



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e. g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

The role of DGCR8 in mammalian pluripotency and innate immunity



THE UNIVERSITY
of EDINBURGH



University
of Glasgow

Lisanne Iris Knol

Submitted for the degree of Doctor of Philosophy

The University of Edinburgh and University of Glasgow

2022



I. DECLARATION

I declare that this thesis has been composed solely by me and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where stated otherwise by reference or acknowledgment, the work presented is entirely my own.

Lisanne Iris Knol

04/09/22

II. CONTRIBUTIONS

The PhD project was supervised by Dr. Sara Macias, with bioinformatic supervision by Dr. Alisdair Ivens. PA-1 *DGCR8* knockout cell lines were obtained by Dr. Pablo Tristán Ramos in the lab of Dr. Sara Heras at GENYO Granada. Sanger sequencing of genomic regions of interest of the PA-1 mutant cell lines was carried out by either Edinburgh Genomics or MRC PPU DNA Sequencing and Services in Dundee. Library preparation and sequencing of RNA was carried out by BGI Genomics, Hong Kong. Creation of BED files and determination of mean read depth for microprocessor processing index calculations was done by Dr. Alasdair Ivens. The results described in section 5.2.1 consist of both preliminary results and experiments published in Witteveldt, Knol, & Macias, (2018). References to this paper refer to work done by Dr. Jeroen Witteveldt and Dr. Sara Macias.

III. ACKNOWLEDGEMENTS

Many thanks to Sara, I am incredibly grateful for the opportunities you offered me, first for an Erasmus exchange, then for inviting me back for a PhD. Your involvement in the project and your visions on science have been very inspiring. Thank you, Al, for all the discussions and for nurturing my enthusiasm for bioinformatics.

I would like to thank everyone in the Macias, Buck and Zamoyska lab for scientific discussions and useful feedback. To Matilda, Patricia, Samir, and Sarah: thanks for all the great experiences both in and outside of the lab.

To my fellow Precision Medicine students Ciara, Heather, Ian, Leire and Sean: it was an honour to start and navigate the PhD landscape together, thanks for all the nice coffee breaks and beers!

To Darren, Michiel, Mike, and Tadgh: thank you for all your support and advice over the years.

To Hanna, Denisie, Juud, Melis and Suus: how do I express my gratitude for over 10 years of friendship and support? I am so glad that, even separated by country, we stuck together while all pursuing PhDs and I am incredibly proud of you. Looking forward to seeing what the future will bring!

IV. LAY SUMMARY

RNAs are important molecules carrying information derived from DNA. While the information encoded in messenger RNAs can be converted into proteins and directly leads to expression of specific genes, several other classes of RNA alter gene expression in a more indirect manner. These RNAs can interact with the DNA, other RNAs, and proteins, thereby stimulating or preventing gene expression.

Micro (mi)RNAs are short single stranded RNAs that prevent the building of proteins by binding to long RNAs, thereby regulating gene expression. Multiple proteins are involved the synthesis of miRNAs, one of which is *DGCR8*. Furthermore, *DGCR8* is one of multiple genes that only have one gene copy instead of two in patients with 22q11.2 deletion syndrome (DS). These patients often show developmental delays, neuropsychiatric disorders, and altered immune responses.

How important *DGCR8* is for both miRNA production and 22q11.2 DS is unclear. Therefore, we studied the impact of *DGCR8* gene depletion on RNA levels, development, and innate immunity. We obtained human cell lines lacking either one or two gene copies of *DGCR8* and compared these to cells with two *DGCR8* gene copies.

In absence of *DGCR8*, levels of most miRNAs are reduced and the overall expression of longer RNAs is drastically different, affecting many cellular processes. Research in mouse embryonic stem cells showed that cells without *DGCR8* are better able to defend themselves against viral infection. This is due to activation of a specific type of the innate immune response that is normally suppressed by miRNAs. Further research should focus on how miRNAs can suppress this immune response in humans.

Furthermore, cells with only one gene copy of *DGCR8* showed changed levels of messenger RNAs that are important for development, indicating that the lack of two *DGCR8* genes contributes to the symptoms observed in 22q11.2 deletion syndrome.

V. ABSTRACT

DGCR8 is an RNA-binding protein involved in canonical micro (mi)RNA biogenesis as part of the nuclear microprocessor complex. While absence of DGCR8 in mammals is lethal, DGCR8 hemizyosity is observed in 22q11.2 deletion syndrome (22q11.2DS), the most frequently observed microdeletion in humans. 22q11.2DS is associated with aberrant immunity, neuropsychiatric disorders, and a range of developmental defects.

To better understand the role of DGCR8 in immunity and development, CRISPR-Cas was used to create *DGCR8* knockout and heterozygous PA-1 pluripotent cell lines. Different types of RNA sequencing revealed an altered transcriptomic landscape. As expected, *DGCR8* knockout cells showed an overall decrease in mature miRNAs compared to wildtype levels, suggested to be caused by impaired primary miRNA processing. On the other hand, only a fraction of miRNAs in *DGCR8* heterozygotes was differentially expressed. Notably, miRNAs in the *miR-105/767* cluster were among the most downregulated among both knockout and heterozygote *DGCR8* mutant cell lines. Pathway enrichment analysis for differently expressed transcripts suggested a defect in pluripotency in *DGCR8* heterozygote cells, indicating that *DGCR8* hemizyosity could have a more profound contribution to the developmental symptoms detected in 22q11.2 DS than previously anticipated.

Lastly, we investigated the role of DGCR8 in antiviral immunity during pluripotency. While embryonic stem cells normally are unable to produce type-I interferons, *Dgcr8*-deficient mouse embryonic stem cells can mount an active interferon response when stimulated via the cytosolic RIG-I-like receptor pathway. This effect was contributed to the central role of *mmu-miR-673* in suppressing the expression of Mitochondrial Antiviral Signalling Protein, leading to increased susceptibility to viral infections. These results highlight the importance of miRNAs in modulating mammalian innate immunity.

VI. LIST OF ABBREVIATIONS

AGO	Argonaut	PACT	Protein activator of interferon induced protein kinase
CAS	CRISPR-associated protein	PAM	Protospacer adjacent motif
cGAMP	Cyclic guanosine monophosphate–adenosine monophosphate	PAMP	Pathogen-associated molecular pattern
cGAS	cGAMP synthetase	Pol	Polymerase
CLIP	Cross-linking and immunoprecipitation	PRR	Pattern recognition receptor
CNA	Copy number alteration	Ran	Ras-related nuclear protein
CRISPR	clustered regularly interspaced short palindromic repeats	RBD	RNA-binding domain
DCL1	Dicer-like 1	RBP	RNA-binding protein
DGCR8	DiGeorge syndrome critical region 8	RIG-I	Retinoic acid-inducible gene I
EXP5	Exportin5	RISC	RNA-induced silencing complex
ESC	Embryonic stem cell	RLR	RIG-I like receptor
HR	Homologous recombination	RNAi	RNA interference
IFN	Interferon	RT	Reverse transcriptase
IPSC	Induced pluripotent stem cell	siRNA	Short interfering RNA
IRF	Interferon regulatory factor	snoRNA	Small nucleolar RNA
JAK	Janus kinase	SNV	Single nucleotide variant
lnc	Long non-coding	Ss	Single stranded
Loqs	Loquacious	STAT	Signal transducer and activator of transcription
MAVS	Mitochondrial antiviral signalling protein	STING	Stimulator of IFN genes
MDA5	Melanoma differentiation-associated gene 5	TBK1	TANK-binding kinase 1
miRNA	MicroRNA	TBX1	T-box transcription factor 1
NHEJ	Non-homologous end joining	TLR	Toll-like receptor
Nt	Nucleotide	TMEV	Theiler's murine encephalomyelitis virus
LGP2	Laboratory of genetics and physiology 2	TRBP	Trans-activation-responsive RNA-binding protein
LncRNA	Long non-coding RNA	TUTase	Terminal uridylyl transferase
LPS	Lipopolysaccharide		
OASL	2'-5'-Oligoadenylate synthetase like		

VII. TABLE OF CONTENTS

I. Declaration.....	3
II. Contributions	4
III. Acknowledgements.....	5
IV. Lay summary	6
V. Abstract.....	7
VI. List of abbreviations.....	8
VII. Table of Contents	9
Chapter 1: Introduction	13
1.1 MicroRNAs.....	14
1.1.1 RNA interference	14
1.1.2 MiRNA biogenesis.....	16
1.1.3 Speed and specificity of miRNA biogenesis.....	17
1.1.4 Non-canonical miRNAs	18
1.1.5 MiRNA sequence and function.....	19
1.2 The microprocessor.....	20
1.2.1 Structure of the microprocessor	20
1.2.2 Pri-miRNA requirements for effective cleaving.....	21
1.2.3 Measuring microprocessor efficacy	27
1.2.4 Noncanonical functions of the microprocessor	28
1.3 Innate antiviral immunity.....	30
1.3.1 The mammalian type-I interferon response	30
1.3.2 Effects of aberrant type-I IFN expression.....	33
1.3.3 Antiviral siRNAs.....	35
1.3.4 The IFN response is lost or gained dependent on pluripotency status.....	36
1.3.5 Possible mechanisms and reasons for attenuated type-I IFN response ...	38
1.4 MiRNAs and the microprocessor in immunity and development	40
1.4.1 The role of miRNAs in innate immunity	40
1.4.2 The role of the microprocessor in innate immunity.....	41
1.4.3 The roles of miRNAs in early development	42
1.4.4 DGCR8 mutants and their effect on development.....	43

1.5 22q11.2 deletion syndrome.....	45
1.6 Aims of this PhD project	48
Chapter 2: Materials and methods	50
2.1 Cell culture	51
2.2 CRISPR-Cas9-based genome editing.....	52
2.2.1 Design and preparation of Cas9-guideRNA mix.....	52
2.2.2 Transfection of cells	52
2.2.3 Clonal expansion of single cell colonies.....	53
2.2.4 Screening of clones	53
2.3 Poly(I:C) transfections.....	54
2.4 Viral infections	54
2.5 Retinoic acid differentiation	55
2.6 Cellular fractionation	56
2.7 RNA extractions	56
2.8 RT-qPCR.....	58
2.9 RT-qPCR for miRNAs	59
2.10 Protein extraction and western blot	59
2.11 Library preparation and sequencing of total, small, and chromatin-associated RNA	60
2.12 Analysis of total RNA-seq.....	62
2.13 Analysis of small RNA-seq.....	62
2.14 Analysis of chromatin-associated RNA-seq and MPI calculation	63
2.15 MirTarRnaSeq	65
Chapter 3: Transcriptomic characterization of <i>DGCR8</i> knockout and heterozygote PA-1 cell lines.....	67
3.1 Introduction	68
3.2 Results.....	69
3.2.1 Creation of PA-1 <i>DGCR8</i> heterozygous cell lines	69
3.2.2 Transcriptomic changes upon <i>DGCR8</i> knockout and heterozygosity	74
3.2.3 Validation of RNA-seq data and concordance between cell lines	76
3.2.4 Functional analysis of RNA-seq data show changes in pluripotency upon <i>DGCR8</i> mutation	79
	10

3.3 Discussion	84
3.3.1 PA-1 <i>DGCR8</i> knockout and heterozygote cells can help further study the role of DGCR8 and the microprocessor in the human context	84
3.3.2 Transcriptomic analyses show a changed RNA landscape upon DGCR8 depletion	85
Chapter 4: The impact of DGCR8 on the the miRNA landscape	88
4.1 Introduction	89
4.2 Results	90
4.2.1 <i>DGCR8</i> mutation changes small RNA levels in PA-1 cells	90
4.2.2 DGCR8 knockout shifts the relative levels of miRNA subtypes	93
4.2.3 Differential expression of miRNAs upon <i>DGCR8</i> mutation	95
4.2.4 Differentially expressed miRNAs in <i>DGCR8</i> heterozygous cells target pathways involved in differentiation and pluripotency	98
4.2.5 The microprocessor index	100
4.2.6 Purification of chromatin-associated RNA and MPI determination	103
4.2.7 Microprocessor efficacy is affected in <i>DGCR8</i> knockout cells	106
4.3 Discussion	110
4.3.1 Knockout of <i>DGCR8</i> transforms the miRNA landscape	110
4.3.2 <i>DGCR8</i> heterozygosity leads to limited changes in miRNA levels compared to wildtype cells	112
4.3.3 The MPI as a determinant of microprocessor cleaving	113
Chapter 5: The role of DGCR8 in innate immunity and development	116
5.1 Introduction	117
5.2 Results	118
5.2.1 <i>mmu-miR-673</i> downregulates MAVS and increases viral sensitivity in mESCs	118
5.2.2 Absence of <i>DGCR8</i> increases the type-I IFN response in PA-1 cells	123
5.2.3 Depletion of <i>DGCR8</i> leads to changes in steady state levels of pluripotency factors	126
5.2.4 <i>DGCR8</i> heterozygote and knockout cells show aberrant differentiation	128
5.3 Discussion	132
5.3.1 <i>DGCR8</i> and miRNAs are involved in downregulation of the RLR-pathway in the type-I IFN response	132

5.3.2 <i>DGCR8</i> heterozygosity leads to an aberrant state of pluripotency	133
Chapter 6: Discussion	136
6.1 Combining RNA-seq, small RNA-seq and MPI determination in the same model	137
6.2 The role of <i>DGCR8</i> and small RNAs in antiviral immunity and differentiation	138
6.3 Final remarks	139
VIII. Supplementary table	141
IX. Bibliography.....	146

CHAPTER 1: INTRODUCTION

1.1 MICRORNAS

1.1.1 RNA interference

RNA interference is a major mechanism to negatively regulate gene expression at the post-transcriptional level, driven by complementarity between short non-coding RNAs and specific target RNAs in presence of an effector protein of the Argonaut (AGO) family (Azlan, Dzaki and Azzam, 2016; Dexheimer and Cochella, 2020). Since its discovery, RNAi has been shown to play major roles in biological pathways in eukaryotes.

