wang, Y., Medvid, R., Melton, C., Jaenisch, R. and Blelloch, R. (2007). DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nat Genet* **39**, 380–5.

Wang, Y., Sheng, G., Juranek, S., Tuschl, T. and Patel, D. J. (2008c). Structure of the guide-strand-containing argonaute silencing complex. *Nature* **456**, 209–13.

Wianny, F. and Zernicka-Goetz, M. (2000). Specific interference with gene function by double-stranded RNA in early mouse development. *Nat Cell Biol* **2**, 70–5.

Wienholds, E., Kloosterman, W. P., Miska, E., Alvarez-Saavedra, E., Berezikov, E., de Bruijn, E., Horvitz, H. R., Kauppinen, S. and Plasterk, R. H. A. (2005). MicroRNA expression in zebrafish embryonic development. *Science* **309**, 310–1.

Wightman, B., Ha, I. and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell* **75**, 855–62.

Winter, J., Jung, S., Keller, S., Gregory, R. I. and Diederichs, S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat Cell Biol* **11**, 228–34.

Wu, C., Orozco, C., Boyer, J., Leglise, M., Goodale, J., Batalov, S., Hodge, C. L., Haase, J., Janes, J., Huss, J. W. and Su, A. I. (2009a). BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol* **10**, R130.

Wu, C.-I., Shen, Y. and Tang, T. (2009b). Evolution under canalization and the dual roles of microRNAs: a hypothesis. *Genome Res* **19**, 734–43.

Xin, M., Small, E. M., Sutherland, L. B., Qi, X., McAnally, J., Plato, C. F., Richardson, J. A., Bassel-Duby, R. and Olson, E. N. (2009). MicroRNAs miR-143 and miR-145 modulate cytoskeletal dynamics and responsiveness of smooth muscle cells to injury. *Genes & Development* **23**, 2166–78.

Xu, P., Vernooy, S. Y., Guo, M. and Hay, B. A. (2003). The Drosophila microRNA Mir-14 suppresses cell death and is required for normal fat metabolism. *Curr Biol* **13**, 790–5.

Xu, Y. and Szoka, F. C. (1996). Mechanism of DNA release from cationic liposome/DNA complexes used in cell transfection. *Biochemistry* **35**, 5616–23.

Yang, J.-S., Maurin, T., Robine, N., Rasmussen, K. D., Jeffrey, K. L., Chandwani, R., Papapetrou, E. P., Sadelain, M., O'Carroll, D. and Lai, E. C. (2010). Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. *Proc Natl Acad Sci USA* **107**, 15163–8.

Yi, R., Qin, Y., Macara, I. G. and Cullen, B. R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes & Development* **17**, 3011–6.

Yoo, A. S., Staahl, B. T., Chen, L. and Crabtree, G. R. (2009). MicroRNA-mediated switching of chromatin-remodelling complexes in neural development. *Nature* **460**, 642–6.

Zabner, J., Fasbender, A. J., Moninger, T., Poellinger, K. A. and Welsh, M. J. (1995). Cellular and molecular barriers to gene transfer by a cationic lipid. *J Biol Chem* **270**, 18997–9007.

Zamore, P. D., Tuschl, T., Sharp, P. A. and Bartel, D. P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**, 25–33.

Zhang, G., Gurtu, V. and Kain, S. R. (1996). An enhanced green fluorescent protein allows sensitive detection of gene transfer in mammalian cells. *Biochem Biophys Res Commun* **227**, 707–11.

Zhao, C., Sun, G., Li, S. and Shi, Y. (2009). A feedback regulatory loop involving microRNA-9 and nuclear receptor TLX in neural stem cell fate determination. *Nat Struct Mol Biol* **16**, 365–71.

Zhao, X., He, X., Han, X., Yu, Y., Ye, F., Chen, Y., Hoang, T., Xu, X., Mi, Q.-S., Xin, M., Wang, F., Appel, B. and Lu, Q. R. (2010). MicroRNA-mediated control of oligodendrocyte differentiation. *Neuron* **65**, 612–26.

Zipprich, J. T., Bhattacharyya, S., Mathys, H. and Filipowicz, W. (2009). Importance of the C-terminal domain of the human GW182 protein TNRC6C for translational repression. *RNA (New York, NY)* **15**, 781–93.

Appendix A

Supplementary Data

Table A.1: The list of putative neuron-specific genes

Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol
11419	Accn2	11488	Adam11	11496	Adam22	11518	Add1
11519	Add2	11674	Aldoa	11676	Aldoc	11735	Ank3
11739	Slc25a4	11769	Ap1s1	11771	Ap2a1	11772	Ap2a2
11773	Ap2m1	11775	Ap3b2	11789	Apc	11829	Aqp4
11838	Arc	11842	Arf3	11899	Astn1	11931	Atp1b1
11932	Atp1b2	11938	Atp2a2	11941	Atp2b2	11964	Atp6v1a
11966	Atp6v1b2	11972	Atp6v0d1	11973	Atp6v1e1	11975	Atp6v0a1
11980	Atp8a1	11981	Atp9a	12032	Bcan	12217	Bsn
12286	Cacna1a	12287	Cacna1b	12293	Cacna2d1	12294	Cacna2d3
12295	Cacnb1	12297	Cacnb3	12298	Cacnb4	12300	Cacng2
12313	Calm1	12314	Calm2	12315	Calm3	12322	Camk2a
12323	Camk2b	12361	Cask	12386	Ctnna2	12554	Cdh13
12558	Cdh2	12561	Cdh4	12568	Cdk5	12569	Cdk5r1
12669	Chrm1	12704	Cit	12709	Ckb	12716	Ckmt1
12799	Cnp	12805	Cntn1	12933	Crmp1	12934	Dpysl2
12950	Hapln1	13004	Ncan	13116	Cyp46a1	13175	Dclk1
13191	Dctn1	13196	Asap1	13199	Ddn	13384	Mpp3
13385	Dlg4	13401	Dmwd	13426	Dync1i1	13429	Dnm1
13476	Reep5	13480	Dpm1	13483	Dpp6	13527	Dtna
13609	S1pr1	13628	Eef1a2	13806	Eno1	13807	Eno2
13821	Epb4.1l1	13823	Epb4.1l3	13829	Epb4.9	13838	Epha4
13855	Epn2	13858	Eps15	14007	Cugbp2	14073	Faah
14086	Fscn1	14226	Fkbp1b	14360	Fyn	14394	Gabra1
14395	Gabra2	14396	Gabra3	14397	Gabra4	14400	Gabrb1
14401	Gabrb2	14402	Gabrb3	14415	Gad1	14432	Gap43
14457	Gas7	14545	Gdap1	14567	Gdi1	14571	Gpd2
14580	Gfap	14586	Gfra2	14645	Glul	14660	Gls
14677	Gnai1	14680	Gnal	14681	Gnao1	14682	Gnaq
14687	Gnaz	14688	Gnb1	14697	Gnb5	14702	Gng2
14704	Gng3	14708	Gng7	14758	Gpm6b	14768	Lancl1
14799	Gria1	14800	Gria2	14802	Gria4	14810	Grin1
14811	Grin2a	14812	Grin2b	15165	Hcn1	15275	Hk1
15441	Hp1bp3	15444	Hpca	15505	Hsph1	15512	Hspa2
15519	Hsp90aa1	15568	Elavl1	15571	Elavl3	15572	Elavl4
15898	Icam5	16438	Itpr1	16443	Itsn1	16485	Kcna1
16490	Kcna2	16497	Kcnab1	16498	Kcnab2	16499	Kcnab3
16500	Kcnb1	16508	Kcnd2	16531	Kcnma1	16536	Kcnq2
16560	Kif1a	16563	Kif2a	16568	Kif3a	16572	Kif5a
16574	Kif5c	16593	Klc1	16594	Klc2	16646	Kpna1
16653	Kras	16728	L1cam	16832	Ldhb	17136	Mag
17196	Mbp	17441	Mog	17449	Mdh1	17754	Mtap1a
17755	Mtap1b	17756	Mtap2	17758	Mtap4	17760	Mtap6
17761	Mtap7	17762	Mapt	17876	Myef2	17918	Myo5a
17957	Napb	17967	Ncam1	17968	Ncam2	18039	Nefl
18040	Nefm	18082	Nipsnap1	18117	Cox4nb	18125	Nos1
18164	Nptx1	18189	Nrxn1	18190	Nrxn2	18191	Nrxn3
18195	Nsf	18223	Numbl	18377	Omg	18415	Hspa4l
18479	Pak1	18483	Palm	18488	Cntn3	18526	Pcdh10
18555	Cdk16	18574	Pde1b	18578	Pde4b	18641	Pfkl
18642	Pfkm	18648	Pgam1	18717	Pip5k1c	18739	Pitpnm1
18746	Pkm2	18749	Prkacb	18752	Prkcc	18754	Prkce
18795	Plcb1	18798	Plcb4	18807	Pld3	18823	Plp1
18845	Plxna2	18952	Sept4	19055	Ppp3ca	19056	Ppp3cb
19084	Prkar1a	19085	Prkar1b	19139	Prps1	19242	Ptn
19261	Sirpa	19266	Ptprd	19280	Ptprs	19281	Ptprt
19283	Ptprz1	19290	Pura	19291	Purb	19317	Qk
19339	Rab3a	19346	Rab6	19387	Rangap1	19418	Rasgrf2
19679	Pitpnm2	19878	Rock2	19894	Rph3a	20168	Rtn3
20191	Ryr2	20192	Ryr3	20320	Nptn	20361	Sema7a
20362	Sept8	20404	Sh3gl2	20511	Slc1a2	20512	Slc1a3
20604	Sst	20614	Snap25	20616	Snap91	20740	Spna2
20741	Spnb1	20743	Spnb3	20817	Srpk2	20907	Stx1a
20910	Stxbp1	20927	Abcc8	20964	Syn1	20965	Syn2
20974	Syngr3	20977	Syp	20979	Syt1	20980	Syt2
21367	Cntn2	21402	Skp1a	21672	Prdx2	21838	Thy1
21960	Tnr	22031	Traf3	22142	Tuba1a	22143	Tuba1b
22151	Tubb2a	22152	Tubb3	22153	Tubb4	22223	Uchl1
22317	Vamp1	22318	Vamp2	22342	Lin7b	22393	Wfs1

Continued on Next Page...

The list of putative neuron-specific genes

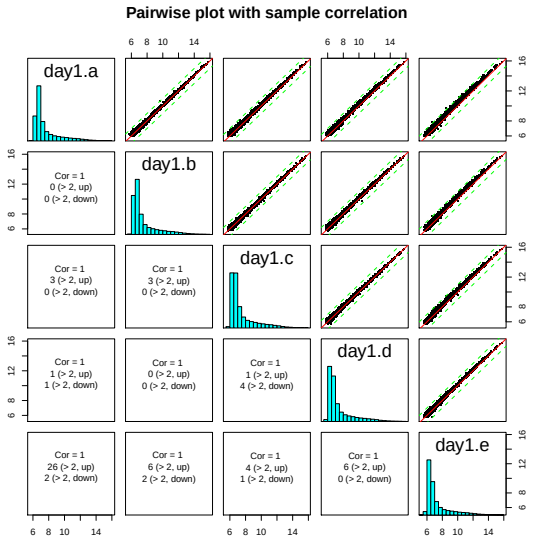
Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol
22628	Ywhag	22629	Ywhah	22631	Ywhaz	23792	Adam23
23859	Dlg2	23881	G3bp2	23936	Lynx1	23945	Mgll
23950	Dnajb6	23966	Odz4	23969	Pacsin1	24012	Rgs7
24050	Sept3	26372	Clcn6	26395	Map2k1	26413	Mapk1
26422	Nbea	26556	Homer1	26557	Homer2	26562	Ncdn
26757	Dpysl4	26874	Abcd2	26875	Pclo	26913	Gprin1
26932	Ppp2r5e	26950	Vsnl1	27062	Cadps	27204	Syn3
27373	Csnk1e	27801	Zdhhc8	27984	Efhd2	28185	Tomm70a
29873	Cspg5	30785	Cttnbp2	30948	Bin1	30957	Mapk8ip3
50791	Magi2	50876	Tmod2	50932	Mink1	50997	Mpp2
51792	Ppp2r1a	52389	Gpr123	52398	Sept11	52589	Ncald
52637	Cisd1	52822	Rufy3	52882	Rgs7bp	53310	Dlg3
53420	Syt5	53612	Vti1b	53623	Gria3	53870	Cntn6
53872	Caprin1	53972	Ngef	54161	Copg	54195	Gucy1b3
54216	Pcdh7	54376	Cacng3	54393	Gabbr1	54401	Ywhab
54403	Slc4a4	54411	Atp6ap1	54418	Fmn2	54525	Syt7
54637	Praf2	54712	Plxnc1	55992	Trim3	56013	Srcin1
56077	Dgke	56149	Grasp	56177	Olfm1	56320	Dbn1
56323	Dnajb5	56370	Tagln3	56421	Pfkp	56438	Rbx1
56455	Dynll1	56462	Mtch1	56491	Vapb	56508	Rapgef4
56526	Sept6	56541	Habp4	56637	Gsk3b	56695	Pnkd
56710	Dbc1	56737	Alg2	56808	Cacna2d2	56839	Lgi1
56876	Nelf	57138	Slc12a5	57340	Jph3	57440	Ehd3
57743	Sec61a2	57754	Cend1	57874	Ptplad1	58175	Rgs20
58234	Shank3	58244	Stx6	58994	Smpd3	64009	Syne1
64011	Nrgn	64051	Sv2a	64297	Gprc5b	64933	Ap3m2
65079	Rtn4r	65945	Clstn1	66049	Rogdi	66082	Abhd6
66098	Chchd6	66237	Atp6v1g2	66335	Atp6v1c1	66797	Cntnap2
66958	Tmx2	67166	Arl8b	67252	Cap2	67295	Rab3c
67306	Fam164a	67412	6330407J23Rik	67433	Ccdc127	67445	C1qtnf4
67453	Slc25a46	67564	Tmem35	67602	Necap1	67792	Rgs8
67801	Pllp	67826	Snap47	67834	Idh3a	67900	1700020C11Rik
67972	Atp2b1	68032	Tmem85	68166	Spire1	68203	Diras2
68267	Slc25a22	68404	Nrn1	68507	Ppfia4	68524	Wipf2
68585	Rtn4	68724	Arl8a	69219	Ddah1	69399	1700025G04Rik
69605	Lnp	69635	Dapk1	69642	2310046A06Rik	69683	2310044H10Rik
69807	Trim32	69894	2010107G23Rik	69908	Rab3b	69981	Tmem30a
70495	Atp6ap2	70549	Tln2	70620	Ube2v2	70762	Dclk2
71146	Golga7b	71302	Arhgap26	71435	Arhgap21	71764	C2cd2l
71770	Ap2b1	71803	Slc25a18	71835	Lancl2	71902	Cand1
72097	2010300C02Rik	72168	Aifm3	72325	1300018I17Rik	72685	Dnajc6
72727	B3gat3	72821	Scn2b	72832	Crtac1	72927	Hepacam
72948	Tppp	72961	Slc17a7	73072	BC068157	73094	Sgip1
73178	Wasl	73242	2610110G12Rik	73420	1700054N08Rik	73442	Hspa12a
73710	Tubb2b	73728	Psd	73825	Klraq1	73834	Atp6v1d
73991	Atl1	74006	Dnm1l	74012	Rap2b	74053	Grip1
74103	Nebl	74205	Acsl3	74256	Cyld	74342	Lrrtm1
74998	Rab11fip2	75029	Purg	75607	Wnk2	75734	Mff
75770	Brsk2	75786	Ckap5	75914	Exoc6b	76089	Rapgef2
76108	Rap2a	76156	Fam131b	76179	Usp31	76192	Abhd12
76217	Jakmip2	76441	Daam2	76499	Clasp2	76580	Mib2
76686	Clip3	76740	Efr3a	76742	Snx27	76787	Ppfia3
76809	Bri3bp	76820	Fam49a	76884	Cyfip2	76960	Bcas1
77480	Kidins220	77531	Anks1b	77573	Vps33a	77579	Myh10
77629	Sphkap	78283	Mtap7d2	78506	Efha2	78779	Spata2L
78808	Stxbp5	78830	Slc25a12	80286	Tusc3	80297	Spnb4
80334	Kcnip4	80906	Kcnip2	80987	Nckipsd	81840	Sorcs2
83767	Wasf1	93739	Gabarapl2	93765	Ube2n	94040	Clmn
94047	Cecr6	94229	Slc4a10	94280	Sfxn3	94282	Sfxn5
97387	Strn4	98660	Atp1a2	98732	Rab3gap2	99010	Lpcat4
99512	Wdr47	100732	Mapre3	103466	Nt5dc3	103967	Dnm3
104001	Rtn1	104015	Synj1	104027	Synpo	104082	Wdr7
104418	Dgkz	104718	Ttc7b	104886	Rab15	105298	Epdr1
105445	Dock9	105689	Mycbp2	105853	Mal2	106042	Prickle1
107065	Lrrtm2	107831	Bai1	108030	Lin7a	108068	Grm2
108069	Grm3	108071	Grm5	108083	Pip4k2b	108100	Baiap2
108123	Napg	108124	Napa	108664	Atp6v1h	108686	Ccdc88a
109676	Ank2	109934	Abr	110012	Gm16517	110279	Bcr
110391	Qdpr	110876	Scn2a1	110891	Slc8a2	116837	Rims1
116838	Rims2	117148	Necab2	118452	Baalc	140559	Igsf8

Continued on Next Page...

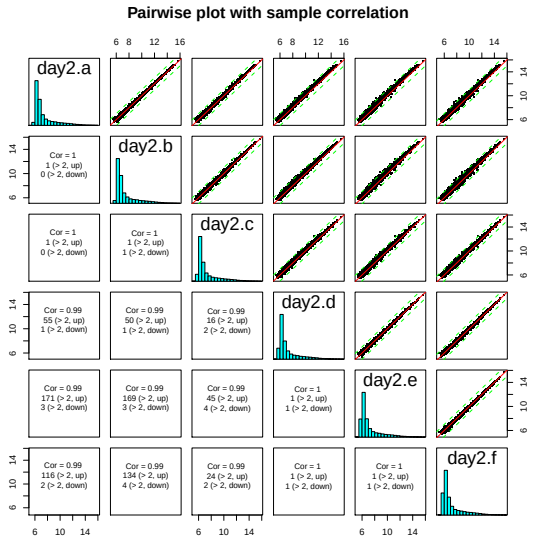
The list of putative neuron-specific genes

Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol
140579	Elmo2	140580	Elmo1	170731	Mfn2	170790	Mlc1
192197	Bcas3	194590	Reps2	207393	Elfn2	207565	Camkk2
207615	Wdr37	207728	Pde2a	208158	Map6d1	208869	Dock3
208898	Unc13c	210274	Shank2	210933	Bai3	211446	Exoc3
212307	Mapre2	213056	Fam126b	213469	Lgi3	213582	Mtap9
213990	Agap3	214230	Pak6	215690	Nav1	215707	Ccdc92
216028	Lrrtm3	216049	Zfp365	216739	Acsl6	216810	Tom1l2
216831	AU040829	216856	Nlgn2	216963	Git1	216965	Taok1
217219	Fam171a2	217480	Dgkb	217692	Sipa1l1	217882	AW555464
218035	Vps41	218038	Amph	218194	Phactr1	218440	Ankrd34b
218461	Pde8b	223435	Trio	223601	Fam49b	224020	Pi4ka
224617	Tbc1d24	224813	Gm88	224997	Dlgap1	225362	Reep2
225849	Ppp2r5b	226525	Rasal2	226751	Cdc42bpa	226778	Mark1
226977	Actr1b	227634	Camsap1	227937	Pkp4	228550	Itpka
228836	Dlgap4	228858	Gdap1l1	229521	Syt11	229709	Ahcyl1
229759	Olfm3	229791	D3Bwg0562e	229877	Rap1gds1	230085	N28178
230235	6430704M03Rik	230868	Igsf21	230904	Fbxo2	231148	Ablim2
231570	A830010M20Rik	231760	Rimbp2	231876	Lmtk2	232227	Iqsec1
232232	Hdac11	232333	Slc6a1	232813	Shisa7	232975	Atp1a3
233071	Snx26	234267	Gpm6a	234353	Psd3	234663	Dync1li2
235044	BC018242	235072	Sept7	235106	Ntm	235339	Dlat
235380	Dmxl2	235402	Lingo1	235431	Coro2b	235604	Camkv
236915	Arhgef9	237459	Cdk17	238130	Dock4	238276	Akap5
238988	Erc2	240058	Cpne5	240121	Fsd1	240185	9430020K01Rik
241263	Gpr158	241520	Fam171b	241589	D430041D05Rik	241638	RP23-100C5.8
241656	Pak7	241688	6330439K17Rik	241727	Snph	241770	Rims4
242481	Palm2	242667	Dlgap3	243043	Kctd8	243300	6430598A04Rik
243312	Elfn1	243499	Lrrtm4	243548	Prickle2	243621	Iqsec3
243743	Plxna4	244310	Dlgap2	244723	Olfm2	245643	Frmpd3
245666	Iqsec2	245684	Cnksr2	245877	Mtap7d1	245880	Wasf3
259302	Srgap3	260297	Prrt1	267019	Rps15a	268566	Gphn
268709	Fam107a	268890	Lsamp	268932	Caskin1	269060	Dagla
269109	Dpp10	269116	Nfasc	269180	Inpp4a	269295	Rtn4rl2
269774	Aak1	269854	Nat14	270058	Mtap1s	270192	Rab6b
271564	Vps13a	319278	A230050P20Rik	319504	Nrcam	319613	5730410E15Rik
319807	3110047P20Rik	319984	Jph4	320271	Scai	320365	Fry
320707	Atp2b3	320772	Mdga2	320840	Negr1	320873	Cdh10
327814	Ppfia2	329152	Hecw2	329165	Abi2	330319	Wipf3
330369	Fbxo41	330790	Hapln4	330814	Lphn1	330908	Opcml
330914	Grit	331461	Il1rapl1	347722	Agap1	360213	Trim46
380684	Nefh	380702	Shisa6	380768	Gm1568	381813	Prmt8
381979	Brsk1	382018	Unc13a	406218	Panx2	433904	Ociad2
545156	Kalrn	545389	Cep170	546071	Mast3	668212	Efr3b
100039795	Ildr2						

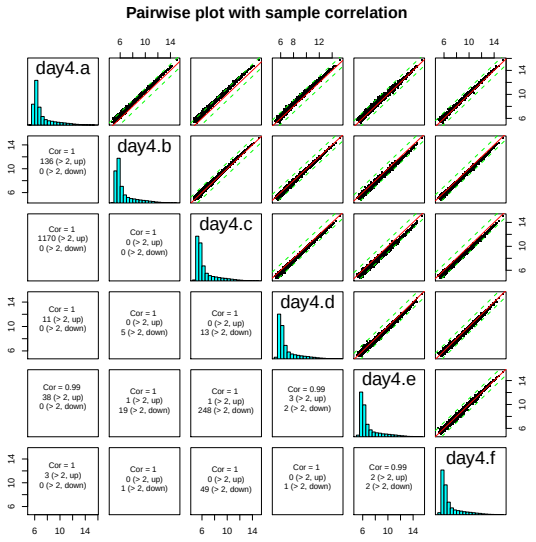
(a) 1DIV replicates



(b) 2DIV replicates



(c) 4DIV replicates



(d) 8DIV replicates

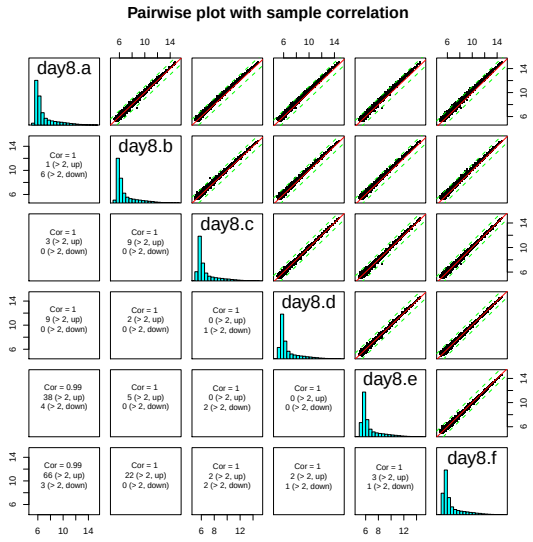
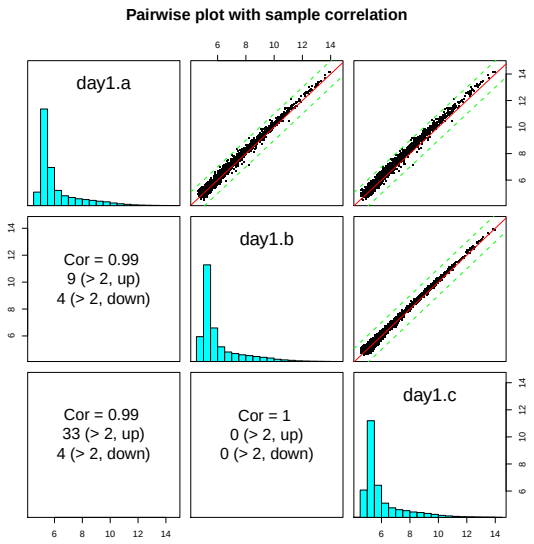
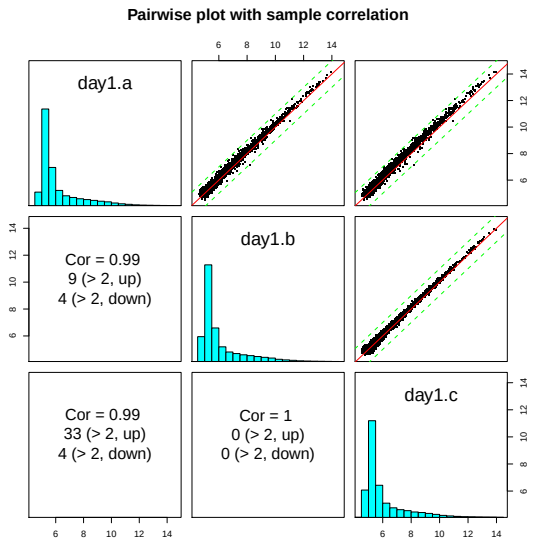


Figure A.1: Pairwise correlation of raw mRNA microarray probe intensities in profiles of hippocampal cultures.
Figure A.1a - correlation of replicates at 1 day of *in vitro* development (1DIV); Figure A.1b - at 2DIV; Figure A.1c - at 4DIV; Figure A.1d - at 8DIV. The plots were produced using *lumi* package (Du et al., 2008). The analysis of microarray data is described in Methods (section 2.7).

