University of Natural Resources
and Applied Life Sciences, Vienna

Universität für Bodenkultur Wien
University of Natural Resources and Life Sciences, Vienna

Department of Biotechnology

Institute of Biotechnology

Master s thesis

microRNAs as tools for CHO cell engineering: In silico annotation of novel microRNAs from updated genomic data

Diendorfer Andreas Bernhard, BSc

Vienna, 2015

Acknowledgements

First and foremost I offer my sincere gratitude to Prof. Nicole Borth and Dr. Matthias Hackl, who both supported me throughout my thesis with their patience, knowledge and advice. The trust and encouragement I received allowed me to approach questions and problems in creative ways, giving me the opportunity to learn and understand science from a new point of view.

I would like to thank all members of our team, especially Vaibhav Jadhav, Gerald Klanert and Manuel Reithofer, for their ideas, suggestions and discussions about my thesis and other not so scientific topics. It was a pleasure working with so many enthusiastic scientists. For their continuous support and fuss-free help in cell culture, my special thanks to Vedrana Dizdarevic and Victoria Friesz.

The team of TAmiRNA allowed me to further study the ins and outs of bioinformatic NGS analysis. Thank you for this great and interesting opportunity!

Laura M. P. Taylor helped me with proofreading both this thesis and the published paper. Her prompt and detailed support is greatly appreciated.

Finally, I thank my family and friends for their patience and encouragement. A special *thank you* to Tamara Rumpold for her time and love.

“It s dangerous to go alone! Take this.”

— The Legend of Zelda, *Nintendo*

Abstract

microRNAs (miRNAs) are short, non-coding RNA molecules that play a powerful and critical role in post-transcriptional control of gene expression. These ~ 22 nucleotides long sequences detect mRNA and induce mRNA down-regulation by either cleavage or inhibition of translation. Given their in uence on a cells metabolic network, they are highly interesting tools in cell engineering.

Chinese hamster ovary (CHO) cell lines are the working horses of pharmaceutical biotechnology when it comes to the production of therapeutic biomolecules like antibodies or enzymes. Understanding the roles miRNAs play on the metabolic playground is a promising field when aiming to optimize the production of these important molecules. miRNAs provide us novel strategies to in uence growth, productivity, and other parameters.

The last years gave rise to a great amount of knowledge, but still leave many questions about miRNAs unanswered. With the advent of new methods like next-generation sequencing, this field can be explored on a new level. Genomic sequences allow the discovery of novel miRNAs with in silico experiments and miRNA-transcriptome sequencing provide the biological evidence of these findings.

The aim of this study was to identify novel miRNAs in CHO cells to further expand the list of possible miRNA-engineering targets. We used recently published genomic data of CHO or *Cricetulus griseus* (the originating species of CHO) cells and next-generation sequencing to annotate evolutionary conserved miRNAs. The identified miRNAs were added to miRBase, a centralized registry for miRNAs.

Zusammenfassung

MicroRNAs sind kurze regulatorische RNAs, die im zellulären Stoffwechsel eine wichtige Rolle spielen. Sie binden sequenzspezifisch mRNA und führen damit entweder zum Abbau der mRNA oder inhibieren die Proteinbiosynthese an den Ribosomen. Ihr Ein uss auf den komplexen Stoffwechsel macht sie zu einem interessanten Werkzeug und Forschungsobjekt in der Optimierung von biotechnologischen Prozessen.

CHO-Zellen (Chinese Hamster Ovary) sind die meistgenutzte tierische Zelllinie in der Produktion von therapeutischen Proteinen wie Antikörpern oder Enzymen. Den Ein uss und die Rolle von microRNA in diesem Produktionssystem zu verstehen, ist ein wichtiger Teil der Forschung zur Optimierung von CHO Bioprozessen.

Die letzten Jahre brachten viele neue Errungenschaften auf dem Gebiet der Sequenzierung. Kostengünstige und schnelle Techniken (next-generation sequencing) ermöglichten die Sequenzierung von CHO oder *Cricetulus griseus* (die Hamsterart, aus der die ursprünglich Zelllinie isoliert wurde) Genomen und eröffneten damit viele neue Möglichkeiten zur Erforschung von microRNAs.

Ziel dieser Studie war die Annotation von neuen microRNAs in CHO Zellen, basierend auf den kürzlich veröffentlichten CHO Genomen. Evolutionär konservierte microRNA Sequenzen aus anderen Spezies wurden in den Genomen gesucht und deren Expression mittels Next-generation Sequencing bestätigt. Die so gefundenen, bisher unbekannten, microRNAs wurden in miRBase, einer zentralen microRNA Datenbank, veröffentlicht.

What's it all about

The advantages in sequencing technologies and other genetic methods led to a great increase in possible targets and tools in the treatment of diseases. To test and exploit these findings, production systems are needed that are capable of building large bio-molecules (proteins, enzymes, oligo-nucleotides ...) which are possible new drugs. Cells provide an amazing set of tools to build and modify complex molecules. The use of them as *cell factories* is one of the major fields in biotechnology.

Cells carry blueprints for all proteins and enzymes in their genomic package, either as deoxyribonucleic acid (DNA) in the nucleus or as extra chromosomal DNA in plasmids, mitochondria, and chloroplasts. They follow the given instructions and build proteins according to the genetic code. In a process called protein bio-synthesis the DNA is first transcribed to ribonucleic acid (RNA) – cells don t want to mess with their only original copy of blue prints – and then translated to proteins, where the sequence of amino acids is defined in the genetic code.

Many cells then begin to further process the protein. These "post-translational modifications" include the addition of functional groups (e. g. glycosylation) or introducing structural changes (e. g. disulfide bonds) and are highly important for the biological function of the finished protein. While the production of proteins from a given DNA sequence can be performed in much simpler organisms like bacteria, these post-translational modifications are only performed in eukariotes. This is where mammalian cell culture comes into play. The use of closely related cells (to humans) in biotechnology allows the production of complex human (or human-like) molecules like antibodies and enzymes. The downsides of mammalian cells, in contrast to the long used bacterial or fungal systems, are the higher cultivation requirements (mammalian cells are not suited for a life outside an organism) and the significantly lower productivity.

To further expand the knowledge about the way cells produce molecules, a deeper insight in the molecular pathways, their regulations and networks is needed. This includes

one of the regulation mechanisms that effects the transcribed RNA: RNAi (RNA interference), where small RNA molecules like microRNAs (miRNAs) regulate the amount of RNA available for the protein bio-synthesis. The story of these players, their role in metabolic engineering and their discovery will be told in the following chapters.

For more insight and details on cells and what we think to know about them, please see *Molecular Biology of the Cell* by Alberts et al. (2008).

Contents

Abbreviations

Ago argonaute protein
cDNA complementary DNA
CHO Chinese hamster ovary
DHFR dihydrofolate reductase
DNA deoxyribonucleic acid
dsRBD double-stranded RNA-binding domain
dsRNA double-stranded RNA
EMA European Medicines Agency
FDA Food and Drug Administration
GS glutamin synthethase
LNA locked nucleic acid
miRISC miRNA-induced silencing complex
miRNA microRNA
mRNA messenger RNA
PCR polymerase chain reaction
piRNA piwi-interacting RNA
pri-miRNA primary miRNA
qRT-PCR quantitative reverse transcription PCR
RISC RNA-induced silencing complex
RNA ribonucleic acid
RNAi RNA interference
siRNA short interfering RNA
tPA human tissue plasminogen activator

1 CHO microRNA-omics

1.1 Chinese hamster ovary cells

Chinese hamster ovary (CHO) cells are important mammalian hosts for the production of biopharmaceuticals. The cell line was first isolated by Puck (1958) from ovary cells of *Cricetulus griseus* and has since developed to a successful and widely used production system. Multiple different cell lines were later derived from this ancestral cell line. CHO-K1, CHO-S and DG44, for example, are the sources of many recombinant protein producing CHO cell lines. Their genetic differences, caused by extensive mutagenesis and selection, were first shown by Wurm and Hacker (2011). The adoption of CHO in the biotech industry began with the first dihydrofolate reductase (DHFR) deficient cell lines. These DHFR mutations (one or both alleles are mutated or deleted) allowed the isolation of high copy number clones and thus a significantly higher productivity of cells. DXB11 (or DUKX) is a cell line derived from CHO-K1, with a deletion of one DHFR allele and an inactivating mutation in the second (Urlaub & Chasin, 1980). DG44, where both alleles are deleted, was isolated shortly after (Urlaub et al., 1983). The family tree of some important CHO cell lines is shown in Figure 1.1.

CHO cells were used as production system in the first approved mammalian derived therapeutic protein tPA (human tissue plasminogen activator) by Kaufman et al. (1985) and approved in 1987 (Wurm, 2004). Nowadays, more than 70% of all recombinant therapeutic proteins are produced using CHO cells as hosts (Jayapal & Wlaschin, 2007). CHO cells play an important role in the biopharmaceutical market and a selection of some CHO derived approved biopharmaceuticals is shown in Table 1.1.

As mammalian cells, CHO cells are capable of post-translational modifications to proteins that are both bioactive and compatible within humans. This allows the production of many biopharmaceutical molecules without any further modifications (Wurm, 2004), which is a safety aspect when producing glycoproteins for humans (Ghaderi et al., 2012). To overcome the downside of low specific production in most mammalian systems, CHO

Table 1.1: A selection of biopharmaceuticals produced in CHO cells, showing the wide application of them for different categories of agents (Datta et al., 2013).

Trade name	Generic name	Category	Biological Importance	FDA Approval
Zaltrap	Ziva ibercept	Rec. protein	Colon cancer	2012
Eylea	A ibercept	Rec. protein	AMD	2011
Actemra	Tocilizumab	Antibodies	Rheumatoid arthritis	2010
Recothrom	Thrombin α	Blood factors	Coagulation Factor	2008
Xyntha	Factor VIII	Blood factors	Hemophilia A	2008
Herceptin	Trastuzumab	Antibodies	Breast cancer	2008
VectibixTM	Panitumumab	Antibodies	Colorectal cancer	2006
Avastin	Bevacizumab	Antibodies	Colon or rectum cancer	2004
Amevive	Alefacept	Dimeric fusion protein	Chronic plaque psoriasis	2003
Xolair	Omalizumab	Antibodies	Asthma treatment	2003
Rebif	Interferon β-1a	Interferons	Multiple sclerosis	2002
Humira	Adalimumab	Antibodies	Human IgG1 mAB	2002
Aranesp	Darbepoetin α	EPO	EPO (anemia)	2001
Follistim	Follitropin β	Hormones	Infertility	1997
Rituxan	Rituximab	Antibodies	B-cell non-Hodgkins lymphoma	1997
Activase	Alteplase	Blood factors	Myocardial infarction	1987

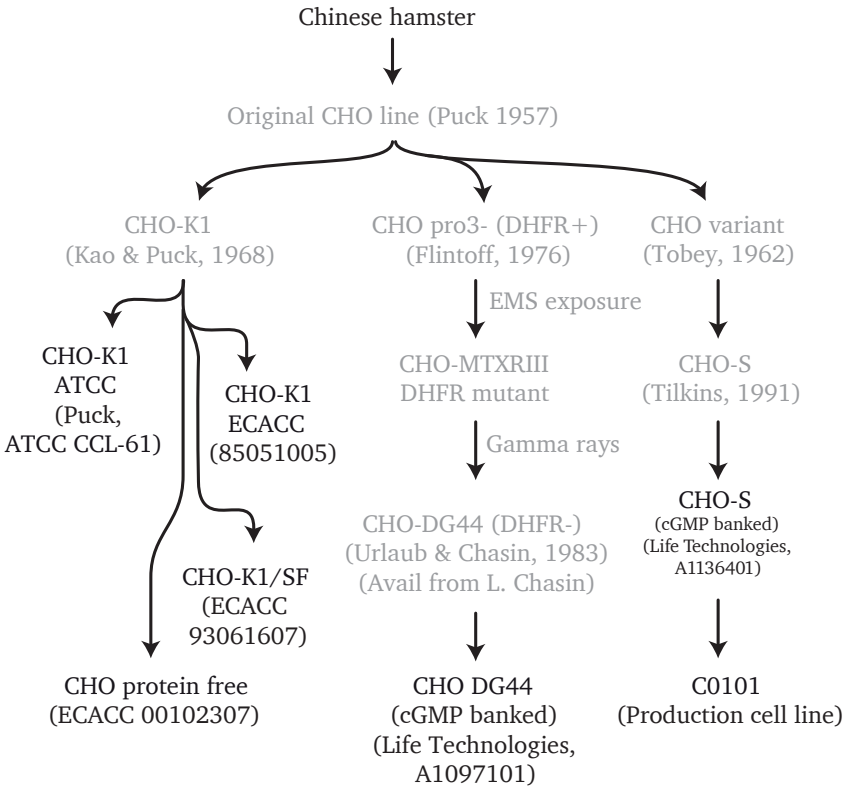

Figure 1.1: Family tree of some important CHO cell lines. If known, the name and year of isolation is given. CHO-K1, DG44 and CHO-S are now important cell lines, used in research and production. (Lewis et al., 2013)

cells can reach good levels when engaging methods such as DHFR or glutamin synthethase (GS) mediated gene amplification systems. In regulatory aspects, CHO cells have a long record as safe hosts and can be adapted to serum free growth, making approval by regulatory agencies like the FDA (Food and Drug Administration) or EMA (European Medicines Agency) easier (J. Y. Kim et al., 2012).