RNAi can be classified into three different categories: micro (mi)RNAs, short-interfering (si)RNAs, and PIWI-interacting (pi)RNAs (Farazi, Juranek and Tuschl, 2008). Both mi- and siRNAs interact with AGO family proteins to form an RNA-induced silencing complex (RISC), whereas piRNAs are larger and primarily expressed in the germline, and associate with the PIWI subfamily of AGO proteins to target transposable elements (Girard *et al.*, 2006; Siomi *et al.*, 2011).

While siRNAs have been mainly studied in the context of antiviral immunity in plants and invertebrates, miRNAs are most known for their roles in development in nematodes and disease in mammals (Giraldez *et al.*, 2005; Kanellopoulou *et al.*, 2005; Obbard *et al.*, 2009; Mendell and Olson, 2012; Sabin *et al.*, 2013; Gruber *et al.*, 2014; Lin and Gregory, 2015; Li *et al.*, 2016; Ressel *et al.*, 2019).

While piRNAs have a distinct biogenesis pathway that has only partially been elucidated, the biogenesis and mechanisms of miRNAs and siRNAs are better understood and partially overlap (Luteijn and Ketting, 2013; Li, Zhang and Liu, 2021). The biogenesis of miRNAs consist of several post-transcriptional processing steps and will be discussed below.

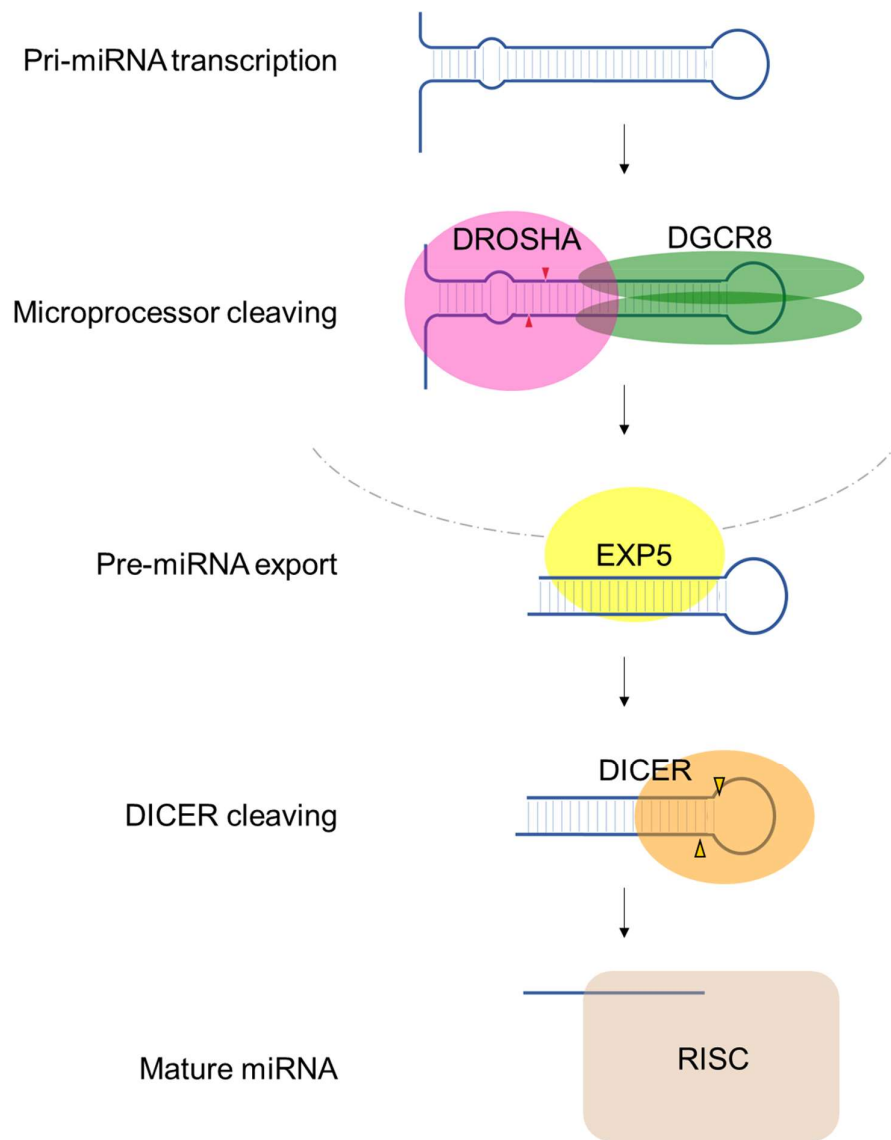


Figure 1.1. Overview of miRNA biogenesis. After transcription, pri-miRNAs are cleaved by the microprocessor, consisting of DROSHA and two copies of DGCR8. The resulting pre-miRNA is subsequently exported out of the cytoplasm by Exportin 5, where it is cleaved by DICER. The now matured miRNA can be loaded on RISC to regulate genes expression.

1.1.2 MiRNA biogenesis

MiRNA genes are spread across the genome and can be nested within other genes. In that case, miRNA genes most commonly occur in introns or untranslated regions, however overlap with exons have also been reported (Olena and Patton, 2010; Hammond, 2015). Typically miRNAs are transcribed as long precursor molecules, known as primary (pri-)miRNAs by RNA-polymerase II (Lee *et al.*, 2002; Y. Lee *et al.*, 2004). In most cases, pri-miRNAs code for multiple miRNAs in clusters. Characteristic of pri-miRNAs are the presence of stem-loop or hairpin structures that contain the sequence of the mature miRNA (figure 1.1).

These structures can be recognised and cleaved a nuclear complex called the microprocessor, consisting of the proteins DROSHA and DGCR8. DROSHA contains two RNase-III domains that specifically cleave dsRNAs, and DGCR8 reportedly assists in binding and positioning of the pri-miRNA (Denli *et al.*, 2004; Gregory *et al.*, 2004; Han *et al.*, 2004; Nguyen *et al.*, 2015).

Microprocessor cleaving results in the release of the hairpin structure with a characteristic 2-nucleotide (nt) 3' overhang and is discussed in more detail in section 1.2. After nuclear processing, the hairpin, now termed precursor (pre-)miRNA, is exported to the cytoplasm by Exportin 5 (EXP5) and its cofactor (Ras-related Nuclear protein) RanGTP (Bohnsack, Czapinski and Görlich, 2004; Okada *et al.*, 2009). EXP5 prefers hairpins with a 2-nt overhang, and therefore nuclear export functions as a quality control step of miRNA biogenesis.

In the cytoplasm, pre-miRNAs are further processed by the RNase-III endonuclease DICER (Dicer-1 in *Drosophila melanogaster*) (Hutvagner *et al.*, 2001; Ketting *et al.*, 2001; Han *et al.*, 2004; Zhang *et al.*, 2004). DICER cleaves in a similar manner to DROSHA, and there are indications their RNase domains have evolved from class III RNases early in metazoan evolution (Shabalina and Koonin, 2008; Kwon *et al.*, 2016). Additional co-factors, such as Trans-Activation-Responsive RNA-Binding Protein (TRBP) and Protein Activator Of Interferon Induced Protein Kinase (PACT) in

humans, and Loquacious (Loqs) in *Drosophila melanogaster*, are also important for guiding successful Dicer-mediated processing (Chendrimada *et al.*, 2005; Haase *et al.*, 2005; Fukunaga *et al.*, 2012; Meister, 2013; Wilson *et al.*, 2015). DICER cleaves approximately 22-nt from the 5' and 3' ends of the pre-miRNA hairpin, which results in the formation of a mature miRNA duplex of 20 to 24-nt length and 2-nt 3' overhangs on both strands (MacRae *et al.*, 2006; MacRae, Zhou and Doudna, 2007).

Subsequently, the RNA duplex is unwound and one strand is selected by the AGO protein, initiating formation of an RNA-induced silencing complex (RISC) (Kwak and Tomari, 2012). Mammals have four AGO proteins (AGO 1-4), and expression of only one of those is enough to facilitate miRNA-mediated gene silencing (Su *et al.*, 2009). Only AGO2 has endonuclease activity upon encountering complementary sequences, leading to direct degradation of target RNAs (Carmell *et al.*, 2004; Meister *et al.*, 2004). However, most miRNAs downregulate gene expression by stimulating mRNA transcript degradation, preventing translation or relocation away from ribosomes (Huntzinger and Izaurralde, 2011; Eichhorn *et al.*, 2014; Jonas and Izaurralde, 2015; Zhang *et al.*, 2018; Eisen *et al.*, 2020).

1.1.3 Speed and specificity of miRNA biogenesis

Several studies have researched the kinetics of miRNA biogenesis and its effect on miRNA expression in cell lines. In *Drosophila melanogaster*, the rate of miRNA production ranged from 17 to over 200 miRNAs per minute, with a turnover rate ranging from minutes to days (Reichholf *et al.*, 2019). Loading of mature miRNAs onto AGO was found to be the most time-consuming step, creating a bottleneck that results in many miRNAs being degraded prematurely (Diederichs and Haber, 2007; Janas *et al.*, 2012; Reichholf *et al.*, 2019; Bose *et al.*, 2020). A possible reason for this is to prevent non-specific loading of small RNAs to detrimental effect.

Furthermore, *in vitro* DROSHA cleaving studies and bioinformatic analysis showed a correlation between both pri-miRNA expression and the efficacy of pri-miRNA cleaving with mature miRNA expression (Feng *et al.*, 2011; Conrad *et al.*, 2014). The latter indicates microprocessor activity is important for the specificity of the expressed miRNAs.

1.1.4 Non-canonical miRNAs

Several mammalian miRNAs circumvent processing via the canonical pathway described above. Firstly, mirtrons are microprocessor-independent miRNAs and are transcribed as the intronic sequences instead of originating from pri-miRNAs. Splicing releases the sequences, which then fold into hairpin structures and join the canonical biogenesis pathway as pre-miRNAs to be processed by DICER (Berezikov *et al.*, 2007; Okamura *et al.*, 2007; Westholm and Lai, 2011). Secondly, capped miRNAs derive from short transcribed sequences by either RNA-polymerase II or polymerase III. While the 5' end is capped after the start of transcription, its 3' end is determined by transcription termination, and is microprocessor-independent (Xie *et al.*, 2013). After DICER processing the 3' strand is selected as guide miRNA, since AGO cannot access the 5' cap. Finally, the miRNA *miR-451* is processed by the microprocessor but circumvents DICER processing and is instead cleaved by AGO2 (Dueck *et al.*, 2012).

In addition, several other short RNAs, originating from tRNAs, small nucleolar RNA, and transposable elements, can be loaded on AGO proteins and exert miRNA-like functions (Ender *et al.*, 2008). Other AGO-associated RNAs, termed Agotrons, bypass both the microprocessor and DICER processing steps (Hansen *et al.*, 2016). This recently discovered group also is a descendant of introns and is transported to the cytoplasm by a yet to discover mechanism and loaded on Ago unprocessed (Daugaard and Hansen, 2017).

1.1.5 MiRNA sequence and function

Target specificity by miRNAs depends on complementarity between part of miRNAs and their targets. It is mainly determined by binding of a 6-mer at the 5' terminus of the miRNA, called the seed sequence (Lewis *et al.*, 2003; Bartel, 2009). In addition, downstream sequences that are complementary to the target sequence can either aid if the seed-target match is imperfect, or allow for expansion of the target repertoire for miRNAs with the same seed (Moore *et al.*, 2015; Broughton *et al.*, 2016). MiRNA families are miRNAs that originate mostly through duplications of miRNA genes, and therefore show sequence similarity (Alvarez-Saavedra and Horvitz, 2010; Luo *et al.*, 2018).

Precise cleavage by DROSHA and DICER is essential for functional miRNA biogenesis, since a shift in cleaving of a single nucleotide will change the miRNA seed sequence and thereby affect target affinity (Lee *et al.*, 2002; Lu *et al.*, 2005; Kim, Kim and Kim, 2016). MiRNAs with non-canonical 3' and 5' ends are termed isomirs, and can originate either by alternative cleaving, or by post-transcriptional modifications. Sometimes these modifications are necessary for nuclear export, for example in the case of *let-7* family members (Heo *et al.*, 2012). In other cases isomirs are detrimental and associated with several types of cancer (C. W. Wu *et al.*, 2018; Bofill-De Ros *et al.*, 2019).

The distribution of mature miRNAs expression is highly skewed, with the 5 most expressed miRNAs representing 50% of all detectable miRNAs in cells, and they show strong spatiotemporal specificity (Sood *et al.*, 2006; Alberti and Cochella, 2017; De Rie *et al.*, 2017; McCall *et al.*, 2017). Overall, the evolutionary oldest miRNAs are usually the most widely expressed (Patel and Capra, 2017). Target redundancy due to a common seed sequence within a family or expression of multiple miRNAs with the same target can make repression of target gene expression more robust (Schirle, Sheu-Gruttadauria and MacRae, 2014; Sheu-Gruttadauria and MacRae, 2017). MiRNAs need to accumulate to high levels in the

cell in order to be effective, and have to be associated with AGO in order to mediate target repression (Ameres and Zamore, 2013; Denzler *et al.*, 2014; Flores *et al.*, 2014). Methods such as crosslinking, ligation, and sequencing of hybrids (CLASH) identify miRNA-target interactions by immunoprecipitation of AGO proteins and therefore are thought to represent functional miRNAs in the experimental conditions (Helwak *et al.*, 2013). While repression of targets by miRNAs does not result in a complete knockdown, miRNAs have shown to be essential in mammals, and play important roles in several diseases (Pasquinelli *et al.*, 2000; Reinhart *et al.*, 2000; Lin and Gregory, 2015). In section 1.3 the role of miRNAs in development and immunity is discussed in more detail.