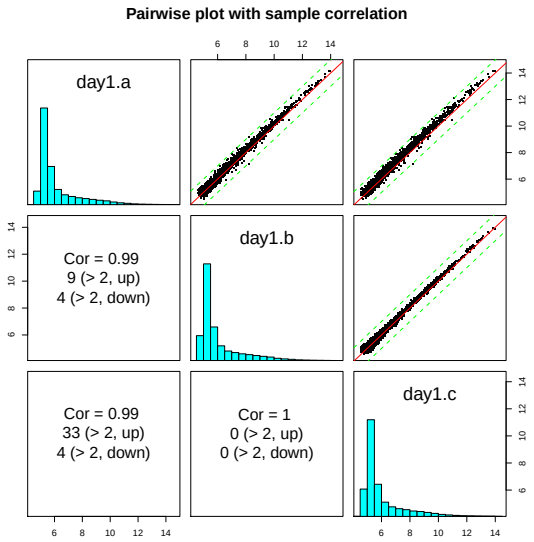
(a) 1DIV replicates



(b) 2DIV replicates



(c) 4DIV replicates



(d) 8DIV replicates

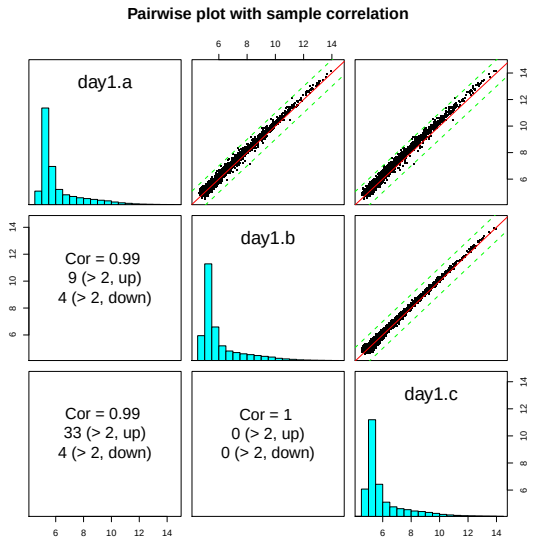
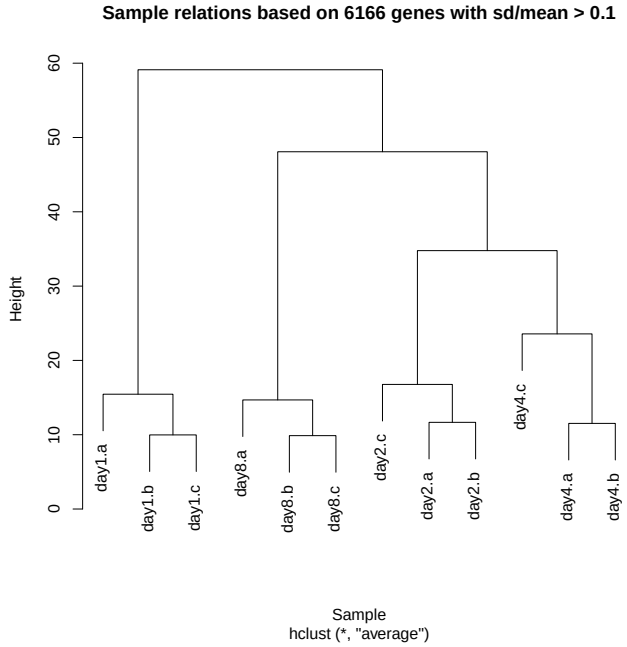


Figure A.2: Pairwise correlation of raw mRNA microarray probe intensities in profiles of hippocampal cultures.
Figure A.2a - correlation of replicates microarray profiling of RNA from replicates at 1 day of *in vitro* development (1DIV); Figure A.2b - at 2DIV; Figure A.2c - at 4DIV; Figure A.2d - at 8DIV. The plots were produced using *lumi* package (Du et al., 2008), see Methods, section 2.7. The analysis of microarray data is described in Methods (section 2.7).

(a) Hippocampal cultures



(b) Forebrain cultures

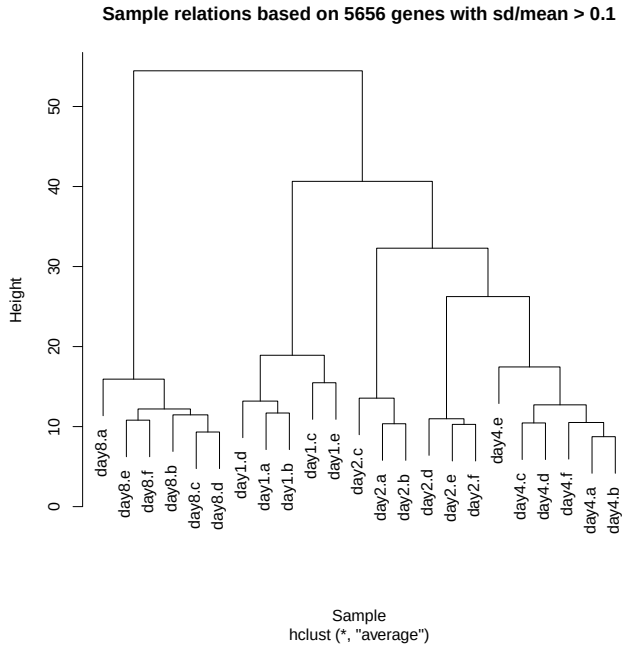
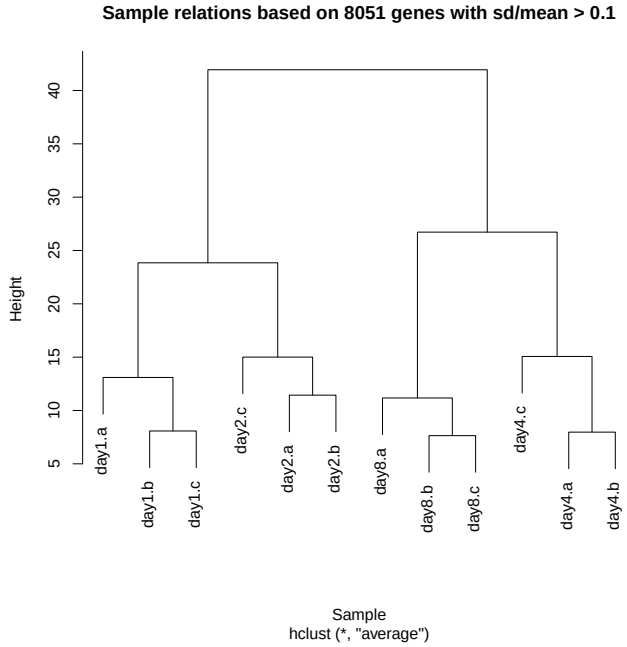


Figure A.3: Sample relation between raw mRNA microarray profiles between replicates of hippocampal and forebrain cultures.
Figure A.3a - replicates of hippocampal cultures; Figure A.3b - replicates of forebrain cultures. The plots were produced using *lumi* package (Du et al., 2008). The analysis of microarray data is described in Methods (section 2.7).

(a) Hippocampal cultures



(b) Forebrain cultures

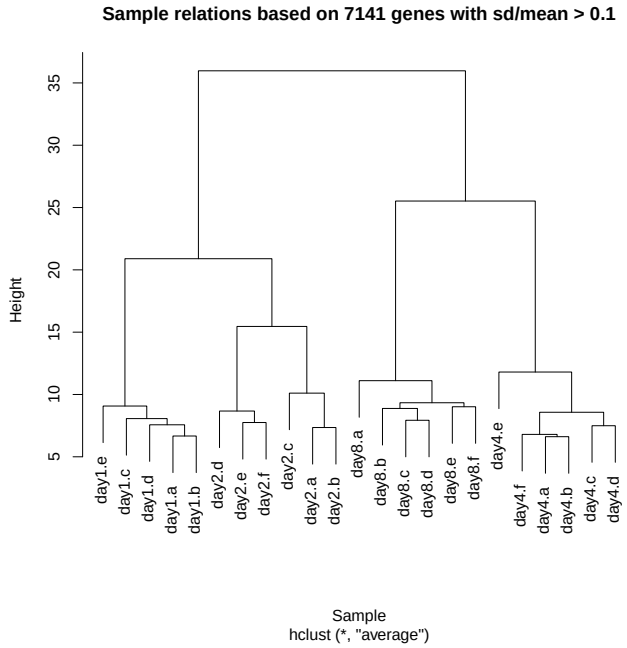


Figure A.4: Sample relation between normalised mRNA microarray profiles between replicates of hippocampal and forebrain cultures.
Figure A.4a - hippocampal cultures experimental replicates; Figure A.4b - forebrain cultures replicates. The plots were produced using *lumi* package (Du et al., 2008), see Methods, section 2.7. The analysis of microarray data is described in Methods (section 2.7).

Table A.2: Top 40 most enriched GO terms ("Biological process" type) in developmentally downregulated genes.
q - number of genes of a GO term that was among the downregulated genes, m - total number of genes of a GO term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	1086	1651	1.05e-83
GO:0010467	gene expression	1012	1545	3.24e-75
GO:0044260	cellular macromolecule metabolic process	1507	2537	7.65e-73
GO:0006807	nitrogen compound metabolic process	1124	1793	5.86e-68
GO:0043170	macromolecule metabolic process	1560	2709	4.96e-62
GO:0009059	macromolecule biosynthetic process	888	1412	5.09e-52
GO:0034645	cellular macromolecule biosynthetic process	880	1401	4.01e-51
GO:0006350	transcription	680	1029	1.03e-49
GO:0010556	regulation of macromolecule biosynthetic process	659	1010	4.42e-45
GO:0010468	regulation of gene expression	672	1040	7.9e-44
GO:0045449	regulation of transcription	612	936	4.95e-42
GO:0019219	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	632	975	1.18e-41
GO:0009889	regulation of biosynthetic process	665	1037	1.19e-41
GO:0051171	regulation of nitrogen compound metabolic process	635	981	1.43e-41
GO:0031326	regulation of cellular biosynthetic process	664	1036	1.82e-41
GO:0016070	RNA metabolic process	547	821	7.25e-41
GO:0060255	regulation of macromolecule metabolic process	719	1156	1.96e-38
GO:0044237	cellular metabolic process	1700	3185	3.05e-37
GO:0031323	regulation of cellular metabolic process	719	1180	4.27e-34
GO:0080090	regulation of primary metabolic process	701	1148	1.4e-33
GO:0044249	cellular biosynthetic process	973	1695	1.44e-32
GO:0044238	primary metabolic process	1690	3206	2.04e-32
GO:0019222	regulation of metabolic process	746	1252	6.29e-31
GO:0009058	biosynthetic process	986	1741	6.5e-30
GO:0006396	RNA processing	211	274	1.21e-28
GO:0016071	mRNA metabolic process	160	201	9.84e-25
GO:0006259	DNA metabolic process	185	244	8.57e-24
GO:0051276	chromosome organization	179	235	1.87e-23
GO:0007049	cell cycle	257	376	2.01e-21
GO:0006325	chromatin organization	154	200	3.76e-21
GO:0006397	mRNA processing	141	179	4.04e-21
GO:0008152	metabolic process	1817	3604	5.23e-21
GO:0006996	organelle organization	380	604	5.39e-21
GO:0008380	RNA splicing	116	141	1.94e-20

Continued on Next Page...

ID	Term Description	q	m	P
GO:0022403	cell cycle phase	142	184	9.44e-20
GO:0000279	M phase	125	157	1.41e-19
GO:0000278	mitotic cell cycle	126	160	5.6e-19
GO:0022402	cell cycle process	159	217	4.35e-18
GO:0000087	M phase of mitotic cell cycle	98	119	1.63e-17
GO:0000280	nuclear division	98	119	1.63e-17

Table A.3: Top 40 most enriched GO terms ("Biological process" type) in developmentally upregulated genes.
q - number of genes of a GO term that was among the upregulated genes, m - total number of genes of a GO term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0006810	transport	652	1196	2.64e-21
GO:0051234	establishment of localization	655	1206	7.27e-21
GO:0051179	localization	718	1360	9.92e-19
GO:0006811	ion transport	192	293	1.22e-16
GO:0006629	lipid metabolic process	188	299	1.09e-13
GO:0006812	cation transport	145	223	1.86e-12
GO:0019226	transmission of nerve impulse	104	148	2.19e-12
GO:0044255	cellular lipid metabolic process	135	205	2.73e-12
GO:0044281	small molecule metabolic process	319	577	1.87e-11
GO:0023052	signaling	626	1244	7.91e-11
GO:0030001	metal ion transport	122	189	2.23e-10
GO:0007268	synaptic transmission	85	121	2.36e-10
GO:0006836	neurotransmitter transport	44	53	9.05e-10
GO:0007267	cell-cell signaling	109	168	1.33e-09
GO:0006066	alcohol metabolic process	108	167	2.08e-09
GO:0050877	neurological system process	166	280	2.39e-09
GO:0003008	system process	193	336	4.37e-09
GO:0005975	carbohydrate metabolic process	126	203	4.46e-09
GO:0007154	cell communication	262	480	7.24e-09
GO:0023060	signal transmission	495	985	1.57e-08
GO:0015672	monovalent inorganic cation transport	85	128	1.75e-08
GO:0023046	signaling process	495	986	1.87e-08
GO:0008610	lipid biosynthetic process	89	137	3.98e-08
GO:0032787	monocarboxylic acid metabolic process	73	111	3.2e-07
GO:0055114	oxidation reduction	165	295	6.49e-07
GO:0055085	transmembrane transport	129	223	9.96e-07
GO:0007610	behavior	97	162	2.82e-06

Continued on Next Page...

ID	Term Description	q	m	P
GO:0006631	fatty acid metabolic process	53	78	3.01e-06
GO:0046483	heterocycle metabolic process	91	151	3.9e-06
GO:0065008	regulation of biological quality	255	493	4.56e-06
GO:0006814	sodium ion transport	35	47	6.09e-06
GO:0050801	ion homeostasis	74	119	6.31e-06
GO:0001505	regulation of neurotransmitter levels	32	42	6.73e-06
GO:0042180	cellular ketone metabolic process	127	225	6.86e-06
GO:0019637	organophosphate metabolic process	58	89	7.96e-06
GO:0006873	cellular ion homeostasis	69	110	8.37e-06
GO:0015837	amine transport	27	34	9.32e-06
GO:0015849	organic acid transport	30	39	9.42e-06
GO:0046942	carboxylic acid transport	30	39	9.42e-06
GO:0019725	cellular homeostasis	91	154	1.21e-05

Table A.4: Top 40 most enriched GO terms ("Cellular compartment" type) in developmentally downregulated genes.
q - number of genes of a GO term that was among the downregulated genes, m - total number of genes of a GO term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0005634	nucleus	1502	2267	2.38e-133
GO:0044428	nuclear part	368	498	5.21e-43
GO:0043226	organelle	2207	4281	1.26e-41
GO:0043229	intracellular organelle	2205	4279	2.46e-41
GO:0044424	intracellular part	2521	5009	5.34e-41
GO:0005622	intracellular	2575	5141	2.23e-40
GO:0043227	membrane-bounded organelle	2015	3870	2.87e-39
GO:0043231	intracellular membrane-bounded organelle	2013	3867	4.21e-39
GO:0005694	chromosome	187	226	2.78e-33
GO:0044427	chromosomal part	164	194	1.53e-31
GO:0031981	nuclear lumen	204	278	2.49e-23
GO:0070013	intracellular organelle lumen	225	319	7.29e-22
GO:0043233	organelle lumen	225	320	1.4e-21
GO:0031974	membrane-enclosed lumen	233	335	2.54e-21
GO:0043228	non-membrane-bounded organelle	525	885	4.09e-21
GO:0043232	intracellular non-membrane-bounded organelle	525	885	4.09e-21
GO:0032991	macromolecular complex	680	1194	4.98e-21
GO:0030529	ribonucleoprotein complex	188	262	9.4e-20
GO:0044422	organelle part	734	1323	1.06e-18
GO:0044446	intracellular organelle part	728	1316	4.15e-18

Continued on Next Page...

ID	Term Description	q	m	P
GO:0005681	spliceosomal complex	72	80	9.21e-18
GO:0000775	chromosome, centromeric region	69	77	7.87e-17
GO:0005654	nucleoplasm	146	201	2.38e-16
GO:0044451	nucleoplasm part	132	179	8.9e-16
GO:0000785	chromatin	78	95	3.52e-14
GO:0000776	kinetochore	41	45	5.03e-11
GO:0044454	nuclear chromosome part	47	58	1.09e-08
GO:0000502	proteasome complex	32	36	3.14e-08
GO:0000228	nuclear chromosome	49	63	6.37e-08
GO:0005667	transcription factor complex	66	92	9.2e-08
GO:0032993	protein-DNA complex	38	47	3.24e-07
GO:0000792	heterochromatin	28	32	5.06e-07
GO:0005730	nucleolus	57	79	5.08e-07
GO:0043234	protein complex	462	888	9.71e-07
GO:0005657	replication fork	16	16	2.28e-06
GO:0000790	nuclear chromatin	27	32	3.69e-06
GO:0000786	nucleosome	33	42	6.44e-06
GO:0034399	nuclear periphery	19	21	1.4e-05
GO:0005635	nuclear envelope	49	71	2.35e-05
GO:0005819	spindle	24	29	2.55e-05

Table A.5: Top 40 most enriched GO terms ("Cellular compartment" type) in developmentally upregulated genes.
q - number of genes of a GO term that was among the upregulated genes, m - total number of genes of a GO term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0016020	membrane	1627	2810	4.21e-92
GO:0044425	membrane part	1337	2226	1.59e-85
GO:0031224	intrinsic to membrane	1167	1914	2.04e-77
GO:0016021	integral to membrane	1141	1877	4.62e-74
GO:0005886	plasma membrane	628	1023	8.58e-39
GO:0044459	plasma membrane part	303	486	7.99e-20
GO:0044444	cytoplasmic part	1053	2092	2.76e-17
GO:0005783	endoplasmic reticulum	256	423	1.34e-14
GO:0045202	synapse	123	175	5.21e-14
GO:0044456	synapse part	79	104	2.88e-12
GO:0005576	extracellular region	251	429	4.92e-12
GO:0031226	intrinsic to plasma membrane	105	153	3.87e-11
GO:0005887	integral to plasma membrane	101	148	1.52e-10

Continued on Next Page...

ID	Term Description	q	m	P
GO:0030054	cell junction	149	239	2.75e-10
GO:0005624	membrane fraction	136	219	2.42e-09
GO:0005626	insoluble fraction	139	225	2.66e-09
GO:0000267	cell fraction	153	253	3.66e-09
GO:0043005	neuron projection	77	113	2.59e-08
GO:0005773	vacuole	79	117	3.21e-08
GO:0030136	clathrin-coated vesicle	37	46	1.46e-07
GO:0031410	cytoplasmic vesicle	131	221	2.55e-07
GO:0031982	vesicle	133	225	2.58e-07
GO:0008021	synaptic vesicle	31	37	2.7e-07
GO:0000323	lytic vacuole	68	103	1.05e-06
GO:0005764	lysosome	68	103	1.05e-06
GO:0030135	coated vesicle	42	57	1.62e-06
GO:0045211	postsynaptic membrane	46	65	3.38e-06
GO:0030424	axon	44	62	4.9e-06
GO:0005794	Golgi apparatus	208	396	2.28e-05
GO:0016023	cytoplasmic membrane-bounded vesicle	74	121	2.33e-05
GO:0031988	membrane-bounded vesicle	76	125	2.44e-05
GO:0031225	anchored to membrane	32	44	4.46e-05
GO:0043025	neuronal cell body	34	48	6.3e-05
GO:0044297	cell body	34	48	6.3e-05
GO:0005737	cytoplasm	1568	3493	7.23e-05
GO:0019717	synaptosome	31	43	7.75e-05
GO:0044421	extracellular region part	115	208	0.00011
GO:0030665	clathrin coated vesicle membrane	15	17	0.000128
GO:0030425	dendrite	28	39	0.000194
GO:0042995	cell projection	152	289	0.000262

Table A.6: Top 25 most enriched KEGG terms in developmentally downregulated genes. q - number of genes of a KEGG term that was among the downregulated genes, m - total number of genes of a KEGG term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
03040	Spliceosome	73	81	3.04e-20
03030	DNA replication	27	28	8.61e-10
04110	Cell cycle	57	75	1.14e-09
03050	Proteasome	27	30	4.54e-08
03440	Homologous recombination	20	21	3.16e-07
03420	Nucleotide excision repair	29	36	2.08e-06
03018	RNA degradation	28	36	1.11e-05

Continued on Next Page...

ID	Term Description	q	m	P
05322	Systemic lupus erythematosus	36	50	1.31e-05
03430	Mismatch repair	17	19	2.21e-05
00240	Pyrimidine metabolism	40	59	4.28e-05
03022	Basal transcription factors	17	20	8.97e-05
03410	Base excision repair	17	22	0.000772
03020	RNA polymerase	15	19	0.0011
05222	Small cell lung cancer	30	50	0.00695
04120	Ubiquitin mediated proteolysis	51	94	0.00922
04623	Cytosolic DNA-sensing pathway	15	22	0.0113
00310	Lysine degradation	15	24	0.0331
03010	Ribosome	20	34	0.0337
04115	p53 signaling pathway	20	34	0.0337
04670	Leukocyte transendothelial migration	28	51	0.04
04114	Oocyte meiosis	34	64	0.0437
04620	Toll-like receptor signaling pathway	24	43	0.0444
00230	Purine metabolism	46	91	0.0557
05200	Pathways in cancer	82	171	0.0568
04621	NOD-like receptor signaling pathway	13	22	0.078

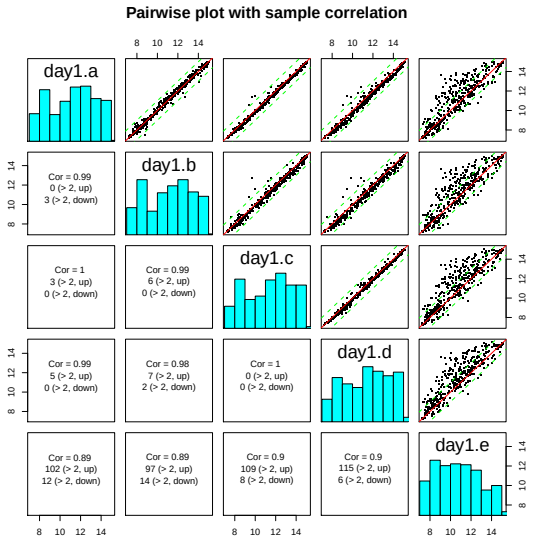
Table A.7: Top 25 most enriched KEGG terms in developmentally upregulated genes. q - number of genes of a KEGG term that was among the upregulated genes, m - total number of genes of a KEGG term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
01100	Metabolic pathways	303	564	1.46e-06
04142	Lysosome	53	80	8.44e-05
04020	Calcium signaling pathway	48	71	8.5e-05
04080	Neuroactive ligand-receptor interaction	44	64	9.14e-05
00600	Sphingolipid metabolism	21	26	0.000201
00640	Propanoate metabolism	15	17	0.00028
00010	Glycolysis / Gluconeogenesis	22	28	0.000293
04514	Cell adhesion molecules (CAMs)	29	42	0.00134
04720	Long-term potentiation	29	42	0.00134
00982	Drug metabolism - cytochrome P450	15	19	0.00272
00564	Glycerophospholipid metabolism	26	38	0.00292
00561	Glycerolipid metabolism	19	26	0.00345
00511	Other glycan degradation	7	7	0.00374
00603	Glycosphingolipid biosynthesis - globo series	7	7	0.00374
00980	Metabolism of xenobiotics by cytochrome P450	14	18	0.00485
00071	Fatty acid metabolism	15	20	0.00636

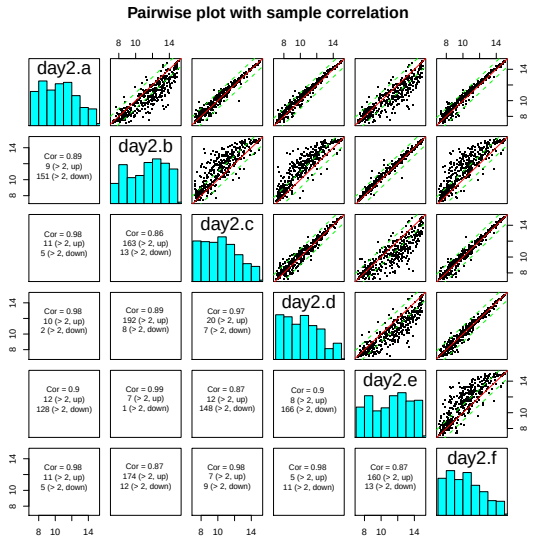
Continued on Next Page...