Considering the popularity and role of CHO cells in the production of biopharmaceuticals, the optimization of growth, product quality, and titer is a major goal. Over the last 25 years these parameters were mainly optimized by the modification of media, feeding strategies, and biotechnological processes. More recently, with the availability of genomic sequences, engineering strategies have emerged as an alternative route to improve cell line performance (Lim et al., 2010; Xiao et al., 2014), although deeper insight into genetic,

transcriptional, and translational regulation is required to obtain full control over cellular metabolism (Jadhav et al., 2013; Kildegaard et al., 2013).

1.2 microRNA

microRNAs (miRNAs) are small, non-coding ribonucleic acids (RNAs) that regulate the level of specific messenger RNA (mRNA) transcripts. They play a crucial role in the regulation of a cell s metabolism and are thus interesting targets for cell line engineering in biotechnological processes. miRNAs and other classes of small RNAs detect a target mRNA sequence in its 3 -UTR by Watson-Crick pairing. They guide macromolecules to the identified mRNA causing inhibition of translation or cleavage of the mRNA.

In 1993 the first miRNA *lin-4* was discovered in *Caenorhabditis elegans* and its regulatory effect on mRNA (through base-pairing in the 3 -UTR) was proposed (R. C. Lee et al., 1993; Wightman et al., 1993). This marked the starting point of miRNAs as a separate class of the previously discovered double-stranded RNAs (dsRNAs) as triggers of RNA interference (RNAi), a general concept of RNA regulating mRNA (Fire et al., 1998). Besides miRNAs there are other categories of small RNAs. Short interfering RNAs (siRNAs) differ in the biogenesis and silencing mechanism compared to miRNAs and were discovered as defense mechanisms against viruses in plants (Hamilton, 1999). Piwi-interacting RNAs (piRNAs) function mainly in the germline and appear to originate from single stranded precursors (Ishizu et al., 2012).

The miRBase registry (Griffiths-Jones, 2004) is a database organizing and listing known miRNAs. In its current version (release 21), there are more than 35,000 mature sequences annotated (all sequences and additional data can be downloaded from ftp://mirbase.org/pub/mirbase/21).

The following roundup of miRNA biogenesis and function (as outlined in Figure 1.2) will focus on animal miRNAs, although miRNAs are also found in plants, but with differences in their molecular pathways.

1.2.1 microRNA biogenesis

miRNAs are encoded in the genome, either as clusters of distinct mRNAs or together with proteins, where the miRNA sequence is placed on an intron (mirtrons). These sequences are

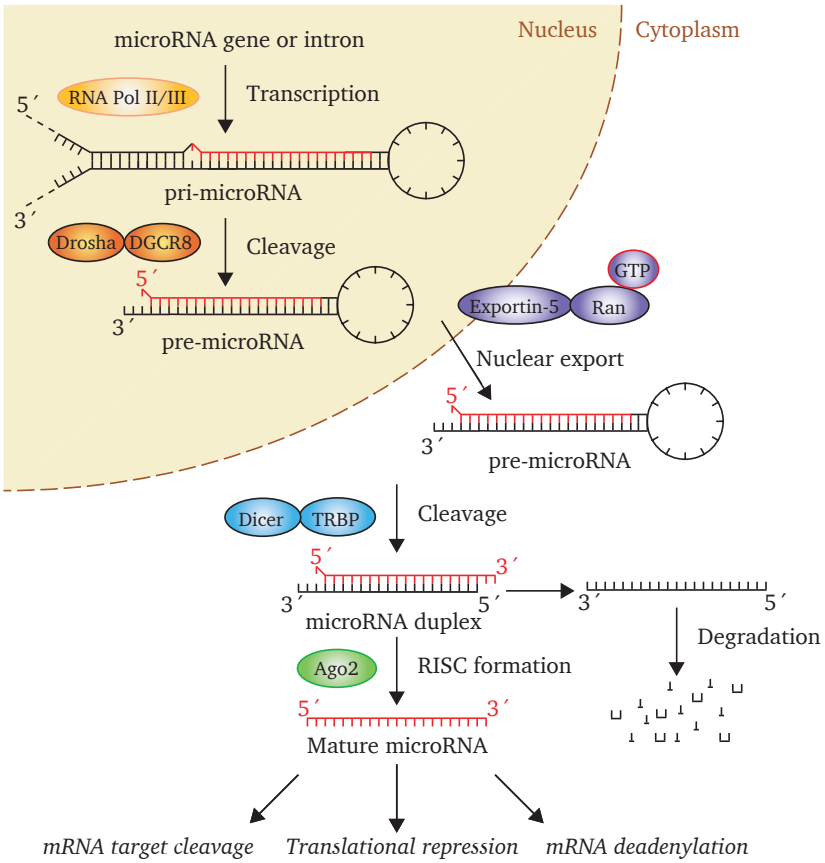

Figure 1.2: Biogenesis pathway of miRNAs from Winter et al. (2009) showing the steps from transcription to the mature miRNA associated with the RISC. miRNAs are encoded in the DNA either as genes or in introns of other proteins. They are transcribed by RNA Pol II and cleaved by Drosha to form the characteristic pre-miRNA stem-loop motif. After export of the pre-miRNA it gets cleaved again by Dicer. The mature miRNA is associated with the RISC and can then fulfill its biological function.

transcribed by RNA polymerase II (which also transcribes mRNA for proteins), polyadenylated, and capped (V. N. Kim, 2005) which results in a primary miRNA (pri-miRNA). Animal pri-miRNAs have a stem of about 33 bp (pairing with mismatches), a terminal loop, and flanking segments (Bartel, 2004). Two consecutive processing steps trim the pri-miRNA to the mature miRNA.

The first step is performed in the nucleus by *Drosha*, a member of the RNAse III family (V. N. Kim, 2005). Drosha depends on a protein cofactor containing two double-stranded RNA-binding domains (dsRBDs) that associates with the ribonuclease and forms the *Mi-*

croprocessor complex (Denli et al., 2004). Sequences originally transcribed together with mRNA take an alternative route, using splicing to form the structural features of pre-miRNA (Okamura et al., 2007; Ruby et al., 2007). An exemplary miRNA (miR-9), the sequence and the structure of the pre-miRNA is shown in Figure 1.3.

Following the first step, the pre-miRNA is transferred to the cytoplasm by Exportin 5 together with Ran, a GTPase moving RNA and proteins through nuclear pores (Yi et al., 2003). In a second processing step, the pre-miRNA is cut by *Dicer* to the final ~22 bp length mature miRNA. This ribonuclease III contains a PAZ (Piwi/Argonaute/Zwille) domain that determines the cleavage site (V. N. Kim, 2005). The double stranded mature miRNA-miRNA* duplex is then ready for association with the RNA-induced silencing complex (RISC).

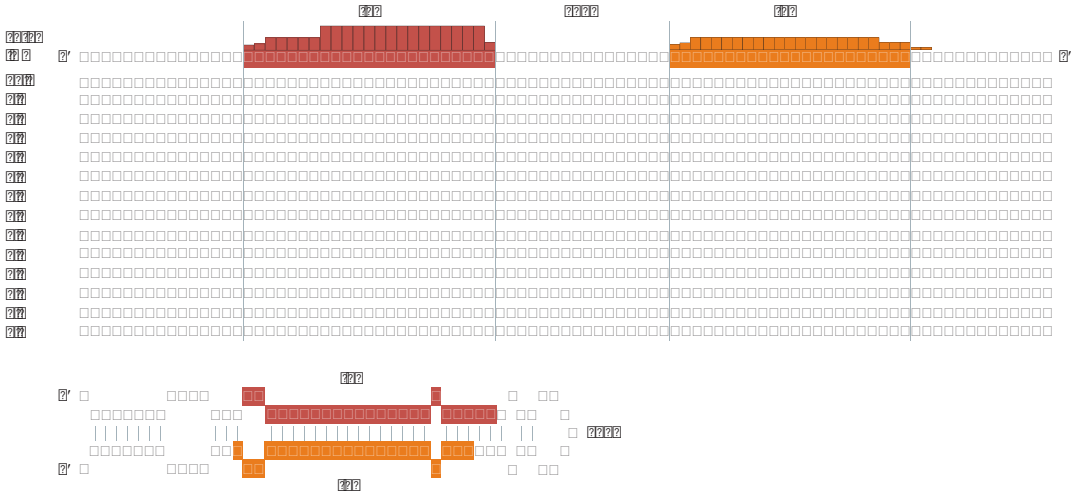

Figure 1.3: Heterogeneity of miRNA sequence and general structure of miRNA precursors (bottom) illustrated on the example of miR-9. The frequency of each variance is given on the left and the bar graph on top of the miRNA sequence shows the relative frequency of each nucleotide (Pritchard et al., 2012).

miRNA biogenesis is strongly regulated and a great number of miRNAs are regulated by the targets that they regulate. This form of a double-negative feedback loop indicates the high importance of the tight regulation of miRNA biogenesis. miRNAs binding off-target mRNAs will inevitable be down-regulated in this double-negative feedback loop. As the actual target (missing the negative regulation of the miRNA) will be kept at high levels, the miRNA biogenesis is repressed (Carthew & Sontheimer, 2009).

1.2.2 microRNA association

Following the double cleavage and export of the pri-miRNA, the mature miRNA is rapidly unwound and associates with an argonaute protein (Ago). Only one strand of the dsRNA will be incorporated and bound to Ago. Previous nomenclature termed this strand *miRNA* while the other strand was denoted as *miRNA**. When it was discovered that either strand can be active depending on cellular state (e. g. Yang et al., 2011), the nomenclature changed to *5p-miRNA* and *3p-miRNA* (Griffiths-Jones, 2011; Kozomara & Griffiths-Jones, 2014).

Together with multiple other proteins (e. g. Gemin3, Gemin4, Mov10, GW182), the complex of Dicer/Ago/miRNA builds the miRNA-induced silencing complex (miRISC). In this setup, the miRNA acts as an adaptor to specifically recognize and regulate mRNAs based on their sequence. The recognition is based on Watson-Crick base pairing between the miRNA and mRNA with possible mismatches and bulges. A so called *seed region* of the miRNA nucleotides 2–8 determines a great part of the specificity and binds without mismatches (Carthew & Sontheimer, 2009). A miRNA can detect mRNA in its 3 -UTR, but is not limited to only one mRNA. Each miRNA can find targets in multiple mRNAs and a mRNA can have multiple targets for miRNAs (Hobert, 2008). The miRISC complex is capable of either cleavage (degradation) of the detected mRNA or promotes repression of mRNA translation.

1.2.3 Post-transcriptional regulation by miRNAs

Perfect or near perfect complementary of miRNA-mRNA was considered to cause Ago-catalyzed cleavage preferentially to repression of mRNA translation. But discoveries by Brodersen et al. (2008) indicated that another fundamental factor must play a critical role in the decision of whether cleavage of repression occurs, which has yet to be found (Carthew & Sontheimer, 2009; Stroynowska-Czerwinska et al., 2014).

Repression of mRNA translation is thought to happen either at translation initiation or elongation, but the exact mechanism is a matter of current research. Repression at the initiation may be caused by impairment of cap recognition (Zdanowicz et al., 2009) or by blockage of the 60S ribosome subunit joining (B. Wang et al., 2008). Regarding inhibition of protein synthesis at the post-initiation level, there are three suggested mechanisms: (1) elongation blocking (Maroney et al., 2006), (2) ribosome drop-off (Petersen et al., 2006), and (3) proteolysis of the nascent protein (Olsen & Ambros, 1999). However, these three seem to be of rather limited significance (Stroynowska-Czerwinska et al., 2014).