1.2 THE MICROPROCESSOR

1.2.1 Structure of the microprocessor

The microprocessor complex is involved in the canonical miRNA biogenesis pathway, cleaving long pri-miRNAs into short hairpin structures and thereby facilitating nuclear export and further maturation (Lee *et al.*, 2002; Denli *et al.*, 2004; Gregory *et al.*, 2004; Han *et al.*, 2004). It is a heterotrimer, consisting of two copies of the RNA binding protein DGCR8 and the nuclease DROSHA (Nguyen *et al.*, 2015). Both DROSHA and DGCR8 consist of multiple functional domains that interact with each other to accurately cleave pri-miRNAs. DGCR8 consists of an ssRNA-binding heme domain (Rhed), two dsRNA binding domains (dsRBDs) and a C-terminal domain (CTD) (figure 1.2A) (Yeom *et al.*, 2006; Partin *et al.*, 2020). DROSHA contains a central domain, two RNase-III domains, and a dsRBD (figure 1.2B) (Kwon *et al.*, 2016; Jin *et al.*, 2020; Partin *et al.*, 2020).

Processing of pri-miRNAs occurs can occur both co- and post-transcriptionally and precedes splicing (Kim and Kim, 2007; Morlando *et al.*, 2008; Louloui *et al.*, 2017). Upon encountering a pri-miRNA, initial contact of the microprocessor is thought to

be initiated by DGCR8 due its stronger RNA binding affinity compared to DROSHA (Nguyen *et al.*, 2015; Kwon *et al.*, 2019). The Rhed domain, dependent on hemin, is responsible for interaction between the two DGCR8 proteins and facilitates correct positioning of the microprocessor relative to the apical loop (Barr *et al.*, 2012; Weitz *et al.*, 2014; Partin *et al.*, 2017; Nguyen *et al.*, 2018; Dang *et al.*, 2020). Once in position, the microprocessor spans the full length of the hairpin (figure 1.2C) (Kwon *et al.*, 2016). Three out of four dsRBDs of the two DGCR8 molecules interact with the apical hairpin stem, albeit in different conformations (Partin *et al.*, 2020). The CTDs of both DGCR8 molecules interact with DROSHA and are essential for microprocessor stability (Nguyen *et al.*, 2015, 2018; Kwon *et al.*, 2016; Partin *et al.*, 2020). With the central domain of DROSHA at the basal end, the microprocessor acts as a molecular ruler, with dsRBDs of both DGCR8 and DROSHA proposed to be involved in stem measurement, resulting in DROSHA cleaving the hairpin ~13 nt from the 5' basal junction (figure 1.3A) (Zeng, Yi and Cullen, 2005; Kwon *et al.*, 2019; Partin *et al.*, 2020). Due to the structural conformational restraints of DROSHA, the two RNase-III domains are positioned in a way that cleaving results in a 2 nt overlap on the hairpin (Jin *et al.*, 2020).

Both members of the microprocessor have been reported to be essential for pri-miRNA cleaving *in vivo*, since knockdown of either DGCR8 or DROSHA leads to depletion of miRNA expression (Gregory *et al.*, 2004; Han *et al.*, 2004). *In vitro*, however, DROSHA is capable of cleaving pri-miRNAs without DGCR8, albeit less effective and specific, and it aggregates easily (Nguyen *et al.*, 2015). Therefore, a minimal chimeric protein resulting from the fusion of the CTD of DGCR8 to DROSHA can be used for cleaving studies (Kwon *et al.*, 2016).

1.2.2 Pri-miRNA requirements for effective cleaving

Although there are over 400,000 RNA hairpin structures found in the human transcriptome, between ~500 and 3000 of these structures are annotated as pri-

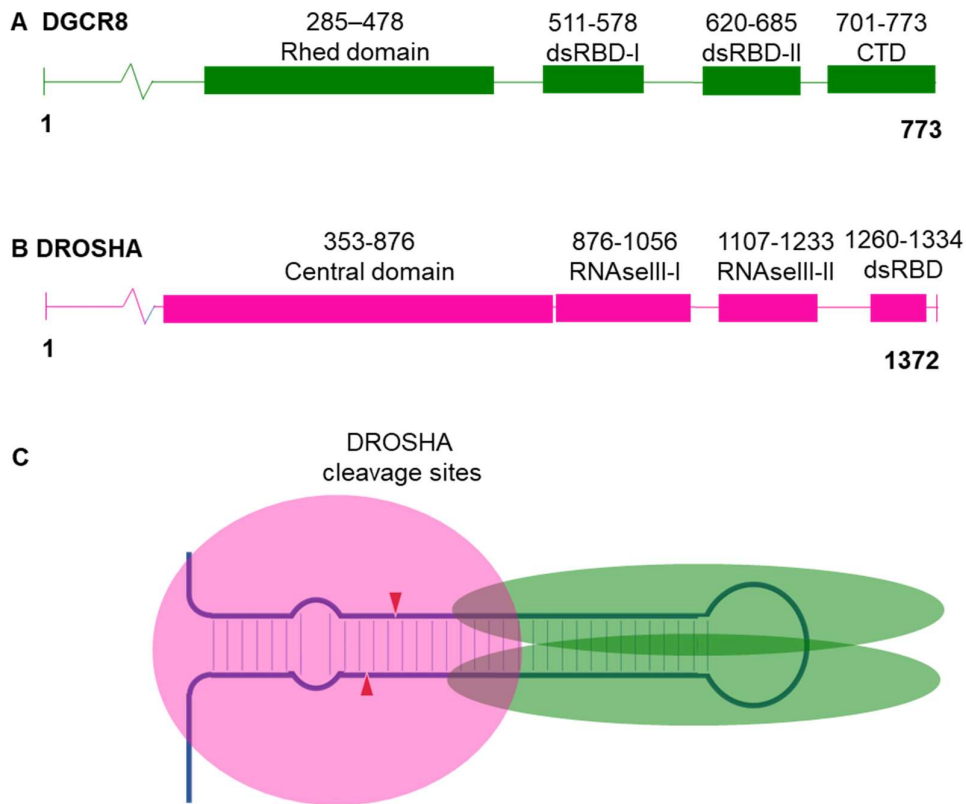


Figure 1.2. Components of the microprocessor. A. Domains of human DGCR8. B. Domains of human DROSHA. C. Structure of the microprocessor in contact with pri-miRNA.

miRNAs (Friedländer *et al.*, 2014; Fromm *et al.*, 2020; Kang *et al.*, 2021). A study in which exogenous miRNA precursors were transfected in cells showed that miRNA precursors are most likely to be processed when originating from or closely related to the model species, indicating convergent evolution of miRNA biogenesis (Lee *et al.*, 2003; Fang and Bartel, 2015; Kang *et al.*, 2021). Even within species there is a wide variety in efficacy of miRNA biogenesis. Over the years, efficacy of microprocessor cleaving has been studied extensively, and several factors affecting pri-miRNA cleaving have been found (figure 1.3).

Stem and hairpin length

The stem length of a pri-miRNA hairpin is an important factor in microprocessor recognition. Since DROSHA and DGCR8 bind to opposite sides of the hairpin, there are restraints on stem length for effective cleaving (Kwon *et al.*, 2019). A range of stem lengths is tolerated, with an optimum of ~ 35 bp. Changing the stem length of existing hairpins can lead to alternative DROSHA cleaving (figure 1.3A) (Fang and Bartel, 2015; Roden *et al.*, 2017). Restrictions for loop size of the hairpin seem less stringent, with edited loops consisting of 3-23 nt still being effectively processed (Zeng and Cullen, 2003). However, loops consisting of 8-16 nt are enriched in existing pri- miRNAs, suggesting this range is optimal for microprocessor recognition (Roden *et al.*, 2017).

Sequence motifs

Apart from structural requirements, several nucleotide sequence patterns are strongly enriched in pri-miRNA hairpins and are beneficial for processing: a 5' basal UG motif; a UGU motif at the 5' apical loop junction; a mismatched GHG motif in the basal stem; and a CNNC motif in the 3' flanking region (figure 1.3B). Depending on the features of the hairpin, motifs can either enhance processing of hairpins already having favourable characteristics or be essential in facilitating processing of suboptimal hairpins that otherwise would not be cleaved (Fang and Bartel, 2015). Furthermore, inverse cleaving of the hairpin occurs more frequently in pri-miRNAs without sequencing motifs, indicating that they also play a role in hairpin orientation (Kim *et al.*, 2021).

The three most basal motifs, UG, GHG and CNNC, act as an anchor for DROSHA interaction and therefore their position can directly influence the location of DROSHA cleaving (Fang and Bartel, 2015; Kim *et al.*, 2018, 2021; Kwon *et al.*, 2019). The UG motif is situated over the 5' basal junction of the pri-miRNA hairpin. DROSHA directly interacts with this site, and the UG motif is predicted to assist in determining the directionality of microprocessor loading, thereby improving

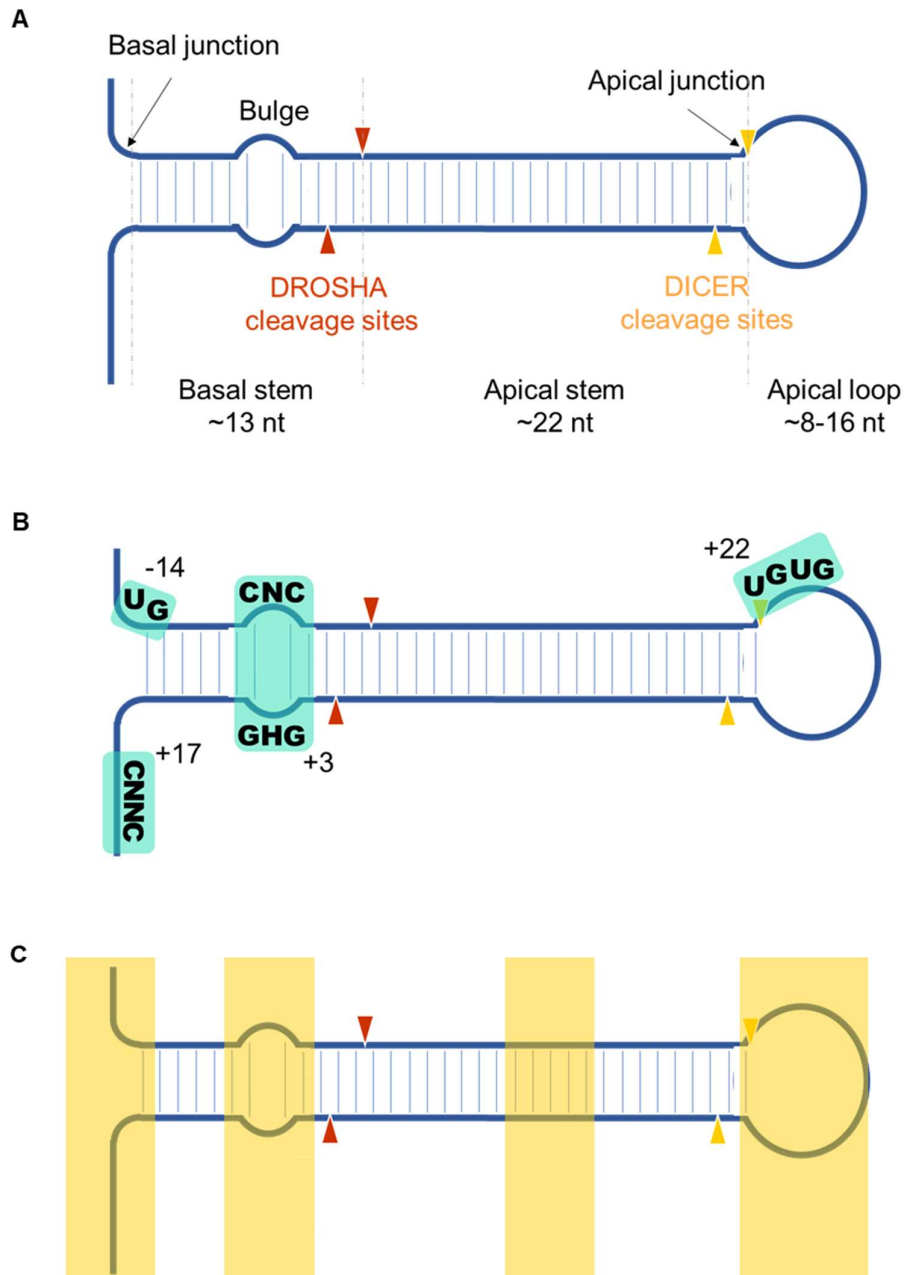


Figure 1.3. Factors impacting pri-miRNA processing. A. Optimal stem and hairpin length. B. Motifs associated with effective processing. The location of the motifs is presented as the number of nucleotides upstream or downstream from the closest DROSHA cleavage site. Figure adapted from (Fang and Bartel, 2015) C. Tight and loose (in yellow) regions associated with effective processing. Figure adapted from (Kang et al., 2021).

microprocessor efficacy (Auyeung *et al.*, 2013; Nguyen *et al.*, 2015). The GHG motif, located 3 nt downstream of the 3' DROSHA processing site, is present in only a small fraction of hairpins but has been reported to enhance microprocessor cleaving. The CNNC motif is situated 17 nt downstream of the 3' DROSHA cleavage site, and can be recognised by the splicing factor SRSF3, which subsequently recruits DROSHA to the basal junction (Kim *et al.*, 2018).