ID	Term Description	q	m	P
04916	Melanogenesis	32	51	0.00778
03320	PPAR signaling pathway	20	29	0.0078
00062	Fatty acid elongation in mitochondria	6	6	0.00832
04540	Gap junction	34	55	0.00854
00280	Valine, leucine and isoleucine degradation	21	31	0.0088
00604	Glycosphingolipid biosynthesis - ganglio series	8	9	0.00908
00910	Nitrogen metabolism	8	9	0.00908
04260	Cardiac muscle contraction	28	44	0.00956
04730	Long-term depression	24	38	0.0182

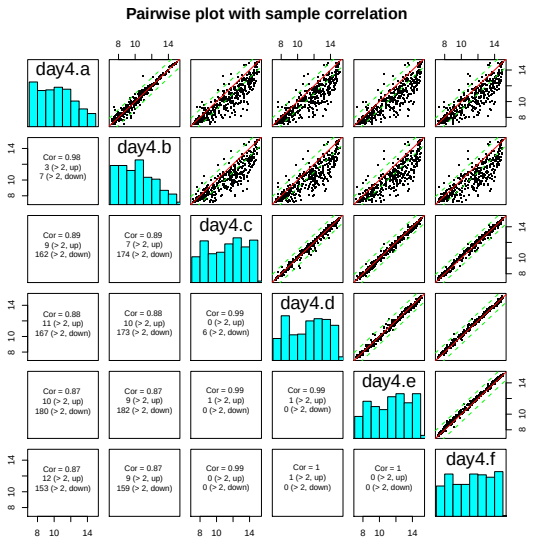
(a) 1DIV replicates



(b) 2DIV replicates



(c) 4DIV replicates



(d) 8DIV replicates

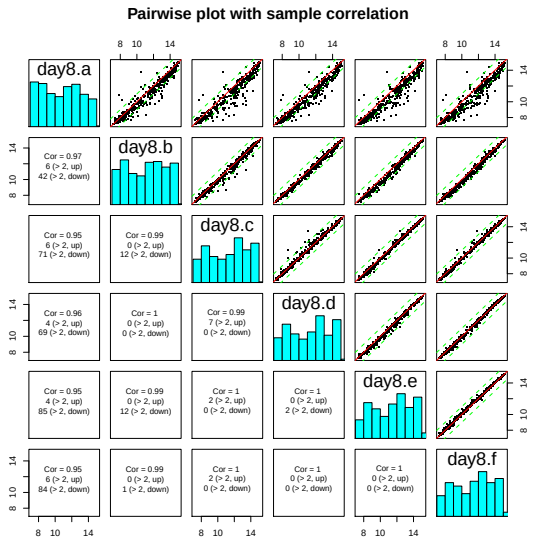
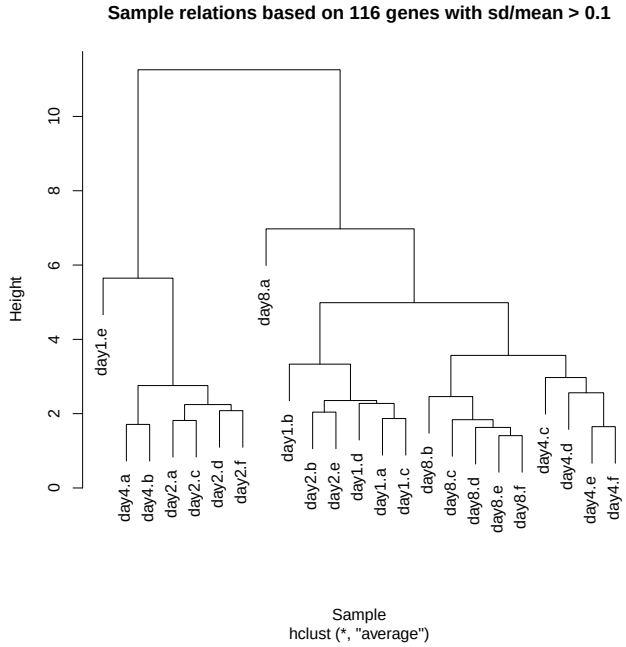


Figure A.5: Pairwise correlation of raw miRNA microarray probe intensities in profiles of forebrain cultures.
Figure A.1a - correlation of replicates at 1 day of *in vitro* development (1DIV); Figure A.1b - from 2DIV; Figure A.1c - from 4DIV; Figure A.1d - from 8DIV. The plots were produced using *lumi* package (Du et al., 2008). The analysis of microarray data is described in Methods (section 2.7).

(a)



(b)

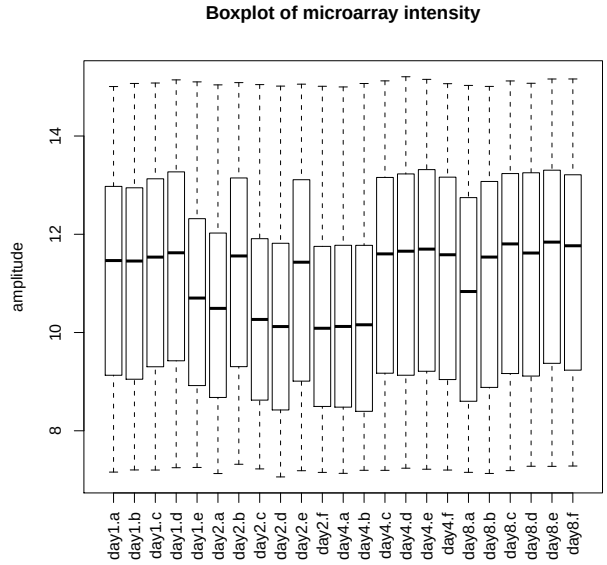


Figure A.6: Relationship between replicate raw miRNA microarray profiles.
Figure A.6a - sample relation between raw miRNA microarray profiles between replicates of forebrain cultures; Figure A.6b - miRNA microarray intensities of replicates of forebrain cultures. The plots were produced using *lumi* package (Du et al., 2008). The analysis of microarray data is described in Methods (section 2.7).

Table A.8: Three categories of miRNAs in development of primary forebrain cultures. miRNA identifiers, miRBase Release 13 (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006, 2008), for each member of the three categeries of miRNAs with distinct modes of expression are given together with the rank (**"Rank"**) of expression at the 8DIV timepoint in the development of cultures.

Steady state	**Rank**	**Downregulated**	**Rank**	**Upregulated**	**Rank**
mmu-miR-9	1	mmu-miR-15b	29	mmu-miR-690	17
mmu-miR-103	2	mmu-miR-99a	30	mmu-miR-24	24
mmu-let-7a	3	mmu-miR-21	39	mmu-miR-434-3p	31
mmu-miR-125b-5p	4	mmu-miR-135a	41	mmu-miR-376b	33
mmu-let-7b	5	mmu-miR-20a	43	mmu-miR-7a	34
mmu-miR-711	6	mmu-miR-135b	47	mmu-miR-218	35
mmu-miR-16	7	mmu-miR-93	52	mmu-miR-709	36
mmu-miR-26a	8	mmu-miR-106b	53	mmu-miR-22	38
mmu-miR-137	9	mmu-miR-149	56	mmu-miR-551b	40
mmu-miR-124	10	mmu-miR-99b	65	mmu-miR-410	44
mmu-let-7d	11	mmu-miR-706	66	mmu-miR-128	45
mmu-let-7g	12	mmu-miR-30e	68	mmu-miR-331-3p	46
mmu-miR-191	13	mmu-miR-335-5p	70	mmu-miR-342-3p	48
mmu-miR-9*	14	mmu-miR-25	72	mmu-miR-30d	50
mmu-let-7c	15	mmu-miR-20b	78	mmu-miR-487b	51
mmu-let-7f	16	mmu-miR-92a	81	mmu-miR-139-5p	54
mmu-miR-125a-5p	18	mmu-miR-15a	83	mmu-miR-127	55
mmu-miR-30c	19	mmu-miR-98	86	mmu-miR-129-3p	59
mmu-miR-17	20	mmu-miR-195	88	mmu-miR-379	60
mmu-let-7i	21	mmu-miR-350	89	mmu-miR-382	61
mmu-miR-181a	22	mmu-miR-18a	91	mmu-miR-138	62
mmu-let-7e	23	mmu-miR-27b	93	mmu-miR-154	63
mmu-miR-181b	25	mmu-miR-301a	94	mmu-miR-338-3p	67
mmu-miR-693-5p	26	mmu-miR-674	95	mmu-miR-132	69
mmu-miR-720	27	mmu-miR-101a	96	mmu-miR-298	74
mmu-miR-100	28	mmu-miR-744	100	mmu-miR-326	76
		mmu-miR-19b	102	mmu-miR-434-5p	80
		mmu-miR-27a	103	mmu-miR-323-3p	84
		mmu-miR-106a	106	mmu-miR-328	85
		mmu-miR-28	110	mmu-miR-495	90
		mmu-miR-181a-1*	116	mmu-miR-324-5p	92
		mmu-miR-374	117	mmu-miR-409-3p	97
		mmu-miR-672	122	mmu-miR-369-5p	99
		mmu-miR-19a	123	mmu-miR-369-3p	104
		mmu-miR-204	130	mmu-miR-668	105
		mmu-miR-322	132	mmu-miR-29c	111

Continued on Next Page...

Three types of miRNAs in development of primary forebrain cultures

Steady state	Rank	Downregulated	Rank	Upregulated	Rank
		mmu-miR-23a	135	mmu-miR-185	112
		mmu-miR-701	137	mmu-miR-376c	114
		mmu-miR-466a-3p	145	mmu-miR-676	115
		mmu-miR-210	149	mmu-miR-592	119
		mmu-miR-423-3p	150	mmu-miR-29a	120
		mmu-miR-126-3p	156	mmu-miR-541	121
		mmu-miR-130b	159	mmu-miR-134	125
		mmu-miR-30a*	160	mmu-miR-187	126
		mmu-miR-451	163	mmu-miR-376b*	127
		mmu-miR-219	167	mmu-miR-129-5p	128
		mmu-miR-351	169	mmu-miR-146b	129
		mmu-miR-297a	175	mmu-miR-673-5p	133
		mmu-miR-685	177	mmu-miR-136	134
		mmu-miR-615-3p	180	mmu-miR-329	136
		mmu-miR-339-5p	181	mmu-miR-376a*	139
		mmu-miR-126-5p	183	mmu-miR-29b	140
		mmu-miR-144	190	mmu-miR-337-3p	141
		mmu-miR-489	194	mmu-miR-7b	142
		mmu-miR-542-3p	195	mmu-miR-222	147
		mmu-miR-450a-5p	196	mmu-miR-370	148
		mmu-miR-192	197	mmu-miR-667	152
		mmu-miR-679	198	mmu-miR-378	154
		mmu-miR-203	202	mmu-miR-330*	162
		mmu-miR-503	203	mmu-miR-485	165
		mmu-miR-345-5p	204	mmu-miR-539	166
		mmu-miR-17*	205	mmu-miR-433	168
		mmu-miR-467a*	206	mmu-miR-433*	173
		mmu-miR-322*	214	mmu-miR-496	174
		mmu-miR-215	216	mmu-miR-485*	176
		mmu-miR-199a-5p	223	mmu-miR-365	184
		mmu-miR-146a	225	mmu-miR-383	191
		mmu-miR-761	226	mmu-miR-31	192
		mmu-miR-145	228	mmu-miR-543	200
		mmu-miR-155	232	mmu-miR-666-5p	207
		mmu-miR-450b-3p	238	mmu-miR-221	210
		mmu-miR-122	245	mmu-miR-700	211
		mmu-miR-142-3p	249	mmu-miR-770-3p	218
		mmu-miR-223	255	mmu-miR-377	224
		mmu-miR-301b	260	mmu-miR-501-3p	230

Continued on Next Page...

Three types of miRNAs in development of primary forebrain cultures

Steady state	Rank	Downregulated	Rank	Upregulated	Rank
		mmu-miR-703	264	mmu-miR-702	231
		mmu-miR-199a-3p	265	mmu-miR-296-5p	233
		mmu-miR-143	267	mmu-miR-182	235
		mmu-miR-325*	268	mmu-miR-183	239
		mmu-miR-142-5p	269	mmu-miR-715	242
		mmu-miR-302b	274	mmu-miR-689	247
		mmu-miR-214	275	mmu-miR-380-3p	248
		mmu-miR-697	276	mmu-miR-211	252
		mmu-miR-199b*	279	mmu-miR-378*	254
		mmu-miR-483*	283	mmu-miR-488*	256
		mmu-miR-450b-5p	288	mmu-miR-412	257
		mmu-miR-448	289	mmu-miR-380-5p	258
		mmu-miR-704	297	mmu-miR-431	259
		mmu-miR-224	298	mmu-miR-676*	272
		mmu-miR-217	312	mmu-miR-760	273
		mmu-miR-200a	329	mmu-miR-133a	277
		mmu-miR-452	336	mmu-miR-300	278
		mmu-miR-150	338	mmu-miR-24-1*	281
		mmu-miR-363	340	mmu-miR-206	292
		mmu-miR-216a	342	mmu-miR-208a	294
		mmu-miR-464	344	mmu-miR-686	296
		mmu-miR-698	347	mmu-miR-201	302
		mmu-miR-10a	349	mmu-miR-499	306
		mmu-miR-675-3p	358	mmu-miR-681	307
				mmu-miR-196b	315
				mmu-miR-705	321
				mmu-miR-692	322
				mmu-miR-196a	326
				mmu-miR-717	343
				mmu-miR-680	354

Table A.9: Putative direct targets of miR-124

Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol
11370	Acadvl	11443	Chrnb1	11491	Adam17	11518	Add1
11600	Angpt1	11637	Ak2	11666	Abcd1	11736	Ankfy1
11744	Anxa11	11747	Anxa5	11867	Arpc1b	11928	Atp1a1
11974	Atp6v0e	12039	Bckdha	12042	Bcl10	12161	Bmp6
12192	Zfp36l1	12321	Calu	12334	Capn2	12350	Car3
12389	Cav1	12443	Ccnd1	12476	Cd151	12499	Entpd5
12521	Cd82	12753	Clock	12826	Col4a1	12831	Col5a1
12837	Col8a1	12908	Crat	13559	E2f5	13610	S1pr3
13617	Ednra	13650	Rhbdf1	13731	Emp2	13846	Ephb4
13866	Erbb2	14020	Evi5	14082	Fadd	14085	Fah
14252	Flot2	14275	Folr1	14314	Fstl1	14375	Xrcc6
14420	Galc	14450	Gart	14595	B4galt1	14678	Gnai2
14726	Pdpn	14792	Lpcat3	15077	Hist2h3c1	15894	Icam1
16007	Cyr61	16009	Igfbp3	16206	Lrig1	16211	Kpnb1
16362	Irf1	16404	Itga7	16412	Itgb1	16561	Kif1b
16589	Uhmk1	16651	Sspn	16784	Lamp2	16848	Lfng
16854	Lgals3	16889	Lipa	16905	Lmna	17083	Tmed1
17127	Smad3	17129	Smad5	17150	Mfap2	17158	Man2a1
17196	Mbp	17216	Mcm2	17217	Mcm4	17242	Mdk
17865	Mybl2	17886	Myh9	17997	Nedd1	18018	Nfatc1
18028	Nfib	18029	Nfic	18032	Nfix	18073	Nid1
18140	Uhrf1	18176	Nras	18201	Nsmaf	18212	Ntrk2
18230	Nxn	18451	P4ha1	18553	Pcsk6	18810	Plec
18824	Plp2	18933	Prrx1	19027	Sypl	19193	Pipox
19205	Ptbp1	19247	Ptpn11	19248	Ptpn12	19250	Ptpn14
19294	Pvrl2	19334	Rab22a	19340	Rab3d	19356	Rad17
19376	Rab34	19697	Rela	19724	Rfx1	19729	Rag1ap1
20130	Rras	20187	Ryk	20249	Scd1	20397	Sgpl1
20416	Shc1	20481	Ski	20496	Slc12a2	20648	Snta1
20848	Stat3	20917	Suclg2	20971	Sdc4	21367	Cntn2
21413	Tcf4	21415	Tcf7l1	21766	Tex261	21859	Timp3
21871	Atp6v0a2	21873	Tjp2	21915	Dtymk	22031	Traf3
22092	Rsph1	22117	Tst	22158	Tulp3	22169	Cmpk2
22271	Upp1	22319	Vamp3	22352	Vim	22401	Zmat3
22403	Wisp2	22695	Zfp36	23885	Gmcl1	23959	Nt5e
23972	Papss2	24044	Scamp2	26416	Mapk14	26425	Nubp1
26433	Plod3	26457	Slc27a1	26564	Ror2	26754	Cops5
27041	G3bp1	27081	Zfp275	27401	Skp2	27410	Abca3
28146	Serp1	28193	Reep3	29875	Iqgap1	30934	Tor1b
30935	Tor3a	50496	E2f6	50918	Myadm	52009	Hn1l
52398	Sept11	52428	Rhpn2	52538	Acaa2	52585	Dhrs1
52840	Dbndd2	53330	Vamp4	53376	Usp2	53378	Sdcbp
53415	Htatip2	53599	Cd164	53623	Gria3	53860	Sept9
54325	Elovl1	54720	Rcan1	56016	Hebp2	56212	Rhog
56248	Ak3	56309	Mycbp	56332	Amotl2	56356	Gltp
56369	Apip	56494	Gosr2	56517	Slc22a21	56520	Nme4
56709	Dnajb12	56722	Litaf	56741	Igdcc4	57267	Apba3
57315	Wdr46	58809	Rnase4	60595	Actn4	65960	Twsg1
66153	Fbxo36	66395	Ahnak	66500	Slc30a7	66523	2810004N23Rik
66616	Snx9	66659	Acp6	66717	Ccdc96	66853	Pnpla2
66859	Slc16a9	66913	Kdelr2	66990	Tmem134	67145	Tomm34
67213	Cmtm6	67374	Jam2	67603	Dusp6	67605	Akt1s1
67843	Slc35a4	67951	Tubb6	67980	Gnpda2	67991	Nacc2
68041	Mid1ip1	68066	Slc25a39	68226	Efcab2	68270	Lrrc50
68465	Adipor2	68520	Zfyve21	68539	Tmem109	68581	Tmed10
68606	Ppm1f	68682	Slc44a2	68738	Acss1	68794	Flnc
69274	Ctdspl	69683	2310044H10Rik	69737	Ttl	70024	Mcm10
70218	Kif18b	70417	Megf10	70435	Inf2	70461	Crtc3
70806	D19Ertd652e	70984	4931406C07Rik	71409	Fmnl2	71567	Mcm9
71602	Myo1e	71712	Dram1	71766	Raver1	71801	Plekhf2
71918	Zcchc24	71943	Tom1l1	71946	Endod1	71956	Rnf135
72157	Pgm2	72287	Plekhf1	72792	2810459M11Rik	73284	Ddit4l
73827	1110012D08Rik	74098	0610037L13Rik	74105	Gga2	74533	Gzf1
75452	Ascc2	75556	1700026D08Rik	75563	Dnali1	75599	Pcdh1
75646	Rai14	75659	Wdr54	75723	Amotl1	76044	Ncapg2
76178	6330578E17Rik	76251	0610007P08Rik	76263	Gstk1	76491	Abhd14b
76566	Fam101b	76893	Lass2	76895	Bicd2	77034	2510039O18Rik
77056	Tmco4	77446	Heg1	77559	Agl	77569	Limch1

Continued on Next Page...

Putative direct targets of miR-124

Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol
77579	Myh10	78388	Mvp	78619	Zfp449	78829	Tsc22d4
78926	Gas2l1	79202	Tnfrsf22	80290	Gpr146	80888	Hspb8
81840	Sorcs2	81879	Tcfcp2l1	81910	Rrbp1	83436	Plekha2
83675	Bicc1	99003	Qser1	99889	Arfip1	100072	Camta1
101543	Wtip	102462	Imp3	102626	Mapkapk3	102644	Oaf
102693	Phldb1	104027	Synpo	105501	Abhd4	106581	Itfg3
106639	Vmac	106840	Unc119b	107976	Bre	108657	Rnpepl1
108673	Ccdc86	108682	Gpt2	108705	Pttg1ip	108735	Sft2d2
109145	Gins4	109154	Mlec	109333	Pkn2	109672	Cyb5
110379	Sec13	110826	Etfb	114774	Pawr	116701	Fgfrl1
116972	Fam57a	117150	Pip4k2c	140570	Plxnb2	140579	Elmo2
170625	Snx18	170748	BC017612	171212	Galnt10	171286	Slc12a8
207175	Cetn4	209195	Clic6	209378	Itih5	209601	4922501L14Rik
211945	Plekhh1	212647	Aldh4a1	214345	Lrrc1	214968	Sema6d
215280	Wipf1	215751	BC013529	217365	Nploc4	217430	Pqlc3
217684	4933426M11Rik	217718	Nek9	218503	Fcho2	218630	Ccno
218952	Fermt2	219148	Fam167a	219189	1300010F03Rik	223693	Tmem184b
224143	Ktelc1	225164	Mib1	226162	Dpcd	226265	Eno4
226519	Lamc1	227292	Ctdsp1	227737	Fam129b	228775	Trib3
228942	Cbln4	228966	Ppp1r3d	229096	Ythdf3	229285	Spg20
230709	Zmpste24	230751	Oscp1	230779	Serinc2	230789	Fam76a
230967	BC046331	231452	Sdad1	233033	Samd4b	233315	Mtmr10
234839	Fam38a	237806	Dnahc9	239273	Abcc4	240660	Tmem20
242553	Kank4	242585	Slc35d1	242687	Wasf2	242785	Klhl21
244152	Tsku	244631	Pskh1	246257	Ovca2	246316	Lgi2
259302	Srgap3	268935	Scube3	269593	Luzp1	269941	Chsy1
269999	Orai3	286940	Flnb	319710	Frmd6	319939	Tns3
320184	Lrrc58	320404	Itpkb	320736	E130203B14Rik	326618	Tpm4
329274	Fam163a	330171	Kctd10	330222	Sdk1	330695	Ctxn1
338365	Slc41a2	382030	Tmem188	382406	Wdr51b	414801	Itprip
432572	Cytsb	544963	Iqgap2	100045343	LOC100045343	100046855	LOC100046855
100047738	LOC100047738	100047856	LOC100047856	100048877	LOC100048877		

Table A.10: Putative direct targets of miR-143

Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol
11438	Chrna4	11652	Akt2	11676	Aldoc	11692	Gfer
11717	Ampd3	11745	Anxa3	11778	Ap3s2	11799	Birc5
11932	Atp1b2	12125	Bcl2l11	12192	Zfp36l1	12297	Cacnb3
12366	Casp2	12400	Cbfb	12476	Cd151	12521	Cd82
12633	Cflar	12702	Socs3	12704	Cit	12751	Tpp1
12798	Cnn2	12856	Cox17	12908	Crat	12972	Cryz
13010	Cst3	13135	Dad1	13386	Dlk1	13445	Cdk2ap1
13481	Dpm2	13644	Efs	13649	Egfr	13728	Mark2
13870	Ercc1	13972	Gnb1l	14043	Ext2	14082	Fadd
14086	Fscn1	14219	Ctgf	14230	Fkbp10	14548	Mrps33
14605	Tsc22d3	14697	Gnb5	14701	Gng12	14726	Pdpn
14739	S1pr2	14755	Pigq	14768	Lancl1	14793	Cdca3
15039	H2-T22	15239	Hgs	15277	Hk2	16011	Igfbp5
16432	Itm2b	16594	Klc2	16796	Lasp1	16801	Arhgef1
16885	Limk1	17126	Smad2	17132	Maf	17390	Mmp2
17391	Mmp24	17828	Muted	17865	Mybl2	18011	Neurl1a
18018	Nfatc1	18140	Uhrf1	18174	Slc11a2	18563	Pcx
18591	Pdgfb	18595	Pdgfra	19039	Lgals3bp	19079	Prkab1
19192	Psme3	19725	Rfx2	19763	Ring1	20111	Rps6ka1
20446	St6galnac2	20511	Slc1a2	20681	Sox8	20779	Src
20922	Supt4h1	20974	Syngr3	21429	Ubtf	21766	Tex261
21853	Timeless	22022	Tpst2	22319	Vamp3	22320	Vamp8
22401	Zmat3	22670	Trim26	23936	Lynx1	23969	Pacsin1
24068	Sra1	26362	Axl	26433	Plod3	26894	Cops7a
27081	Zfp275	27276	Plekhb1	27366	Txnl4a	27965	Spg21
28000	Prpf19	28035	Usp39	50918	Myadm	51875	Tmem141
52004	Cdk2ap2	52064	Coq5	52276	Cdca8	52585	Dhrs1
52683	Ncaph2	52838	Dnlz	53598	Dctn3	55963	Slc1a4
56233	Hdac7	56374	Tmem59	56491	Vapb	56542	Ick
56722	Litaf	57028	Pdxp	57434	Xrcc2	57436	Gabarapl1
57776	Ttyh1	58238	Fam181b	60441	Mrpl38	64075	Smoc1
66078	Tsen34	66179	1110031I02Rik	66191	Ier3ip1	66236	1500011B03Rik

Continued on Next Page...