In contrast to these mechanisms of mRNA translational down-regulation, there are also conditions in which miRNAs induce up-regulation. It was shown, that upon cell cycle arrest in G0/G1, AU-rich elements in TNFα together with Ago turn into a translation activation signal (Vasudevan et al., 2007). Another example would be miR-122, a liver specific miRNA that stimulates the translation of hepatitis C virus (HCV) caused by increased ribosome loading after miRNA binding (Henke et al., 2008).

1.2.4 mRNA degradation by miRNAs

mRNA degradation is a widespread effect of miRNA activity and is caused by mRNA deadenylation (Eulalio et al., 2009). The major effect of decreased protein activity is related to lower mRNA levels (Guo et al., 2010), making this miRNA in uence the most interesting. Ago and GW182, two proteins of the miRISC complex, play a crucial role in the deadenylation of mRNA. But in addition, other proteins are of importance. It was proposed, that PABP, Pan2-Pan3 and Ccr4-not mediate the shortening of the polyA tail (Wahle & Winkler, 2013). The deadenylation is linked to the general mRNA decay in cells (Behm-Ansmant et al., 2006).

Deadenylation may be followed by decapping of the mRNA transcript by Dcp1-Dcp2 and the exonuclease Xrn1 or by 3 –5 cytoplasmic exonucleases (Houseley & Tollervey, 2009).

Recent studies suggest, that mRNA repression is proceeded by deadenylation and mRNA decay (Béthune et al., 2012). Meijer et al. (2013) showed that translational inhibition by eIF4A2 is required for and followed by mRNA degradation.

1.3 CHO microRNA engineering

The recently gained knowledge about miRNAs, and especially their regulatory role on metabolic pathways, makes them an interesting target for cell engineering. The goals of these efforts are optimization of growth, viability and apoptosis, productivity, stability, and product quality. Engineering cell metabolism on the miRNA level gives the opportunity of controlling it on a networking level, as miRNAs not only regulate one target, but often multiple related proteins.

The availability of genomic data for CHO in the recent years (Brinkrolf et al., 2013; Lewis et al., 2013; Xu et al., 2011, see Table 1.2) made miRNAs available as tools in CHO

cell engineering. The current state of genomic data is available online at http://www.CHOgenome.org (Hammond et al., 2012). In addition, other tools and engineering strategies are now possible: (1) CHO-specific mRNA, (2) complementary DNA (cDNA) or whole genome microarrays, as well as (3) sequence alignment, and (4) primer design tools for gene cloning and site-specific gene knock-in or knock-out. The possibilities of cell line engineering with miRNAs were reviewed by Jadhav et al. (2013); the most crucial findings will be recapped in the following sections.

Table 1.2: Currently available genomic data sets for CHO or *C. griseus* (CGR). K1-BB is a currently unpublished genome assembly. Accession IDs refer to GenBank (http://www.ncbi.nlm.nih.gov/genbank/), where these data sets can be downloaded.

	Lewis et al. (2013)	Xu et al. (2011)	Brinkrolf et al. (2013)	K1-BB
Genome size	2.3 Gbp	2.3 Gbp	2.1 Gbp	2.98 Gbp
Scaffolds	7,468	109,151	28,749	11,400,490
x Coverage	89.1	130.0	70.0	17.1
Accession ID	AMDS01	AFTD01	APMK01	–
Source	CGR	CHO	CGR	CHO

1.3.1 Cell growth

The association of miRNAs with cell cycle arrest, cell proliferation, and cell death became clear, when miRNAs were linked to cancer phenotypes (e. g. Blenkiron et al., 2007). Further studies of differential expression gave rise to *oncomirs*, which regulate pathways related to cell cycle, cell growth, and cell death (Y. S. Lee & Dutta, 2006). There are two classes of miRNAs associated with cell growth: (1) oncomirs, which are oncogenic miRNAs and stimulate growth. A very prominent member of this category is miR-21, an oncogenic miRNA regulating multiple genes relevant for key pathways like JAK/STAT, MAPK, PPAR, and cell cycle related pathways (Frezzetti et al., 2011; Hatley et al., 2010). The other class are (2) tumor suppressor miRNAs, which down-regulate oncogenes such as let-7 family members, in uencing RAS (Johnson et al., 2005).

To facilitate the effects of miRNAs on cell growth, it has to be considered, that for an ideal bioprocess, growth is only needed and wanted at the beginning. After the initial growth phase, cells should utilize their resources for product creation and not for a further increase in bio mass. The oncogenic (growth-enhancing) and tumor suppressing (growth-

repressing) capability of miRNAs could both be of interest in these regards (Müller et al., 2008). Hernández Bort et al. (2012) found 10 miRNAs that are down-regulated during stationary phase, when looking for differentially expressed miRNAs during a batch culture of CHO-K1 cells. miR-7 was identified by Barron et al. (2011) as down-regulated during a temperature shift from 37 °C to 31 °C.

The relevance of miRNAs on growth was also shown by Hackl et al. (2014), when levels of Dicer (a main factor of miRNA biogenesis) were correlated to culture conditions. During exponential growth phase, Dicer levels were 3 fold higher in fast growing CHO cells ($\mu \sim 1.0\ d^{-1}$) than in slower growing ones ($\mu \sim 0.5\ d^{-1}$). In addition, the down-regulation of Dicer with siRNA lead to slower growing cells.

1.3.2 Apoptosis and cell viability

Apoptosis is a fundamental mechanism in cells (especially in tissue). The programmed cell death ensures that only healthy cells remain alive and thus ensures the survival of the organism. In modern mammalian bioprocesses, where high density growth is favored, this natural physiologic function is a problem (Müller et al., 2008). Conditions (e. g. nutrient limitation, shear or oxidative stress, osmolality and hypoxia) in bioreactors necessary for higher yields, better product quality or simply because of economic or process-control reasons, may lead cells to programmed cell death. The effects of apoptosis in an ongoing bioprocess are lowered product quality, byproduct and debris accumulation, and in general lower yields (Gammell, 2007). This makes apoptosis prevention an interesting and important objective in bioprocess engineering (Hacker et al., 2009; Jadhav et al., 2013).

Multiple miRNAs have been identified to be linked to enhancing cell life by inhibition of apoptosis (mainly in cancer cells). Members of the miR-15a/16 cluster were shown to induce programmed cell death (by targeting Bcl-xL and further activating Caspase-3/7) and reduce the viability of several cancers (Druz et al., 2013). Silencing of the previously mentioned miR-21 leads to increased apoptosis in glioblastoma cells (Chan et al., 2005; Si et al., 2007).

Regarding CHO cells, Druz et al. (2012) found members of the miR-297-669 cluster to be up-regulated in apoptopic conditions caused by nutrient depletion. miR-466h-5p could be linked to multiple anti-apoptopic genes (bcl2l2, dad1, birc6, stat5a, and smo) and was up-regulated when cells were limited in glucose supply.

Recently published CHO genome and miRNA-transcriptome data showed multiple miRNAs to be present in CHO that are linked to apoptopic processes (miR-1, let-7 family, miR-7b, miR-10a, miR-15/16 cluster, miR-93, miR-107, miR-144, miR-200b, miR-210, miR-214, miR-218, and miR-708). Together with miRNAs of the miR-297-699 cluster, these miRNAs are possible targets for apoptosis pathway engineering (Jadhav et al., 2013).

1.3.3 Energy metabolism

Metabolic balance is a crucial factor in a bioprocess. The maintenance of physiological conditions is controlled by a complex regulatory network. The first link between metabolic regulation and miRNAs was reported by Jopling (2012) where miR-122 was found to regulate genes from the cholesterol biosynthesis. Recent studies also suggest mitochondria as targets for metabolic engineering using miRNAs (Carrer et al., 2012; Sripada et al., 2012).

Another approach is the use of miRNAs levels as reporters that monitor biologically controlled bioprocesses or even to fine tune and modulate them (Jadhav et al., 2013).

1.3.4 Productivity and product quality

The control and regulation of cellular processes is a highly complex but also very crucial aspect in the production of biopharmaceutical molecules like antibodies. Although there are no studies yet showing methods of altering these metabolic pathways with miRNAs, multiple groups reported findings of miRNAs associated with productivity or product quality of cell lines. miR-221 and miR-222 were found to be down-regulated in cell lines producing recombinant human IgG, compared to the primordial DG44 cell line (Lin et al., 2011).

Glycosylation is a key feature of CHO cells, allowing them to be used in the production of pharmaceutical proteins with human or human-like glycosylation patterns. The possible consequences of wrong glycosylation patterns (e. g. immunologic reactions) makes them a carefully watched quality criterion. Serino et al. (2012) found that miR-148b modulates the expression of β-1,3-galactosyltransferase-1 (C1GALT1), which plays a critical role in O-glycosylation.

1.3.5 microRNA tools for cell engineering

The previous sections showed that many principal roles of miRNAs in cells, relevant to bioprocess engineering, could be identified. But as the function of a miRNA only affects mRNA, the presence and expression levels of mRNAs in a cell is a key factor when applying miRNAs as tools in bioprocess optimization. The state of the mRNA transcriptome differs highly, depending on cultivation conditions, cell types, and other factors (Shu et al., 2012).

Jadhav et al. (2013) summarized the following list of tools that are currently available for transient and stable overexpression or knockdown of miRNAs.

Antagomirs

Antagomirs (antagonistic miRNAs) are oligonucleotides designed to silence endogenous miRNAs in target cells. They have antisense sequences and down-regulate miRNAs by Watson-Crick binding. To further increase the target affinity, antagomirs often have chemical modifications like locked nucleic acid (LNA) and 2-O-methyl or 2-O-methoxyethyl (Jadhav et al., 2013; Krützfeldt et al., 2005).

They are easy and quick tools for transient silencing of miRNAs, but possible downsides are the need of an efficient delivery system, especially when applied in fed-batch or continuous bioprocesses. For long-term effects, the expensive synthetic antagomirs would have to be supplied in large amounts and with the media feed to avoid dilution and wash-out (Stein et al., 2010).

miRNA sponges

Another approach of miRNA silencing are *sponges*. These synthetic molecules provide a vast amount of miRNA binding sites (like in the mRNA the target miRNA regulates) and thus *absorb* all miRNAs, which now no longer effect mRNA transcripts (Ebert et al., 2007). The exible design of sponges allows them to target and silence not only one miRNA, but entire miRNA families.

Mimics and vector based expression

"miRNA gain-of-function" is achieved by either supplying double stranded synthetic miRNA equivalents (*mimics* as they mimic miRNA function on mRNA; Z. Wang, 2011) or by using the endogenous miRNA maturation pathway after producing miRNA precursors from plasmids. These artificially added miRNAs aim to have the same effect as the endogenous miRNAs and target mRNAs.

1.4 microRNA discovery and detection

1.4.1 Discovery of novel microRNAs

Given all the information and potential of miRNAs in bioprocess engineering, the question about the discovery of novel miRNAs remains. It is important to find novel miRNAs and target-mRNA, to further expand the possible field of application and to understand the role of miRNAs in the cellular system. Regarding CHO miRNA annotation, Hackl et al. (2012, 2011) added most of the available CHO-miRNAs in miRBase v21.

For the annotation of novel miRNAs, Ambros et al. (2003, p. 278) defined multiple criteria to avoid misannotation of miRNA-like structures (e. g. siRNAs or fragments of other RNAs):

Expression criteria

(a) Detection of a distinct 22-nt RNA transcript by hybridization to a size-fractionated RNA sample (ordinarily by the Northern blotting method).

(b) Identi cation of the 22-nt sequence in a library of cDNAs made from size-fractionated RNA. Such sequences must precisely match the genomic sequence of the organism from which they were cloned (except as noted below).

Biogenesis criteria

(c) Prediction of a potential fold-back precursor structure that contains the 22-nt miRNA sequence within one arm of the hairpin. In this criterion, the hairpin must be the folding alternative with the lowest free energy, as predicted by mfold or another conventional RNA-folding program, and must include at least 16 bp involving the rst 22 nt of the miRNA and the other arm of the hairpin. It should not contain large internal loops or

bulges, particularly not large asymmetric bulges. In animals, these fold-back precursors are usually about 60—80 nt, whereas in plants, they are more variable, and may include up to a few hundred nucleotides.