The UGU motif, also often occurring as UGUG or GUG, sits at the junction of the 5' stem and the apical loop (Auyeung *et al.*, 2013). DGCR8 binds to this region with its Rhed domain as first contact between microprocessor and pri-miRNA (Kwon *et al.*, 2019; Dang *et al.*, 2020). Although deleting this motif decreased binding to the microprocessor, shifting the UGUG motif does not lead to alternative Drosha cleaving *in vitro*, indicating the motif is not influencing DROSHA fidelity (Kwon *et al.*, 2019; Dang *et al.*, 2020).

Structure of the hairpin stem

Mutagenesis studies looking at processing efficacy showed that the stem of pri-miRNA hairpins is the region most sensitive to mutations (Rice *et al.*, 2020). While an *in vitro* study indicated that Shannon entropy, related to the number of potential alternative structures, was an important determinant of microprocessor efficacy (Rice *et al.*, 2020). However, these findings were not replicated *in vivo* (Kang *et al.*, 2021). Instead, hairpins with a stable lower basal stem were found to be processed more effectively and show higher mature miRNA levels. Furthermore, a combination of tight and loose regions in the stem, as determined by local free energy, proved most favourable for effective processing (figure 1.3C) (Kang *et al.*, 2021; Kim *et al.*, 2021). These regions ensure optimal contact with the microprocessor and are enriched in conserved miRNAs. They are caused by deviations from Watson and Crick base pairing in the stem, resulting in bulges. Bulges 6 nt upstream and 10-12 nt downstream from the 5' DROSHA cleavage site generally increase processing and these structures are enriched in existing pri-

miRNAs (Han *et al.*, 2006; Fang and Bartel, 2015; Roden *et al.*, 2017; Kang *et al.*, 2021). On the other hand, the presence of asymmetric bulges reduced processing efficacy (Roden *et al.*, 2017; Li *et al.*, 2020; Nguyen *et al.*, 2020; Rice *et al.*, 2020).

Auxiliary factors and cluster assistance

Apart from characteristics of the hairpin sequence itself, several auxiliary factors are associated with processing of specific pri-miRNAs. In ~14% of pri-miRNA hairpins, the apical loop has a conserved structure (Michlewski *et al.*, 2008). These loops were shown to be the target of several RBPs, among which HnRNPA1. Binding of this protein either improved pri-miRNA processing due to structural rearrangement of the hairpin, or decreased processing by blocking access to stimulatory factors, depending on the pri-miRNA (Michlewski and Cáceres, 2010). Several other proteins and long non-coding (lnc)RNAs have been reported to positively or negatively influence microprocessor cleaving, either by recruiting the microprocessor or by stimulating processing; by blocking the microprocessor; or by unknown mechanisms (Michlewski and Cáceres, 2019). When mechanisms are described, they are often reported for single or a small subset of hairpins and remain the focus of ongoing research.

Around 37% of pri-miRNAs are transcribed as part of clusters, and recent research revealed that these clusters are enriched for pri-miRNAs with suboptimal hairpins for microprocessor cleaving (Altuvia *et al.*, 2005; Shang *et al.*, 2020). It is proposed that neighbouring pri-miRNA hairpins more suitable for cleaving recruit the microprocessor and the suboptimal hairpins get cleaved through local transfer (Shang *et al.*, 2020). The protein ERH has been proposed to be important in this cluster assistance within pri-miRNA hairpins up to 900 bp apart (Fang and Bartel, 2020; Kwon *et al.*, 2020).

In conclusion, although previous studies focussed on structure, motifs and folding energy have contributed to a more comprehensive view of pri-miRNA processing, the model for explaining variations in processing is still incomplete. A better

understanding of the features for effective processing allows design of artificial pri-miRNAs resulting in miRNAs with specific targets or improve or decrease the processing of specific pri-miRNAs, offering potential opportunities for clinical interventions in cases where miRNA expression is dysregulated or unbalanced.

1.2.3 Measuring microprocessor efficacy

To this date, a variety of methods have been applied to study the microprocessor and pri-miRNA cleaving. X-ray crystallography and cryo-EM structures of DGCR8 and DROSHA, separately or in complex, with a variety of substrates, elucidated the different protein domains that interact with specific regions of the hairpin (Sohn *et al.*, 2007; Senturia *et al.*, 2010; Jin *et al.*, 2020; Partin *et al.*, 2020).

Studies focussing on pri-miRNAs initially focused on a single or small subset of pri-miRNAs, and mutagenesis of specific sequences were useful in finding novel or miRNA-specific factors influencing processing. In recent years, the focus has shifted to high-throughput methods aiming to determine general features in all pri-miRNAs instead of a small subset, offering an unbiased approach. Mature miRNA expression is not necessarily the response of effective microprocessor cleaving, therefore studying processing of pri-miRNAs with low mature expression might offer new insights. Furthermore, a recent high-throughput study found that processing efficacy correlated with the number in the miRNA name, indicating a publication bias favouring well expressed miRNAs (Rice *et al.*, 2020).

As an alternative to wet lab studies, prediction of features contributing to cleaving can be predicted *in silico*. Comparisons of hairpin structures in the human genome by *in silico* folding studies confirmed known and proposed novel features that are enriched in pri-miRNA hairpins (Roden *et al.*, 2017).

Microprocessor effectiveness is often biochemically examined *in vitro* (Fang and Bartel, 2015; Rice *et al.*, 2020). While these studies are useful for determining

processing in optimal conditions, they do not consider the complexity of other auxiliary factors required for *in vivo* processing. Furthermore, they could lead to artificial results, such as microprocessor cleavage of mirtrons and cleaving of only a single arm of that pri-miRNA hairpin (Rice *et al.*, 2020; Kim *et al.*, 2021). For this reason, several methods aim to study microprocessor efficacy in cells, either by transfection of exogenous miRNA precursors or by studying endogenous cleaving (Conrad *et al.*, 2014; Kang *et al.*, 2021).

The microprocessor index (MPI) is an estimator for *in vivo* pri-miRNA processing efficacy (Conrad *et al.*, 2014). Chromatin-associated RNA is extracted and sequenced, and the read depth of sequences mapping to pri-miRNA hairpins are determined. In canonical miRNA biogenesis, the pre-miRNA is released from the pri-miRNA precursor by the microprocessor and subsequently exported out of the nucleus. For pri-miRNAs that are cleaved co-transcriptionally, the remaining pri-miRNA flanks are retained in the nucleus and are still associated with chromatin after processing. The MPI relies on this mechanism and determines processing efficacy by comparing the ratio of reads mapping to pre-miRNAs with the reads mapping to their direct flanks. In case of effective processing, pre-miRNA sequences should have less read depth than their flanks, while in ineffective processing read depth is evenly spread over the pri-miRNA. Because the MPI reveals processing in cells as opposed to *in vitro* cleavage studies, this approach is particularly suitable for directly studying the consequences of nuclear processing on the expression of mature miRNAs.

1.2.4 Noncanonical functions of the microprocessor

Both DROSHA and DGCR8 are involved in cellular processes besides miRNA biogenesis. Cross-linking and immunoprecipitation (CLIP) studies with DGCR8 and DROSHA showed that the microprocessor can cleave RNA classes other than pri-miRNAs (Macias *et al.*, 2012; Kim, Jeong and Kim, 2017). Interestingly, the

microprocessor can cleave the mRNA of *DGCR8* itself, thereby creating a negative feedback loop (Han *et al.*, 2009; Karginov *et al.*, 2010). Furthermore, both coding and noncoding genes were found to be associated with the microprocessor, with 30% of targets being repetitive elements (Macias *et al.*, 2012). Additional study revealed that the microprocessor protects against retrotransposition of LINE-1 and Alu elements, and thereby contributes to genome stability (Heras *et al.*, 2013).

In addition to mediating cleaving of RNAs, both DROSHA and DGCR8 have been reported to directly alter transcription by binding to chromatin. Chromatin immunoprecipitation (ChIP) analysis suggested that the microprocessor associates with promotor regions of specific genes (Gromak *et al.*, 2013). Knockdown of *DROSHA* resulted in decreased transcription of associated genes. This led to a model of the microprocessor directly interacting with Pol II and thereby positively regulating gene expression. Besides stimulation of transcription, the microprocessor has also been shown to mediate transcription termination of long non-coding RNAs that contain miRNA sequences, thereby preventing transcriptional interference (Wagschal *et al.*, 2012; Dhir *et al.*, 2015).

Apart from functioning as part of the microprocessor, DGCR8 was found to be involved in cleaving of snoRNAs in a DROSHA-independent manner, suggesting that DROSHA is not the only nuclease DGCR8 can associate with (Macias *et al.*, 2012). Further research revealed that DGCR8 recruits the exosome to snoRNAs and thereby influence snoRNA turnover (Macias *et al.*, 2015). Furthermore, in human embryonic stem cells (hESCs) DGCR8 was found to associate with heterochromatin independently of DROSHA and proposed to maintain its stability, with loss of intact DGCR8 resulting in decreased proliferation and cellular senescence (Deng *et al.*, 2019).

Besides the roles described here, DGCR8 and the microprocessor play a role in inflammation and differentiation, which will be discussed in section 1.4.

1.3 INNATE ANTIVIRAL IMMUNITY

1.3.1 The mammalian type-I interferon response

In mammalian antiviral immunity, the type-I interferon (IFN) response plays a central and crucial role. It relies on recognition of pathogen-associated molecular patterns (PAMPs) by specialised pattern recognition receptors (PRRs), inducing a signalling pathway leading to the expression of type-I IFNs, a group of cytokines. There are 8 subtypes of mammalian type-I IFNs (IFN- α , IFN β , IFN- κ , IFN- δ , IFN- ϵ , IFN- τ , IFN- ω , and IFN- ζ), but IFN- α and IFN β are expressed by most cell types and have by far been the most studied (Pestka, Krause and Walter, 2004; Swiecki and Colonna, 2011; McNab *et al.*, 2015). Secreted type-I IFNs can stimulate the IFN response in both infected and neighbouring cells, initiating a signalling cascade that leads to an overall shift in cellular metabolism caused by expression of Interferon-stimulated genes (ISGs), resulting in an antiviral state.

There are several different types of PRRs, recognizing different PAMPs and situated at different locations in the cell (figure 1.4). In the cytosol, the DExD/H-box helicase RIG-I-like receptors (RLRs) retinoic acid-inducible gene I (RIG-I, also called DEAD Box Polypeptide 58) and Melanoma Differentiation-Associated Gene 5 (MDA5) are specialised in recognizing dsRNAs, which could indicate viral infection due to a dsRNA genome or as an intermediate of viral replication. While MDA5 mainly recognises long dsRNA strands between 0.5 and 7 kb, RIG-I preferably binds to sequences containing a 5'-triphosphate and diphosphate cap, suggesting that they are involved in recognition of different viruses (Kato *et al.*, 2006; Peisley *et al.*, 2012; Reikine, Nguyen and Modis, 2014). A third RLR, LGP2, is thought to have a regulatory function, but studies about its mechanism of action show contradicting results (Sato *et al.*, 2010; Malur, Gale and Krug, 2012; Goubau, Deddouche and Reis e Sousa, 2013; Bruns and Horvath, 2015; Quicke, Diamond and Suthar, 2017). Recognition of dsRNAs by RLRs results in oligomerisation of the RLRs along the nucleic acid strand. These polymers associate with the adapter Mitochondrial

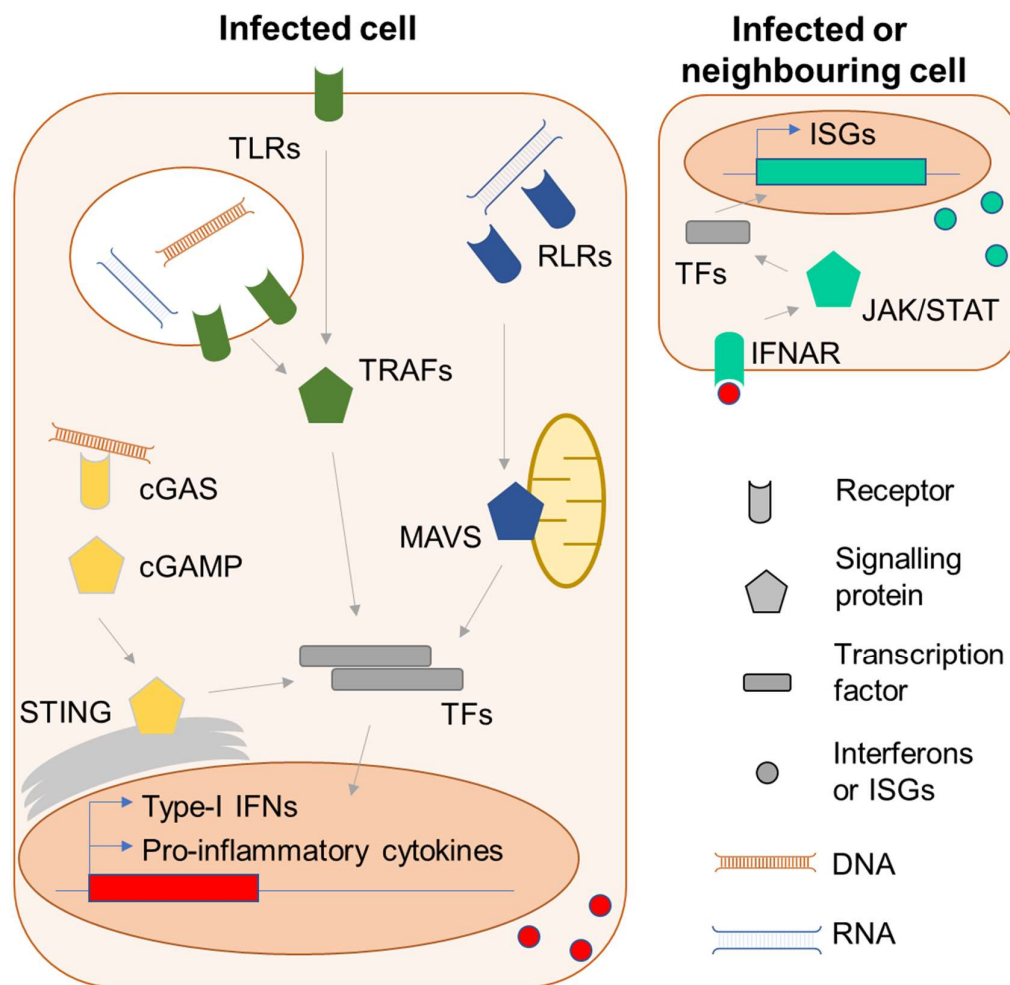


Figure 1.4. Overview of the mammalian IFN response. The cytosolic dsRNA-sensing pathway is displayed in blue; The cytosolic DNA-sensing pathway is displayed in yellow; The TLR pathway is displayed in green. The second wave of IFN signalling is displayed in turquoise. Figure based on (McNab *et al.*, 2015)

Antiviral Signalling Protein (MAVS, also known as IPS-1, IFN- β promoter stimulator 1) and subsequently lead to activation of TANK-binding kinase 1 (TBK1), which phosphorylates transcription factors to drive type I IFN gene expression (figure 1.4). RLRs are ubiquitously expressed in the cytoplasm in absence of infection, but their expression is additionally stimulated by IFN signalling, resulting in a positive feedback loop (Yoneyama *et al.*, 2005).