Putative direct targets of miR-143

Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol
66241	Tmem9	66278	1810013D10Rik	66314	Tpd52l2	66442	Spc25
66588	Cmpk1	66840	Wdr45l	66855	Tcf25	66962	2310047B19Rik
67213	Cmtm6	67513	2610002J02Rik	67657	Rabl3	67693	2310003F16Rik
67695	2310016E02Rik	67792	Rgs8	67843	Slc35a4	67861	Akr1b10
67916	Ppap2b	68087	Dcakd	68338	Golt1a	68505	1110014N23Rik
68597	1110021J02Rik	68666	Svop	68713	Ifitm1	68889	Ubac2
68910	Zfp467	69035	Zdhhc3	69195	Tmem121	69310	Pacrg
69549	2310009B15Rik	69928	Apitd1	69961	2810432D09Rik	70225	Ppil3
70296	Tbc1d13	70612	5730494N06Rik	70686	Dusp16	71448	Tmem80
71452	Ankrd40	71711	Mus81	71726	Smug1	71803	Slc25a18
71909	Haus5	71918	Zcchc24	72106	Jmjd8	72151	Rfc5
72500	Ier5l	72514	Fgfbp3	72775	Fance	72792	2810459M11Rik
73095	Slc25a42	74105	Gga2	74137	Nuak2	74244	Atg7
74342	Lrrtm1	74763	Nat15	75007	Fam63a	75104	Mmd2
75146	Tmem180	75210	Prr3	75495	Morn5	76156	Fam131b
76626	Msi2	76799	2510006D16Rik	76854	Gper	76877	Rab36
77031	Slc9a8	77254	Yif1b	78339	Ttyh3	78829	Tsc22d4
78935	Saal1	81489	Dnajb1	94044	Bcl2l13	94047	Cecr6
94282	Sfxn5	99237	Tm9sf4	99543	Olfml3	100169	Phactr4
101095	Zfp282	102626	Mapkapk3	102693	Phldb1	103743	Tmem98
104418	Dgkz	104479	Ccdc117	105245	Txndc5	105352	Dusp22
105675	Ppif	107522	Ece2	107976	Bre	108037	Shmt2
108912	Cdca2	109006	Ciapin1	109154	Mlec	109648	Npy
109674	Ampd2	111241	Hmga1-rs1	116972	Fam57a	117146	Ube3b
117592	B3galt6	140499	Ube2j2	170460	Stard5	170625	Snx18
171508	Creld1	192185	Nadk	192231	Hexim1	207819	4930539E08Rik
208501	1810043H04Rik	209773	Dennd2a	211286	Cln5	211535	Ccdc114
211798	Mfsd9	212127	2810046L04Rik	212508	Mtg1	212996	Wbscr17
213491	D4Ertd22e	214058	Megf11	214895	Lman2l	214932	Cecr5
217715	Eif2b2	219151	Scara3	223690	Ankrd54	224139	Golgb1
226970	Arhgef4	227619	Man1b1	229504	Isg20l2	230514	Leprot
231863	Fbxl18	235431	Coro2b	235584	Dusp7	237988	Cdr2l
243867	Fbxo46	244152	Tsku	245828	Trappc1	246104	Rhbdl3
246257	Ovca2	252864	Dusp15	268417	Zkscan17	268420	Alkbh5
319757	Smo	320394	Cenpt	330260	Pon2	378702	Serf2
381045	Ccdc58	381921	Taok2	384009	Glipr2	100113398	Adat3

Table A.11: Putative direct targets of miR-145

Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol
11421	Ace	11492	Adam19	11676	Aldoc	11758	Prdx6
11765	Ap1g1	11799	Birc5	11932	Atp1b2	11988	Slc7a2
12014	Bach2	12153	Bmp1	12388	Ctnnd1	12400	Cbfb
12450	Ccng1	12545	Cdc7	12753	Clock	12842	Col1a1
12848	Cops2	13063	Cycs	13132	Dab2	13204	Dhx15
13837	Epha3	13990	Smarcad1	14007	Cugbp2	14057	Sfxn1
14062	F2r	14086	Fscn1	14199	Fhl1	14402	Gabrb3
14571	Gpd2	14583	Gfpt1	14634	Gli3	14725	Lrp2
14772	Grk4	14783	Grb10	15077	Hist2h3c1	15270	H2afx
15530	Hspg2	16007	Cyr61	16011	Igfbp5	16151	Ikbkg
16526	Kcnk2	16561	Kif1b	16568	Kif3a	16574	Kif5c
16579	Kifap3	16589	Uhmk1	17129	Smad5	17196	Mbp
17294	Mest	17389	Mmp16	17920	Myo6	17967	Ncam1
17975	Ncl	17995	Ndufv1	18003	Nedd9	18027	Nfia
18028	Nfib	18176	Nras	18212	Ntrk2	18424	Otx2
18569	Pdcd4	18595	Pdgfra	18762	Prkcz	18823	Plp1
18858	Pmp22	19055	Ppp3ca	19108	Prkx	19206	Ptch1
19212	Pter	19244	Ptp4a2	19285	Ptrf	19290	Pura
19291	Purb	19302	Pxmp3	19655	Rbmx	19714	Rev3l
19731	Rgl1	19820	Rlim	20166	Rtkn	20174	Ruvbl2
20239	Atxn2	20249	Scd1	20356	Sema5a	20358	Sema6a
20397	Sgpl1	20463	Cox7a2l	20529	Slc31a1	20682	Sox9
20689	Sall3	20747	Spop	20913	Stxbp4	20947	Swap70
21346	Tagln2	21844	Tiam1	21981	Ppp1r13b	22034	Traf6
22042	Tfrc	22422	Wnt7b	22687	Zfp259	22718	Zfp60
23972	Papss2	26398	Map2k4	26401	Map3k1	26432	Plod2
26897	Acot1	26951	Zw10	27205	Podxl	27360	Add3
27406	Abcf3	27418	Mkln1	28075	Pppde2	28193	Reep3

Continued on Next Page...

Putative direct targets of miR-145

Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol
29806	Limd1	30930	Vps26a	50755	Fbxo18	50789	Fbxl3
50793	Orc3l	50876	Tmod2	50912	Exosc10	52357	Wwc2
52398	Sept11	52874	D19Bwg1357e	53945	Slc40a1	54161	Copg
54403	Slc4a4	55946	Ap3m1	55989	Nop58	56150	Mad2l1
56426	Pdcd10	56430	Clip1	56496	Tspan6	56554	Raet1d
56709	Dnajb12	56771	Med20	56805	Zbtb33	57434	Xrcc2
57776	Ttyh1	58998	Pvrl3	59008	Anapc5	64297	Gprc5b
64602	Ireb2	66074	Tmem167	66084	Rmnd1	66101	Ppih
66161	Pop4	66420	Polr2e	66427	Cyb5b	66586	Crls1
66868	Mfsd1	66869	Zfp869	66882	Bzw1	66884	Appbp2
66953	Cdca7	67064	Chmp1b	67072	Cdc37l1	67073	Pi4k2b
67211	Armc10	67273	Ndufa10	67370	Zfp606	67414	Mfn1
67433	Ccdc127	67452	Pnpla8	67528	Nudt7	67544	Fam120b
67603	Dusp6	67629	Spc24	67784	Plxnd1	67920	Mak16
68038	Chid1	68145	Etaa1	68187	Fam135a	68268	Zdhhc21
68275	Rpa1	68581	Tmed10	68659	Fam198b	68693	Hnrnpul2
68738	Acss1	69035	Zdhhc3	69207	Sfrs11	69692	Hddc2
69718	Ipmk	69727	Usp46	69821	Mterfd2	70052	Prpf4
70296	Tbc1d13	70375	Ica1l	70417	Megf10	70454	Cenpl
71435	Arhgap21	71436	Flrt3	71770	Ap2b1	71782	Ankle2
71914	Antxr2	72102	Dusp11	72141	Adpgk	72181	Nsun4
72238	Tbc1d5	72349	Dusp3	72425	2410042D21Rik	72459	Htatsf1
72519	Tmem55a	72759	Tmem135	72792	2810459M11Rik	72823	Pard3b
72925	March1	73122	Tgfbrap1	73230	Bmper	73284	Ddit4l
73296	Rhobtb3	74030	Rin2	74186	Ccdc3	74492	Kbtbd13
74498	Gorasp1	74868	Tmem65	75678	Ippk	75805	Nln
76178	6330578E17Rik	76267	Fads1	76376	Slc24a2	76574	Mfsd2a
76742	Snx27	77106	Tmem181a	77574	Fam115a	77976	Nuak1
78100	8430410K20Rik	78593	Nrip3	78655	Eif3j	78748	Rassf10
78937	Avl9	79464	Lias	80890	Trim2	81840	Sorcs2
81879	Tcfcp2l1	83397	Akap12	83671	Sytl2	83675	Bicc1
83997	Slmap	93871	Brwd1	94040	Clmn	94190	Ophn1
94282	Sfxn5	98660	Atp1a2	100317	AU040320	100434	Slc44a1
100710	Pds5b	100986	Akap9	101565	6330503K22Rik	102247	Agpat6
103135	Usp52	103537	Mbtd1	104625	Cnot6	105638	Dph3
106369	Ypel1	106840	Unc119b	107029	Me2	107568	Wwp1
108071	Grm5	108652	Slc35b3	108653	Rimklb	108760	Galntl1
108897	Aif1l	108912	Cdca2	109079	Sephs1	109232	Sccpdh
109552	Sri	109801	Glo1	116914	Slc19a2	117198	Ivns1abp
117600	Srgap1	140740	Sec63	170753	Zfp704	170822	Usp33
194401	Mical3	207214	Larp4	208659	Fam20a	208718	Dis3l2
210004	B3gntl1	211286	Cln5	211535	Ccdc114	211914	Asap2
211945	Plekhh1	211949	Spsb4	213056	Fam126b	213391	Rassf4
213582	Mtap9	214137	Arhgap29	214944	Mobkl2b	215008	Vezt
216440	Os9	216578	Papolg	216825	Usp22	216965	Taok1
217431	Nol10	217653	C79407	217864	Rcor1	217893	Pacs2
218397	Rasa1	218454	Lhfpl2	219148	Fam167a	223739	5031439G07Rik
224836	Usp49	224997	Dlgap1	225283	Rprd1a	226144	Erlin1
226151	Fam178a	226432	Ipo9	226470	Zbtb41	226562	Bat2l2
226781	Slc30a10	227682	Trub2	228357	Lrp4	228829	Phf20
228876	Zfp334	229709	Ahcyl1	230234	BC026590	230753	Thrap3
230857	Ece1	231238	Sel1l3	231834	Snx8	232164	Paip2b
232431	Gprc5a	233103	4931406P16Rik	233271	Luzp2	233315	Mtmr10
233532	Rsf1	234549	Heatr3	234734	Aars	235132	Zbtb44
235542	Ppp2r3a	236920	Stard8	237082	Nxt2	241075	Plekhm3
241589	D430041D05Rik	242362	Manea	242585	Slc35d1	242687	Wasf2
243725	Ppp1r9a	245446	Slitrk4	246154	Vasn	246196	Zfp277
259302	Srgap3	268882	Fbxo45	269003	Sap130	286940	Flnb
320508	Cachd1	320705	Bend6	320713	Mysm1	321022	Cdv3
326618	Tpm4	329641	6030405A18Rik	329828	AI464131	329941	Col8a2
330260	Pon2	407786	Taf9b	432572	Cytsb	100037258	Dnajc3
100039795	Ildr2	100043555	LOC100043555				

Table A.12: Putative direct targets of miR-25

Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol
11302	Aatk	11886	Asah1	11906	Zfhx3	12036	Bcat2
12111	Bgn	12125	Bcl2l11	12577	Cdkn1c	13138	Dag1

Continued on Next Page...

Putative direct targets of miR-25

Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol
13356	Dgcr2	13640	Efna5	13654	Egr2	14020	Evi5
14025	Bcl11a	14057	Sfxn1	14165	Fgf10	14167	Fgf12
14584	Gfpt2	14675	Gna14	15166	Hcn2	15529	Sdc2
16001	Igf1r	16011	Igfbp5	16579	Kifap3	16909	Lmo2
17158	Man2a1	17918	Myo5a	18027	Nfia	18029	Nfic
18032	Nfix	18046	Nfyc	18417	Cldn11	18549	Pcsk2
18555	Cdk16	18627	Per2	18640	Pfkfb2	18709	Pik3r2
18717	Pip5k1c	18738	Pitpna	19084	Prkar1a	19085	Prkar1b
19277	Ptpro	19326	Rab11b	19352	Rabggtb	19679	Pitpnm2
19822	Rnf4	20320	Nptn	20393	Sgk1	20544	Slc9a1
20595	Smn1	20621	Snn	20652	Soat1	20719	Serpinb6a
20965	Syn2	20970	Sdc3	21685	Tef	21841	Tia1
21858	Timp2	21885	Tle1	22348	Slc32a1	22401	Zmat3
22418	Wnt5a	22644	Rnf103	22781	Ikzf4	23792	Adam23
26373	Clcn7	26398	Map2k4	26419	Mapk8	27362	Dnajb9
50781	Dkk3	51813	Ccnc	52882	Rgs7bp	54418	Fmn2
56248	Ak3	56468	Socs5	56613	Rps6ka4	56747	Sez6l
56876	Nelf	57431	Dnajc4	59046	Arpp19	64378	Gpr88
64933	Ap3m2	65247	Asb1	65964	B230120H23Rik	66114	Dnajc30
66259	Camk2n1	66310	Dpy30	66686	Dcbld1	66701	Spryd4
66756	4933411K20Rik	66878	Riok3	66894	Wwp2	66902	Mtap
67117	Dynlt3	67398	Srpr	67602	Necap1	68202	Ndufa5
68365	Rab14	68514	Efha1	68520	Zfyve21	68659	Fam198b
69046	Isca1	70052	Prpf4	70465	Wdr77	70599	Ssfa2
71063	Zfp597	71722	Cic	71982	Snx10	72007	Fndc3b
72056	1810055G02Rik	72344	Usp36	72536	Tagap	73744	Man2c1
74106	Dcaf6	74158	Josd1	74197	Gtf2e1	74349	Fam160a2
74513	Neto2	74519	Cyp2j9	74769	Pik3cb	75901	Dcp1a
76366	Mtif3	76477	Pcolce2	76740	Efr3a	76788	Klhdc10
78937	Avl9	81535	Sgpp1	83922	Tsga14	94249	Slc24a3
94282	Sfxn5	98682	Mfsd6	98952	Fam102a	99738	Kcnc4
100383	Bsdc1	101476	Plekha1	102323	Dcun1d2	103768	Tubg2
103850	Nt5m	104111	Adcy3	105171	Arrdc3	106564	Ppcs
107767	Scamp1	108652	Slc35b3	109205	Sobp	114896	Afg3l1
117197	Cno	140919	Slc17a6	170625	Snx18	195209	Gm22
209318	Gps1	210106	Papd7	211255	Kbtbd7	211499	Tmem87a
213056	Fam126b	216134	Pdxk	216558	Ugp2	216742	Fnip1
216965	Taok1	217119	Xylt2	217351	Tnrc6c	218454	Lhfpl2
218503	Fcho2	223254	Farp1	223693	Tmem184b	225280	Ino80c
225642	Grp	226252	Fam160b1	226856	Lpgat1	227723	Bat2l
228942	Cbln4	228983	Osbpl2	228998	Arfgap1	229541	Dennd4b
230514	Leprot	232947	Lrrc68	234734	Aars	235281	Scn3b
235283	Gramd1b	241494	Zfp385b	245944	Vps54	252966	Cables2
268860	Abat	269401	Znf512b	269593	Luzp1	270076	Gcdh
319642	Rab9b	319832	Tmem229a	319865	E130114P18Rik	320492	A830018L16Rik
328329	Mast4	333789	N4bp2	338467	Morc3	353047	Plekhm1
382985	Rrm2b	403187	Opa3	432450	Nkain2	433100	AA388235

Table A.13: Putative direct targets of cel-miR-67

Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol
11566	Adss	11610	Agtrap	11736	Ankfy1	11765	Ap1g1
12043	Bcl2	12050	Bcl2l2	12055	Bcl7c	12166	Bmpr1a
12177	Bnip3l	12457	Ccrn4l	12555	Cdh15	12633	Cflar
12805	Cntn1	12807	Hps3	12915	Atf6b	12937	Pcdha6
12953	Cry2	13014	Cstb	13063	Cycs	13135	Dad1
13193	Dcx	13199	Ddn	13204	Dhx15	13478	Dpagt1
13483	Dpp6	13527	Dtna	13690	Eif4g2	13709	Elf1
13844	Ephb2	13855	Epn2	14007	Cugbp2	14042	Ext1
14057	Sfxn1	14105	Sfrs13a	14230	Fkbp10	14365	Fzd3
14388	Gab1	14615	Gjc1	14674	Gna13	14696	Gnb4
14718	Got1	14897	Trip12	15200	Hbegf	15312	Hmgn1
15525	Hspa4	16009	Igfbp3	16210	Impact	16362	Irf1
16561	Kif1b	16570	Kif3c	16589	Uhmk1	16653	Kras
16889	Lipa	17535	Mre11a	17918	Myo5a	18000	Sept2
18011	Neurl1a	18082	Nipsnap1	18140	Uhrf1	18167	Npy2r
18212	Ntrk2	18616	Peg3	18744	Pja1	18768	Pkib
18799	Plcd1	18986	Pou2f1	19043	Ppm1b	19046	Ppp1cb

Continued on Next Page...

Putative direct targets of cel-miR-67

Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol
19055	Ppp3ca	19058	Ppp3r1	19070	Mobkl3	19072	Prep
19084	Prkar1a	19159	Cyth3	19245	Ptp4a3	19267	Ptpre
19288	Ptx3	19290	Pura	19349	Rab7	19384	Ran
19645	Rb1	19726	Rfx3	19729	Rag1ap1	19820	Rlim
19891	Rpa2	20168	Rtn3	20224	Sar1a	20250	Scd2
20382	Sfrs2	20399	Sh2b1	20587	Smarcb1	20617	Snca
20652	Soat1	20842	Stag1	20843	Stag2	20853	Stau1
20887	Sult1a1	20932	Surf4	20935	Surf6	20971	Sdc4
21912	Tspan7	21991	Tpi1	22129	Ttc3	22350	Ezr
22380	Wbp4	22764	Zfx	22793	Zyx	24017	Rnf13
24030	Mrps12	24116	Whsc2	26398	Map2k4	26416	Mapk14
27041	G3bp1	27055	Fkbp9	27096	Trappc3	28028	Mrpl50
28146	Serp1	29806	Limd1	29820	Tnfrsf19	29861	Dpf1
29864	Rnf11	30058	Timm8a1	50754	Fbxw7	50996	Pdcd7
52206	Anapc4	52468	Ctdsp2	52666	D10Ertd610e	52850	Sgsm1
52882	Rgs7bp	53323	Ube2k	53380	Psmd10	53619	Blcap
54151	Cyhr1	54473	Tollip	54484	Mkrn1	54613	St3gal6
56048	Lgals8	56248	Ak3	56309	Mycbp	56351	Ptges3
56367	Scoc	56386	B4galt6	56418	Ykt6	56433	Vps29
56459	Sae1	57743	Sec61a2	57912	Cdc42se1	58194	Sh3kbp1
58239	Dexi	58242	Nudt11	58243	Nap1l5	59069	Tpm3
64010	Sav1	64050	Yeats4	64143	Ralb	64297	Gprc5b
65973	Asph	66046	Ndufb5	66052	Sdhc	66140	Fam33a
66191	Ier3ip1	66194	Pycrl	66246	Osgep	66335	Atp6v1c1
66566	2310079N02Rik	66648	5730494M16Rik	66700	Vps24	66849	Ppp1r2
66884	Appbp2	66892	Eif4e3	66923	Pbrm1	66953	Cdca7
66966	Trit1	67027	Mkrn2	67070	Lsm14a	67130	Ndufa6
67181	Dullard	67238	2810453I06Rik	67245	Peli1	67326	1700037H04Rik
67388	1110008F13Rik	67414	Mfn1	67529	Fgfr1op2	67590	Tctn3
67738	Ppid	67808	Tprgl	67887	Tmem66	67889	Rbm18
67897	Rnmt	67933	Hcfc2	68050	Akirin1	68149	Otub2
68272	Rbm28	68364	0610030E20Rik	68477	Rmnd5a	68558	Ankra2
68861	1190002N15Rik	68874	Klhdc9	68969	Eif1b	69053	1810013L24Rik
69109	Fam58b	69136	Tusc1	69150	Snx4	69227	2810407C02Rik
69329	1700003M02Rik	69372	Mocs3	69538	Antxr1	70093	Ube2q1
70369	Bag5	70556	Slc25a33	70584	Pak4	70612	5730494N06Rik
71778	Klhl5	71900	Tmem106b	71952	2410016O06Rik	71963	Cdca4
71978	Ppp2r2a	72075	Ogfr	72124	Seh1l	72139	2610044O15Rik
72170	Chchd4	72193	Sfrs2ip	72195	Supt7l	72542	Pgam5
72552	Hsdl1	72585	Lypd1	72685	Dnajc6	72792	2810459M11Rik
72993	Appl1	73137	Prrc1	73713	Rbm20	74022	Glyr1
74030	Rin2	74256	Cyld	74340	Ahcyl2	74356	4931428F04Rik
74450	Pank2	74479	Snx11	74493	Tnks2	74763	Nat15
74769	Pik3cb	75625	Mageh1	75678	Ippk	75710	Rbm12
75723	Amotl1	75769	4833424O15Rik	75778	Them4	75956	Srrm2
76007	Zmym2	76252	Atp6v0e2	76302	Pcnp	76308	Rab1b
76626	Msi2	76688	Arfrp1	76893	Lass2	76958	2210418O10Rik
77305	Wdr82	77781	Epm2aip1	78408	Fam131a	78757	Rictor
78808	Stxbp5	78938	Fbxo34	80509	Med8	80909	Gatsl2
80986	Ckap2	81879	Tcfcp2l1	93683	Glce	93739	Gabarapl2
97884	B3galnt2	98741	Kcnb2	99311	Commd7	99887	Tmem56
99889	Arfip1	103266	AI597468	103694	Tmed4	104318	Csnk1d
104625	Cnot6	104725	1110002B05Rik	105000	Dnalc1	106298	Rrn3
106522	Pkdcc	106840	Unc119b	106894	Hmgxb3	107566	Arl2bp
107581	Col16a1	107823	Whsc1	107885	Mthfs	108958	Fam73b
108960	Irak2	109006	Ciapin1	109689	Arrb1	109711	Actn1
110809	Sfrs1	110959	Nudt19	111241	Hmga1-rs1	116731	Pcdha1
116873	Stim2	117109	Pop5	140904	Caln1	170459	Stard4
193813	Mcfd2	207806	Gm608	208292	Zfp871	209268	Igsf1
211286	Cln5	211739	Vstm2a	213056	Fam126b	213464	Rbbp5
213541	Ythdf2	214162	Mll1	214579	Aldh5a1	214952	Rhot2
216119	A130042E20Rik	216549	Aftph	216792	A230051G13Rik	216987	Utp6
217732	2310044G17Rik	217864	Rcor1	217893	Pacs2	218772	Rarb
218975	Mapk1ip1l	219022	Ttc5	219181	Akap11	223752	Gramd4
224105	Pak2	224129	Adcy5	224647	D17Wsu92e	225215	Rsl24d1
225280	Ino80c	225363	Etf1	226043	Cbwd1	226144	Erlin1
226744	Cnst	226844	Mfsd7b	227619	Man1b1	227682	Trub2
228071	Sestd1	228714	Csrp2bp	228812	Pigu	228880	Zmynd8
229517	Slc25a44	229593	Golph3l	229615	Pias3	230235	6430704M03Rik
230709	Zmpste24	230917	Tmem201	231070	Insig1	231724	Rad9b

Continued on Next Page...

Putative direct targets of cel-miR-67

Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol
231834	Snx8	231997	Fkbp14	232784	Zfp212	233271	Luzp2
234728	Ftsjd1	234736	Rfwd3	235574	Atp2c1	237859	Ccdc55
241263	Gpr158	242291	Impad1	242297	Fam110b	242384	Lingo2
242687	Wasf2	242800	Ttc34	242864	Napepld	244058	Rgma
244631	Pskh1	244810	AW551984	245468	Pnma3	245555	C77370
246316	Lgi2	268697	Ccnb1	269582	Clspn	269639	Zfp512
272589	Tbcel	277414	Trp53i11	319370	Fam100b	319468	Ppm1h
320184	Lrrc58	320333	D830030K20Rik	320472	Ppm1e	320495	Ipcef1
329739	Fam102b	330050	Fam185a	380614	Intu	381038	Parl
381280	Hjurp	381511	Pdp1	384763	Zfp667	504193	Cbx6-Nptxr
545260	Arsi	668661	2410002F23Rik	100041567	Gm10060	100042480	Nhsl2
100043555	LOC100043555	100047834	LOC100047834				

Table A.14: Putative direct targets of miR-434-3p

Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol	Entrez ID	Symbol
11745	Anxa3	11799	Birc5	12444	Ccnd2	12450	Ccng1
12476	Cd151	12505	Cd44	12527	Cd9	12661	Chl1
12830	Col4a5	12842	Col1a1	13356	Dgcr2	13617	Ednra
13618	Ednrb	13649	Egfr	14168	Fgf13	14282	Fosb
14609	Gja1	14613	Gja5	14735	Gpc4	15529	Sdc2
15568	Elavl1	16601	Klf9	17973	Nck1	18041	Nfs1
19038	Ppic	19303	Pxn	20377	Sfrp1	20512	Slc1a3
20621	Snn	20689	Sall3	20818	Srprb	20848	Stat3
20922	Supt4h1	21417	Zeb1	21804	Tgfb1i1	21813	Tgfbr2
22042	Tfrc	22319	Vamp3	23827	Bpnt1	23873	Faim
23947	Mid2	26362	Axl	27058	Srp9	27273	Pdk4
27428	Shroom3	30057	Timm8b	52276	Cdca8	53374	Chst3
53901	Rcan2	56078	Car5b	56248	Ak3	56291	Styx
56397	Morf4l2	56516	Rbms2	60599	Trp53inp1	65960	Twsg1
66273	1810020D17Rik	66467	Gtf2h5	66471	Anp32e	66628	Thg1l
66870	Serbp1	66905	Plin3	67145	Tomm34	67468	Mmd
68420	Ankrd13a	68659	Fam198b	68801	Elovl5	69241	Polr2d
73569	Vgll3	73828	Dcaf4	74148	1300001I01Rik	75616	2810008M24Rik
75646	Rai14	76626	Msi2	78232	Trappc6b	78808	Stxbp5
80860	Ghdc	81879	Tcfcp2l1	99237	Tm9sf4	103266	AI597468
103724	Tbc1d10a	105245	Txndc5	107272	Psat1	107566	Arl2bp
108735	Sft2d2	109801	Glo1	110460	Acat2	114774	Pawr
116914	Slc19a2	117149	Tirap	171567	Nme7	192216	Tmem47
208936	Adamts18	209357	Gtf2h3	213673	9530068E07Rik	214944	Mobkl2b
219140	Spata13	223453	Dap	225280	Ino80c	226562	Bat2l2
229534	Pbxip1	232157	Mobkl1b	232313	Gxylt2	233406	Prc1
240725	Sulf1	268697	Ccnb1	319613	5730410E15Rik	330192	Vps37b
384009	Glipr2	432879	Gm5465	504193	Cbx6-Nptxr	100041103	LOC100041103

Table A.15: Top 25 most enriched KEGG terms in targets of cel-miR-67.
q - number of genes of a KEGG term that was among the predicted targets, m - total number of genes of a KEGG term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
04210	Apoptosis	8	48	6e-04
05014	Amyotrophic lateral sclerosis (ALS)	5	33	0.0102
04722	Neurotrophin signaling pathway	8	76	0.0116
04012	ErbB signaling pathway	6	51	0.0168
05222	Small cell lung cancer	6	51	0.0168
04660	T cell receptor signaling pathway	6	52	0.0184
04914	Progesterone-mediated oocyte maturation	6	52	0.0184

Continued on Next Page...