(d) Phylogenetic conservation of the 22-nt miRNA sequence and its predicted fold-back precursor secondary structure. The conserved hairpin should meet the same minimal pairing requirements as in criterion C, but need not be the lowest free energy folding alternative.

(e) Detection of increased precursor accumulation in organisms with reduced Dicer function.

To annotate a novel sequence as miRNA, a combination of these criteria has to be fulfilled.

Recently, miRNAs are mostly annotated by the use of computational methods. This was made possible by the great improvements in sequencing technologies, providing accurate and fast genomic data. In these approaches, miRNA-precursors are predicted using various bioinformatic tools (e. g. machine learning algorithms for the recognition of stem-loop Dicer substrates). All these in silico applications involve the risk of random hits, making it essential to confirm miRNA annotation candidates with biological evidence from miRNA detection experiments (Kozomara & Griffiths-Jones, 2014).

1.4.2 Detection methods for microRNAs

The detection of miRNAs is challenging due to multiple reasons: (1) their small size, (2) sequence similarities among each other (*isomirs* which sometimes differ as little as a single nucleotide), (3) lack of a common sequence (such as a polyA tail), (4) low levels ($\sim$0.01 % w of total RNA), and (5) tissue or cell cycle/process stage-specific expression (Planell-Saguer & Rodicio, 2013). The predominant and interesting tools for miRNA detection will be presented in the following paragraphs, highlighting their strengths and limitations (Pritchard et al., 2012; Rooij, 2011).

Profiling and differential expression analysis of hundreds of miRNAs is especially challenging, as the short length and variance in GC content leads to a wide range of melting temperatures T_m (Pritchard et al., 2012).

An overview and the basic principles of the following methods is given in Figure 1.4.

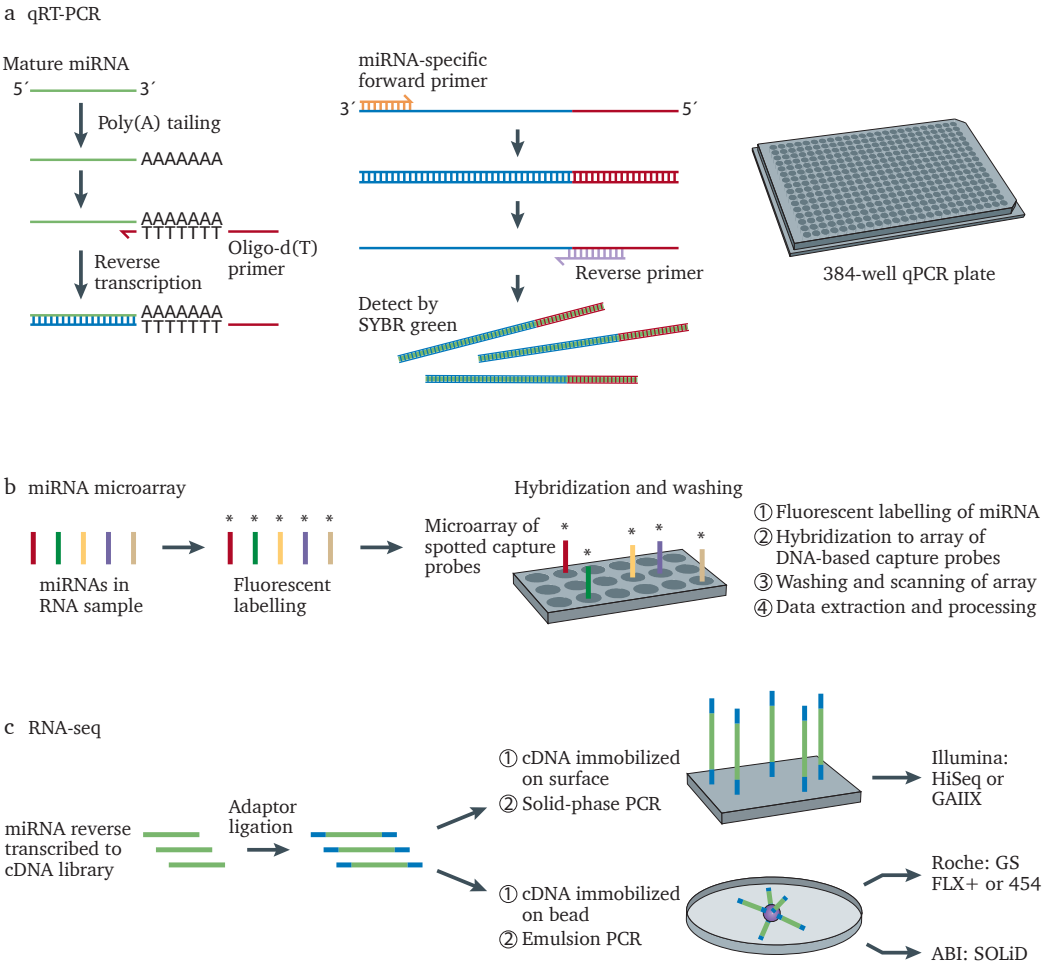

Figure 1.4: Commonly used miRNA detection methods: (a) Quantitative reverse transcription PCR (qRT-PCR), multiplying a specific miRNA with primers and detection of SYBR green signals over the course of multiple replication cycles, allowing the calculation of fold changes in contrast to a standard. (b) miRNA microarrays, a hybridization based method, where target probes are spotted on a glass surface, capturing labeled miRNAs on defined locations on the array. (c) RNA-seq or next-generation sequencing, allows the analysis of the whole miRNA-transcriptome, including novel miRNAs. (Adapted from Pritchard et al., 2012)

Quantitative reverse transcription PCR

As the multiplication reaction in a polymerase chain reaction (PCR) is performed by DNA-polymerase, a PCR for RNA starts with the *reverse transcription* of RNA to cDNA. During

the following *real time* PCR, the amount of synthesized DNA is monitored with the help of a deoxyribonucleic acid (DNA) binding molecule like SYBR green.

qRT-PCR can be performed in every laboratory equipped for real-time PCR, making this method a good choice for small scale miRNA analysis. For the measurement of multiple miRNAs, the qRT-PCR protocol can be optimized by the use of pre-plated PCR primers.

To overcome the problem of different temperatures (T_M) of each miRNA for primer annealing, special primers are available from different vendors, having LNAs incorporated to normalize T_M (Pritchard et al., 2012).

miRNA microarrays

For the parallel analysis of multiple miRNAs, microarrays were one of the first used methods. A glass carrier is spotted with probes for the miRNAs of interest in a grid pattern. Each spot contains probes for only one miRNA and the position of every spot and its probe content is known. miRNAs get labeled with a uorophore and are captured by their specific probes. This results in a miRNA concentration dependent signal on the spot, which can be measured using a microarray reader (basically a scanner that measures the uorophores intensity at each spot).

Labeling of the miRNAs is often done using T4 RNA ligase. To avoid the risk of circularization (miRNAs have a 5 phosphate from Dicer cleavage which could cause an intramolecular ligation reaction), most protocols involve an initial dephosphorylation step. The labeling process in general is prone to substrate sequence bias, which means that some sequences have a higher affinity to the labeling process (the T4 RNA ligase), causing a higher signal which does not correlate to the miRNA level (Nelson et al., 2008).

As with qRT-PCR, variance of T_M can be reduced by the use of LNA modified capture probes. These specificity problems do not make microarrays an ideal tool for absolute measurements, but still a very valuable player in the often (especially in clinical research) performed differential expression (where relative differences in miRNA expression between multiple states are evaluated) analysis (Pritchard et al., 2012).

RNA-seq

Next-generation sequencing paved the way of novel miRNA detection methods. After RNA extraction and purification, a small RNA cDNA library is generated. This library is then completely sequenced, meaning that every molecule should get sequenced (*massively parallel sequencing*).

With this approach, not only can known miRNAs be analyzed (hybridization or primer based methods both require the knowledge of the sequence in question), but also (putative) novel miRNAs can be found. Ambros et al. (2003) and Ruby et al. (2006) established criteria for the annotation of small RNA as miRNAs (see listing in subsection 1.4.1).

In general, the number of reads per sequence (*read count*) allows the estimation of the relative abundance of a miRNA. Data normalization (making read counts of different experiments comparable) and quality control (e. g. purging of low quality reads) is an important step in the bioinformatic processing of RNA-seq experiments (Pritchard et al., 2012).

A remaining problem is the sequence-specific bias, introduced by enzymatic steps in the small RNA library preparation. This leads to a higher or lower abundance of some miRNAs, not representing the ratio in the RNA sample (Linsen et al., 2009; Tian et al., 2010).

RNA-seq is still a very expensive method, but with techniques like *barcoding* (allowing to multiplex many samples in one run) and newer instruments, it is becoming an interesting alternative to the well established hybridization or PCR based methods.

Encoded hydrogel particles

A novel approach of qualitative and quantitative miRNA analysis was recently introduced by Fire y Bioworks (outlined in Figure 1.5). Hydrogel particles, produced with optical liquid stamping, are equipped with a miRNA binding site at the center of each particle, specific for one miRNA. The miRNA binds to the particle probe together with an universal adapter. Probes with miRNAs associated are further modified to bind a uorescence reporter. After the final step, only probes that had a miRNA bound now have a reporter, resulting in a miRNA dependent uorescence signal.

Two barcode regions on the edges of each particle have predefined uorescence signals on other wavelengths. The signal intensity of each barcode region allows the identification

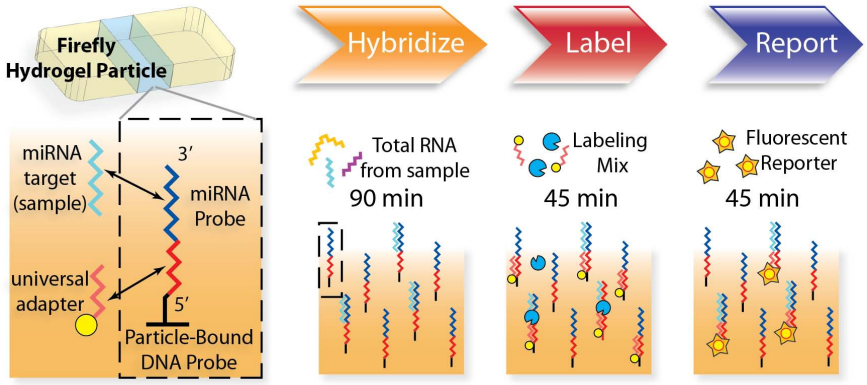

Figure 1.5: Fire y microRNA Assay: Barcoded hydrogel particles are used for the quantification of miRNA. The uorescence reporter areas (two on the ends of each particle acting as barcode and one in the center with miRNA depended signal levels) is read with a ow cytometer together with the barcode which is used for the identification of the particle and thus the associated miRNA. (Fire y Bioworks, 2014)

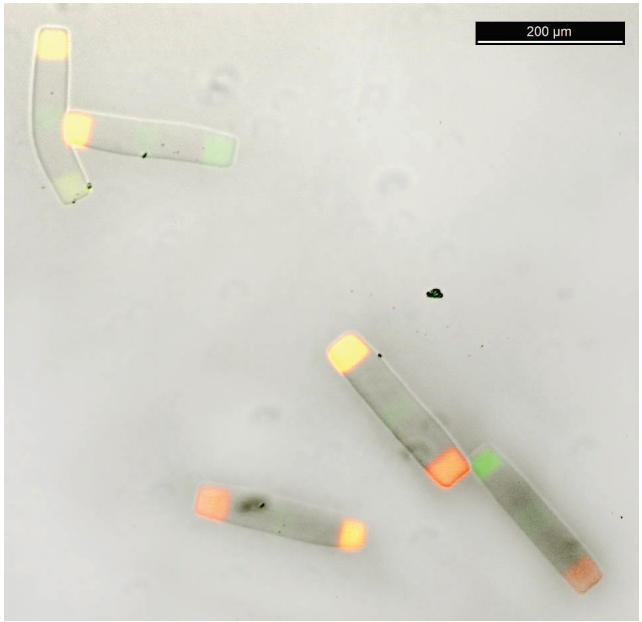

Figure 1.6: Fluorescence microscopic image of labeled Fire y hydrogel particles. The two barcode areas at the ends are clearly visible while the central miRNA binding site gives only a slight signal on these particles. [1]

[1]Picture acquired with technical assistance by Theresa Böck.

of the particle and thus the association of particle and miRNA. A microscopic image of labeled particles is given in Figure 1.6 showing the three distinct uorescence regions on each particle.