Apart from RLRs, several other DExD/H-box helicases have been proposed to be involved in viral nucleic acid sensing and regulation of type-I IFNs (Goubau, Deddouché and Reis e Sousa, 2013; McNab *et al.*, 2015). cGAMP synthetase (cGAS) recognises cytosolic DNA and subsequently catalyses cyclic guanosine monophosphate–adenosine monophosphate (cGAMP) production, which function as a ligand for stimulator of IFN genes (STING).

Besides cytoplasmic PRRs, viral nucleic acids can be detected by a class of transmembrane proteins called Toll-Like Receptors (TLRs) (Thompson *et al.*, 2011). In endosomes, TLR 3 and 9 can recognise dsRNA and unmethylated CpG DNA, respectively, while TLR 7 and 8 recognise single stranded RNAs (ssRNAs). Apart from functioning in antiviral immunity, specialised TLRs can also recognise hallmarks of bacterial infection, such as lipopolysaccharide (LPS) for TLR4. Downstream signalling from TLRs occurs via either myD88 or TRAF adaptor proteins, depending on the type of TLR (McNab *et al.*, 2015).

Ultimately, activation of PRRs in the first wave of type-I IFN signalling leads to activation of the transcription factors NF- κ B, Interferon Regulatory Factor (IRF)3 and/or IRF7 and subsequent expression and secretion of type-I IFNs and pro-inflammatory cytokines. As a result, type-I IFNs can be recognized by the heterodimer receptor complex consisting of IFNAR1 and IFNAR2 on the transmembrane of IFN-producing and their surrounding cells. This initiates the second wave of the type-I IFN response, which is mediated by Janus Kinase (JAK)/Signal Transducer and Activator of Transcription (STAT) signalling and results in the expression of ISGs. There are hundreds of different reported ISGs, the majority of which have unknown functions and which expression is highly dependent on the nature and condition of infection (Schoggins *et al.*, 2011; Schoggins, 2014). ISGs with known functions are involved, among others, in pathogen recognition, translation inhibition and post-transcriptional modifications and contribute to an overall antiviral state (Bowie and Unterholzner, 2008). Since

some ISGs are also involved in the first wave of in the first wave of type-I IFN signalling, a positive feedback loop is created, strengthening the antiviral response.

Apart from PRRs associated with the type-I IFN response, there are several other molecules associated with antiviral activity. Upon encountering dsRNA, The cytoplasmic Protein Kinase R (PKR) undergoes autophosphorylation and induces an overall translational shutoff, impacting both viral and endogenous transcripts (Zhang *et al.*, 2001; Dalet, Gatti and Pierre, 2015). Exceptions to this shutoff are IFNs and several ISGs, therefore viral replication is prevented but antiviral activity is stimulated (Watson, Bellora, & Macias, 2019). Furthermore, the 2'-5'-Oligoadenylate Synthetase Like (OASL)/RNaseL pathway can directly recognise and cleave dsRNA (Han *et al.*, 2012). Although these pathways are not part of the first or second waves of type-I IFN signalling, they are ISGs and their actions contribute to the IFN response (Malathi *et al.*, 2007).

Besides inducing a cell-intrinsic response in infected tissues, expression of type-I IFNs also leads to activation of the adaptive immune system. By stimulating antigen presentation on innate immune cells, such as dendritic cells and macrophages, effector T cells are activated (Marrack, Kappler and Mitchell, 1999; Ivashkiv and Donlin, 2014). Furthermore, type-I IFNs lead to B-cell activation and proliferation, thereby stimulating pathogen-specific antibody production and long term immunological memory (Braun, Caramalho and Demengeot, 2002).

1.3.2 Effects of aberrant type-I IFN expression

As described above, activation of the type-I IFN response has an anti-proliferative effect and results in the release of pro-inflammatory cytokines. This response is necessary to clear viral infections, but should be strictly regulated, since uncontrolled or excessive signalling leads to immunopathology.

A class of genetic diseases called interferonopathies demonstrates the importance of controlled IFN expression. Gain of function mutations in the RLR genes, MDA5 and RIGI, result in interferonopathies, characterised by intrinsic expression of IFN- β in absence of infection and are associated with developmental disorders (Funabiki *et al.*, 2014; Oda *et al.*, 2014; Rice *et al.*, 2014; Rodero and Crow, 2016; Gorman *et al.*, 2017; Chung *et al.*, 2018). Furthermore, due to their role in stimulating antigen presentation and expression of pro-inflammatory cytokines, IFN signalling is associated with autoimmune disorders such as rheumatoid arthritis, systemic lupus erythematosus, and Sjögren's syndrome (Meyer, 2009; Hall and Rosen, 2010; Ivashkiv and Donlin, 2014). Currently, several treatments targeting the type-I IFN pathway are being used or under development to treat these disorders (Kirou and Gkrouzman, 2013; Chaichian, Wallace and Weisman, 2019; Duguet *et al.*, 2021).

Recent studies concerning infection with SARS-CoV-2 highlight the contrasting effects of IFN signalling. On the one hand, congenital mutations that lead to aberrant type-I IFN signalling or absence of an IFN signature are associated with severe COVID-19 infection (Hadjadj *et al.*, 2020; Zhang *et al.*, 2020). On the other hand, excessive immune signalling can induce a “cytokine storm” with detrimental consequences (Ragab *et al.*, 2020; Akamatsu *et al.*, 2021). Treatment with IFN β in hospitalised patients showed either positive or no effects on disease severity and morbidity, depending on the study (Alavi Darazam *et al.*, 2021; Kalil *et al.*, 2021; WHO Solidarity Trial Consortium, 2021).

In conclusion, while type-I IFN signalling is essential in mounting an initial antiviral response, continuous or inappropriate timing of signalling is ineffective or detrimental for recovery. Therefore, the expression of IFNs and ISGs should be tightly regulated. Intervening in IFN signalling holds potential in assisting with viral clearance and for reducing immunopathology in autoimmunity and severe infections.

1.3.3 Antiviral siRNAs

As mentioned before, siRNAs are a class of small RNAs involved in RNAi. While humans and most nematodes only have one gene coding for DICER, most arthropods have two variants, Dicer-1 and Dicer-2 (Gao *et al.*, 2014). While Dicer-1 is similar in structure to mammalian DICER, Dicer-2 contains an extra helicase domain and is involved in the biogenesis of siRNAs (Y. S. Lee *et al.*, 2004). Dicer-2 can directly cleave cytoplasmic dsRNAs, targeting stem-loop structures and dsRNA hybrids originating from repetitive sequences (Okamura *et al.*, 2008; Watanabe *et al.*, 2008; Piatek and Werner, 2014). Mature siRNAs can be loaded onto RISC containing Ago2, which cleaves endogenous RNAs with exact complementarity (Chung *et al.*, 2008). By cleaving these repetitive sequences, including transposons, Dicer-2 contributes to genome stability. However, Dicer proteins can also function as antiviral factors. In plants, fungi and invertebrates, by directly cleaving the viral nucleic acids, Dicer displays a dual antiviral function: firstly by directly destroying the viral genome and preventing translation, and secondly by creating virus-derived siRNAs (vsiRNAs) that guide AGO-mediated cleaving of the viral sequences (Hamilton and Baulcombe, 1999; Ding and Voinnet, 2007; Dang *et al.*, 2011; Bronkhorst and Van Rij, 2014). Disruption of this RNAi pathway leads to increased viral susceptibility in *D. melanogaster*, indicating its importance as an immune protein (Ding and Voinnet, 2007).

Initially, mammals were thought to have no siRNAs, since the presence of dsRNAs in the cytoplasm could trigger an immune response. Indeed, several studies have failed to detect vsiRNAs in human cells (Backes *et al.*, 2014; Bogerd *et al.*, 2014; Schuster *et al.*, 2019). Since then, it has been shown that an N-terminal helicase domain of human DICER inhibits its activity, while deletion of this domain enables siRNA synthesis (Ma *et al.*, 2008; Bogerd *et al.*, 2014). However, endo-siRNAs have since been reported in oocytes, mouse ESCs, and somatic cells, targeting both mRNAs and transposable elements (Babiarz *et al.*, 2008; Tam *et al.*, 2008; Watanabe *et al.*, 2008; L. Chen *et al.*, 2012)

In mice oocytes, an N-terminal truncated isoform of Dicer lacking part of the helicase domain, caused by expression of a retrotransposon-derived promoter, has been detected (Flemr *et al.*, 2013). This oocyte-specific Dicer shows higher cleavage activity and can mediate siRNA synthesis. Recently another endogenous DICER isoform, lacking the central Hel2i domain, was detected in both human and mice (Poirier *et al.*, 2021). This isoform was expressed at low levels in diverse human and mouse cell types, but at higher levels in stem cells, and expression was correlated with resistance to viral infection. Interestingly, this isoform was also able to process pre-miRNAs, indicating that the presence of the Hel2i domain is not essential for miRNA biogenesis.

Furthermore, vsiRNAs have been detected in infection with mutant virus lacking RNAi-suppressing properties, in undifferentiated mouse cells and in somatic mouse cells (Li *et al.*, 2013, 2016; Maillard *et al.*, 2013). It has been suggested that antiviral RNAi may only be of relevance in mammals in absence of more effective antiviral mechanisms, such as an interferon response. However, no consensus has been reached on this topic as of yet.

1.3.4 The IFN response is lost or gained dependent on pluripotency status

As described above, almost all mammalian cells have active type-I IFN signalling upon infection or stimulation with viral mimics. A striking exception are certain cell types associated with pluripotency and early development.

Both human and mouse (ESC) lines derived from pre-implantation blastocysts are unable to produce type-I IFNs upon stimulation. H9 hESCs failed to express *IFNb* after stimulation with high concentrations of the dsRNA mimic poly I:C, while HeLa cells, in comparison, expressed *IFNb* even at low poly I:C concentrations (Chen, Yang and Carmichael, 2010). Compared to HeLa, hESCs showed less cell death and less stress granule formation after poly I:C transfection, suggesting that the cells are protected from the toxic effect of cytokine release that comes with IFN signalling

(Chen, Yang and Carmichael, 2010). Similarly, mouse ESCs also showed no detectable IFN β upon poly I:C transfection (Wang *et al.*, 2013). Neither stimulation with LPS, TNF α or viral infection resulted in type-I IFN production in mESCs, suggesting that not only the RLR pathway, but also other pathways in the first wave of IFN signalling are inactive (D'Angelo *et al.*, 2017)

Interestingly, after differentiation into trophoblast cells with BMP4, H9 cells were able to express *IFNb* upon stimulation (Hong and Carmichael, 2013). Likewise, mESC-derived fibroblasts were able to express *ifnb* mRNA after either poly I:C transfection or infection with La Crosse virus (D'Angelo *et al.*, 2016). This suggests that the ability to express IFNs is acquired during embryonic development and differentiation. In mice, expression of IFNs was detected post-implantation of the trophoblast, on the 7th day after fertilisation (Barlow, Randle and Burke, 1984).

Despite the inability to produce IFNs, it is unclear whether mESCs are capable of responding to exogenous IFN, with studies showing contradicting results, indicating a potentially difference in regulation of the second wave of type-I IFN signalling between human and mouse (Whyatt *et al.*, 1993; Hong and Carmichael, 2013; Eggenberger *et al.*, 2019).

The lack of IFN production in mammalian cells is not limited to ESCs, and was first reported in murine teratocarcinoma cells (Burke, Graham and Lehman, 1978). Importantly, induced pluripotent stems cells (iPSCs), created by reprogramming of somatic cells, were also shown to have an aberrant IFN response. IMR-90 human fibroblast cells reprogrammed to iPSCs showed a strongly attenuated IFNb expression compared to regular IMR-90 cells upon poly I:C transfection (Chen, Yang and Carmichael, 2010). Furthermore, the same cells and iPSCs de-differentiated from foreskin fibroblasts fail to express ISGs upon incubation with exogenous IFN- β , suggesting that the second wave of IFN signalling is inactive as well (Hong and Carmichael, 2013; Eggenberger *et al.*, 2019).