ID	Term Description	q	m	P
00562	Inositol phosphate metabolism	4	26	0.0202
04370	VEGF signaling pathway	5	42	0.0272
05210	Colorectal cancer	6	59	0.0323
04710	Circadian rhythm - mammal	2	8	0.04
05223	Non-small cell lung cancer	4	36	0.0581
00100	Steroid biosynthesis	2	11	0.0725
04664	Fc epsilon RI signaling pathway	4	40	0.0797
04912	GnRH signaling pathway	5	57	0.0828
04662	B cell receptor signaling pathway	4	43	0.0982
05200	Pathways in cancer	11	175	0.101
04120	Ubiquitin mediated proteolysis	7	97	0.101
04620	Toll-like receptor signaling pathway	4	45	0.112
05212	Pancreatic cancer	4	47	0.126
05211	Renal cell carcinoma	4	48	0.133
04114	Oocyte meiosis	5	66	0.133
00534	Heparan sulfate biosynthesis	2	16	0.139
04650	Natural killer cell mediated cytotoxicity	4	50	0.148
05216	Thyroid cancer	2	17	0.153

Table A.16: Top 25 most enriched KEGG term in targets of miR-124. q - number of genes of a KEGG term that was among predicted targets, m - total number of genes of a KEGG term in a test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
04510	Focal adhesion	13	121	0.00211
04520	Adherens junction	7	46	0.00347
04530	Tight junction	8	68	0.00907
05416	Viral myocarditis	5	31	0.0104
05222	Small cell lung cancer	7	56	0.0105
05220	Chronic myeloid leukemia	6	49	0.0192
04810	Regulation of actin cytoskeleton	11	127	0.0223
05221	Acute myeloid leukemia	5	38	0.024
05412	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	5	42	0.0354
05216	Thyroid cancer	3	18	0.0421
04512	ECM-receptor interaction	5	45	0.0458
04670	Leukocyte transendothelial migration	6	60	0.0465
05213	Endometrial cancer	4	33	0.0549
02010	ABC transporters	3	20	0.0552
00240	Pyrimidine metabolism	6	63	0.0568

Continued on Next Page...

ID	Term Description	q	m	P
05212	Pancreatic cancer	5	49	0.0623
04650	Natural killer cell mediated cytotoxicity	5	51	0.0716
04622	RIG-I-like receptor signaling pathway	4	38	0.0839
04514	Cell adhesion molecules (CAMs)	5	54	0.087
04660	T cell receptor signaling pathway	5	55	0.0925
04130	SNARE interactions in vesicular transport	3	26	0.104
05200	Pathways in cancer	13	204	0.11
04115	p53 signaling pathway	4	42	0.111
04370	VEGF signaling pathway	4	42	0.111
04662	B cell receptor signaling pathway	4	42	0.111

Table A.17: Top 25 most enriched KEGG terms in targets of miR-143.
q - number of genes of a KEGG term that was among the predicted targets, m - total number of genes of a KEGG term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
04010	MAPK signaling pathway	9	172	0.0342
04810	Regulation of actin cytoskeleton	7	124	0.0422
04142	Lysosome	5	82	0.0624
00510	N-Glycan biosynthesis	3	36	0.0671
00563	Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	2	17	0.0719
05210	Colorectal cancer	4	61	0.0746
03440	Homologous recombination	2	20	0.0956
04140	Regulation of autophagy	2	20	0.0956
05218	Melanoma	3	42	0.0966
04540	Gap junction	4	67	0.0976
00130	Ubiquinone and other terpenoid-quinone biosynthesis	1	4	0.101
00460	Cyanoamino acid metabolism	1	4	0.101
00680	Methane metabolism	1	4	0.101
00534	Heparan sulfate biosynthesis	2	21	0.104
04520	Adherens junction	3	45	0.113
05214	Glioma	3	46	0.119
00561	Glycerolipid metabolism	2	23	0.121
04370	VEGF signaling pathway	3	47	0.125
01100	Metabolic pathways	21	626	0.127
04920	Adipocytokine signaling pathway	3	48	0.131
00750	Vitamin B6 metabolism	1	6	0.148
00051	Fructose and mannose metabolism	2	26	0.148
05200	Pathways in cancer	8	203	0.161
04130	SNARE interactions in vesicular transport	2	28	0.167

Continued on Next Page...

ID	Term Description	q	m	P
04930	Type II diabetes mellitus	2	31	0.195

Table A.18: Top 25 most enriched KEGG terms in targets of miR-145.
q - number of genes of a KEGG term that was among the predicted targets, m - total number of genes of a KEGG term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
04010	MAPK signaling pathway	14	166	0.000891
04340	Hedgehog signaling pathway	5	31	0.00307
04360	Axon guidance	8	80	0.00435
04144	Endocytosis	10	117	0.00463
00450	Selenoamino acid metabolism	3	16	0.0144
01040	Biosynthesis of unsaturated fatty acids	3	18	0.02
04614	Renin-angiotensin system	2	7	0.0205
00620	Pyruvate metabolism	3	20	0.0267
05016	Huntington's disease	8	115	0.0345
04720	Long-term potentiation	4	44	0.0558
05217	Basal cell carcinoma	3	34	0.101
04210	Apoptosis	4	55	0.107
04062	Chemokine signaling pathway	6	99	0.107
03018	RNA degradation	3	38	0.129
04912	GnRH signaling pathway	4	62	0.148
04540	Gap junction	4	64	0.16
04520	Adherens junction	3	42	0.161
03440	Homologous recombination	2	22	0.164
04114	Oocyte meiosis	4	66	0.173
05020	Prion diseases	2	23	0.176
04662	B cell receptor signaling pathway	3	44	0.177
04020	Calcium signaling pathway	5	92	0.187
04622	RIG-I-like receptor signaling pathway	3	47	0.203
00920	Sulfur metabolism	1	7	0.21
04260	Cardiac muscle contraction	3	48	0.211

Table A.19: Top 25 most enriched KEGG terms in targets of miR-25.
q - number of genes of a KEGG term that was among the predicted targets, m - total number of genes of a KEGG term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
05218	Melanoma	5	42	0.00181

Continued on Next Page...

ID	Term Description	q	m	P
04914	Progesterone-mediated oocyte maturation	5	54	0.00553
00770	Pantothenate and CoA biosynthesis	2	8	0.0118
04664	Fc epsilon RI signaling pathway	4	45	0.0151
04960	Aldosterone-regulated sodium reabsorption	3	26	0.0174
04930	Type II diabetes mellitus	3	28	0.0213
04012	ErbB signaling pathway	4	52	0.0245
04210	Apoptosis	4	52	0.0245
04910	Insulin signaling pathway	5	83	0.0316
05210	Colorectal cancer	4	57	0.033
04810	Regulation of actin cytoskeleton	6	117	0.0381
04010	MAPK signaling pathway	7	151	0.0414
04620	Toll-like receptor signaling pathway	4	63	0.0453
04512	ECM-receptor interaction	3	39	0.0504
00250	Alanine, aspartate and glutamate metabolism	2	19	0.0619
04070	Phosphatidylinositol signaling system	3	45	0.0715
05214	Glioma	3	45	0.0715
05212	Pancreatic cancer	3	47	0.0793
04510	Focal adhesion	5	111	0.0888
04514	Cell adhesion molecules (CAMs)	3	50	0.0916
04722	Neurotrophin signaling pathway	4	82	0.0987
00600	Sphingolipid metabolism	2	25	0.1
04666	Fc gamma R-mediated phagocytosis	3	54	0.109
05200	Pathways in cancer	7	190	0.111
00040	Pentose and glucuronate interconversions	1	6	0.123

Table A.20: Top 25 most enriched KEGG terms in targets of miR-434-3p.
q - number of genes of a KEGG term that was among the predicted targets, m - total number of genes of a KEGG term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
04640	Hematopoietic cell lineage	3	23	0.00456
04512	ECM-receptor interaction	3	35	0.0149
04115	p53 signaling pathway	3	36	0.0161
03060	Protein export	2	17	0.0262
00620	Pyruvate metabolism	2	19	0.0323
05212	Pancreatic cancer	3	47	0.0326
00240	Pyrimidine metabolism	3	53	0.0443
05210	Colorectal cancer	3	56	0.0508
04510	Focal adhesion	4	102	0.0652
00730	Thiamine metabolism	1	5	0.0733

Continued on Next Page...

ID	Term Description	q	m	P
00750	Vitamin B6 metabolism	1	5	0.0733
04144	Endocytosis	4	107	0.0751
04020	Calcium signaling pathway	3	68	0.0813
03420	Nucleotide excision repair	2	32	0.083
00072	Synthesis and degradation of ketone bodies	1	6	0.0874
00920	Sulfur metabolism	1	6	0.0874
05218	Melanoma	2	38	0.111
00900	Terpenoid backbone biosynthesis	1	8	0.115
04520	Adherens junction	2	39	0.116
05200	Pathways in cancer	5	182	0.135
00910	Nitrogen metabolism	1	10	0.141
04630	Jak-STAT signaling pathway	2	45	0.147
00532	Chondroitin sulfate biosynthesis	1	12	0.167
04012	ErbB signaling pathway	2	52	0.184
03020	RNA polymerase	1	14	0.192

Table A.21: Top 25 most enriched KEGG terms in the induced by transfection set.
q - number of genes of a KEGG term that was among the predicted targets, m - total number of genes of a KEGG term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
04115	p53 signaling pathway	21	38	3.88e-09
04512	ECM-receptor interaction	18	37	6.75e-07
04510	Focal adhesion	34	107	2.69e-06
04110	Cell cycle	26	79	2.12e-05
05200	Pathways in cancer	46	188	0.000115
05412	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	14	40	0.000898
05322	Systemic lupus erythematosus	19	66	0.00179
05222	Small cell lung cancer	14	48	0.00619
05215	Prostate cancer	16	58	0.00641
04060	Cytokine-cytokine receptor interaction	22	89	0.0065
05221	Acute myeloid leukemia	11	35	0.00806
00071	Fatty acid metabolism	8	23	0.012
00980	Metabolism of xenobiotics by cytochrome P450	9	28	0.0138
05210	Colorectal cancer	15	59	0.0177
05213	Endometrial cancer	10	34	0.0184
05220	Chronic myeloid leukemia	12	45	0.0226
05410	Hypertrophic cardiomyopathy (HCM)	12	45	0.0226
05217	Basal cell carcinoma	10	35	0.0226

Continued on Next Page...

KEGG terms enriched in targets of mockTr

ID	Term Description	q	m	P
00531	Glycosaminoglycan degradation	5	13	0.0294
04010	MAPK signaling pathway	33	167	0.0331
04142	Lysosome	18	82	0.0418
05414	Dilated cardiomyopathy	12	49	0.0422
04640	Hematopoietic cell lineage	8	29	0.0482
04621	NOD-like receptor signaling pathway	8	30	0.0578
04080	Neuroactive ligand-receptor interaction	19	93	0.0701

Table A.22: Top 25 most enriched KEGG terms in the Ago HITS-CLIP set.
q - number of genes of a KEGG term that was in the Ago HITS-CLIP set, m - total number of genes of a KEGG term in the test universe, P - P-valueof enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
05010	Alzheimer's disease	34	103	0.00011
04720	Long-term potentiation	16	42	0.00142
04810	Regulation of actin cytoskeleton	36	127	0.00186
04360	Axon guidance	24	78	0.00334
00100	Steroid biosynthesis	6	10	0.00343
00190	Oxidative phosphorylation	25	83	0.0038
04142	Lysosome	24	80	0.00481
04512	ECM-receptor interaction	15	45	0.00859
04510	Focal adhesion	32	121	0.0102
05016	Huntington's disease	30	113	0.012
00600	Sphingolipid metabolism	9	25	0.0233
05214	Glioma	13	42	0.0264
04350	TGF-beta signaling pathway	14	47	0.03
00603	Glycosphingolipid biosynthesis - globo series	4	8	0.0383
04070	Phosphatidylinositol signaling system	13	45	0.0453
04530	Tight junction	18	68	0.0473
00010	Glycolysis / Gluconeogenesis	9	28	0.048
04960	Aldosterone-regulated sodium reabsorption	8	25	0.0623
04666	Fc gamma R-mediated phagocytosis	14	52	0.0662
05211	Renal cell carcinoma	14	53	0.0759
05012	Parkinson's disease	19	79	0.0972
04730	Long-term depression	11	41	0.099
05215	Prostate cancer	16	65	0.103
04662	B cell receptor signaling pathway	11	42	0.114
04540	Gap junction	14	58	0.138

Table A.23: Top 40 most enriched GO terms ("Biological process" type) in targets of cel-miR-67.
q - number of genes of a GO term that was among the predicted targets, m - total number of genes of a GO term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0009790	embryonic development	23	333	0.000734
GO:0031399	regulation of protein modification process	9	75	0.000819
GO:0019079	viral genome replication	2	2	0.00111
GO:0032268	regulation of cellular protein metabolic process	13	148	0.00131
GO:0051246	regulation of protein metabolic process	15	186	0.00137
GO:0048592	eye morphogenesis	6	41	0.00219
GO:0001932	regulation of protein amino acid phosphorylation	7	58	0.00298
GO:0010224	response to UV-B	2	3	0.00325
GO:0019058	viral infectious cycle	2	3	0.00325
GO:0043374	CD8-positive, alpha-beta T cell differentiation	2	3	0.00325
GO:0044267	cellular protein metabolic process	58	1257	0.00545
GO:0048872	homeostasis of number of cells	7	65	0.00567
GO:0006582	melanin metabolic process	2	4	0.00636
GO:0048169	regulation of long-term neuronal synaptic plasticity	3	13	0.00818
GO:0006464	protein modification process	36	724	0.0097
GO:0006413	translational initiation	4	26	0.0101
GO:0016032	viral reproduction	2	5	0.0104
GO:0022415	viral reproductive process	2	5	0.0104
GO:0030168	platelet activation	2	5	0.0104
GO:0035020	regulation of Rac protein signal transduction	2	5	0.0104
GO:0043412	macromolecule modification	37	754	0.0107
GO:0007398	ectoderm development	6	57	0.0114
GO:0006633	fatty acid biosynthetic process	5	42	0.0123
GO:0043687	post-translational protein modification	31	614	0.0131
GO:0019748	secondary metabolic process	4	28	0.0132
GO:0032270	positive regulation of cellular protein metabolic process	5	43	0.0136
GO:0000082	G1/S transition of mitotic cell cycle	3	16	0.0149
GO:0048168	regulation of neuronal synaptic plasticity	3	16	0.0149
GO:0051656	establishment of organelle localization	3	16	0.0149
GO:0009994	oocyte differentiation	2	6	0.0152
GO:0048599	oocyte development	2	6	0.0152
GO:0019941	modification-dependent protein catabolic process	18	310	0.0156
GO:0043632	modification-dependent macromolecule catabolic process	18	310	0.0156
GO:0048513	organ development	38	801	0.0162
GO:0030218	erythrocyte differentiation	4	30	0.0167
GO:0031401	positive regulation of protein modification process	4	31	0.0187

Continued on Next Page...

ID	Term Description	q	m	P
GO:0009653	anatomical structure morphogenesis	29	585	0.0203
GO:0048598	embryonic morphogenesis	11	163	0.0203
GO:0007173	epidermal growth factor receptor signaling pathway	2	7	0.0208
GO:0014047	glutamate secretion	2	7	0.0208

Table A.24: Top 40 most enriched GO terms ("Biological process" type) in targets of miR-124.
q - number of genes from a GO term that was among the predicted targets, m - total number of genes of a GO term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0009653	anatomical structure morphogenesis	47	585	4.79e-07
GO:0007275	multicellular organismal development	80	1244	4.88e-07
GO:0048856	anatomical structure development	72	1077	4.99e-07
GO:0001944	vasculature development	21	160	5.19e-07
GO:0032502	developmental process	84	1360	1.29e-06
GO:0048731	system development	67	1003	1.42e-06
GO:0048513	organ development	57	801	1.43e-06
GO:0001568	blood vessel development	20	158	1.74e-06
GO:0048514	blood vessel morphogenesis	18	135	2.72e-06
GO:0023034	intracellular signaling pathway	33	412	2.98e-05
GO:0048869	cellular developmental process	55	844	3.05e-05
GO:0061061	muscle structure development	16	136	4.78e-05
GO:0032501	multicellular organismal process	88	1593	6.23e-05
GO:0030154	cell differentiation	51	791	8.16e-05
GO:0048646	anatomical structure formation involved in morphogenesis	21	222	8.9e-05
GO:0051146	striated muscle cell differentiation	10	63	0.000106
GO:0009888	tissue development	27	331	0.000119
GO:0048518	positive regulation of biological process	44	660	0.000125
GO:0048522	positive regulation of cellular process	40	589	0.000176
GO:0000082	G1/S transition of mitotic cell cycle	5	16	0.000225
GO:0014706	striated muscle tissue development	11	84	0.000285
GO:0055001	muscle cell development	8	47	0.000319
GO:0007517	muscle organ development	13	115	0.000362
GO:0030029	actin filament-based process	13	117	0.000429
GO:0060537	muscle tissue development	11	88	0.00043
GO:0042692	muscle cell differentiation	10	75	0.000463
GO:0001501	skeletal system development	15	150	5e-04
GO:0030048	actin filament-based movement	4	11	0.000524
GO:0032787	monocarboxylic acid metabolic process	14	137	0.000616

Continued on Next Page...

ID	Term Description	q	m	P
GO:0007155	cell adhesion	22	277	0.00073
GO:0022610	biological adhesion	22	277	0.00073
GO:0055002	striated muscle cell development	7	42	0.000861
GO:0051234	establishment of localization	70	1309	0.0011
GO:0007519	skeletal muscle tissue development	7	44	0.00115
GO:0060538	skeletal muscle organ development	7	44	0.00115
GO:0001525	angiogenesis	11	99	0.00117
GO:0009987	cellular process	233	5490	0.00124
GO:0006631	fatty acid metabolic process	11	100	0.00127
GO:0044281	small molecule metabolic process	41	673	0.00134
GO:0048705	skeletal system morphogenesis	8	58	0.00136

Table A.25: Top 40 most enriched GO terms ("Biological process" type) in targets of miR-143.
q - number of genes of a GO term that was among the predicted targets, m - total number of genes of a GO term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0010720	positive regulation of cell development	5	28	0.000571
GO:0045743	positive regulation of fibroblast growth factor receptor signaling pathway	2	3	0.00183
GO:0008219	cell death	21	428	0.00234
GO:0016265	death	21	433	0.00268
GO:0046546	development of primary male sexual characteristics	4	24	0.00272
GO:0050769	positive regulation of neurogenesis	4	25	0.00317
GO:0006497	protein amino acid lipidation	4	27	0.00423
GO:0046661	male sex differentiation	4	27	0.00423
GO:0001501	skeletal system development	10	150	0.00429
GO:0000087	M phase of mitotic cell cycle	9	128	0.00469
GO:0000280	nuclear division	9	128	0.00469
GO:0007067	mitosis	9	128	0.00469
GO:0043066	negative regulation of apoptosis	9	128	0.00469
GO:0045597	positive regulation of cell differentiation	7	84	0.0049
GO:0043069	negative regulation of programmed cell death	9	131	0.00545
GO:0048285	organelle fission	9	132	0.00573
GO:0040036	regulation of fibroblast growth factor receptor signaling pathway	2	5	0.00591
GO:0060548	negative regulation of cell death	9	133	0.00602
GO:0042158	lipoprotein biosynthetic process	4	31	0.00702
GO:0006915	apoptosis	19	411	0.00705

Continued on Next Page...

ID	Term Description	q	m	P
GO:0012501	programmed cell death	19	415	0.00779
GO:0008584	male gonad development	3	17	0.00807
GO:0007271	synaptic transmission, cholinergic	2	6	0.00872
GO:0042157	lipoprotein metabolic process	4	34	0.00976
GO:0009791	post-embryonic development	5	53	0.0101
GO:0001503	ossification	6	74	0.0102
GO:0000278	mitotic cell cycle	10	171	0.0105
GO:0010165	response to X-ray	2	7	0.012
GO:0060325	face morphogenesis	2	7	0.012
GO:0042981	regulation of apoptosis	13	257	0.0122
GO:0001701	in utero embryonic development	9	150	0.0128
GO:0043067	regulation of programmed cell death	13	261	0.0138
GO:0010941	regulation of cell death	13	262	0.0142
GO:0001649	osteoblast differentiation	4	38	0.0144
GO:0008543	fibroblast growth factor receptor signaling pathway	3	21	0.0147
GO:0006970	response to osmotic stress	2	8	0.0157
GO:0060323	head morphogenesis	2	8	0.0157
GO:0060324	face development	2	8	0.0157
GO:0060348	bone development	6	82	0.0164
GO:0051094	positive regulation of developmental process	7	109	0.0192

Table A.26: Top 40 most enriched GO terms ("Biological process" type) in targets of miR-145.
q - number of genes of a GO term that was among the predicted targets, m - total number of genes of a GO category in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0048731	system development	48	1003	0.000418
GO:0016337	cell-cell adhesion	9	81	0.000592
GO:0051301	cell division	15	200	0.000812
GO:0000724	double-strand break repair via homologous recombination	3	8	0.00127
GO:0000725	recombinational repair	3	8	0.00127
GO:0000087	M phase of mitotic cell cycle	11	128	0.00135
GO:0000280	nuclear division	11	128	0.00135
GO:0007067	mitosis	11	128	0.00135
GO:0000279	M phase	13	169	0.00142
GO:0048285	organelle fission	11	132	0.00173
GO:0007275	multicellular organismal development	54	1244	0.0018
GO:0048856	anatomical structure development	48	1077	0.00197
GO:0022403	cell cycle phase	14	198	0.00211

Continued on Next Page...

ID	Term Description	q	m	P
GO:0060710	chorio-allantoic fusion	2	3	0.00254
GO:0021904	dorsal/ventral neural tube patterning	3	10	0.00259
GO:0006310	DNA recombination	6	48	0.00267
GO:0032501	multicellular organismal process	65	1593	0.00275
GO:0060348	bone development	8	82	0.00275
GO:0048872	homeostasis of number of cells	7	65	0.00287
GO:0032502	developmental process	57	1360	0.00295
GO:0022402	cell cycle process	15	234	0.00381
GO:0000278	mitotic cell cycle	12	171	0.00457
GO:0007049	cell cycle	22	410	0.00467
GO:0001501	skeletal system development	11	150	0.00469
GO:0007399	nervous system development	23	439	0.00514
GO:0021532	neural tube patterning	3	13	0.00579
GO:0001503	ossification	7	74	0.00593
GO:0016043	cellular component organization	46	1085	0.0064
GO:0048592	eye morphogenesis	5	41	0.0067
GO:0048873	homeostasis of number of cells within a tissue	3	14	0.00721
GO:0050804	regulation of synaptic transmission	5	42	0.00743
GO:0042592	homeostatic process	17	301	0.00759
GO:0051969	regulation of transmission of nerve impulse	5	43	0.00821
GO:0006826	iron ion transport	3	15	0.00882
GO:0031644	regulation of neurological system process	5	44	0.00905
GO:0051216	cartilage development	5	46	0.0109
GO:0009953	dorsal/ventral pattern formation	4	30	0.011
GO:0023052	signaling	55	1399	0.013
GO:0019226	transmission of nerve impulse	10	150	0.013
GO:0048878	chemical homeostasis	11	174	0.0137

Table A.27: Top 40 most enriched GO terms ("Biological process" type) in targets of miR-25.
q - number of genes of a GO term that was among the predicted targets, m - total number of genes of a GO term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0006814	sodium ion transport	6	57	0.000454
GO:0006366	transcription from RNA polymerase II promoter	13	323	0.00432
GO:0055085	transmembrane transport	11	259	0.00572
GO:0006357	regulation of transcription from RNA polymerase II promoter	12	302	0.00667
GO:0033059	cellular pigmentation	2	8	0.00796

Continued on Next Page...