Particles are read using a ow cytometer. They get aligned with the sheath uid and read in sequence: barcode area 1, probe area, barcode area 2. The orientation of the particle is random, but does not matter as the software will take this into account.

Analysis of the generated data is performed using a software tool from Fire y Bioworks. It calculates the amount of miRNA based on the barcode and probe signal, allowing a relative quantification of up to 96 miRNAs in parallel (Chapin et al., 2011).

Although this method seems promising, the accuracy depends on the ow cytometer model used and does not work on all devices. It is aimed to provide a quick and easy protocol for clinical laboratories, as they often already have ow cytometers for cellular analysis and could use it to profile miRNAs in patient samples as tools for biomarkers.

2 Annotation of additional evolutionary conserved microRNAs in CHO cells from updated genomic data

The manuscript was submitted to Biotechnology and Bioengineering on July 28th 2014, a revision was provided on November 10th 2014, was accepted on January 7th 2015, and published online on April 8th 2015 (DOI: 10.1002/bit.25539).

COMMUNICATION TO THE EDITOR

BIOTECHNOLOGY and BIOENGINEERING

Annotation of Additional Evolutionary Conserved microRNAs in CHO Cells From Updated Genomic Data

Andreas B. Diendorfer,[1] Matthias Hackl,[1] Gerald Klanert,[2] Vaibhav Jadhav,[1] Manuel Reithofer,[1] Fabian Stiefel,[3] Friedemann Hesse,[3] Johannes Grillari,[1] Nicole Borth[1,2]

[1]Department of Biotechnology, BOKU University of Natural Resources and Life Sciences, Vienna, Austria; telephone: +43476546803; fax: +43476546675; e-mail: andreas.diendorfer@boku.ac.at

[2]ACIB GmbH, Austrian Centre of Industrial Biotechnology, Graz, Austria

[3]Biberach University of Applied Sciences, Biberach, Germany

ABSTRACT: MicroRNAs are small non-coding RNAs that play a critical role in post-transcriptional control of gene expression. Recent publications of genomic sequencing data from the Chinese Hamster (CGR) and Chinese hamster ovary (CHO) cells provide new tools for the discovery of novel miRNAs in this important production system. Version 20 of the miRNA registry miRBase contains 307 mature miRNAs and 200 precursor sequences for CGR/CHO. We searched for evolutionary conserved miRNAs from miRBase v20 in recently published genomic data, derived from Chinese hamster and CHO cells, to further extend the list of known miRNAs. With our approach we could identify several hundred miRNA sequences in the genome. For several of these, the expression in CHO cells could be verified from multiple next-generation sequencing experiments. In addition, several hundred unexpressed miRNAs are awaiting further confirmation by testing for their transcription in different Chinese hamster tissues.

Biotechnol. Bioeng. 2015;9999: 1–7.

© 2015 The Authors. Biotechnology and Bioengineering Published by Wiley Periodicals, Inc.

KEYWORDS: microRNA; Chinese hamster ovary cell; next-generation sequencing

Introduction

Chinese hamster ovary (CHO) cells are important mammalian hosts for the production of biopharmaceuticals. Over the last 25 years, the optimization of growth, product quality and titer was mainly driven by the modification of media, feeding strategies, and biotechnological processes. More recently, with the availability of genomic sequences, cell engineering strategies have emerged as an alternative route to improve cell line performance (Lim et al., 2010; Xiao et al., 2014), although deeper insight into genetic, transcriptional and translational regulation is required to obtain full control over cellular metabolism (Jadhav et al., 2013; Kildegaard et al., 2013).

MicroRNAs (miRNAs) are small (18–24 nucleotides long) non-coding RNAs (ncRNA) that are transcribed as a primary transcript (pri-miRNA) by RNA polymerase II. These transcripts are cleaved in the nucleus by the RNAase Drosha, to produce a pre-miRNA with a characteristic stem-loop secondary structure and a size of about 50–80 nt. For the final miRNA sequence, pre-miRNAs are processed by the enzyme Dicer and incorporated into the miRNA-induced silencing complex (miRISC) and bind to mRNA transcripts in their 3′-UTR. Binding of the complex triggers either inhibition of translation or mRNA cleavage. By this process, miRNAs are capable of negatively regulating protein translation (Gregory et al., 2004) and therefore constitute an important layer in the post-transcriptional control of gene expression (Hobert, 2008).

In 2011, the first systematic approaches to miRNA annotation in CHO cells were reported (Hackl et al., 2011; Johnson et al., 2011). Subsequent studies were focused on providing insights into the importance of miRNAs for molecular pathways relevant to cell culture engineering, such as growth (Jadhav et al., 2014), apoptosis (Druz et al., 2011), and protein secretion (Barron et al., 2011; Loh et al., 2014).

Recent publications of genomic sequencing data from CHO cells (Xu et al., 2011) as well as that of *Cricetulus griseus* (CGR) (Brinkrolf et al., 2013; Lewis et al., 2013) now have provided a basis for refining the annotation of: (i) known expressed cgr-miRs, (ii) the identification of additional expressed miRNAs in CHO cells, and (iii) the discovery of miRNAs without current evidence for transcription in CHO cell lines.

For the annotation of novel microRNAs, (Ambros et al. 2003) specified five conditions, of which a reasonable combination has to

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Correspondence to: A. B. Diendorfer and N. Borth
Contract grant sponsor: FWF, Austrian Science Fund
Contract grant number: W1224 Biotop
Contract grant sponsor: Austrian Center of Industrial Biotechnology
Received 28 July 2014; Revision received 10 November 2014; Accepted 7 January 2015
Accepted manuscript online xx Month 2015;
Article first published online in Wiley Online Library (wileyonlinelibrary.com).
DOI 10.1002/bit.25539

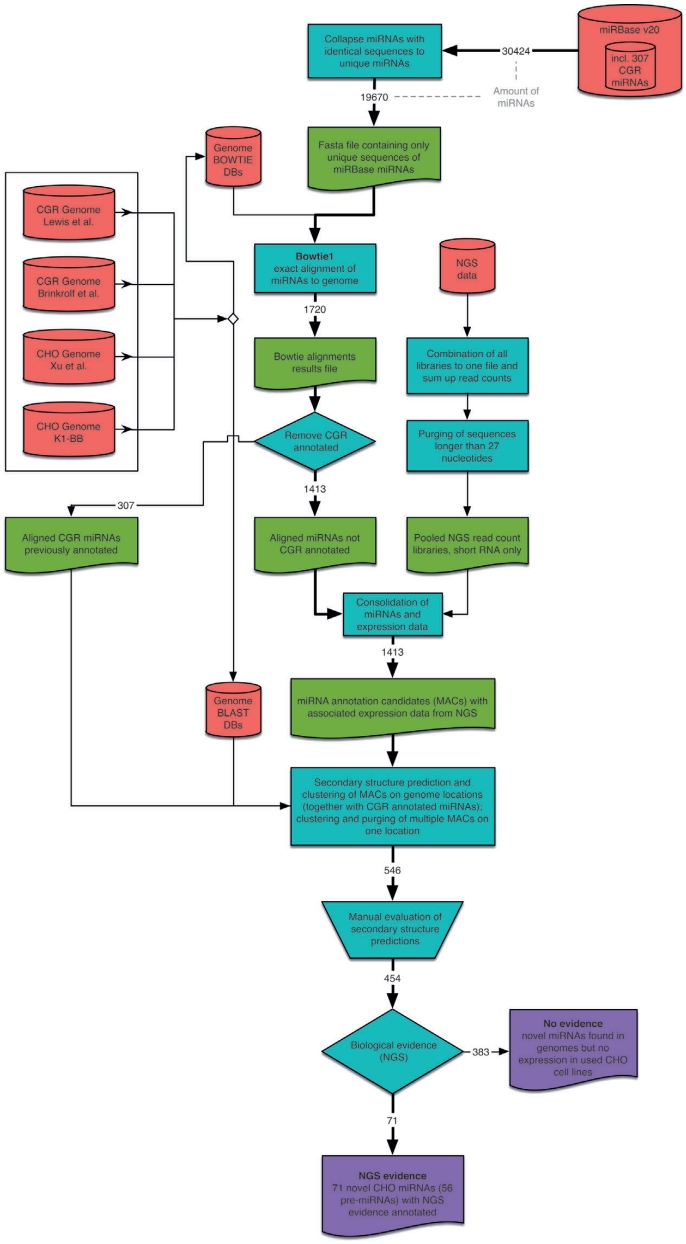

Figure 1. Bioinformatic pipeline flowchart. The main path of analysis is highlighted with a bold arrow from the top right to the bottom. Numbers on the arrows describe the amount of miRNAs processed in this step.

be met to count as valid microRNA. In this study we based our criteria on the identification of a distinct ~22 nt RNA transcript and the phylogenetic conservation of mature and pre-miRNA sequences.

To expand the list of miRNAs available as possible engineering tools, we searched for evolutionary conserved miRNA sequences from other species in four genomic datasets, identified their genomic locations and hairpin sequences, and confirmed the

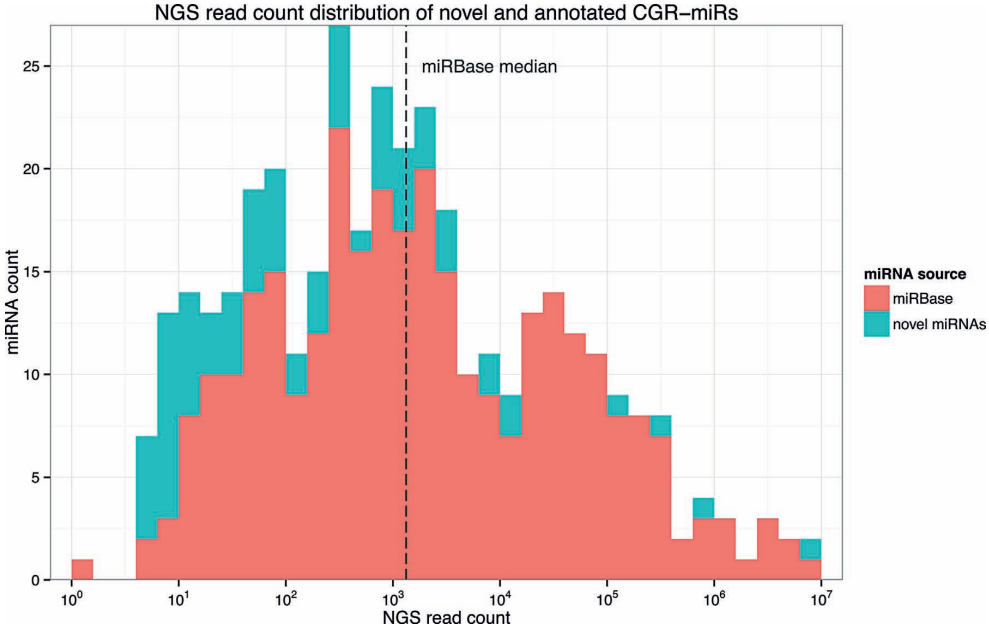

Figure 2. Read count distribution of miRBase annotated CGR and novel microRNAs showing a higher abundance of low read count microRNAs in the novel data set (77.8 % of novel miRNAs are below the miRBase median read count).

expression of some of these in CHO, using next-generation transcriptome sequencing results. This improvement and expansion of sequence and expression information of cgr-miRs will be useful for further functional investigation of miRNAs, to gain a better understanding of post-transcriptional regulation in cellular pathways, and to explore the potential function of silent CHO miRNAs, for which no expression evidence could be found yet.

miRBase Version 20 (Griffiths-Jones, 2004) contains 307 mature cgr-miRs on 200 precursors. In order to identify and annotate novel cgr-miRs, we based our search on evolutionary conservation of miRNAs by aligning all mature miRNAs to the available CGR/CHO genomic data, applying the workflow outlined in Figure 1.