These findings suggests that the activation of type-I IFNs response is reversible and dependent on the pluripotency status.

1.3.5 Possible mechanisms and reasons for attenuated type-I IFN response

Multiple mechanisms have been proposed that explain the absence of type-I IFN production during pluripotency. As described above, dsRNA can be sensed in multiple ways: by RLRs, TLR3, PKR or RNaseL. In hESC H9 cells, RNaseL and MDA5 protein are barely expressed compared to HeLa cells (Chen, Yang and Carmichael, 2010). Furthermore, while unphosphorylated (inactive) PKR levels are unaffected, phosphorylated PKR cannot be detected in H9 cells after poly I:C stimulation, indicating that it is not initiating its antiviral properties. Both *TLR3* and its downstream partner *TRIF* are expressed at extremely low levels, suggesting this pathway is not active either. Furthermore, STING is not expressed, suggesting multiple pathways leading to IFN expression, upon viral DNA or RNA sensing, are downregulated. In mESCs, RLRs were reported to be expressed at low levels or to be inactive compared to fibroblasts, thereby being unable to initiate signalling (Wang *et al.*, 2013). However, how this downregulation of signalling molecules is mediated is not known.

Several alternative innate immune mechanisms have been proposed to function in absence of IFN signalling, including RNAi. It has been suggested that ESCs are protected against viral infection by the maternal placenta, or by maternal IFNs, and therefore do not need an IFN response (Hong and Carmichael, 2013; Delorme-Axford, Sadovsky and Coyne, 2014; Wang *et al.*, 2014). Alternatively, several ISGs were found to be intrinsically expressed at high levels in hESCs, thereby circumventing the need for first and second wave IFN signalling (X. Wu *et al.*, 2018). Finally, it has been hypothesised that activation of the type-I IFN response in stem cells might have undesirable consequences during early development (Eggenberger *et al.*, 2019).

IFN activation by endogenous PRR substrates

The dsRNA sensing pathway is most often described for its role in antiviral immunity. However, knockdown of the RNA editing enzyme ADAR 1 in neuronal progenitor cells leads to MDA5-dependent IFN activation, indicating that it is possible for endogenous nucleotides to bind PRRs and initiate immune signalling (Chung *et al.*, 2018). Since ESCs show both retrotransposon and endogenous retrovirus expression, there is a risk of activating the IFN response (Peaston *et al.*, 2004; van den Hurk *et al.*, 2007; Kano *et al.*, 2009; Grow *et al.*, 2015).

Interestingly, siRNAs derived from endogenous transcripts have been reported in mouse oocytes and mESCs (Babiarz *et al.*, 2008; Tam *et al.*, 2008; Watanabe *et al.*, 2008). Additionally, the reports of vsiRNA presence in several pluripotent cell models suggest that antiviral RNAi might be functioning as an alternative immune mechanism in the pluripotent context. In line with this, inactivation of the IFN response in differentiated mouse cells led to Dicer and Ago-dependent gene silencing mediated by dsRNA (Maillard *et al.*, 2016). Furthermore, pre-treatment with virus-specific dsRNA protected these cells against viral infection. These findings suggest that antiviral RNAi and the type-I IFN response are mutually exclusive, or that the presence of the IFN response masks antiviral RNAi.

Cell death and disruption of differentiation

As described above, differentiated H9 cells are able to produce type I IFNs when stimulated, indicating that the ability is acquired sometime around implantation of the blastocysts and trophoblast formation (Barlow, Randle and Burke, 1984; Hong and Carmichael, 2013). During this stage, cells are rapidly dividing and follow a tightly regulated differentiation pattern. IFN signalling, with the release of inflammatory and anti-proliferating cytokines, would inhibit successful differentiation (Hertzog, Hwang and Kola, 1994).

Alternatively, it has been proposed that IFN signalling could result in disruption of differentiation. mESCs that were transfected with poly I:C showed a temporary

decrease in pluripotency markers and an increase in differentiation markers that correlated with a weak expression of *ifnb* at the transcript level, suggesting that even slight IFN signalling could dysregulate pluripotency pathways (Wang *et al.*, 2013). Eggenberger *et al.* forced iPSCs to express IFNs using an inducible IRF7 mutant and subsequently differentiated the cells into the different germ layers. Compared to non-IFN expression iPSCs, transcriptional changes were observed in germ layer markers, particularly for ectoderm (Eggenberger *et al.*, 2019). These observations suggest that pluripotency and IFN signalling are mutually exclusive. The mechanism behind IFN activation during differentiation or repression during pluripotency have yet to be elucidated.

1.4 MiRNAs and the microprocessor in immunity and development

1.4.1 The role of miRNAs in innate immunity

Several miRNAs have been predicted or are confirmed to directly mediate the production of type-I IFNs by targeting transcripts involved in innate immunity, among which in the first or second wave of type-I IFN signalling (reviewed in Li and Shi, 2013; Sedger, 2013; Forster, Tate and Hertzog, 2015; Nejad, Stunden and Gantier, 2018). For example, both *hsa-miR-125a* and *b* family members have been reported to target *MAVS* (Hsu *et al.*, 2017). On the other hand, for several miRNAs the expression levels change upon IFN induction. As of 2015, 36 miRNAs were known to be IFN regulated in humans (Forster, Tate and Hertzog, 2015). Interestingly, many of these miRNAs target transcripts that are themselves involved in the IFN response, and are therefore thought to function as a negative feedback response (Taganov *et al.*, 2006; O'Connell *et al.*, 2007; Y. Chen *et al.*, 2013).

While some miRNAs have known targets and have been extensively studied, most miRNAs have not been thoroughly studied. Adding to complexity, many miRNAs have the capability to target multiple transcripts involved in the same pathway or sometimes have both immunostimulatory and immunosuppressant functions, further obscuring the relevance of each miRNA during infection.

In the context of antiviral immunity, miRNAs can also indirectly affect viral sensitivity, for example by targeting ERK/MAPK signalling (Santhakumar *et al.*, 2010). In addition, given the importance of miRNAs in immunity, several viruses have been reported to inhibit miRNA function, by preventing RISC assembly or stimulation of miRNA degradation (Bennasser, Yeung and Jeang, 2006; Backes *et al.*, 2012). On the other hand, specific miRNAs have been reported to have pro-viral functions, either by directly interacting with the host genome or by targeting endogenous factors (Jopling *et al.*, 2005; Chiang, Liu and Rice, 2013; O'Connor, Vanicek and Murphy, 2014; Ziv *et al.*, 2018).

Studying mutations of miRNA biogenesis factors bypasses the complexities of relevance of individual miRNAs. These studies have proved valuable in establishing biologically relevant relationships between miRNAs and the type-I IFN response. A homozygous hypomorphic Dicer mutation in mice resulted in the upregulation of several ISGs, including *Oas1a* and *Ifit1*, as measured by microarray (Ostermann *et al.*, 2012). These findings coincided with a decrease in miRNA expression, implicating an overall role for miRNAs in IFN signalling. Similarly, knockdown of DICER in human endometrial cancer cell lines resulted in increased steady state levels of *IFNb* and several ISGs, suggesting increased first wave IFN signalling (Chiappinelli *et al.*, 2012).

1.4.2 The role of the microprocessor in innate immunity

The microprocessor itself has also been directly implicated with innate immunity. It has been suggested that the DROSHA, a nuclear protein, is able to translocate to the

cytoplasm to directly cleave viral RNAs, thereby directly decreasing viral load (Shapiro *et al.*, 2012, 2014).

Importantly, work in our lab has shown that the type-I IFN response negatively influences microprocessor cleaving, as measured by the MPI (Witteveldt, Ivens and Macias, 2018). Transfection of HeLa cells with poly I:C resulted in a decrease in processing in 37% of the pri-miRNAs, while only 8% of pri-miRNAs showed increased processing. Treating cells with exogenous IFN β did not result in altered miRNA biogenesis, indicating that this repression is associated exclusively with the first wave of type-I IFN signalling. The UG and CNNC motif were enriched among pri-miRNAs that showed decreased processing upon poly I:C transfection. The UGU motif did not show an enrichment, suggesting that DROSHA or SRSF3 affinity for pri-miRNAs might be decreased, since these are known to bind to the UG and CNNC motifs, respectively (Auyeung *et al.*, 2013; Nguyen *et al.*, 2015; Kwon *et al.*, 2018). Many of the miRNAs whose biogenesis were negatively affected have been implicated to be negative regulators of the IFN response (Witteveldt, Ivens and Macias, 2018). A functional explanation could be that downregulation of these miRNAs are necessary to elicit a strong immune response.

1.4.3 The roles of miRNAs in early development

MiRNAs have been reported to be essential for animal embryonic development, and are regulated with great temporal and spatial specificity (Alberti *et al.*, 2018; Dexheimer and Cochella, 2020). Extensive efforts have elucidate the roles of several individual miRNAs during development in mice (reviewed in DeVeale, Swindlehurst-Chan and Blelloch, 2021). Most of the broadly conserved miRNAs regulate transcripts involved in essential developmental processes, and loss of these miRNAs leads to aberrant development or physiology. For example, members of the *miR-302* family are highly expressed in early embryonic differentiation and are essential for neuronal development. Knockout of all *miR-302* family members is

embryonically lethal due to neural tube close defects and aberrant development of cardiomyocytes (Parchem *et al.*, 2015). *Let-7*, the first miRNA reported to be involved in developments in nematodes, regulates the timing of differentiation and is thought to have an opposing function towards *miR-302* for neuronal development (Reinhart *et al.*, 2000; Melton, Judson and Blelloch, 2010). Furthermore, maintenance and exit of pluripotency via expression or repression of the central regulators *SOX2*, *NANOG* and *OCT4* is heavily dependent on miRNA function (reviewed in Greve, Judson and Blelloch, 2013).

Surprisingly, miRNAs do not play a big role in gene regulation in mouse oocytes (Suh *et al.*, 2010). Recent research showed low levels of miRNAs compared to mRNAs and suggest the miRNA concentrations are diluted during oocyte growth (Kataruka *et al.*, 2020). Fertilization results in an overall genome reprogramming as part of the oocyte-to-zygote transition (Svoboda and Flemr, 2010). This includes a switch from maternal to zygotic miRNAs, with the transcription of most pri-miRNAs starting after the first zygotic cell division (Tang *et al.*, 2007).

The importance of miRNAs in early development is reflected in the severity of the phenotypes observed upon depletion of miRNA biogenesis factors. No cases of complete knockouts of miRNA biogenesis factors have been reported in the germline of humans, presumably due to an embryonically lethal phenotype.

1.4.4 DGCR8 mutants and their effect on development

The mechanism behind the phenotypes caused by DGCR8 depletion is not always clear, and therefore canonical and non-canonical functions of DGCR8 cannot always be distinguished. In mice, *Dgcr8* knockout cells survived until day 6.5 after fertilisation, while Dicer-deficient mouse oocytes failed to undergo the first cell division, suggesting important roles for both miRNAs and miRNA-independent functions of Dicer (Wang *et al.*, 2007; Suh *et al.*, 2010; reviewed in Greve, Judson and Blelloch, 2013).

In contrast to animal models, mouse embryonic cell lines without functional *Dgcr8* or *Dicer* are viable and show regular expression of pluripotency markers (Kanellopoulou *et al.*, 2005; Wang *et al.*, 2007). However, upon stimulation of differentiation, the cells fail to silence these pluripotency markers and are unable to differentiate normally.

Human DGCR8 has been related to inflammation and proliferation via a non-canonical mechanism. Expression of an N-terminal truncated DGCR8 protein in hESCs did not affect miRNA biogenesis or expression of pluripotency factors but affected proliferation of mesenchymal stem cells (Deng *et al.*, 2019). In the same study, overexpression of *Dgcr8* in mice alleviated osteoarthritis by reducing the presence of inflammatory factors. Mutated *Dgcr8*, without functional dsRBDs, had the same effect, suggesting a miRNA-independent mechanism, and was suggested to be due to DGCR8-mediated stability of heterochromatin.

The role of *Dgcr8* and miRNAs in neuronal development has been studied using several mouse models overexpressing, heterozygous for, or conditional knockouts of *Dgcr8*, *Drosha* or *Dicer*. Several miRNAs were found to be important in hippocampus development (Pons-Espinal *et al.*, 2017). Conditional *Dgcr8* ablation led to premature differentiation of progenitor cells into neurons and altered neuron morphology (Stark *et al.*, 2008; Fénelon *et al.*, 2011; Ouchi *et al.*, 2013; Luhur *et al.*, 2014; Marinaro *et al.*, 2017). In some instances, *Drosha* or *Dicer* knockout did not show such a strong phenotype, indicating functions of *Dgcr8* independent of miRNA biogenesis into neuronal development. While overexpression of *Dgcr8* initially repressed neurogenesis, it stimulated proliferation of neuronal progenitor cells at a later stage, indicating different mechanisms during differentiation (Hoffmann *et al.*, 2018). Conditional *Dgcr8* knockouts showed defects in synaptic transmission, showing the effect is not just in neuronal development, but also in mature neurons (Hsu *et al.*, 2012)

Interestingly, *Dgcr8* heterozygosity altered the expression of several genes associated with schizophrenia, which is common in patients with 22q11.2 deletion syndrome (DS) (Ouchi *et al.*, 2013). Since *DGCR8* is one of the genes that is hemizygous in 22q11.2 DS, this indicates a potential role for *DGCR8* in symptoms seen in patients.