ID	Term Description	q	m	P
GO:0065007	biological regulation	63	2781	0.00906
GO:0007272	ensheathment of neurons	3	25	0.00912
GO:0008366	axon ensheathment	3	25	0.00912
GO:0048536	spleen development	2	9	0.0101
GO:0015672	monovalent inorganic cation transport	7	141	0.0118
GO:0050801	ion homeostasis	7	142	0.0122
GO:0065008	regulation of biological quality	18	579	0.0123
GO:0006013	mannose metabolic process	2	10	0.0125
GO:0060441	branching involved in lung morphogenesis	2	10	0.0125
GO:0019228	regulation of action potential in neuron	3	29	0.0138
GO:0030850	prostate gland development	3	29	0.0138
GO:0045944	positive regulation of transcription from RNA polymerase II promoter	8	181	0.0141
GO:0030001	metal ion transport	9	219	0.0147
GO:0001655	urogenital system development	5	83	0.015
GO:0042476	odontogenesis	3	30	0.0151
GO:0042592	homeostatic process	11	301	0.0165
GO:0035272	exocrine system development	3	31	0.0165
GO:0006244	pyrimidine nucleotide catabolic process	1	1	0.0175
GO:0006668	sphinganine-1-phosphate metabolic process	1	1	0.0175
GO:0009131	pyrimidine nucleoside monophosphate catabolic process	1	1	0.0175
GO:0009159	deoxyribonucleoside monophosphate catabolic process	1	1	0.0175
GO:0009178	pyrimidine deoxyribonucleoside monophosphate catabolic process	1	1	0.0175
GO:0009223	pyrimidine deoxyribonucleotide catabolic process	1	1	0.0175
GO:0009448	gamma-aminobutyric acid metabolic process	1	1	0.0175
GO:0010447	response to acidity	1	1	0.0175
GO:0016540	protein autoprocessing	1	1	0.0175
GO:0021561	facial nerve development	1	1	0.0175
GO:0021569	rhombomere 3 development	1	1	0.0175
GO:0021571	rhombomere 5 development	1	1	0.0175
GO:0021593	rhombomere morphogenesis	1	1	0.0175
GO:0021594	rhombomere formation	1	1	0.0175
GO:0021604	cranial nerve structural organization	1	1	0.0175
GO:0021610	facial nerve morphogenesis	1	1	0.0175
GO:0021612	facial nerve structural organization	1	1	0.0175
GO:0021658	rhombomere 3 morphogenesis	1	1	0.0175

Table A.28: Top 40 most enriched GO terms ("Biological process" type) in targets of miR-434-3p.
q - number of genes of a GO term that was among the predicted targets, m - total number of genes of a GO category in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0009790	embryonic development	12	333	3.49e-05
GO:0060429	epithelium development	8	145	4.17e-05
GO:0035295	tube development	8	163	9.57e-05
GO:0030855	epithelial cell differentiation	5	56	0.000133
GO:0009888	tissue development	11	331	0.000157
GO:0022612	gland morphogenesis	5	58	0.000158
GO:0009653	anatomical structure morphogenesis	15	585	0.000173
GO:0009887	organ morphogenesis	10	294	0.000265
GO:0002064	epithelial cell development	3	15	0.000288
GO:0048732	gland development	6	103	3e-04
GO:0035239	tube morphogenesis	6	107	0.000368
GO:0002070	epithelial cell maturation	2	4	0.000466
GO:0043009	chordate embryonic development	8	217	0.00067
GO:0001763	morphogenesis of a branching structure	5	81	0.000751
GO:0009792	embryonic development ending in birth or egg hatching	8	221	0.000755
GO:0048729	tissue morphogenesis	6	137	0.00135
GO:0007431	salivary gland development	3	26	0.00153
GO:0048754	branching morphogenesis of a tube	4	59	0.00184
GO:0035050	embryonic heart tube development	2	8	0.00212
GO:0002009	morphogenesis of an epithelium	5	105	0.0024
GO:0035272	exocrine system development	3	31	0.00257
GO:0048598	embryonic morphogenesis	6	163	0.00326
GO:0060442	branching involved in prostate gland morphogenesis	2	10	0.00338
GO:0023034	intracellular signaling pathway	10	412	0.00347
GO:0048523	negative regulation of cellular process	12	559	0.00374
GO:0048513	organ development	15	801	0.00432
GO:0018108	peptidyl-tyrosine phosphorylation	3	39	0.00496
GO:0018212	peptidyl-tyrosine modification	3	39	0.00496
GO:0007399	nervous system development	10	439	0.00543
GO:0007275	multicellular organismal development	20	1244	0.00544
GO:0007154	cell communication	11	513	0.00557
GO:0009069	serine family amino acid metabolic process	2	13	0.00575
GO:0007167	enzyme linked receptor protein signaling pathway	6	185	0.00604
GO:0009968	negative regulation of signal transduction	4	83	0.00633
GO:0023057	negative regulation of signaling process	4	83	0.00633
GO:0034329	cell junction assembly	2	14	0.00667

Continued on Next Page...

ID	Term Description	q	m	P
GO:0032502	developmental process	21	1360	0.0069
GO:0050863	regulation of T cell activation	3	44	0.00696
GO:0048869	cellular developmental process	15	844	0.007
GO:0009966	regulation of signal transduction	8	317	0.00708

Table A.29: Top 40 most enriched GO terms ("Biological process" type) in the induced by transfection set.
q - number of genes of a GO term that was in the mock transfection set, m - total number of genes of a GO term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0048513	organ development	147	801	5.78e-11
GO:0016043	cellular component organization	182	1085	5.58e-10
GO:0048731	system development	171	1003	5.84e-10
GO:0001944	vasculature development	45	160	1.41e-09
GO:0001568	blood vessel development	44	158	3.01e-09
GO:0048518	positive regulation of biological process	121	660	3.97e-09
GO:0048856	anatomical structure development	177	1077	5.26e-09
GO:0007275	multicellular organismal development	198	1244	7.72e-09
GO:0007155	cell adhesion	63	277	9.95e-09
GO:0022610	biological adhesion	63	277	9.95e-09
GO:0032501	multicellular organismal process	240	1593	2.23e-08
GO:0048514	blood vessel morphogenesis	38	135	2.64e-08
GO:0048522	positive regulation of cellular process	108	589	2.98e-08
GO:0032502	developmental process	210	1360	3.18e-08
GO:0008283	cell proliferation	68	326	9.93e-08
GO:0009653	anatomical structure morphogenesis	104	585	2.75e-07
GO:0051301	cell division	47	200	2.79e-07
GO:0007049	cell cycle	79	410	3.11e-07
GO:0050896	response to stimulus	147	911	4.35e-07
GO:0009987	cellular process	673	5490	8.56e-07
GO:0016049	cell growth	20	57	1.26e-06
GO:0048519	negative regulation of biological process	107	626	1.33e-06
GO:0042127	regulation of cell proliferation	50	230	1.44e-06
GO:0001525	angiogenesis	28	99	1.61e-06
GO:0048523	negative regulation of cellular process	96	559	3.95e-06
GO:0065007	biological regulation	368	2781	4.86e-06
GO:0022402	cell cycle process	49	234	5.59e-06
GO:0048869	cellular developmental process	133	844	6.1e-06
GO:0008285	negative regulation of cell proliferation	26	96	9.01e-06

Continued on Next Page...

ID	Term Description	q	m	P
GO:0042221	response to chemical stimulus	59	308	1.22e-05
GO:0008361	regulation of cell size	20	65	1.25e-05
GO:0006260	DNA replication	24	87	1.41e-05
GO:0001558	regulation of cell growth	16	46	1.65e-05
GO:0030154	cell differentiation	124	791	1.69e-05
GO:0050789	regulation of biological process	346	2625	1.73e-05
GO:0090066	regulation of anatomical structure size	29	118	2.17e-05
GO:0010033	response to organic substance	36	161	2.22e-05
GO:0050794	regulation of cellular process	328	2479	2.41e-05
GO:0009888	tissue development	61	331	2.99e-05
GO:0000087	M phase of mitotic cell cycle	30	128	4.17e-05

Table A.30: Top 40 most enriched GO terms ("Biological process" type) in the Ago HITS-CLIP set.
q - number of genes of a GO term that was in the Ago HITS-CLIP set, m - total number of genes of a GO term in test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0051179	localization	348	1469	1.16e-20
GO:0006810	transport	307	1300	8.74e-18
GO:0051234	establishment of localization	307	1309	2.56e-17
GO:0009987	cellular process	946	5490	1.02e-10
GO:0006812	cation transport	73	257	4.13e-08
GO:0006811	ion transport	86	333	2.94e-07
GO:0023052	signaling	277	1399	5.1e-07
GO:0019725	cellular homeostasis	50	173	3.29e-06
GO:0007154	cell communication	116	513	4.52e-06
GO:0055085	transmembrane transport	66	259	1.15e-05
GO:0023060	signal transmission	219	1113	1.66e-05
GO:0023046	signaling process	219	1114	1.77e-05
GO:0007275	multicellular organismal development	241	1244	1.83e-05
GO:0015672	monovalent inorganic cation transport	41	141	2.11e-05
GO:0030001	metal ion transport	57	219	2.34e-05
GO:0065007	biological regulation	491	2781	2.6e-05
GO:0065008	regulation of biological quality	124	579	3.35e-05
GO:0046034	ATP metabolic process	18	44	3.69e-05
GO:0032502	developmental process	258	1360	4.27e-05
GO:0009199	ribonucleoside triphosphate metabolic process	19	49	5.47e-05
GO:0009205	purine ribonucleoside triphosphate metabolic process	19	49	5.47e-05
GO:0055082	cellular chemical homeostasis	37	128	6.03e-05

Continued on Next Page...

ID	Term Description	q	m	P
GO:0007264	small GTPase mediated signal transduction	55	216	6.23e-05
GO:0033036	macromolecule localization	121	570	6.32e-05
GO:0008104	protein localization	108	499	7e-05
GO:0006754	ATP biosynthetic process	17	42	7.1e-05
GO:0032501	multicellular organismal process	295	1593	7.28e-05
GO:0048731	system development	196	1003	7.59e-05
GO:0015031	protein transport	97	440	7.95e-05
GO:0048856	anatomical structure development	208	1077	9.3e-05
GO:0006873	cellular ion homeostasis	36	126	9.83e-05
GO:0009201	ribonucleoside triphosphate biosynthetic process	18	47	0.000103
GO:0009206	purine ribonucleoside triphosphate biosynthetic process	18	47	0.000103
GO:0050801	ion homeostasis	39	142	0.000131
GO:0045184	establishment of protein localization	97	446	0.000136
GO:0009144	purine nucleoside triphosphate metabolic process	19	52	0.00014
GO:0009142	nucleoside triphosphate biosynthetic process	18	48	0.000142
GO:0009145	purine nucleoside triphosphate biosynthetic process	18	48	0.000142
GO:0015985	energy coupled proton transport, down electrochemical gradient	11	22	0.000144
GO:0015986	ATP synthesis coupled proton transport	11	22	0.000144

Table A.31: Top 40 most enriched GO terms ("Cellular compartment" type) in targets of cel-miR-67.
q - number of genes from a GO category that was among the predicted targets, m - total number of genes of a GO category in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0008287	protein serine/threonine phosphatase complex	6	21	4.17e-05
GO:0005783	endoplasmic reticulum	30	469	0.000318
GO:0005955	calcineurin complex	2	3	0.00313
GO:0032153	cell division site	2	4	0.00612
GO:0032155	cell division site part	2	4	0.00612
GO:0044424	intracellular part	196	5358	0.00765
GO:0005794	Golgi apparatus	23	414	0.009
GO:0043231	intracellular membrane-bounded organelle	155	4119	0.01
GO:0043227	membrane-bounded organelle	155	4123	0.0104
GO:0044444	cytoplasmic part	91	2259	0.0131
GO:0005622	intracellular	199	5506	0.0133
GO:0005737	cytoplasm	142	3789	0.0186
GO:0000172	ribonuclease MRP complex	1	1	0.0327
GO:0005655	nucleolar ribonuclease P complex	1	1	0.0327

Continued on Next Page...

ID	Term Description	q	m	P
GO:0005775	vacuolar lumen	1	1	0.0327
GO:0005845	mRNA cap binding complex	1	1	0.0327
GO:0016014	dystrobrevin complex	1	1	0.0327
GO:0030677	ribonuclease P complex	1	1	0.0327
GO:0030681	multimeric ribonuclease P complex	1	1	0.0327
GO:0030685	nucleolar preribosome	1	1	0.0327
GO:0032154	cleavage furrow	1	1	0.0327
GO:0034518	RNA cap binding complex	1	1	0.0327
GO:0000159	protein phosphatase type 2A complex	2	9	0.033
GO:0044452	nucleolar part	2	10	0.0403
GO:0001726	ruffle	3	24	0.0422
GO:0033176	proton-transporting V-type ATPase complex	2	11	0.0483
GO:0043229	intracellular organelle	163	4555	0.0565
GO:0043226	organelle	163	4557	0.0573
GO:0005819	spindle	3	28	0.0621
GO:0002102	podosome	1	2	0.0643
GO:0034708	methyltransferase complex	2	13	0.0656
GO:0035097	histone methyltransferase complex	2	13	0.0656
GO:0032991	macromolecular complex	50	1249	0.0722
GO:0030529	ribonucleoprotein complex	13	256	0.0771
GO:0005624	membrane fraction	12	233	0.0804
GO:0031594	neuromuscular junction	2	15	0.0846
GO:0005662	DNA replication factor A complex	1	3	0.095
GO:0005732	small nucleolar ribonucleoprotein complex	1	3	0.095
GO:0005826	actomyosin contractile ring	1	3	0.095
GO:0031932	TORC2 complex	1	3	0.095

Table A.32: Top 40 most enriched GO terms ("Cellular compartment" type) in targets of miR-124.
q - number of genes from a GO category that was among the predicted targets, m - total number of genes of a GO term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0001725	stress fiber	6	10	5.53e-07
GO:0005886	plasma membrane	77	1179	1.03e-06
GO:0032432	actin filament bundle	6	11	1.18e-06
GO:0042641	actomyosin	6	13	4.1e-06
GO:0016020	membrane	157	3144	2.24e-05
GO:0005623	cell	295	6995	0.000183
GO:0044464	cell part	295	6995	0.000183

Continued on Next Page...

ID	Term Description	q	m	P
GO:0001772	immunological synapse	3	4	0.000216
GO:0005856	cytoskeleton	38	553	0.000296
GO:0030141	secretory granule	7	36	0.000361
GO:0005737	cytoplasm	176	3789	0.000399
GO:0015629	actin cytoskeleton	12	104	0.00059
GO:0044463	cell projection part	9	66	0.000852
GO:0030667	secretory granule membrane	3	6	0.00102
GO:0031088	platelet dense granule membrane	2	2	0.00146
GO:0042827	platelet dense granule	2	2	0.00146
GO:0044459	plasma membrane part	36	567	0.00179
GO:0030054	cell junction	21	274	0.00188
GO:0044420	extracellular matrix part	7	50	0.00273
GO:0044430	cytoskeletal part	21	285	0.00302
GO:0042588	zymogen granule	2	3	0.00427
GO:0005604	basement membrane	6	42	0.00491
GO:0005911	cell-cell junction	9	87	0.00586
GO:0043296	apical junction complex	6	49	0.0105
GO:0044425	membrane part	115	2498	0.0112
GO:0042383	sarcolemma	4	24	0.0122
GO:0005667	transcription factor complex	9	98	0.0125
GO:0016327	apicolateral plasma membrane	6	51	0.0127
GO:0035085	cilium axoneme	2	5	0.0135
GO:0005938	cell cortex	6	54	0.0165
GO:0031594	neuromuscular junction	3	15	0.0179
GO:0043209	myelin sheath	2	6	0.0198
GO:0005819	spindle	4	28	0.0208
GO:0005923	tight junction	5	42	0.0212
GO:0070160	occluding junction	5	42	0.0212
GO:0005626	insoluble fraction	16	240	0.0218
GO:0005576	extracellular region	32	578	0.0218
GO:0044424	intracellular part	223	5358	0.0259
GO:0005887	integral to plasma membrane	12	167	0.0264
GO:0005814	centriole	2	7	0.027

Table A.33: Top 40 most enriched GO terms ("Cellular compartment" type) in targets of miR-143.
q - number of genes of a GO term that was among the predicted targets, m - total number of genes of a GO term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0005794	Golgi apparatus	21	414	0.0012
GO:0044431	Golgi apparatus part	9	119	0.0025
GO:0044444	cytoplasmic part	74	2259	0.00265
GO:0005802	trans-Golgi network	4	28	0.00449
GO:0000775	chromosome, centromeric region	6	65	0.00496
GO:0005737	cytoplasm	112	3789	0.00507
GO:0030496	midbody	2	5	0.00567
GO:0005623	cell	186	6995	0.00929
GO:0044464	cell part	186	6995	0.00929
GO:0016020	membrane	94	3144	0.00966
GO:0016021	integral to membrane	67	2123	0.0111
GO:0031224	intrinsic to membrane	68	2165	0.0117
GO:0030175	filopodium	2	8	0.0151
GO:0043231	intracellular membrane-bounded organelle	116	4119	0.0225
GO:0043227	membrane-bounded organelle	116	4123	0.0233
GO:0005775	vacuolar lumen	1	1	0.0244
GO:0005797	Golgi medial cisterna	1	1	0.0244
GO:0005831	steroid hormone aporeceptor complex	1	1	0.0244
GO:0016942	insulin-like growth factor binding protein complex	1	1	0.0244
GO:0032433	filopodium tip	1	1	0.0244
GO:0045180	basal cortex	1	1	0.0244
GO:0043229	intracellular organelle	126	4555	0.0281
GO:0043226	organelle	126	4557	0.0285
GO:0044425	membrane part	73	2498	0.0432
GO:0000109	nucleotide-excision repair complex	1	2	0.0483
GO:0000214	tRNA-intron endonuclease complex	1	2	0.0483
GO:0005682	U5 snRNP	1	2	0.0483
GO:0031501	mannosyltransferase complex	1	2	0.0483
GO:0031527	filopodium membrane	1	2	0.0483
GO:0033185	dolichol-phosphate-mannose synthase complex	1	2	0.0483
GO:0005869	dynactin complex	1	3	0.0716
GO:0008537	proteasome activator complex	1	3	0.0716
GO:0055037	recycling endosome	1	3	0.0716
GO:0034702	ion channel complex	4	67	0.081
GO:0034707	chloride channel complex	2	20	0.0848
GO:0000776	kinetochore	3	43	0.0872

Continued on Next Page...

ID	Term Description	q	m	P
GO:0022624	proteasome accessory complex	1	4	0.0943
GO:0044424	intracellular part	141	5358	0.0962
GO:0005769	early endosome	2	22	0.0999
GO:0005925	focal adhesion	2	23	0.108

Table A.34: Top 40 most enriched GO terms ("Cellular compartment" type) in targets of miR-145.
q - number of genes of a GO term that was among the predicted targets, m - total number of genes of a GO term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0030175	filopodium	3	8	0.00127
GO:0012506	vesicle membrane	5	40	0.00604
GO:0005623	cell	223	6995	0.00715
GO:0044464	cell part	223	6995	0.00715
GO:0042995	cell projection	17	304	0.0084
GO:0005622	intracellular	181	5506	0.0099
GO:0000775	chromosome, centromeric region	6	65	0.0119
GO:0043209	myelin sheath	2	6	0.012
GO:0030424	axon	6	66	0.0128
GO:0001673	male germ cell nucleus	2	7	0.0165
GO:0030659	cytoplasmic vesicle membrane	4	35	0.0188
GO:0044448	cell cortex part	4	35	0.0188
GO:0005938	cell cortex	5	54	0.0208
GO:0044433	cytoplasmic vesicle part	4	37	0.0227
GO:0043005	neuron projection	8	119	0.0241
GO:0043073	germ cell nucleus	2	9	0.0271
GO:0043229	intracellular organelle	150	4555	0.029
GO:0043226	organelle	150	4557	0.0294
GO:0000176	nuclear exosome (RNase complex)	1	1	0.0294
GO:0005775	vacuolar lumen	1	1	0.0294
GO:0005816	spindle pole body	1	1	0.0294
GO:0005960	glycine cleavage complex	1	1	0.0294
GO:0016939	kinesin II complex	1	1	0.0294
GO:0016942	insulin-like growth factor binding protein complex	1	1	0.0294
GO:0031205	endoplasmic reticulum Sec complex	1	1	0.0294
GO:0032433	filopodium tip	1	1	0.0294
GO:0043231	intracellular membrane-bounded organelle	137	4119	0.0299
GO:0043227	membrane-bounded organelle	137	4123	0.0309
GO:0005871	kinesin complex	2	10	0.0333

Continued on Next Page...

ID	Term Description	q	m	P
GO:0005634	nucleus	84	2403	0.0386
GO:0044463	cell projection part	5	66	0.0445
GO:0044427	chromosomal part	10	185	0.0467
GO:0005694	chromosome	11	213	0.0501
GO:0044424	intracellular part	171	5358	0.0552
GO:0002102	podosome	1	2	0.058
GO:0002139	stereocilia coupling link	1	2	0.058
GO:0002141	stereocilia ankle link	1	2	0.058
GO:0002142	stereocilia ankle link complex	1	2	0.058
GO:0009331	glycerol-3-phosphate dehydrogenase complex	1	2	0.058
GO:0031527	filopodium membrane	1	2	0.058

Table A.35: Top 40 most enriched GO terms ("Cellular compartment" type) in targets of miR-25.
q - number of genes of a GO term that was among the predicted targets, m - total number of genes of a GO term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0005952	cAMP-dependent protein kinase complex	2	6	0.00441
GO:0005942	phosphoinositide 3-kinase complex	2	7	0.0061
GO:0016939	kinesin II complex	1	1	0.0176
GO:0016942	insulin-like growth factor binding protein complex	1	1	0.0176
GO:0060077	inhibitory synapse	1	1	0.0176
GO:0030665	clathrin coated vesicle membrane	2	17	0.0352
GO:0005739	mitochondrion	22	834	0.0353
GO:0005886	plasma membrane	29	1179	0.0375
GO:0000267	cell fraction	9	269	0.048
GO:0001533	cornified envelope	1	3	0.0519
GO:0030118	clathrin coat	2	21	0.052
GO:0005626	insoluble fraction	8	240	0.0614
GO:0016020	membrane	65	3144	0.0642
GO:0042383	sarcolemma	2	24	0.066
GO:0016011	dystroglycan complex	1	4	0.0686
GO:0046540	U4/U6 x U5 tri-snRNP complex	1	4	0.0686
GO:0030662	coated vesicle membrane	2	25	0.071
GO:0030136	clathrin-coated vesicle	3	55	0.0723
GO:0031410	cytoplasmic vesicle	8	249	0.0727
GO:0031982	vesicle	8	253	0.0782
GO:0044425	membrane part	52	2498	0.0913
GO:0000300	peripheral to membrane of membrane fraction	1	6	0.101

Continued on Next Page...

ID	Term Description	q	m	P
GO:0016010	dystrophin-associated glycoprotein complex	1	6	0.101
GO:0030125	clathrin vesicle coat	1	6	0.101
GO:0044445	cytosolic part	2	31	0.103
GO:0045202	synapse	6	184	0.106
GO:0030135	coated vesicle	3	65	0.106
GO:0016023	cytoplasmic membrane-bounded vesicle	5	145	0.113
GO:0014069	postsynaptic density	1	7	0.117
GO:0042470	melanosome	1	7	0.117
GO:0048770	pigment granule	1	7	0.117
GO:0005624	membrane fraction	7	233	0.117
GO:0030117	membrane coat	2	34	0.12
GO:0048475	coated membrane	2	34	0.12
GO:0031988	membrane-bounded vesicle	5	149	0.122
GO:0044421	extracellular region part	8	282	0.124
GO:0016021	integral to membrane	44	2123	0.125
GO:0030659	cytoplasmic vesicle membrane	2	35	0.126
GO:0005615	extracellular space	5	151	0.128
GO:0001750	photoreceptor outer segment	1	8	0.132

Table A.36: Top 40 most enriched GO terms ("Cellular compartment" type) in targets of miR-434-3p.
q - number of genes of a GO term that was among the predicted targets, m - total number of genes of a GO term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0016323	basolateral plasma membrane	4	53	0.00147
GO:0005922	connexon complex	2	9	0.00298
GO:0044459	plasma membrane part	12	567	0.00619
GO:0005912	adherens junction	3	44	0.00792
GO:0005921	gap junction	2	15	0.00837
GO:0070161	anchoring junction	3	46	0.00895
GO:0000439	core TFIIH complex	1	1	0.00934
GO:0000441	SSL2-core TFIIH complex	1	1	0.00934
GO:0032806	carboxy-terminal domain protein kinase complex	1	1	0.00934
GO:0005737	cytoplasm	46	3789	0.0137
GO:0030054	cell junction	7	274	0.014
GO:0005623	cell	74	6995	0.0156
GO:0044464	cell part	74	6995	0.0156
GO:0005925	focal adhesion	2	23	0.0192
GO:0005924	cell-substrate adherens junction	2	24	0.0209

Continued on Next Page...