Thereby, 1,720 unique mature sequences out of 19,670 miRBase v20 entries (all miRBase annotated miRNAs of any species) were aligned to at least one of the four genomes. The 307 already annotated cgr-miRNAs were excluded, leaving 1,413 miRNA sequences as candidates for novel cgr-miRNAs. Clustering of these sequences on genomic locations, to find similar, overlapping sequences that constitute only one possible new miRNA, reduced the number of sequences down to 546. These 546 candidates were

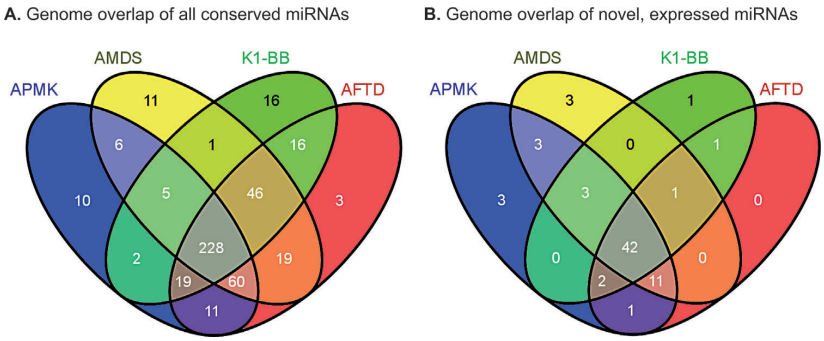

Figure 3. Venn diagrams showing the occurance of (**A**) all evolutionary conserved and (**B**) expressed, novel miRNAs in the four used genomic datasets.

then evaluated by prediction of their in silico secondary structure, whereby 454 sequences showed a pre-miRNA like secondary stem loop structure. These putative novel cgr-miRs are listed in Supplement 1.

For classification as novel cgr-miRs, the next-generation sequencing (NGS) read counts from new and existing datasets (Hackl et al., 2011) of these sequences were examined using more than five read counts as cut-off for the existence of these miRNAs. Thereby, 383 sequences did not show expression under the constraints for valid NGS signals, leaving 71 NGS confirmed miRNAs.

Four of these 71 miRNAs were present on two genomic locations (75 possible novel miRNA locations) with different hairpin sequences (highlighted in blue in Supplement 1). A set of six pairs (miRNA-5p/miRNA-3p) could be matched on the same hairpin, (highlighted in green), giving 69 pre-miRNA sequences. Thirteen of the new mature miRNA sequences were found on already annotated cgr-pre-miRNAs (see column "On hairpin with" in Supplement 1), thus complementing already annotated 5p or 3p miRNAs. Therefore, in summary, our study resulted in a total of 56 novel pre-miRNA and 71 miRNA sequences that were added to the already annotated ones, extending the hamster miRNome to a total of 378 mature sequences (+23.1%), and 256 precursors (+28.0%).

The described process presents an easy and fast pipeline for the discovery of novel miRNAs from genomic data and next-generation sequencing experiments. Recently published genomic data for *C. griseus* allowed the annotation of multiple conserved miRNAs from other species, by definition missing out on possible non-conserved miRNAs present in the Chinese hamster.

In total, 71 novel expressed mature miRNAs and 56 pre-miRNAs were added to the existing data. These novel miRNAs are mostly expressed at low levels, as their read counts were identified to be mainly below the median read count (77.8% below and 22.2% above) of already annotated cgr-miRNAs (Fig. 2). In addition, we provide information on genomic loci of 345 mature miRNAs with no evidence for expression in CHO cells, but conserved homologous sequences in related species. These may be expressed in other tissue and cell types in the Chinese hamster, however , as confirmation of expression is required, they are currently not uploaded to miRBase. The occurrence of both the evolutionary conserved and expressed newly identified miRNAs in the four genomic datasets is shown in Figure 3. In both cases, the majority of miRNAs can be found in three or all genomic datasets.

A list of validated targets for these novel miRNAs (if available) is provided in Supplement 2. We grouped the pathways into cell engineering relevant categories. A list of pathways and miRNAs that possibly influence them is given in Table I. As these include process relevant cellular properties such as growth and apoptosis, the miRNAs may be relevant engineering targets.

Taken together, the cgr-miRNAs identified here will enhance the use of miRNAs as tools for CHO cell engineering. They also pave the way for using miRNAs not transcribed in CHO cells to restore or alter biotechnologically relevant cell line characteristics using the correct hamster primary miRNA sequences for engineering. The positive effects of using endogenous in contrast to chimeric miRNA sequences was shown by Klanert et al. (2014) and highlights the importance of good and comprehensive genomic data for CHO cell engineering.

Material and Methods

Datasets

For the annotation of novel miRNAs, the mature sequences of all miRNAs available in the miRNA database (miRBase version 20 - http://www.mirbase.org; Griffiths-Jones, 2004) were used. This constitutes a set of 30,424 mature miRNAs which was merged down to 19,670 unique sequences.

These miRNAs were aligned to four genome assemblies, sequenced from CHO cells and *C. griseus* (CGR). The genomes were downloaded from GenBank (http://www.ncbi.nlm.nih.gov/genbank/) and are listed with their accession IDs in Table II. In addition, the unpublished genome assembly "K1-BB" was used as described by Hackl et al. (2012).

Next-generation sequencing (NGS) data were obtained by Illumina sequencing as described by Hackl et al. (2011) and is available at the Sequence Read Archive SRA (www.ncbi.nlm.nih.gov/sra/), accession number SRA024456.1. Cell lines sequenced included CHO-DUXB11 (ATCC CRL-9096), CHO-K1 (ECACC-CCL-61), and two derivative recombinant cell lines producing a monoclonal antibody (CHO-K1) and an EpoFc fusion protein (DUXB11). Samples were taken both from cells grown in the presence of 5% FCS and after adaptation to protein free medium. To further extend the amount of cultivation conditions and cell lines we sequenced additional cell lines. CHO-DG44 (passage 13) cells were grown in a 2 L bioreactor in batch culture. Total RNA was extracted with Trizol reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's recommendations. The RNA quality was tested with an Agilent chip (Agilent Technologies, Santa Clara, California). Labelling of the samples followed the standard Illumina TrueSeq protocol for small RNA (12 barcoded samples on one lane). Sequencing of ready to load libraries was conducted by GATC, Germany on an Illumina HiSeq Analyser. The NGS data files of these experiments are available at Gene Expression Omnibus GEO (http://www.ncbi.nlm.nih.gov/geo/), accession number GSE59838 and SRA, accession number SRP044946.

Bioinformatic Processing

Bioinformatic processing followed a pipeline of freely available tools and self-written scripts. The scripts are available upon request from the corresponding author. The pipeline is outlined in Figure 1.

The collapsed mature miRNA sequences from miRBase were aligned against each of the four genome databases using Bowtie v1.0.0 (Langmead et al., 2009) to find exact matches. This produced a list of annotated miRNAs that were present in the genomic sequences (see Supplement 1). MicroRNAs already annotated for CGR were filtered and stored in a separate file. For all other miRNAs that could be aligned to the genomes, the read counts from next-generation sequencing experiments were consolidated into one file. Next-generation sequencing read counts were summed up from different experiments to consider the different culture conditions. A read count of five reads was set as cutoff.

For these miRNA annotation candidates, the up and downstream sequences were extracted in the same length as in the originating species (assuming that not only the mature sequence is conserved,

Table I. Selected pathways possibly influenced by novel miRNAs. Data for miRNA protein interaction was derived from miRWalk (validated targets). DAVID was used to associate the proteins with pathways.

Pathway category	miRNAs
Cell cycle related	mmu-let-7a-1–3p, mmu-miR-106b-5p, mmu-miR-126–5p, mmu-miR-127–3p, mmu-miR-202–5p, mmu-miR-211–5p, mmu-miR-216b-5p, mmu-miR-24–1-5p, mmu-miR-760–3p, hsa-let-7a-3p, hsa-let-7c-5p, hsa-let-7e-3p, hsa-let-7e-5p, hsa-miR-148a-3p, hsa-miR-18b-5p, hsa-miR-192–3p, hsa-miR-20b-5p, hsa-miR-217, hsa-miR-26a-1–3p, hsa-miR-26a-2–3p, hsa-miR-449a, hsa-miR-582–5p
Apoptosis related	mmu-let-7a-1–3p, mmu-miR-106b-5p, mmu-miR-126–5p, mmu-miR-150–5p, mmu-miR-30c-1–3p, hsa-let-7a-3p, hsa-let-7c-5p, hsa-let-7e-3p, hsa-let-7e-5p, hsa-miR-192–3p, hsa-miR-30c-2–3p, hsa-miR-449a, hsa-miR-451a
Cancer related	mmu-let-7a-1–3p, mmu-miR-101a-5p, mmu-miR-106b-5p, mmu-miR-126–5p, mmu-miR-127–3p, mmu-miR-150–5p, mmu-miR-193a-5p, mmu-miR-193a-3p, mmu-miR-211–5p, mmu-miR-216b-5p, mmu-miR-24–1-5p, mmu-miR-296–5p, mmu-miR-30c-1–3p, mmu-miR-760–3p, mmu-miR-99b-3p, mmu-miR-99b-5p, hsa-let-7a-3p, hsa-let-7c-5p, hsa-let-7e-3p, hsa-let-7e-5p, hsa-miR-148a-3p, hsa-miR-18b-5p, hsa-miR-192–3p, hsa-miR-20b-5p, hsa-miR-217, hsa-miR-26a-1–3p, hsa-miR-26a-2–3p, hsa-miR-301b-3p, hsa-miR-30c-2–3p, hsa-miR-449a, hsa-miR-451a, hsa-miR-582–5p
Focal adhesion related	mmu-let-7a-1–3p, mmu-miR-101a-5p, mmu-miR-126–5p, mmu-miR-150–5p, mmu-miR-30c-1–3p, mmu-miR-99b-3p, mmu-miR-99b-5p, hsa-let-7a-3p, hsa-let-7c-5p, hsa-let-7e-3p, hsa-let-7e-5p, hsa-miR-148a-3p, hsa-miR-192–3p, hsa-miR-20b-5p, hsa-miR-217, hsa-miR-26a-1–3p, hsa-miR-26a-2–3p, hsa-miR-30c-2–3p, hsa-miR-449a, hsa-miR-451a, hsa-miR-582–5p
Jak-STAT related	mmu-let-7a-1–3p, hsa-let-7a-3p, hsa-let-7c-5p, hsa-let-7e-3p, hsa-let-7e-5p, hsa-miR-192–3p, hsa-miR-26a-1–3p, hsa-miR-451a align="center"
Cytoskeleton related	mmu-let-7a-1–3p, mmu-miR-101a-5p, mmu-miR-126–5p, mmu-miR-127–3p, mmu-miR-150–5p, mmu-miR-216b-5p, mmu-miR-24–1-5p, mmu-miR-296–5p, mmu-miR-30c-1–3p, mmu-miR-99b-3p, mmu-miR-99b-5p, hsa-let-7a-3p, hsa-let-7e-3p, hsa-miR-192–3p align="center"
MAPK related	mmu-let-7a-1–3p, mmu-miR-101a-5p, mmu-miR-106b-5p, mmu-miR-126–5p, mmu-miR-127–3p, mmu-miR-150–5p, mmu-miR-193a-5p, mmu-miR-193a-3p, mmu-miR-216b-5p, mmu-miR-296–5p, mmu-miR-30c-1–3p, mmu-miR-99b-3p, mmu-miR-99b-5p, hsa-let-7a-3p, hsa-let-7c-5p, hsa-let-7e-3p, hsa-let-7e-5p, hsa-miR-148a-3p, hsa-miR-18b-5p, hsa-miR-192–3p, hsa-miR-217, hsa-miR-26a-1–3p, hsa-miR-26a-2–3p, hsa-miR-451a align="center"

but also the precursor). For this step, a BLAST (Altschul et al., 1990) database was generated for each genome, allowing the fast retrieval of sequences from given positions in the genomes. A secondary structure prediction was performed with RNAfold v2.1.3 from the ViennaRNA package (Lorenz et al., 2011), giving a graphical representation of each folding in addition to its free energy. For each annotation candidate a results file (HTML) was generated, consolidating the secondary structure images and other information for quick evaluation in the next steps.

For most of the identified locations, more than one miRNA sequence from different species were found by clustering all aligned miRNA sequences on distinct locations in each of the four genomes. To ensure that these evolutionary conserved sequences were only considered once, these not completely identical sequences (they differ in length and/or are shifted some nucleotides in their position) were sorted for their NGS read count and the highest was kept as annotation candidate. The secondary structures were manually evaluated for common miRNA characteristics (low minimum free energy, stem-loop motif, miRNA sequence located in the stem, and no large internal loops or bulges) and are shown in Supplement 3.