1.5 22Q11.2 DELETION SYNDROME

22q11.2 deletion syndrome (DS), also called DiGeorge or velocardiofacial syndrome, is a genetic autosomal dominant developmental syndrome, characterised by a microdeletion on chromosome 22, locus q11.2. It is the most common microdeletion in humans, affecting 1:3.000-1:6.000 live births, and is mainly acquired *de novo* due to aberrant chromosomal rearrangement of low copy repeats during meiosis (Repetto *et al.*, 2014; McDonald-McGinn *et al.*, 2015). The observed microdeletion could be as small as 700 kilobases, but 90% of patients have a 3 megabase deletion that leads to hemizyosity of a total of 90-106 genes, including *DGCR8* (McDonald-McGinn *et al.*, 2015; Motahari *et al.*, 2019; Du, de la Morena and van Oers, 2020).

22q11.2DS leads to predisposition for a variety of specific disorders and developmental delays. Since the exact mechanism causing symptoms observed in 22q11.2DS has not yet been elucidated, treatment is symptomatic, and patients have a decreased life expectancy (Repetto *et al.*, 2014). Patients often show characteristic craniofacial features and abnormal development of specific organs, leading to thyroid and kidney defects, congenital heart disease, and cleft palate, with the latter two occurring in 75% and 42% of the patients, respectively (Rozas *et al.*, 2019; Sellier *et al.*, 2014). Furthermore, a range of neurological and behavioural disorders is observed with much higher frequency in 22q11.2 patients, including attention deficit hyperactivity disorder, schizophrenia, autism spectrum disorders,

and anxiety disorders (Schneider et al., 2014; Sellier et al., 2014). Some of these syndromes are attributed to altered morphology and gene expression in the nervous system, as is demonstrated in 22q11.2 mouse models (Fénelon et al., 2011; Stark et al., 2008). This led to studies into neuronal development, using 22q11.2 deleted iPSCs (Lin et al., 2016; Toyoshima et al., 2016; Zhao et al., 2015). Transcriptomics analysis of iPSC-derived neurons showed a ~2-fold decrease in gene expression of genes in the 22q11.2 locus, and aberrant expression of genes involved in brain development (Lin et al., 2016). Furthermore, 22q11.2DS patient-derived neurospheres showed abnormal differentiation (Toyoshima et al., 2016). All in all, this indicates that the symptoms observed in 22q11.2DS have their origin in early neuronal development. Many of the observed defects can specifically be attributed to alteration in cellular structures associated with or derived from neural crest cells (McDonald-McGinn *et al.*, 2015).

Beside neuronal alterations, 22q11.2DS patients often have an underdeveloped thymus, leading to impaired T-cell production (Sullivan *et al.*, 1998; Piliero *et al.*, 2004). Remarkably, 22q11.2DS patients are more susceptible to infectious diseases, but are also more likely to develop allergies and autoimmunity (Sullivan *et al.*, 1997, 1998; Kornfeld *et al.*, 2000; Kawame *et al.*, 2001; Staple *et al.*, 2005).

Existing mouse models for 22q11.2 have a hemizygous deletion of locus 16qA13, which encompasses many of the genes hemizygous in 22q11.2DS. Mice with these deletions exhibit similar symptoms as observed in 22q11.2DS, indicating that overlapping genes are central players in causing symptoms in 22q11.2DS (Stark *et al.*, 2008; Du, de la Morena and van Oers, 2020). However, in humans, the length of the microdeletion does not correlate with the presence and severity of observed symptoms, even between family members (Rozas *et al.*, 2019; Du, de la Morena and van Oers, 2020). Likewise, individual mice with the same genetic background show heterogenic phenotypes, indicating the syndrome may be aggravated by additional mutations outside the 22q11.2 locus or may not be exclusively genetic (Du, de la Morena and van Oers, 2020). Indeed, several single nucleotide variants or copy

number alterations have been found in 22q11.2 patients in genes coding for miRNAs or epigenetic regulators. Targeted deletions in mouse models identified several other genes involved in development, including members of the *Hox* and *Pax* families, as potential aggravators of or causing symptoms similar to that of 22q11.2 (Pohl *et al.*, 2013; Paganini *et al.*, 2017; Du, de la Morena and van Oers, 2020), although their clinical relevance has yet to be established.

The functions of several the hemizygous genes in 22q11.2DS have been studied, with a particular focus on *TBX1* (McDonald-McGinn *et al.*, 2015; Motahari *et al.*, 2019; Du, de la Morena and van Oers, 2020). Homo- or heterozygous deletions in mice resulted in a variety of symptoms frequently seen in 22q11.2DS, such as cleft palate and thymic defects (Lindsay *et al.*, 2001; Merscher *et al.*, 2001; Zhang *et al.*, 2005; Zhang, Huynh and Baldini, 2006). Furthermore, conditional mutants of *Tbx1* affected neural crest cell patterning, which is important for craniofacial development, although *Tbx1* itself is not detectable in neural crest cells (Vitelli *et al.*, 2002; Trainor, 2005; Calmont *et al.*, 2009; McDonald-McGinn *et al.*, 2015).

Compared to *TBX1*, less research into the contribution of *DGCR8* to 22q11.2 syndrome has been carried out. *DGCR8* expression was found to be decreased in patients with 22q11.2DS (Sellier *et al.*, 2014). When mice heterozygous for 16qA13 or *Dgcr8* were studied, an accumulation of pri-miRNAs was found, suggesting that *Dgcr8* heterozygosity could be responsible for the observed alterations in miRNA biogenesis (Stark *et al.*, 2008). More studies have researched mature miRNA expression in 22q11.2DS patient samples or derived cell models. All these studies showed altered expression of mature miRNAs, although the identity of miRNAs different on species, model, tissue, and detection method (Stark *et al.*, 2008; De la Morena *et al.*, 2013; Zhao *et al.*, 2015; Toyoshima *et al.*, 2016). Interestingly, differentially expressed miRNAs in neurons differentiated from 22q11.2-hemizygous iPSCs were predicted to target mRNAs involved neurological and psychological disorders (Zhao *et al.*, 2015).

These results indicate that *DGCR8* is a relevant gene to study in the context of 22q11.2DS.

1.6 AIMS OF THIS PHD PROJECT

MiRNAs are essential in the post-transcriptional control of gene expression and, as such, play important roles in most biological processes. For instance, embryonic development is a setting where the functional relevance of regulation by miRNAs is well-documented. Interestingly, early mammalian embryonic development, from where ESCs are derived, do not have a functional type-I IFN response, raising questions about the potential mechanisms of repression or initiation, and the biological relevance of this response in early development.

This thesis focusses on the relationship between the miRNA biogenesis factor *DGCR8*, early development and innate immunity through the following aims:

1. To study the transcriptome of a human *DGCR8* ^{-/-} cell line.

To our knowledge, as of this moment, no stable knockout of *DGCR8* in a human cell line has been reported to this date, although cell lines containing truncated *DGCR8* have been published (Nguyen et al., 2018; Deng et al., 2019). However, our collaborators have created a human PA-1 *DGCR8* knockout cell line. Since canonical miRNA biogenesis is absent in these cells, the transcriptome is expected to be heavily affected. The effects of *DGCR8* knockout on RNA expression and its functional consequences and cellular metabolism will be discussed in chapter 3. Studying mature miRNA levels will also help to identify microprocessor-independent miRNAs and will be discussed in chapter 4.

2. To create and study the transcriptome of a human *DGCR8* ^{+/-} cell line.

The effect of *DGCR8* heterozygosity is understudied, whereas it has clear clinical relevance in the context of 22q11.2 DS. The creation of *DGCR8* +/- cell lines and the consequences on coding RNA expression are discussed in chapter 3, while the effect on miRNA expression is discussed in chapter 4.

3. To determine the effect of *DGCR8* mutation on miRNA biogenesis efficiency. Inactivating one of both copies of *DGCR8* is expected to affect microprocessor functioning and therefore pri-miRNA processing. To determine microprocessor efficiency, an adapted version of the Microprocessor Processing Index (MPI) (Conrad *et al.*, 2014) will be used. This aim will be discussed in chapter 4.
4. To study the effect of *DGCR8* knockout and heterozygosity on innate immunity and differentiation potential in pluripotent cell lines. Mutating *DGCR8* is expected to have biologically relevant consequences for antiviral immunity and the differentiation potential of pluripotent cell lines. Both mESCs and human pluripotent cell lines are used to study these effects, that will be the focus of chapter 5.

CHAPTER 2: MATERIALS AND METHODS

2.1 CELL CULTURE

All cell lines were grown at 37°C with 5% CO₂ in either T25 or T75 cell culture flasks (Corning) except where indicated. Cells were split 2-3 times per week, by briefly washing with PBS, incubating with 0.025% trypsin (Gibco) in PBS at 37 °C until cell detachment and subsequent resuspension in fresh medium.

PA-1 wildtype and *DGCR8* knockout cell lines were acquired from Sara Heras' lab (GenYo, Granada). Mutant cell lines were created using CRISPR-Cas by inducing two different knicks using the Cas-9 D10A mutant (Ran *et al.*, 2013) in exon 2, the first coding exon of *DGCR8* (Pablo Tristan, Heras lab).

All PA-1 cell lines were cultured in DMEM + Glutamax (ThermoFisher) supplemented with 20% FBS (ThermoFisher), 1x non-essential amino acids (ThermoFisher) and 1x penicillin/streptomycin (ThermoFisher). Although wildtype PA-1 cells can be cultured in 10% serum, the *DGCR8* knockout cells were very sensitive to splitting and showed better adhesion and less cell death in medium with 20% serum. For this same reason, 10 μ M of the apoptosis inhibitor iRock (Sigma) was added to the medium after splitting, keeping all cell lines in the same condition. The iRock was removed within 24 hours and at least 24 hours before treatment or harvesting by replacing the media.

V6.5 wildtype (Novus Biologicals) and v6.5 *Dgcr8* knockout (ThermoFisher) mouse ESC lines were cultured on plates coated with 0.1% Gelatine (Embryomax, Millipore) in DMEM supplemented with 15% FBS (Thermo Fisher), non-essential amino acids (Thermo Fisher), 100 units/mL penicillin and streptomycin, 2mM L-glutamine, 40 μ M β -mercapto-ethanol (all ThermoFisher) and 1,000 units/mL lymphocyte inhibitory factor (LIF; Merck).

2.2 CRISPR-CAS9-BASED GENOME EDITING

2.2.1 Design and preparation of Cas9-guideRNA mix

Genomic targets for Cas9 were selected using the Alt-R CRISPR-Cas9 System software (IDT), based on high efficiency and low off-target scores (table 2.1). 245 pmol crRNA and 245 pmol tracrRNA (both IDT) in were incubated in 25 uL nuclease-free buffer for 5 minutes at 95 °C and allowed to anneal at room temperature until cooled. The resulting guideRNA was incubated with 25 ug Alt-R® S.p. Cas9 Nuclease V3 (IDT) in 22.5 uL buffer T (Neon transfection system, ThermoFisher) for 15-25 minutes at 37°C immediately prior to transfection.

Table 2.1 Overview of genomic mutagenesis in cell lines using CRISPR-Cas9.

Cell line targeted	gDNA target	crRNA + <u>PAM</u> target sequence (5' to 3')	Resulting cell line(s)
v6.5 wildtype	<i>mmu-miR-673</i>	CCAGAGAAAATGTTGCTCCG <u>GGG</u>	<i>mmu-miR-673</i> knockout 1
v6.5 wildtype	<i>mmu-miR-673</i>	ACCAGAGCTGTGAGCCCCTC <u>AGG</u>	<i>mmu-miR-673</i> knockout 2 & 3
PA-1 knockout 1	Human <i>DGCR8</i> exon 2	AGGTAATGGACGTTGGACGT <u>TGG</u>	PA-1 heterozygote 1.1 & 1.2
PA-1 wildtype	Human <i>DGCR8</i> exon 10	ACATCAGCATCGAGGACTCG <u>CGG</u>	PA-1 heterozygote 2

2.2.2 Transfection of cells

Cells were cultured under regular conditions before transfection, and cell medium was replaced 4 hours prior. In case of PA-1 cells, 10 uM iRock was added to the medium. After splitting and washing, 1.2×10^6 cells were resuspended in 80 uL buffer

T (Neon™ Transfection System, ThermoFisher). Subsequently, the Cas9/guideRNA mixture was added to the cells, and, if applicable, 300 pmol of repair template (supplementary table). Cells were then electroporated with 3 pulses of 1600V and a pulse width of 10 milliseconds (Neo Transfection System, ThermoFisher) and subsequently plated in a T25 plate (Corning).

2.2.3 Clonal expansion of single cell colonies

Transfected mESCs were sorted as single cells in 96-well plates 3 days after transfection. Transfected PA-1 cells did not survive sorting, and single cell cultures were obtained by splitting them in 96-well plates using limiting dilution. Cells that successfully grew colonies were split and expanded until able to grow in a 6-well format, upon which the colonies were tested for successful mutagenesis. For PA-1 cells, 10 uM iRock was continually added to the culture medium from transfection up to the point where clones were able to grow in a 24 well format.