ID	Term Description	q	m	P
GO:0030055	cell-substrate junction	2	25	0.0225
GO:0005886	plasma membrane	18	1179	0.0229
GO:0043235	receptor complex	2	28	0.0279
GO:0005856	cytoskeleton	10	553	0.0328
GO:0009986	cell surface	3	75	0.0329
GO:0005587	collagen type IV	1	4	0.0369
GO:0005786	signal recognition particle, endoplasmic reticulum targeting	1	4	0.0369
GO:0030935	sheet-forming collagen	1	4	0.0369
GO:0042719	mitochondrial intermembrane space protein transporter complex	1	4	0.0369
GO:0048500	signal recognition particle	1	4	0.0369
GO:0005892	nicotinic acetylcholine-gated receptor-channel complex	1	5	0.0459
GO:0030496	midbody	1	5	0.0459
GO:0005911	cell-cell junction	3	87	0.0477
GO:0005625	soluble fraction	2	38	0.0488
GO:0016020	membrane	37	3144	0.0536
GO:0005844	polysome	1	7	0.0636
GO:0043231	intracellular membrane-bounded organelle	46	4119	0.0637
GO:0043227	membrane-bounded organelle	46	4123	0.0648
GO:0009897	external side of plasma membrane	2	45	0.066
GO:0043229	intracellular organelle	50	4555	0.0663
GO:0043226	organelle	50	4557	0.0668
GO:0016021	integral to membrane	26	2123	0.0757
GO:0005741	mitochondrial outer membrane	2	49	0.0765
GO:0031968	organelle outer membrane	2	50	0.0793
GO:0019867	outer membrane	2	52	0.0848

Table A.37: Top 40 most enriched GO terms ("Cellular component" type) in the Ago HITS-CLIP set.
q - number of genes of a GO term that was in the Ago HITS-CLIP set, m - total number of genes of a GO term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0005623	cell	1229	6995	4.47e-26
GO:0044464	cell part	1229	6995	4.47e-26
GO:0016020	membrane	654	3144	5.98e-24
GO:0044425	membrane part	529	2498	1.05e-19
GO:0016021	integral to membrane	454	2123	3.91e-17
GO:0031224	intrinsic to membrane	461	2165	4.5e-17

Continued on Next Page...

ID	Term Description	q	m	P
GO:0005737	cytoplasm	708	3789	5.88e-13
GO:0044444	cytoplasmic part	453	2259	6.25e-12
GO:0005783	endoplasmic reticulum	127	469	2.3e-11
GO:0005622	intracellular	950	5506	2.28e-09
GO:0044424	intracellular part	923	5358	1.25e-08
GO:0005624	membrane fraction	66	233	2.9e-07
GO:0005886	plasma membrane	241	1179	5.38e-07
GO:0043005	neuron projection	40	119	6.07e-07
GO:0005626	insoluble fraction	66	240	9.36e-07
GO:0031410	cytoplasmic vesicle	66	249	3.77e-06
GO:0000267	cell fraction	70	269	3.91e-06
GO:0043229	intracellular organelle	781	4555	4.81e-06
GO:0043226	organelle	781	4557	5.22e-06
GO:0031982	vesicle	66	253	6.73e-06
GO:0045202	synapse	49	184	5.76e-05
GO:0005794	Golgi apparatus	92	414	0.000124
GO:0043234	protein complex	185	940	0.000137
GO:0016469	proton-transporting two-sector ATPase complex	12	26	0.000206
GO:0016023	cytoplasmic membrane-bounded vesicle	39	145	0.000255
GO:0043227	membrane-bounded organelle	697	4123	0.00029
GO:0043231	intracellular membrane-bounded organelle	696	4119	0.000314
GO:0031090	organelle membrane	90	414	0.000326
GO:0005789	endoplasmic reticulum membrane	19	55	0.000364
GO:0042175	nuclear envelope-endoplasmic reticulum network	19	55	0.000364
GO:0031988	membrane-bounded vesicle	39	149	0.000463
GO:0005792	microsome	22	69	0.000471
GO:0033017	sarcoplasmic reticulum membrane	4	4	0.000568
GO:0042598	vesicular fraction	22	70	0.000589
GO:0044456	synapse part	30	107	6e-04
GO:0044432	endoplasmic reticulum part	21	66	0.000649
GO:0005829	cytosol	48	197	0.000652
GO:0005856	cytoskeleton	113	553	0.000741
GO:0030136	clathrin-coated vesicle	18	55	0.00107
GO:0071212	subsynaptic reticulum	24	83	0.00128

Table A.38: Top 40 most enriched GO terms ("Cellular compartment" type) in the induced by transfection set.
q - number of genes of a GO term that was among in the mock transfection set, m - total number of genes of a GO term in the test universe, P - P-value of enrichment. See Methods, section 2.10.

Term ID	Term Description	q	m	P
GO:0005578	proteinaceous extracellular matrix	42	135	1.69e-10
GO:0044421	extracellular region part	68	282	1.89e-10
GO:0031012	extracellular matrix	43	146	7.13e-10
GO:0005576	extracellular region	108	578	1.15e-08
GO:0044420	extracellular matrix part	19	50	5.76e-07
GO:0043228	non-membrane-bounded organelle	145	910	1.26e-06
GO:0043232	intracellular non-membrane-bounded organelle	145	910	1.26e-06
GO:0044427	chromosomal part	43	185	1.28e-06
GO:0005604	basement membrane	16	42	4.33e-06
GO:0005694	chromosome	46	213	4.75e-06
GO:0005856	cytoskeleton	94	553	8.36e-06
GO:0032993	protein-DNA complex	18	55	1.33e-05
GO:0015629	actin cytoskeleton	27	104	1.46e-05
GO:0000786	nucleosome	16	49	4.13e-05
GO:0005623	cell	817	6995	0.000137
GO:0044464	cell part	817	6995	0.000137
GO:0000323	lytic vacuole	25	105	0.000139
GO:0005764	lysosome	25	105	0.000139
GO:0005886	plasma membrane	166	1179	0.000283
GO:0005615	extracellular space	31	151	0.000424
GO:0032432	actin filament bundle	6	11	0.000499
GO:0005773	vacuole	26	120	0.000511
GO:0009986	cell surface	18	75	0.00105
GO:0000785	chromatin	21	95	0.00131
GO:0005826	actomyosin contractile ring	3	3	0.00133
GO:0043256	laminin complex	3	3	0.00133
GO:0070938	contractile ring	3	3	0.00133
GO:0016323	basolateral plasma membrane	14	53	0.00138
GO:0005605	basal lamina	4	6	0.00183
GO:0005912	adherens junction	12	44	0.00222
GO:0001725	stress fiber	5	10	0.00252
GO:0009897	external side of plasma membrane	12	45	0.00273
GO:0070161	anchoring junction	12	46	0.00333
GO:0044459	plasma membrane part	83	567	0.00367
GO:0005581	collagen	5	11	0.0042
GO:0005626	insoluble fraction	40	240	0.00477

Continued on Next Page...

ID	Term Description	q	m	P
GO:0005587	collagen type IV	3	4	0.0049
GO:0030935	sheet-forming collagen	3	4	0.0049
GO:0032153	cell division site	3	4	0.0049
GO:0032155	cell division site part	3	4	0.0049

Symbol	Description
miR-124 (**Ranked 11, P ≈ 0.0458**)	
Col4a1	collagen, type IV, alpha 1
Col5a1	collagen, type V, alpha 1
Itga7	integrin alpha 7
Itgb1	integrin beta 1 (fibronectin receptor beta)
Lamc1	laminin, gamma 1
miR-145 (Ranked 97, P ≈ 0.76)	
Col1a1	collagen, type I, alpha 1
miR-25 (**Ranked 14, P ≈ 0.0504**)	
Dag1	dystroglycan 1
Sdc2	syndecan 2
Sdc3	syndecan 3
miR-434-3p (**Ranked 2, P ≈ 0.0149**)	
Cd44	CD44 antigen
Col1a1	collagen, type I, alpha 1
Sdc2	syndecan 2

Table A.39: miRNA targets within "ECM-receptor interaction" KEGG pathway.
The text in parenthesis shows the rank of the enrichment of the "ECM-receptor interaction" among all the KEGG pathways and the P-value of that enrichment. The text is in bold if the enrichment was ranked within the top 25 most enriched pathways.

Symbol	Description
cel-miR-67 (Ranked 55, P ≈ 0.42)	
Adcy5	adenylate cyclase 5
Csnk1d	casein kinase 1, delta
Kras	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
miR-124 (Ranked 100, P ≈ 0.735)	
Nras	neuroblastoma ras oncogene
Tubb6	tubulin, beta 6
miR-143 (**Ranked 10, P ≈ 0.0976**)	
Egfr	epidermal growth factor receptor
Pdgfb	platelet derived growth factor, B polypeptide
Pdgfra	platelet derived growth factor receptor, alpha polypeptide
Src	Rous sarcoma oncogene
miR-145 (**Ranked 16, P ≈ 0.16**)	
Grm5	glutamate receptor, metabotropic 5
Nras	neuroblastoma ras oncogene
Pdgfra	platelet derived growth factor receptor, alpha polypeptide
Prkx	protein kinase, X-linked
miR-25 (Ranked 95, P ≈ 0.728)	
Adcy3	adenylate cyclase 3
miR-434-3p (Ranked 28, P ≈ 0.207)	
Egfr	epidermal growth factor receptor
Gja1	gap junction protein, alpha 1

Table A.40: miRNA targets within "Gap junction" KEGG pathway.
The text in parenthesis shows the rank of the enrichment of the "Gap junction" among all the KEGG pathways and the P-value of that enrichment. The text is in bold if the enrichment was ranked within the top 25 most enriched pathways.

Symbol	Description
cel-miR-67 (**Ranked 4, P ≈ 0.0168**)	
Gab1	growth factor receptor bound protein 2-associated protein 1
Hbegf	heparin-binding EGF-like growth factor
Kras	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
Map2k4	mitogen-activated protein kinase kinase 4
Pak4	p21 protein (Cdc42/Rac)-activated kinase 4
Pik3cb	phosphatidylinositol 3-kinase, catalytic, beta polypeptide
miR-124 (Ranked 76, P ≈ 0.51)	
Erbb2	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian)
Nras	neuroblastoma ras oncogene
Shc1	src homology 2 domain-containing transforming protein C1
miR-143 (Ranked 58, P ≈ 0.464)	
Egfr	epidermal growth factor receptor
Src	Rous sarcoma oncogene
miR-145 (Ranked 79, P ≈ 0.607)	
Map2k4	mitogen-activated protein kinase kinase 4
Nras	neuroblastoma ras oncogene
miR-25 (**Ranked 7, P ≈ 0.0245**)	
Map2k4	mitogen-activated protein kinase kinase 4
Mapk8	mitogen-activated protein kinase 8
Pik3cb	phosphatidylinositol 3-kinase, catalytic, beta polypeptide
Pik3r2	phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 2 (p85 beta)
miR-434-3p (**Ranked 24, P ≈ 0.184**)	
Egfr	epidermal growth factor receptor
Nck1	non-catalytic region of tyrosine kinase adaptor protein 1

Table A.41: miRNA targets within "ErbB signaling pathway" KEGG pathway.
The text in parenthesis shows the rank of the enrichment of the "ErbB signaling pathway" among all the KEGG pathways and the P-value of that enrichment. The text is in bold if the enrichment was ranked within the top 25 most enriched pathways.

Symbol	Description
cel-miR-67 (Ranked 68, P ≈ 0.527)	
Amotl1	angiomotin-like 1
Kras	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
Ppp2r2a	protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), alpha isoform
miR-124 (**Ranked 3, P ≈ 0.00907**)	
Actn4	actinin alpha 4
Amotl1	angiomotin-like 1
Jam2	junction adhesion molecule 2
Myh10	myosin, heavy polypeptide 10, non-muscle
Myh9	myosin, heavy polypeptide 9, non-muscle
Nras	neuroblastoma ras oncogene
Rras	Harvey rat sarcoma oncogene, subgroup R
Tjp2	tight junction protein 2
miR-143 (Ranked 95, P ≈ 0.868)	
Src	Rous sarcoma oncogene
miR-145 (Ranked 86, P ≈ 0.667)	
Nras	neuroblastoma ras oncogene
Prkcz	protein kinase C, zeta
miR-25 (Ranked 76, P ≈ 0.412)	
B230120H23Rik	RIKEN cDNA B230120H23 gene
Cldn11	claudin 11

Table A.42: miRNA targets within "Tight junction" KEGG pathway.
The text in parenthesis shows the rank of the enrichment of the "Tight junction" among all the KEGG pathways and the P-value of that enrichment. The text is in bold if the enrichment was ranked within the top 25 most enriched pathways.

Symbol	Description
cel-miR-67 (Ranked 39, P ≈ 0.26)	
Ccnb1	cyclin B1
Cycs	cytochrome c, somatic
Igfbp3	insulin-like growth factor binding protein 3
miR-124 (**Ranked 23, P ≈ 0.111**)	
Ccnd1	cyclin D1
Cd82	CD82 antigen
Igfbp3	insulin-like growth factor binding protein 3
Zmat3	zinc finger matrin type 3
miR-143 (Ranked 36, P ≈ 0.284)	
Cd82	CD82 antigen
Zmat3	zinc finger matrin type 3
miR-145 (Ranked 49, P ≈ 0.408)	
Ccng1	cyclin G1
Cycs	cytochrome c, somatic
miR-25 (Ranked 50, P ≈ 0.245)	
Rrm2b	ribonucleotide reductase M2 B (TP53 inducible)
Zmat3	zinc finger matrin type 3
miR-434-3p (**Ranked 3, P ≈ 0.0161**)	
Ccnb1	cyclin B1
Ccnd2	cyclin D2
Ccng1	cyclin G1

Table A.43: miRNA targets within "p53 signaling pathway" KEGG pathway.
The text in parenthesis shows the rank of the enrichment of the "p53 signaling pathway" among all the KEGG pathways and the P-value of that enrichment. The text is in bold if the enrichment was ranked within the top 25 most enriched pathways.

Symbol	Description
cel-miR-67 (Ranked 65, P ≈ 0.494)	
Gna13	guanine nucleotide binding protein, alpha 13
Kras	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
Pak4	p21 protein (Cdc42/Rac)-activated kinase 4
Pik3cb	phosphatidylinositol 3-kinase, catalytic, beta polypeptide
Wasf2	WAS protein family, member 2
miR-124 (**Ranked 7, P ≈ 0.0223**)	
Actn4	actinin alpha 4
Arpc1b	actin related protein 2/3 complex, subunit 1B
Iqgap1	IQ motif containing GTPase activating protein 1
Itga7	integrin alpha 7
Itgb1	integrin beta 1 (fibronectin receptor beta)
Myh10	myosin, heavy polypeptide 10, non-muscle
Myh9	myosin, heavy polypeptide 9, non-muscle
Nras	neuroblastoma ras oncogene
Pip4k2c	phosphatidylinositol-5-phosphate 4-kinase, type II, gamma
Rras	Harvey rat sarcoma oncogene, subgroup R
Wasf2	WAS protein family, member 2
miR-143 (**Ranked 2, P ≈ 0.0422**)	
Arhgef1	Rho guanine nucleotide exchange factor (GEF) 1
Arhgef4	Rho guanine nucleotide exchange factor (GEF) 4
Egfr	epidermal growth factor receptor
Gng12	guanine nucleotide binding protein (G protein), gamma 12
Limk1	LIM-domain containing, protein kinase
Pdgfb	platelet derived growth factor, B polypeptide
Pdgfra	platelet derived growth factor receptor, alpha polypeptide
miR-145 (Ranked 52, P ≈ 0.413)	
F2r	coagulation factor II (thrombin) receptor
Nras	neuroblastoma ras oncogene
Pdgfra	platelet derived growth factor receptor, alpha polypeptide
Tiam1	T-cell lymphoma invasion and metastasis 1
Wasf2	WAS protein family, member 2
miR-25 (**Ranked 11, P ≈ 0.0381**)	
Fgf10	fibroblast growth factor 10
Fgf12	fibroblast growth factor 12
Pik3cb	phosphatidylinositol 3-kinase, catalytic, beta polypeptide
Pik3r2	phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 2 (p85 beta)
Pip5k1c	phosphatidylinositol-4-phosphate 5-kinase, type 1 gamma
Slc9a1	solute carrier family 9 (sodium/hydrogen exchanger), member 1
miR-434-3p (Ranked 31, P ≈ 0.226)	
Egfr	epidermal growth factor receptor
Fgf13	fibroblast growth factor 13
Pxn	paxillin

Table A.44: miRNA targets within "Regulation of actin cytoskeleton" KEGG pathway. The text in parenthesis shows the rank of the enrichment of the "Regulation of actin cytoskeleton" among all the KEGG pathways and the P-value of that enrichment. The text is in bold if the enrichment was ranked within the top 25 most enriched pathways.

Symbol	Description
cel-miR-67 (Ranked 76, P ≈ 0.622)	
Bcl2	B-cell leukemia/lymphoma 2
Pak4	p21 protein (Cdc42/Rac)-activated kinase 4
Pik3cb	phosphatidylinositol 3-kinase, catalytic, beta polypeptide
Zyx	zyxin
miR-124 (**Ranked 1, P ≈ 0.00211**)	
Actn4	actinin alpha 4
Capn2	calpain 2
Cav1	caveolin 1, caveolae protein
Ccnd1	cyclin D1
Col4a1	collagen, type IV, alpha 1
Col5a1	collagen, type V, alpha 1
Erbb2	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived [...]
Flnb	filamin, beta
Flnc	filamin C, gamma
Itga7	integrin alpha 7
Itgb1	integrin beta 1 (fibronectin receptor beta)
Lamc1	laminin, gamma 1
Shc1	src homology 2 domain-containing transforming protein C1
miR-143 (Ranked 42, P ≈ 0.346)	
Egfr	epidermal growth factor receptor
Pdgfb	platelet derived growth factor, B polypeptide
Pdgfra	platelet derived growth factor receptor, alpha polypeptide
Src	Rous sarcoma oncogene
miR-145 (Ranked 94, P ≈ 0.738)	
Col1a1	collagen, type I, alpha 1
Flnb	filamin, beta
Pdgfra	platelet derived growth factor receptor, alpha polypeptide
miR-25 (**Ranked 19, P ≈ 0.0888**)	
Igf1r	insulin-like growth factor I receptor
Mapk8	mitogen-activated protein kinase 8
Pik3cb	phosphatidylinositol 3-kinase, catalytic, beta polypeptide
Pik3r2	phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 2 (p85 beta)
Pip5k1c	phosphatidylinositol-4-phosphate 5-kinase, type 1 gamma
miR-434-3p (**Ranked 9, P ≈ 0.0652**)	
Ccnd2	cyclin D2
Col1a1	collagen, type I, alpha 1
Egfr	epidermal growth factor receptor
Pxn	paxillin

Table A.45: miRNA targets within "Focal adhesion" KEGG pathway.
The text in parenthesis shows the rank of the enrichment of the "Focal adhesion" among all the KEGG pathways and the P-value of that enrichment. The text is in bold if the enrichment was ranked within the top 25 most enriched pathways.

Symbol	Description
cel-miR-67 (Ranked 44, P ≈ 0.287)	
Arrb1	arrestin, beta 1
Kras	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
Map2k4	mitogen-activated protein kinase kinase 4
Mapk14	mitogen-activated protein kinase 14
Ntrk2	neurotrophic tyrosine kinase, receptor, type 2
Ppm1b	protein phosphatase 1B, magnesium dependent, beta isoform
Ppp3ca	protein phosphatase 3, catalytic subunit, alpha isoform
Ppp3r1	protein phosphatase 3, regulatory subunit B, alpha isoform (calcineurin B, type I)
miR-124 (Ranked 55, P ≈ 0.315)	
Dusp6	dual specificity phosphatase 6
Flnb	filamin, beta
Flnc	filamin C, gamma
Mapk14	mitogen-activated protein kinase 14
Mapkapk3	mitogen-activated protein kinase-activated protein kinase 3
Nras	neuroblastoma ras oncogene
Ntrk2	neurotrophic tyrosine kinase, receptor, type 2
Rela	v-rel reticuloendotheliosis viral oncogene homolog A (avian)
Rras	Harvey rat sarcoma oncogene, subgroup R
miR-143 (**Ranked 1, P ≈ 0.0342**)	
Cacnb3	calcium channel, voltage-dependent, beta 3 subunit
Dusp16	dual specificity phosphatase 16
Dusp7	dual specificity phosphatase 7
Egfr	epidermal growth factor receptor
Gng12	guanine nucleotide binding protein (G protein), gamma 12
Mapkapk3	mitogen-activated protein kinase-activated protein kinase 3
Pdgfb	platelet derived growth factor, B polypeptide
Pdgfra	platelet derived growth factor receptor, alpha polypeptide
Rps6ka1	ribosomal protein S6 kinase polypeptide 1
miR-145 (**Ranked 1, P ≈ 0.000891**)	
Dusp3	dual specificity phosphatase 3 (vaccinia virus phosphatase VH1-related)
Dusp6	dual specificity phosphatase 6
Flnb	filamin, beta
Ikbkg	inhibitor of kappaB kinase gamma
Map2k4	mitogen-activated protein kinase kinase 4
Map3k1	mitogen-activated protein kinase kinase kinase 1
Nras	neuroblastoma ras oncogene
Ntrk2	neurotrophic tyrosine kinase, receptor, type 2
Pdgfra	platelet derived growth factor receptor, alpha polypeptide
Ppp3ca	protein phosphatase 3, catalytic subunit, alpha isoform
Prkx	protein kinase, X-linked
Rasa1	RAS p21 protein activator 1
Taok1	TAO kinase 1
Traf6	TNF receptor-associated factor 6
miR-25 (**Ranked 12, P ≈ 0.0414**)	
B230120H23Rik	RIKEN cDNA B230120H23 gene
Fgf10	fibroblast growth factor 10
Fgf12	fibroblast growth factor 12
Map2k4	mitogen-activated protein kinase kinase 4
Mapk8	mitogen-activated protein kinase 8
Rps6ka4	ribosomal protein S6 kinase, polypeptide 4
Taok1	TAO kinase 1
miR-434-3p (Ranked 48, P ≈ 0.393)	
Egfr	epidermal growth factor receptor
Fgf13	fibroblast growth factor 13
Tgfbr2	transforming growth factor, beta receptor II

Table A.46: miRNA targets within "MAPK signaling pathway" KEGG pathway.
The text in parenthesis shows the rank of the enrichment of the "MAPK signaling pathway" among all the KEGG pathways and the P-value of that enrichment. The text is in bold if the enrichment was ranked within the top 25 most enriched pathways.

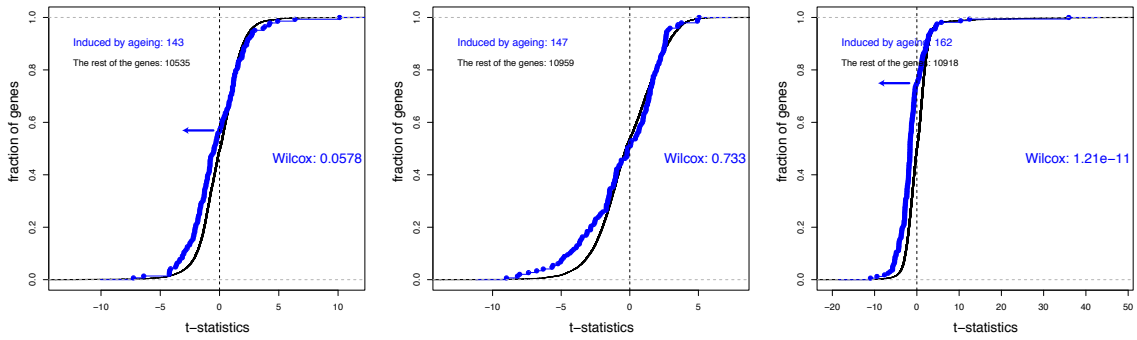
Symbol	Description
cel-miR-67 (**Ranked 9, P ≈ 0.0272**)	
Kras	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
Mapk14	mitogen-activated protein kinase 14
Pik3cb	phosphatidylinositol 3-kinase, catalytic, beta polypeptide
Ppp3ca	protein phosphatase 3, catalytic subunit, alpha isoform
Ppp3r1	protein phosphatase 3, regulatory subunit B, alpha isoform (calcineurin B, type I)
miR-124 (**Ranked 24, P ≈ 0.111**)	
Mapk14	mitogen-activated protein kinase 14
Mapkapk3	mitogen-activated protein kinase-activated protein kinase 3
Nfatc1	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1
Nras	neuroblastoma ras oncogene
miR-143 (**Ranked 18, P ≈ 0.125**)	
Mapkapk3	mitogen-activated protein kinase-activated protein kinase 3
Nfatc1	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1
Src	Rous sarcoma oncogene
miR-145 (Ranked 50, P ≈ 0.408)	
Nras	neuroblastoma ras oncogene
Ppp3ca	protein phosphatase 3, catalytic subunit, alpha isoform
miR-25 (Ranked 56, P ≈ 0.27)	
Pik3cb	phosphatidylinositol 3-kinase, catalytic, beta polypeptide
Pik3r2	phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 2 (p85 beta)
miR-434-3p (Ranked 49, P ≈ 0.424)	
Pxn	paxillin

Table A.47: miRNA targets within "VEGF signaling pathway" KEGG pathway.
The text in parenthesis shows the rank of the enrichment of the "VEGF signaling pathway" among all the KEGG pathways and the P-value of that enrichment. The text is in bold if the enrichment was ranked within the top 25 most enriched pathways.