To investigate possible uses of the novel miRNAs, we looked at validated miRNA-targets in the originating species (*H. sapiens* or *M. musculus*) with miRWalk (Dweep et al., 2011). The list of target proteins were clustered using the functional annotation clustering tool of DAVID (Huang et al., 2009a; Huang et al.,

Table II. Genome references and statistics used for the identification of evolutionary conserved microRNAs.

	Lewis et al. (2013)	Xu et al. (2011)	Brinkrolf et al. (2013)	K1-BB
Genome size (Gbp)	2.3	2.3	2.1	2.98
Scaffolds	7,468	109,151	28,749	11,400,490
x Coverage	89.1	130.0	70.0	17.1
Accession ID	AMDS00000000.1	AFTD00000000.1	APMK00000000.1	—
Source	CGR	CHO	CGR	CHO

2009b) and the KEGG pathway database and then grouped into seven categories (Table II).

GK and VJ are supported by the FWF doctorate program "BioTop", W1224. GK is supported by the Austrian Center of Industrial Biotechnology, a public-private competence center funded by the Austrian FFG.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.

Ambros V, Bartel B, Bartel DP, Krichevsky AM, King KS, Donahue CP, Burge CB, Carrington JC, Chen X, Dreyfuss G, Eddy SR, Griffiths-Jones S. 2003. A uniform system for microRNA annotation. RNA Soc 9:277–279.

Barron N, Kumar N, Sanchez N, Doolan P, Clarke C, Meleady P, O'Sullivan F, Clynes M. 2011. Engineering CHO cell growth and recombinant protein productivity by overexpression of miR-7. J Biotechnol 151:204–211.

Brinkrolf K, Rupp O, Laux H, Kollin F, Ernst W, Linke B, Kofler R, Romand S, Hesse F, Budach WE, Galosy S, Müller D, Noll T, Wienberg J, Jostock T, Leonard M, Grillari J, Tauch A, Goesmann A, Helk B, Mott JE, Pühler A, Borth N. 2013. Chinese hamster genome sequenced from sorted chromosomes. Nat Biotechnol 31:694–695.

Druz A, Chu C, Majors B, Santuary R, Betenbaugh MJ, Shiloach J. 2011. A novel microRNA mmu-miR-466h affects apoptosis regulation in mammalian cells. Biotechnol Bioeng 108:1651–1661.

Dweep H, Sticht C, Pandey P, Gretz N. 2011. MiRWalk–database: Prediction of possible miRNA binding sites by "walking" the genes of three genomes. J Biomed Inform 44:839–847.

Gregory RI, Yan K-P, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, Shiekhattar R. 2004. The microprocessor complex mediates the genesis of microRNAs. Nature 432:235–240.

Griffiths-Jones S. 2004. The microRNA registry. Nucleic Acids Res 32:D109–D111.

Hackl M, Jadhav V, Jakobi T, Rupp O, Brinkrolf K, Goesmann A, Pühler A, Noll T, Borth N, Grillari J. 2012. Computational identification of microRNA gene loci and precursor microRNA sequences in CHO cell lines. J Biotechnol 158:151–155.

Hackl M, Jakobi T, Blom J, Doppmeier D, Brinkrolf K, Szczepanowski R, Bernhart SH, Honer zu, Siederdissen C, Bort J, a H, Wieser M, Kunert R, Jeffs S, Hofacker IL, Goesmann A, Pühler A, Borth N, Grillari J. 2011. Next-generation sequencing of the Chinese hamster ovary microRNA transcriptome: Identification, annotation and profiling of microRNAs as targets for cellular engineering. J Biotechnol 153: 62–75.

Hobert O. 2008. Gene regulation by transcription factors and microRNAs. Science 319:1785–1786.

Huang DW, Sherman BT, Lempicki RA. 2009a. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37:1–13.

Huang DW, Sherman BT, Lempicki RA. 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4:44–57.

Jadhav V, Hackl M, Druz A, Shridhar S, Chung C-Y, Heffner KM, Kreil DP, Betenbaugh M, Shiloach J, Barron N, Grillari J, Borth N. 2013. CHO microRNA engineering is growing up: Recent successes and future challenges. Biotechnol Adv 31:1501–1513.

Jadhav V, Hackl M, Klanert G, Bort JAH, Kunert R, Grillari J, Borth N. 2014. Stable overexpression of miR-17 enhances recombinant protein production of CHO cells. J Biotechnol 175:38–44.

Johnson KC, Jacob NM, Nissom PM, Hackl M, Lee LH, Yap M, Hu W-S. 2011. Conserved microRNAs in Chinese hamster ovary cell lines. Biotechnol Bioeng 108:475–480.

Kildegaard HF, Baycin-Hizal D, Lewis NE, Betenbaugh MJ. 2013. The emerging CHO systems biology era: Harnessing the 'omics revolution for biotechnology. Curr Opin Biotechnol 24:1102–1107.

Klanert G, Jadhav V, Chanoumidou K, Grillari J, Borth N, Hackl M. 2014. Endogenous microRNA clusters outperform chimeric sequence clusters in Chinese hamster ovary cells. Biotechnol J 9:538–544.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25.

Lewis NE, Liu X, Li Y, Nagarajan H, Yerganian G, O'Brien E, Bordbar A, Roth AM, Rosenbloom J, Bian C, Xie M, Chen W, Li N, Baycin-Hizal D, Latif H, Forster J, Betenbaugh MJ, Famili I, Xu X, Wang J, Palsson BO. 2013. Genomic landscapes of Chinese hamster ovary cell lines as revealed by the Cricetulus griseus draft genome. Nat Biotechnol 31:759–765.

Lim Y, Wong NSC, Lee YY, Ku SCY, Wong DCF, Yap MGS. 2010. Engineering mammalian cells in bioprocessing – current achievements and future perspectives. Biotechnol Appl Biochem 55:175–189.

Loh WP, Loo B, Zhou L, Zhang B, Lee DY, Yang Y, Lam KP. 2014. Overexpression of microRNAs enhances recombinant protein production in Chinese Hamster Ovary cells. Biotechnol J 9(9):1140–1151.

Lorenz R, Bernhart SH, Honer zu, Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. 2011. ViennaRNA Package 2.0. Algorithms Mol Biol 6:26.

Xiao S, Shiloach J, Betenbaugh MJ. 2014. Engineering cells to improve protein expression. Curr Opin Struct Biol 26C:32–38.

Xu X, Nagarajan H, Lewis NE, Pan S, Cai Z, Liu X, Chen W, Xie M, Wang W, Hammond S, Andersen MR, Neff N, Passarelli B, Koh W, Fan HC, Wang J, Gui Y, Lee KH, Betenbaugh MJ, Quake SR, Famili I, Palsson BO, Wang J. 2011. The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. Nat Biotechnol 29:735–741.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web-site.

References

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2008). *Molecular Biology of the Cell* (5th Ed.). New York: Garland Science.

Ambros, V., Bartel, B., Bartel, D. P., Krichevsky, A. M., King, K. S., Donahue, C. P., ... Griffiths-Jones, S. (2003). A uniform system for microRNA annotation. *RNA Soc. 9*, 277–279.

Barron, N., Kumar, N., Sanchez, N., Doolan, P., Clarke, C., Meleady, P., ... Clynes, M. (2011). Engineering CHO cell growth and recombinant protein productivity by overexpression of miR-7. *J. Biotechnol. 151*(2), 204–211.

Bartel, D. P. (2004). MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell*, *116*(2), 281–297.

Behm-Ansmant, I., Rehwinkel, J., & Izaurralde, E. (2006). MicroRNAs silence gene expression by repressing protein expression and/or by promoting mRNA decay. *Cold Spring Harb. Symp. Quant. Biol. 71*, 523–530.

Béthune, J., Artus-Revel, C. G., & Filipowicz, W. (2012). Kinetic analysis reveals successive steps leading to miRNA-mediated silencing in mammalian cells. *EMBO Rep. 13*(8), 716–723.

Blenkiron, C., Goldstein, L. D., Thorne, N. P., Spiteri, I., Chin, S.-F., Dunning, M. J., ... Miska, E. A. (2007). MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype. *Genome Biol. 8*(10), R214.

Brinkrolf, K., Rupp, O., Laux, H., Kollin, F., Ernst, W., Linke, B., ... Borth, N. (2013). Chinese hamster genome sequenced from sorted chromosomes. *Nat. Biotechnol. 31*(8), 694–695.

Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y. Y., Sieburth, L., & Voinnet, O. (2008). Widespread translational inhibition by plant miRNAs and siRNAs. *Science*, *320*(5880), 1185–1190.

Carrer, M., Liu, N., Grueter, C. E., Williams, A. H., Frisard, M. I., Hulver, M. W., ... Olson, E. N. (2012). Control of mitochondrial metabolism and systemic energy homeostasis by microRNAs 378 and 378*. *Proc. Natl. Acad. Sci. U. S. A. 109*(38), 15330–15335.

Carthew, R. W. & Sontheimer, E. J. (2009). Origins and mechanisms of miRNAs and siRNAs. *Cell*, *136*(4), 642–655.

Chan, J. A., Krichevsky, A. M., & Kosik, K. S. (2005). MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer Res. 65*(14), 6029–6033.

Chapin, S. C., Appleyard, D. C., Pregibon, D. C., & Doyle, P. S. (2011). Rapid microRNA profiling on encoded gel microparticles. *Angew. Chem. Int. Ed. Engl. 50*(10), 2289–2293.

Datta, P., Linhardt, R. J., & Sharfstein, S. T. (2013). An omics approach towards CHO cell engineering. *Biotechnol. Bioeng. 110*(5), 1255–1271.

Denli, A. M., Tops, B. B. J., Plasterk, R. H. A., Ketting, R. F., & Hannon, G. J. (2004). Processing of primary microRNAs by the microprocessor complex. *Nature*, *432*(7014), 231–235.

Druz, A., Betenbaugh, M., & Shiloach, J. (2012). Glucose depletion activates mmu-miR-466h-5p expression through oxidative stress and inhibition of histone deacetylation. *Nucleic Acids Res. 40*(15), 7291–7302.

Druz, A., Son, Y.-J., Betenbaugh, M., & Shiloach, J. (2013). Stable inhibition of mmu-miR-466h-5p improves apoptosis resistance and protein production in CHO cells. *Metab. Eng. 16*, 87–94.

Ebert, M. S., Neilson, J. R., & Sharp, P. A. (2007). MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. *Nat. Methods*, *4*(9), 721–726.

Eulalio, A., Huntzinger, E., Nishihara, T., Rehwinkel, J., Fauser, M., & Izaurralde, E. (2009). Deadenylation is a widespread effect of miRNA regulation. *RNA*, *15*(1), 21–32.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, *391*(6669), 806–11.

Fire y Bioworks. (2014). Technology overview. Retrieved August 18, 2014, from http://www.fire ybio.com/technology

Frezzetti, D., De Menna, M., Zoppoli, P., Guerra, C., Ferraro, A., Bello, A. M., ... De Vita, G. (2011). Upregulation of miR-21 by Ras in vivo and its role in tumor growth. *Oncogene*, *30*(3), 275–286.

Gammell, P. (2007). MicroRNAs: recently discovered key regulators of proliferation and apoptosis in animal cells: Identification of miRNAs regulating growth and survival. *Cytotechnology*, *53*(1-3), 55–63.

Ghaderi, D., Zhang, M., Hurtado-Ziola, N., & Varki, A. (2012). Production platforms for biotherapeutic glycoproteins. Occurrence, impact, and challenges of non-human sialylation. *Biotechnol. Genet. Eng. Rev. 28*(1), 147–175.

Griffiths-Jones, S. (2004). The microRNA registry. *Nucleic Acids Res. 32*(Database issue), D109–D111.

Griffiths-Jones, S. (2011). What s in a name? - miRBase blog. Retrieved August 14, 2014, from http://www.mirbase.org/blog/2011/04/whats-in-a-name/

Guo, H., Ingolia, N. T., Weissman, J. S., & Bartel, D. P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature*, *466*(7308), 835–840.

Hacker, D. L., De Jesus, M., & Wurm, F. M. (2009). 25 years of recombinant proteins from reactor-grown cells - where do we go from here? *Biotechnol. Adv. 27*(6), 1023–1027.

Hackl, M., Jadhav, V., Jakobi, T., Rupp, O., Brinkrolf, K., Goesmann, A., ... Grillari, J. (2012). Computational identification of microRNA gene loci and precursor microRNA sequences in CHO cell lines. *J. Biotechnol. 158*(3), 151–155.

Hackl, M., Jadhav, V., Klanert, G., Karbiener, M., Scheideler, M., Grillari, J., & Borth, N. (2014). Analysis of microRNA transcription and post-transcriptional processing by Dicer in the context of CHO cell proliferation. *J. Biotechnol. 190*, 76–84.

Hackl, M., Jakobi, T., Blom, J., Doppmeier, D., Brinkrolf, K., Szczepanowski, R., ... Grillari, J. (2011). Next-generation sequencing of the Chinese hamster ovary microRNA transcriptome: Identification, annotation and profiling of microRNAs as targets for cellular engineering. *J. Biotechnol. 153*(1-2), 62–75.

Hamilton, A. J. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*, *286*(5441), 950–952.

Hammond, S., Kaplarevic, M., Borth, N., Betenbaugh, M. J., & Lee, K. H. (2012). Chinese hamster genome database: an online resource for the CHO community at www.CHOgenome.org. *Biotechnol. Bioeng. 109*(6), 1353–1356.

Hatley, M. E., Patrick, D. M., Garcia, M. R., Richardson, J. A., Bassel-Duby, R., Rooij, E. van, & Olson, E. N. (2010). Modulation of K-Ras-dependent lung tumorigenesis by MicroRNA-21. *Cancer Cell*, *18*(3), 282–293.

Henke, J. I., Goergen, D., Zheng, J., Song, Y., Schüttler, C. G., Fehr, C., ... Niepmann, M. (2008). microRNA-122 stimulates translation of hepatitis C virus RNA. *EMBO J. 27*(24), 3300–3310.

Hernández Bort, J. A., Hackl, M., Hö mayer, H., Jadhav, V., Harreither, E., Kumar, N., ... Borth, N. (2012). Dynamic mRNA and miRNA profiling of CHO-K1 suspension cell cultures. *Biotechnol. J. 7*(4), 500–515.

Hobert, O. (2008). Gene regulation by transcription factors and microRNAs. *Science*, *319*(5871), 1785–1786.

Houseley, J. & Tollervey, D. (2009). The many pathways of RNA degradation. *Cell*, *136*(4), 763–776.

Ishizu, H., Siomi, H., & Siomi, M. C. (2012). Biology of PIWI-interacting RNAs: New insights into biogenesis and function inside and outside of germlines. *Genes Dev. 26*(21), 2361–2373.

Jadhav, V., Hackl, M., Druz, A., Shridhar, S., Chung, C.-Y., Heffner, K. M., ... Borth, N. (2013). CHO microRNA engineering is growing up: Recent successes and future challenges. *Biotechnol. Adv. 31*(8), 1501–1513.

Jayapal, K. & Wlaschin, K. (2007). Recombinant protein therapeutics from CHO cells – 20 years and counting. *Chem. Eng. 103*(10), 40–47.

Johnson, S. M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., ... Slack, F. J. (2005). RAS is regulated by the let-7 microRNA family. *Cell*, *120*(5), 635–647.

Jopling, C. (2012). Liver-specific microRNA-122: Biogenesis and function. *RNA Biol. 9*(2), 137–142.

Kaufman, R. J., Wasley, L. C., Spiliotes, A. J., Gossels, S. D., Latt, S. A., Larsen, G. R., & Kay, R. M. (1985). Coamplification and coexpression of human tissue-type plasminogen activator and murine dihydrofolate reductase sequences in Chinese hamster ovary cells. *Mol. Cell. Biol. 5*(7), 1750–1759.

Kildegaard, H. F., Baycin-Hizal, D., Lewis, N. E., & Betenbaugh, M. J. (2013). The emerging CHO systems biology era: harnessing the omics revolution for biotechnology. *Curr. Opin. Biotechnol. 24*(6), 1102–1107.

Kim, J. Y., Kim, Y.-G., & Lee, G. M. (2012). CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Appl. Microbiol. Biotechnol. 93*(3), 917–930.

Kim, V. N. (2005). MicroRNA biogenesis: Coordinated cropping and dicing. *Nat. Rev. Mol. Cell Biol. 6*(5), 376–385.

Kozomara, A. & Griffiths-Jones, S. (2014). miRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res. 42*(Database issue), D68–D73.

Krützfeldt, J., Rajewsky, N., Braich, R., Rajeev, K. G., Tuschl, T., Manoharan, M., & Stoffel, M. (2005). Silencing of microRNAs in vivo with antagomirs . *Nature*, *438*(7068), 685–689.

Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*, *75*(5), 843–854.

Lee, Y. S. & Dutta, A. (2006). MicroRNAs: Small but potent oncogenes or tumor suppressors. *Curr. Opin. Investig. Drugs*, *7*(6), 560–564.

Lewis, N. E., Liu, X., Li, Y., Nagarajan, H., Yerganian, G., O Brien, E., ... Palsson, B. O. (2013). Genomic landscapes of Chinese hamster ovary cell lines as revealed by the Cricetulus griseus draft genome. *Nat. Biotechnol. 31*(8), 759–765.

Lim, Y., Wong, N. S. C., Lee, Y. Y., Ku, S. C. Y., Wong, D. C. F., & Yap, M. G. S. (2010). Engineering mammalian cells in bioprocessing – current achievements and future perspectives. *Biotechnol. Appl. Biochem. 55*(4), 175–189.

Lin, N., Davis, A., Bahr, S., Borgschulte, T., Achtien, K., & Kayser, K. (2011). Profiling highly conserved microrna expression in recombinant IgG-producing and parental Chinese hamster ovary cells. *Biotechnol. Prog. 27*(4), 1163–1171.

Linsen, S. E. V., Wit, E. de, Janssens, G., Heater, S., Chapman, L., Parkin, R. K., ... Cuppen, E. (2009). Limitations and possibilities of small RNA digital gene expression profiling. *Nat. Methods*, *6*(7), 474–476.

Maroney, P. A., Yu, Y., Fisher, J., & Nilsen, T. W. (2006). Evidence that microRNAs are associated with translating messenger RNAs in human cells. *Nat. Struct. Mol. Biol. 13*(12), 1102–1107.

Meijer, H. A., Kong, Y. W., Lu, W. T., Wilczynska, A., Spriggs, R. V., Robinson, S. W., ... Bushell, M. (2013). Translational repression and eIF4A2 activity are critical for microRNA-mediated gene regulation. *Science*, *340*(6128), 82–85.

Müller, D., Katinger, H., & Grillari, J. (2008). MicroRNAs as targets for engineering of CHO cell factories. *Trends Biotechnol. 26*(7), 359–365.

Nelson, P. T., Wang, W.-X., Wilfred, B. R., & Tang, G. (2008). Technical variables in high-throughput miRNA expression profiling: Much work remains to be done. *Biochim. Biophys. Acta*, *1779*(11), 758–765.

Okamura, K., Hagen, J. W., Duan, H., Tyler, D. M., & Lai, E. C. (2007). The mirtron pathway generates microRNA-class regulatory RNAs in drosophila. *Cell*, *130*(1), 89–100.

Olsen, P. H. & 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*(2), 671–680.

Petersen, C. P., Bordeleau, M.-E., Pelletier, J., & Sharp, P. a. (2006). Short RNAs repress translation after initiation in mammalian cells. *Mol. Cell*, *21*(4), 533–542.

Planell-Saguer, M. de & Rodicio, M. C. (2013). Detection methods for microRNAs in clinic practice. *Clin. Biochem. 46*(10-11), 869–878.

Pritchard, C. C., Cheng, H. H., & Tewari, M. (2012). MicroRNA profiling: Approaches and considerations. *Nat. Rev. Genet. 13*(5), 358–369.

Puck, T. T. (1958). Genetics of somatic mammalian cells: III. Long-term cultivation of euploid cells from human and animal subjects. *J. Exp. Med. 108*(6), 945–956.

Rooij, E. van. (2011). The art of microRNA research. *Circ. Res. 108*(2), 219–234.

Ruby, J. G., Jan, C. H., & Bartel, D. P. (2007). Intronic microRNA precursors that bypass Drosha processing. *Nature*, *448*(7149), 83–86.

Ruby, J. G., Jan, C., Player, C., Axtell, M. J., Lee, W., Nusbaum, C., ... Bartel, D. P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. *Cell*, *127*(6), 1193–1207.

Serino, G., Sallustio, F., Cox, S. N., Pesce, F., & Schena, F. P. (2012). Abnormal miR-148b expression promotes aberrant glycosylation of IgA1 in IgA nephropathy. *J. Am. Soc. Nephrol. 23*(5), 814–824.

Shu, J., Xia, Z., Li, L., Liang, E. T., Slipek, N., Shen, D., ... Steer, C. J. (2012). Dose-dependent differential mRNA target selection and regulation by let-7a-7f and miR-17-92 cluster microRNAs. *RNA Biol. 9*(10), 1275–1287.

Si, M.-L., Zhu, S., Wu, H., Lu, Z., Wu, F., & Mo, Y.-Y. (2007). miR-21-mediated tumor growth. *Oncogene*, *26*(19), 2799–2803.

Sripada, L., Tomar, D., & Singh, R. (2012). Mitochondria: one of the destinations of miRNAs. *Mitochondrion*, *12*(6), 593–599.

Stein, C. A., Hansen, J. B., Lai, J., Wu, S., Voskresenskiy, A., Høg, A., ... Koch, T. (2010). Efficient gene silencing by delivery of locked nucleic acid antisense oligonucleotides, unassisted by transfection reagents. *Nucleic Acids Res. 38*(1), e3.

Stroynowska-Czerwinska, A., Fiszer, A., & Krzyzosiak, W. J. (2014). The panorama of miRNA-mediated mechanisms in mammalian cells. *Cell. Mol. Life Sci. 71*(12), 2253–2270.

Tian, G., Yin, X., Luo, H., Xu, X., Bolund, L., Zhang, X., ... Li, N. (2010). Sequencing bias: Comparison of different protocols of microRNA library construction. *BMC Biotechnol. 10*(1), 64.

Urlaub, G. & Chasin, L. A. (1980). Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity. *Proc. Natl. Acad. Sci. U. S. A. 77*(7), 4216–4220.

Urlaub, G., Käs, E., Carothers, A. M., & Chasin, L. A. (1983). Deletion of the diploid dihydrofolate reductase locus from cultured mammalian cells. *Cell*, *33*(2), 405–412.

Vasudevan, S., Tong, Y., & Steitz, J. A. (2007). Switching from repression to activation: microRNAs can up-regulate translation. *Science*, *318*(5858), 1931–1934.

Wahle, E. & Winkler, G. S. (2013). RNA decay machines: Deadenylation by the Ccr4-not and Pan2-Pan3 complexes. *Biochim. Biophys. Acta*, *1829*(6-7), 561–570.

Wang, B., Yanez, A., & Novina, C. D. (2008). MicroRNA-repressed mRNAs contain 40S but not 60S components. *Proc. Natl. Acad. Sci. U. S. A. 105*(14), 5343–5348.

Wang, Z. (2011). The guideline of the design and validation of MiRNA mimics. *Methods Mol. Biol. 676*, 211–223.

Wightman, B., Ha, I., & Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell*, *75*(5), 855–862.

Winter, J., Jung, S., Keller, S., Gregory, R. I., & Diederichs, S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat. Cell Biol. 11*(3), 228–234.

Wurm, F. M. (2004). Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat. Biotechnol. 22*(11), 1393–1398.

Wurm, F. M. & Hacker, D. L. (2011). First CHO genome. *Nat. Biotechnol. 29*(8), 718–720.

Xiao, S., Shiloach, J., & Betenbaugh, M. J. (2014). Engineering cells to improve protein expression. *Curr. Opin. Struct. Biol. 26C*, 32–38.

Xu, X., Nagarajan, H., Lewis, N. E., Pan, S., Cai, Z., Liu, X., ... Wang, J. (2011). The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat. Biotechnol. 29*(8), 735–741.

Yang, J.-S., Phillips, M. D., Betel, D., Mu, P., Ventura, A., Siepel, A. C., ... Lai, E. C. (2011). Widespread regulatory activity of vertebrate microRNA* species. *RNA*, *17*(2), 312–326.

Yi, R., Qin, Y., Macara, I. G., & Cullen, B. R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev. 17*(24), 3011–3016.

Zdanowicz, A., Thermann, R., Kowalska, J., Jemielity, J., Duncan, K., Preiss, T., ... Hentze, M. W. (2009). Drosophila miR2 primarily targets the m7GpppN cap structure for translational repression. *Mol. Cell*, *35*(6), 881–888.

List of Figures

List of Tables