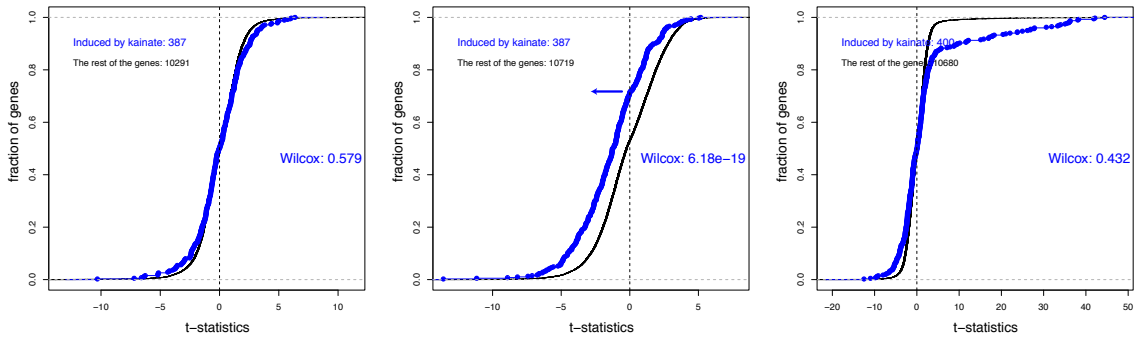
Symbol	Description
cel-miR-67 (**Ranked 19, P ≈ 0.112**)	
Map2k4	mitogen-activated protein kinase kinase 4
Mapk14	mitogen-activated protein kinase 14
Pik3cb	phosphatidylinositol 3-kinase, catalytic, beta polypeptide
Tollip	toll interacting protein
miR-124 (Ranked 44, P ≈ 0.242)	
Fadd	Fas (TNFRSF6)-associated via death domain
Mapk14	mitogen-activated protein kinase 14
Rela	v-rel reticuloendotheliosis viral oncogene homolog A (avian)
Traf3	TNF receptor-associated factor 3
miR-143 (Ranked 89, P ≈ 0.807)	
Fadd	Fas (TNFRSF6)-associated via death domain
miR-145 (Ranked 41, P ≈ 0.329)	
Ikbkg	inhibitor of kappaB kinase gamma
Map2k4	mitogen-activated protein kinase kinase 4
Traf6	TNF receptor-associated factor 6
miR-25 (**Ranked 13, P ≈ 0.0453**)	
Map2k4	mitogen-activated protein kinase kinase 4
Mapk8	mitogen-activated protein kinase 8
Pik3cb	phosphatidylinositol 3-kinase, catalytic, beta polypeptide
Pik3r2	phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 2 (p85 beta)

Table A.48: miRNA targets within "Toll-like receptor signaling pathway" KEGG pathway.
The text in parenthesis shows the rank of the enrichment of the "Toll-like receptor signaling pathway" among all the KEGG pathways and the P-value of that enrichment. The text is in bold if the enrichment was ranked within the top 25 most enriched pathways.

(a) **miR-25** vs. ageing (b) **miR-143** vs. ageing (c) **miR-145** vs. ageing



(d) **miR-25** vs. kainate (e) **miR-143** vs. kainate (f) **miR-145** vs. kainate



(g) **miR-25** vs. transf. (h) **miR-143** vs. transf. (i) **miR-145** vs. transf.

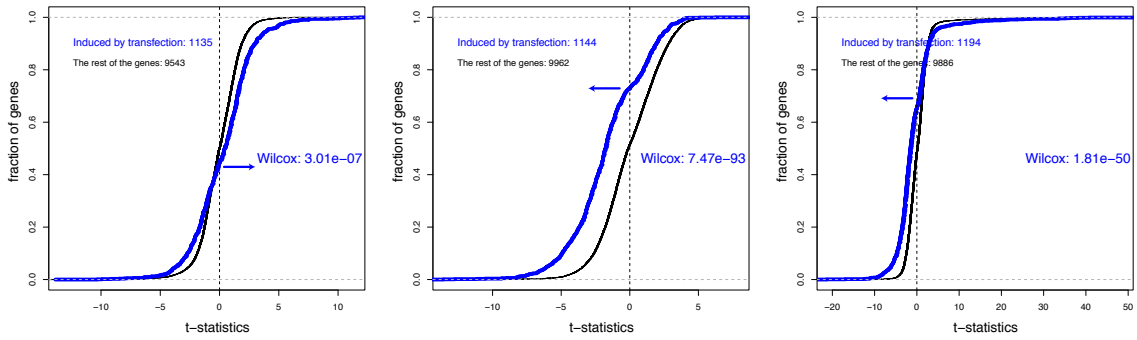
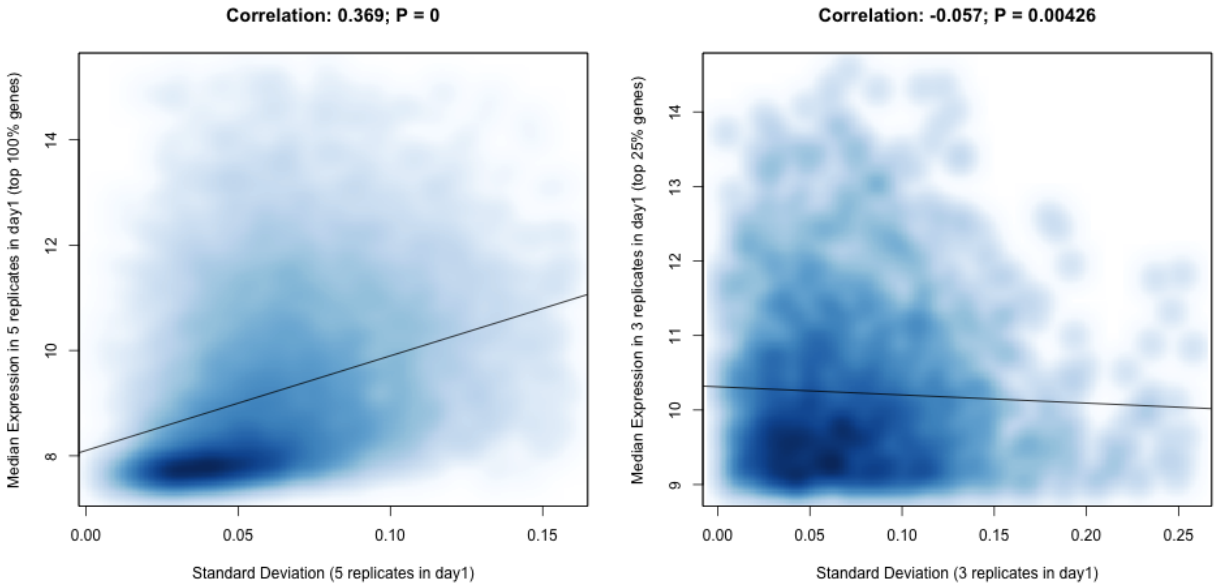


Figure A.7: Regulation of stress induced genes by miR-25, miR-143 and miR-145. The y-axes show the cumulative fraction of genes, the x-axes show the fold change t-statistics (Methods, section 2.7). Genes significantly induced (differential expression $P < 0.05$) by one of the three stresses (the ageing, kainate or transfection stresses) are shown as the blue line/points. The rest of the genes (except 0.01% most highly up- and downregulated genes, which were not plotted for the purpose of better scaling) is shown as the black lines. The text in the plot areas shows: 1) The number of genes *induced by* a stress that were expressed in the miRNA transfection experiments; 2) The number of other expressed genes (*The rest of the genes*); 3) The Wilcoxon test P-value for the difference in medians of the fold change t-statistics for the stress induced genes and the rest of the genes (*Wilcox*). The blue arrows show the direction of the shift in experiments where the Wilcoxon test P-value was significant ($P < 0.05$). The titles of the subfigures show: The names of the perturbed miRNAs (in bold) and the name of the stress experiment where the stress induced genes were identified.

(a) HP cultures at 1DIV; 100% of genes

(b) HP cultures at 1DIV; top 25% of genes



(c) HP cultures at 8DIV; 100% of genes

(d) HP cultures at 8DIV; top 25% of genes

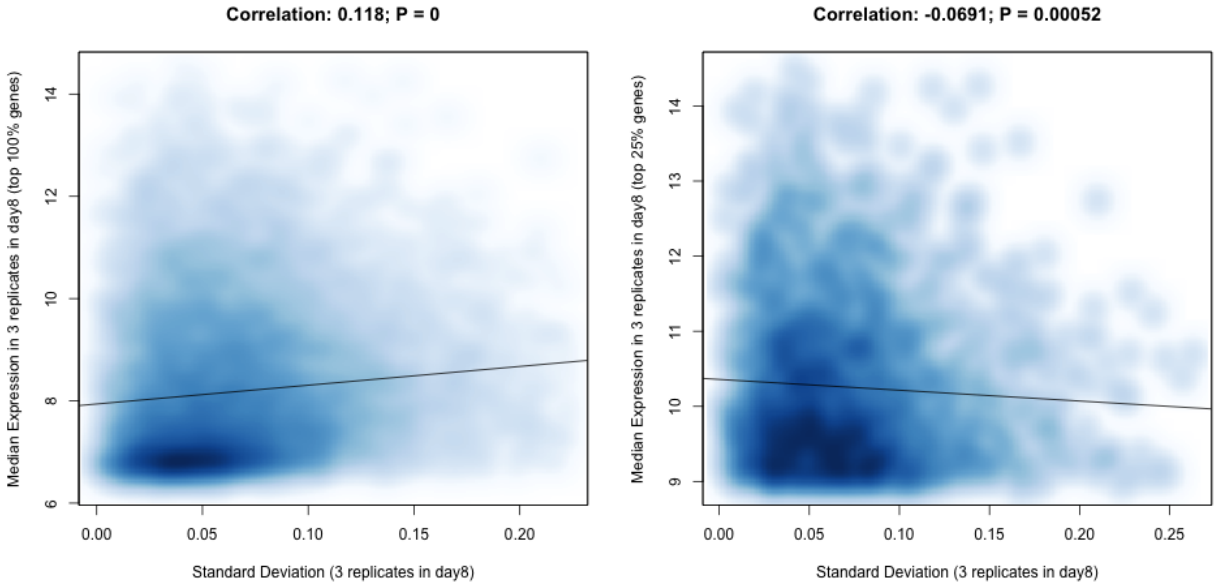
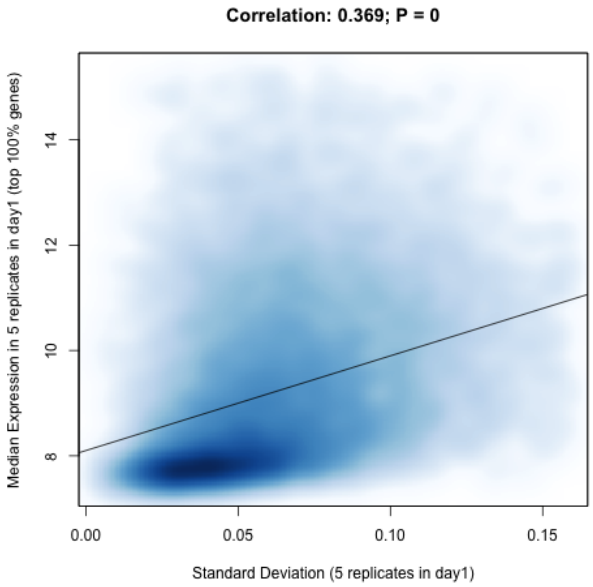
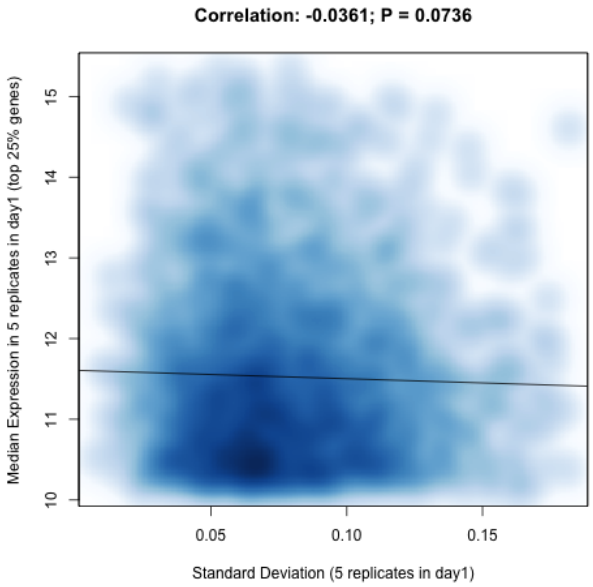


Figure A.8: Correlation of expression and standard deviation in hippocampal primary cultures.

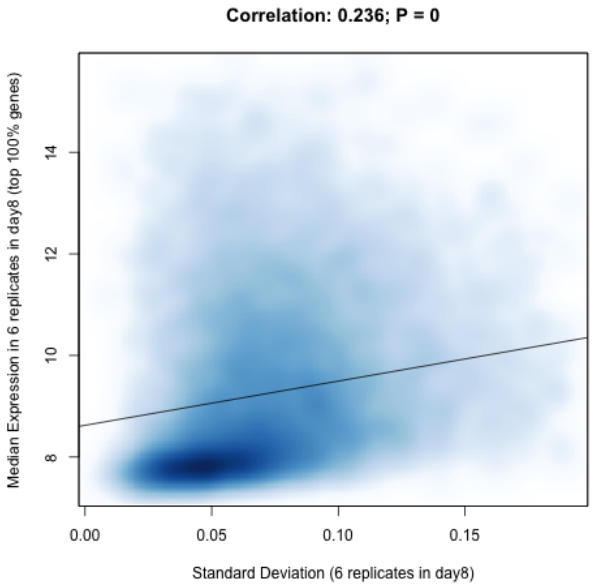
(a) FB cultures at 1DIV; 100% of genes



(b) FB cultures at 1DIV; top 25% of genes



(c) FB cultures at 8DIV; 100% of genes



(d) FB cultures at 8DIV; top 25% of genes

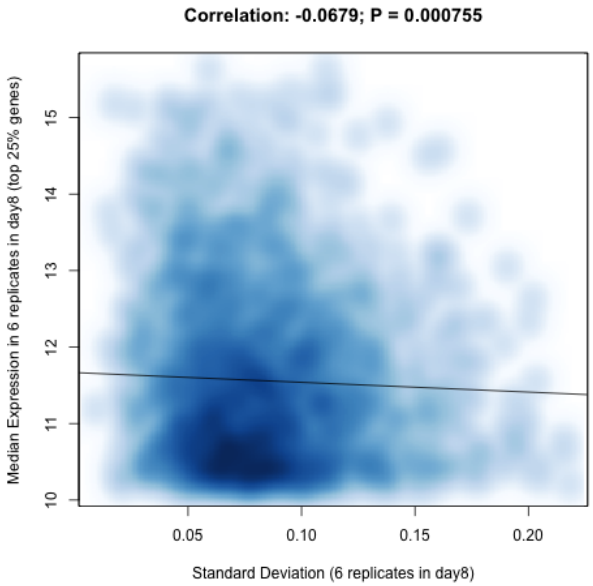
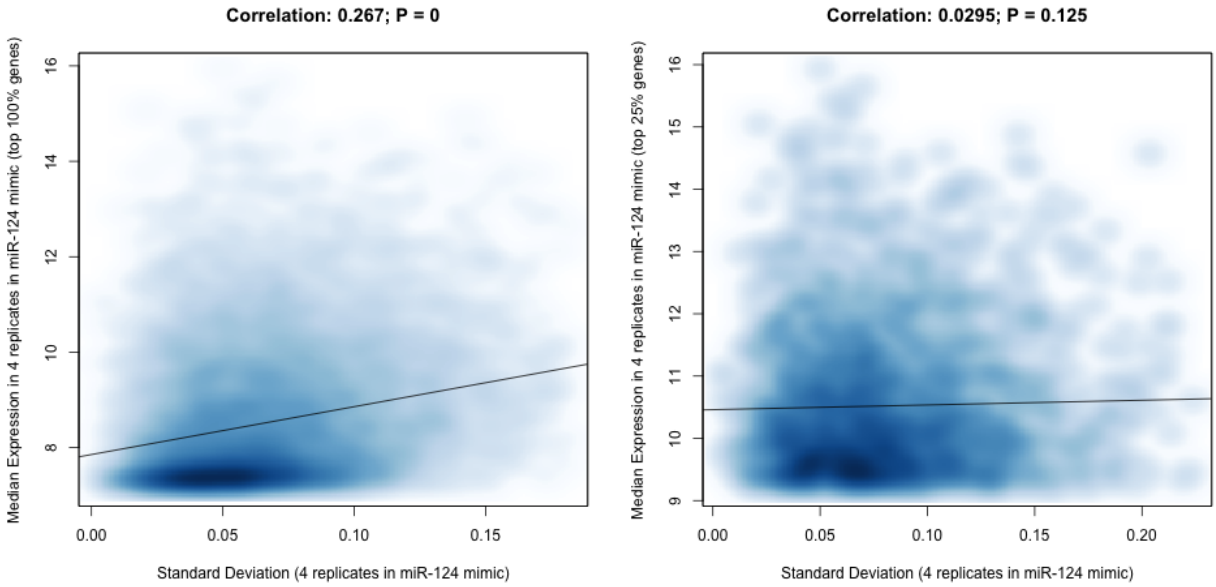


Figure A.9: Correlation of expression and standard deviation in forebrain primary cultures.

(a) Overexpression of miR-124; 100% of genes

(b) Overexpression of miR-124; top 25% of genes



(c) Inhibition of miR-124; 100% of genes

(d) Inhibition of miR-124; top 25% of genes

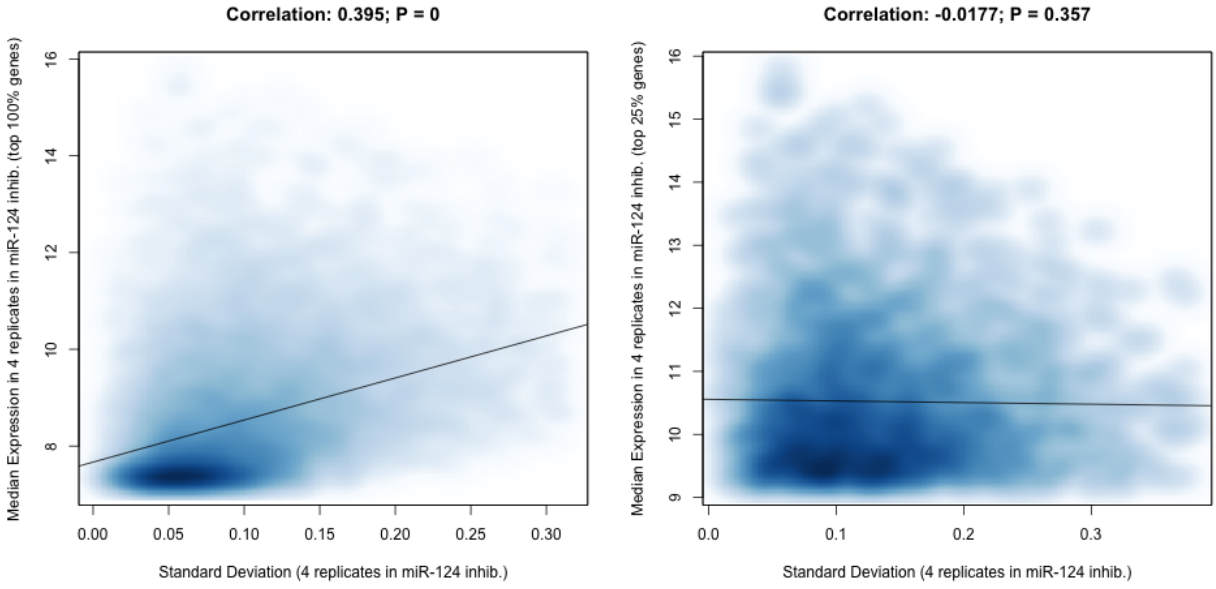
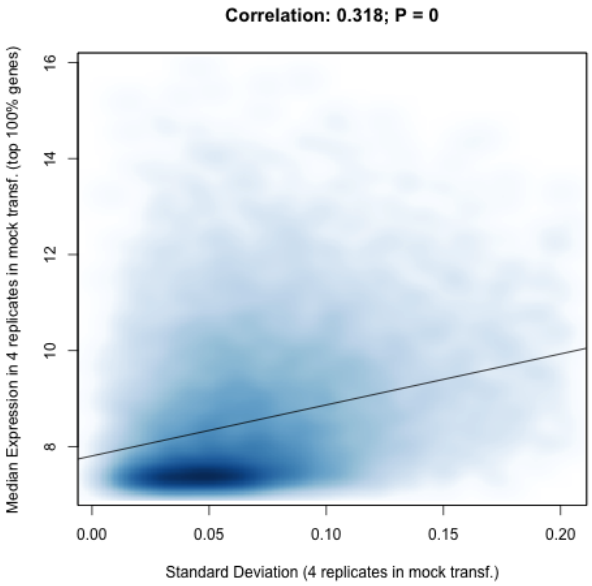
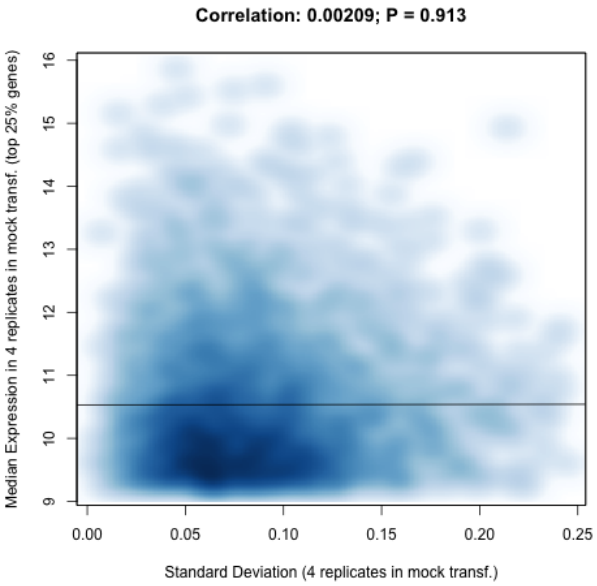


Figure A.10: Correlation of expression and standard deviation in miR-124 overexpression and inhibition.

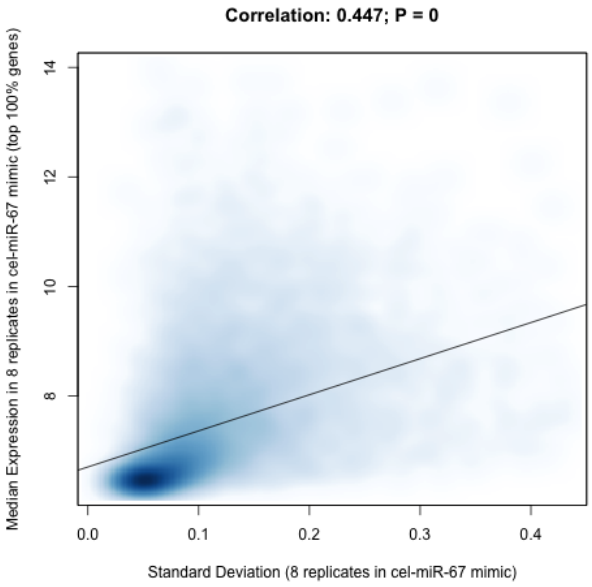
(a) mock transfection; 100% of genes



(b) mock transfection; top 25% of genes



(c) Overexpression cel-miR-67; 100% of genes



(d) Overexpression cel-miR-67; top 25% of genes

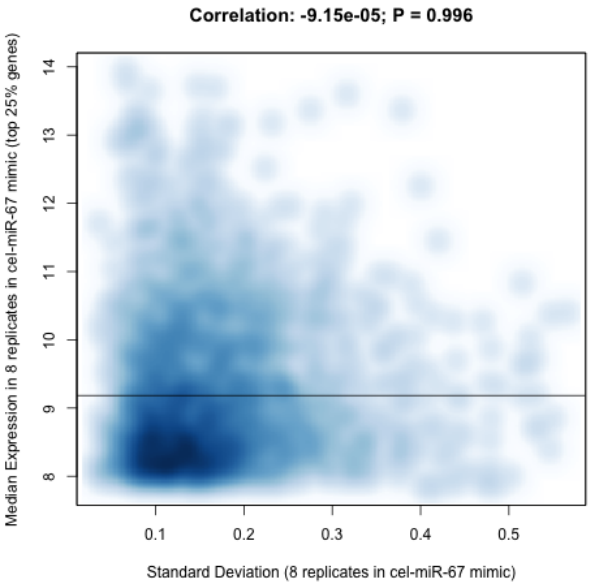


Figure A.11: Correlation of expression and standard deviation in mock transfection and cel-miR-67 overexpression.

Appendix B

Publications and presentations of this work

Publications:

- **S.A Manakov**, A. Morton, S.G Grant, A.J Enright. A neuronal transcriptome response involving stress pathways is buffered by neuronal microRNAs. *In preparation.*
- **S.A Manakov**, S.G Grant, A.J Enright. Reciprocal regulation of microRNA and mRNA profiles in neuronal development and synapse formation. *BMC Genomics* (2009) vol. 10 (1) pp. 419
- *Methods of this work were used in the following:*
  D Santhakumar, T Forster, N.N Laqtom, R Fragkoudis, P Dickinson, C Abreu-Goodger, **S.A Manakov**, N.R Choudhury, S.J Griffiths, A Vermeulen, A.J Enright, B Dutiae, A Kohle, P Ghazalb, A.H. Buck. Combined agonist–antagonist genome-wide functional screening identifies broadly active antiviral microRNAs. *Proceedings of the National Academy of Sciences* (2010) vol. 107 (31) pp. 13830

Presentations:

- **S.A Manakov**, A.J Enright, S.G Grant. Activity of evolutionary distinct classes of microRNAs in neuronal development. *The Society for Neuroscience annual meeting, San Diego, U.S.A.* (2010) Poster presentation
- **S.A Manakov**, S.G Grant, A.J Enright. Role of miRNAs in the genetic program of synaptogenesis. *The Society for Neuroscience annual meeting, Washington D.C., U.S.A.* (2008) Slide presentation