2.2.4 Screening of clones

Genomic DNA from the clones of interest was extracted by incubating cells in lysis buffer (30mM Tris HCl; 10mM EDTA; 0.1% SDS; 0.5% Tween 20; 10ug/ml Proteinase K) for 15 minutes at room temperature at a concentration of 20ul/100,000 cells. Subsequently, the mixture was transferred to PCR tubes and heated to 57 °C for 10 minutes and then at 98°C for 10 minutes. From this mixture, 0.2 uL was taken for amplification of the sequence of interest by PCR, using Taq polymerase (NEB) according to manufacturer's instructions. Primers were designed to anneal >200 nucleotides from the predicted site of mutagenesis (table B). Amplification and size of fragments were confirmed by 1% agarose gel electrophoresis. Fragments of interest were excised and purified using the QIAquick Gel Extraction Kit (Qiagen), ligated into pGEM-T easy vector (Promega), transformed into DH5α cells (NEB)

using heat shock (NEB High Efficiency Transformation Protocol) and incubated overnight on LB agar plates (10 g/L tryptone; 5 g/L yeast extract; 10 g/L NaCl; 1.5% agar) containing 100 µg/mL ampicillin. White colonies were selected and grown overnight in LB medium (10 g/L tryptone; 5 g/L yeast extract; 10 g/L NaCl; 100 µg/mL ampicillin). Plasmids were isolated with the QIAprep Spin Miniprep Kit (Qiagen) and sent for sequencing using the plasmid sequencing service from either Edinburgh Genomics or MRC PPU (Dundee). The returned sequencing results were analysed using Serial Cloner 2.6.1.

2.3 POLY(I:C) TRANSFECTIONS

For PA-1 cells, poly(I:C) transfections were carried out in a 6-well format, using 1 µg/mL high-molecular weight poly(I:C) (Invivogen) and 8 µl Lipofectamine 2000 Reagent (Invitrogen) according to manufacturer's instructions. Cells were harvested 6 hours after transfection.

mESCs were transfected with using 2.5 µg/mL high-molecular weight poly(I:C) (Invivogen) and 2 µl Lipofectamine 2000 Reagent (Invitrogen) in a 24-well format and were harvested 4 hours after transfection.

2.4 VIRAL INFECTIONS

Viral infections were carried out using Theiler's murine encephalomyelitis virus (TMEV) GDVII (obtained from Peter Simmonds' lab, Oxford). Viral stocks were prepared by Jeroen Witteveld with an estimated concentration of 35 million pfu/ml for wildtype v6.5 cells (Witteveldt, Knol and Macias, 2019). These stocks were diluted either 10, 100 or 1,000-fold and directly added to 60% confluent mESCs in 24-well plates to an estimated MOI of 1.5, 0.15 or 0.015, respectively. The virus-containing medium was replaced after 1 hour and cells were harvested 24 hours

after infection by removing the medium and directly adding the required amount of TRI reagent (Sigma). The plates were subsequently frozen at -80 °C until continuation of RNA extraction, described in section 2.7.

2.5 RETINOIC ACID DIFFERENTIATION

Differentiation of PA-1 cell lines took place in coated 6-well plates. For this, the surface of plates was covered with 10 µg/mL poly-L-ornithine (Sigma) and incubated for 2 hours at 37 °C. To improve cell annealing, plates were subsequently rinsed twice with sterile water and coated with 20 µg/mL laminin (Invitrogen) in depleted PBS (without Ca⁺⁺ and Mg⁺⁺; ThermoFisher) for 2 hours at 37 °C. After rinsing the plates once with depleted PBS they were stored overnight at 4 °C in depleted PBS prior to cell plating.

Cells were resuspended in normal culture media without iRock and seeded on the coated plates. After 4 hours, cells for timepoint day 0 were harvested and the medium of the remaining cells was replaced by differentiation medium, and later on replaced every two days until harvesting. The differentiation medium consisted of MEM + Glutamax (ThermoFisher), 10% KnockOut Serum replacement medium (Invitrogen), 1x non-essential amino acids (Thermo Fisher), 1x penicillin and streptomycin (ThermoFisher) and 1 µM of trans-only retinoic acid (Sigma). Every two days, images of the cells were taken with the EVOS M7000 with 20x objective. At the time of harvesting, wells containing 80% confluent cells were collected, pelleted, and stored at -80 °C until processing. In order to obtain enough material for RNA extractions, two wells of cells were pooled per sample and cell lines were plated at specific concentrations on day 0 (table 2.2).

Table 2.2 Overview of cell numbers plated for differentiation assay

Cell lines	Day of harvesting				
	0	2	4	6	8
PA-1 wild type	5x10 ⁵	2x10 ⁵	2x10 ⁵	1x10 ⁵	5x10 ⁴
PA-1 <i>DGCR8</i> heterozygotes	5x10 ⁵	2x10 ⁵	1.5x10 ⁵	1x10 ⁵	1x10 ⁵
PA-1 <i>DGCR8</i> knockouts	5x10 ⁵	5x10 ⁵	2x10 ⁵	1.5x10 ⁵	5x10 ⁴

2.6 CELLULAR FRACTIONATION

To obtain cytoplasmic, nucleoplasmic and chromatin-associated cellular fractions of PA-1 cells, 8x10⁶ PA-1 cells grown under regular culturing conditions were trypsinised, resuspended and washed with PBS. Subsequently, cells were lysed for 5 minutes on ice in mild lysis buffer (500 µL: 10 mM Tris pH 7.4; 150 mM NaCl; 0.075% NP-40) prior to layering over sucrose buffer (1 mL: 10 mM Tris pH 7.4; 150 mM NaCl; 24% sucrose). The mixture was then centrifuged for 5 minutes at 3,500 g and the supernatant was collected as the cytoplasmic fraction. After washing with PBS, the pellet containing the nuclei was resuspended in glycerol buffer (250 µl : 20 mM Tris pH 7.4; 75 mM NaCl; 0.5 mM EDTA; 50% Glycerol) mixed with nuclear lysis buffer (250 µl: 10 mM Tris pH 7.4, 7.5 mM MgCl₂, 0.2 mM EDTA, 0.3 M NaCl, 1 M Urea, 1% NP-40) and vortexed for 4 seconds. The mixture was then incubated on ice for 2 minutes and spun down for 2 minutes at maximum speed in a cooled centrifuge. The supernatant was collected as the nucleoplasmic fraction, while the pellet contained the insoluble chromatin. The latter was washed and resuspended in 500 µL PBS and sonicated (Bioruptor: 5 times 20 sec on/off intervals) prior to shearing the chromatin and release bound RNA and proteins.

2.7 RNA EXTRACTIONS

For both total and small RNA-seq and RT-qPCRs of steady state RNA levels, RNA was extracted from a 60% confluent T75 cell flask (Gibco) with the miRNeasy Mini Kit

(Qiagen), following the protocol “Purification of Total RNA, Including Small RNAs, from Animal Cells”, including the additional on-column DNase digestion.

For RNA extractions following poly(I:C) transfection, differentiation and all mESC RNA purifications, phenol-chloroform extraction was used. Cells were resuspended in TRI reagent (Sigma, 10^6 cells/mL), mixed with 1/5 volume chloroform incubated at room temperature for 15 minutes and centrifuged at $>12,000 \times g$ for 15 minutes at 4°C . Following the resulting phase separation, the upper aqueous phase was collected and mixed with an equal volume of isopropanol to precipitate the RNA. After 5 minutes incubation, the samples were centrifuged at $>12,000 \times g$ for ≥ 15 minutes at 4°C . The isopropanol was removed, and the pellet was washed once with 900 μL 70% ethanol and air dried before resuspension in nuclease-free H_2O .

For RNA extraction of the subcellular fractions, 250 μL of each fraction was mixed with 750 μL Tri-LS Reagent (Sigma) and RNA was extracted according to the manufacturer’s protocol and resuspended in 10 μL H_2O . The chromatin-associated RNA was subsequently DNase treated using RQ1 DNase (Promega) according to manufacturer’s instructions at a total volume of 100 μL . After 15 minutes of DNase incubation at 37°C , the solution was mixed with equal volume of Acid-Phenol:Chloroform (ThermoFisher) and vortexed until the solution turned white. After centrifuging at maximum speed in a cooled centrifuge, 100 μL of the top aqueous solution was mixed with 250 μL Ethanol and 10 μL 3M sodium acetate pH 5.2 (ThermoFisher). The mixture was then spun down at maximum speed in a cooled centrifuge prior to removing the supernatant. The RNA pellet was then washed with 70% ethanol and spin centrifuged for 5 minutes in a cool centrifuge at maximum speed. The RNA pellet was then air dried and resuspended in H_2O .

After purification, RNA concentration and purity was measured on a nanodrop photospectrometer (ThermoFisher).

2.8 RT-qPCR

For qPCR of samples stemming from total RNA of PA-1 cells, 750-1500 ng RNA was reverse transcribed using the Transcriptor Universal cDNA master (Roche) with random hexamers according to the manufacturer's protocol. For PA-1 cellular fractions, 200 ng RNA from each fraction was used for the reverse transcriptase reaction. Prior to qPCR, the cDNA mixture was diluted 1/5 (for differentiation assays and cellular fractions) or 1/10 (all other assays). QPCR was carried out in a 384-well format using LightCycler480 SYBR Green I Master Mix (Roche) on a LightCycler 480 Instrument II (Roche) according to manufacturers' instructions.

For mESCs, 200-500 ng RNA was used for cDNA synthesis. The RNA was incubated with 250 ng random hexamers for 5 minutes at 70°C and cooled on ice immediately afterwards. Subsequently, cDNA was synthesised using M-MLV (Promega) according to manufacturer's instructions but without addition of ribonuclease inhibitor, in a total volume of 12.5 µL. The resulting cDNA was diluted 1/10 and specific transcripts were quantified in a 96-well plate format on a StepOnePlus light cycler (ThermoFisher) using GoTaq mastermix (Promega).

All novel primers used for qPCR were designed using Serial Cloner (v2.6.1) or the NIH primer blast tool (<https://www.ncbi.nlm.nih.gov/tools/primer-blast>) with a T_m around 60°C (supplementary table). Validation of qPCR products occurred by inspection on an agarose gel or melt curve analysis.

All qPCR reactions were carried out in technical duplicates. Samples that showed a big variety (~ >1 CT value) between technical duplicates were excluded from subsequent analysis. Quantification of transcripts was carried out using the second derivative method. The average relative expression of household gene transcripts *GAPDH* and *ACTINB* were used as normalisers for PA-1 samples. For mESCs, murine *Gapdh* was used as normaliser.

2.9 RT-QPCR FOR miRNAs

For quantification of miRNAs in PA-1 clones, the miRCURY LNA miRNA PCR system (Qiagen) was used. For each RT reaction with the miRCURY LNA RT Kit, 200 ng RNA was used. The resulting cDNA was stored undiluted at -20°C and diluted 1/60 immediately before the qPCR reaction, which was carried out using the miRCURY LNA SYBR® Green PCR Kit (Qiagen) with corresponding primers on a LightCycler® 480 Instrument II (Roche) following the miRCURY LNA protocol. No spike-ins were used in the assay.

All qPCR reactions were carried out in technical duplicates and analysed with the second derivative method. Since the used normaliser U6 showed wide unexplained variability between samples and the RNA was measured and normalised prior to RT, no normaliser was used in this case. An alternative normaliser should be considered for future studies.

2.10 PROTEIN EXTRACTION AND WESTERN BLOT

Cells were lysed in chilled lysis buffer (approximately 10^6 cells/100 µL: 50 mM Tris pH 7.5; 150 mM NaCl; 1 mM EDTA; 1% Triton X-100; 0.5 mM DTT; 0.2 mM PMSF; 1 tablet protease inhibitor cocktail (Roche)/10 mL) for 30 minutes on ice. Afterwards the samples were centrifuged in a cooled centrifuge at $>12,000 \times g$ for 10 minutes and the supernatant containing proteins was transferred to a new chilled tube.

Protein concentration was measured using the Qubit™ Protein Assay Kit (ThermoFisher) and normalised protein samples were mixed with recommended ratios of reducing agent and LDS sample buffer (Novex, ThermoFisher). Samples were heated to 70°C for 10 minutes prior to loading on a 4-12% Bis-Tris gel (NuPAGE, Invitrogen) in MOPS buffer (Invitrogen). Contrastingly, as cellular fractions were already lysed, they were directly mixed with reducing agent and LDS sample buffer and loaded on a gel in equal volumes.

Proteins from PA-1 cells were transferred to PVDF membranes for 105 minutes at 100V in a gel tank containing 1X NuPAGE Transfer Buffer (Invitrogen) with 10% methanol, while mESC samples were transferred to a nitrocellulose membrane by semi-dry transfer on an iBlot2 (ThermoFisher). Prior to blocking, the membrane contain cellular fractions was stained with Ponceau to visualise successful transfer and equal loading. The membrane was then rinsed with sterile H₂O using agitation.

Membranes were subsequently blocked for ≥30 minutes in PBS-0.05 % Tween containing 5% milk with agitation at room temperature and incubated with primary antibody in PBS-0.05 % Tween containing 5% milk (table 2.3) overnight at 4° C with agitation. Afterwards, membranes were washed with PBS-0.05 % Tween and incubated with secondary antibody (table 2.3) in PBS-0.05 % Tween containing 5% milk for 4 hours at room temperature with agitation. Chemiluminescence using horseradish peroxidase was carried out using either Pierce ECL (ThermoFisher; for highly abundant proteins) or SuperSignal (ThermoFisher; for lowly abundant proteins) on a ChemiDoc imager (Bio-Rad) and analysed with Image Lab software (v6.0.1; Bio-Rad).

2.11 LIBRARY PREPARATION AND SEQUENCING OF TOTAL, SMALL, AND CHROMATIN-ASSOCIATED RNA

To determine transcriptome and miRNA biogenesis changes, total RNA-seq, small RNA-seq and chromatin-associated RNA-seq were carried out for three cell lines, with four replicates for each: *DGCR8* wild type, heterozygote (cell line 1.1) and knockout (cell line 1).

Library preparation, including adaptor ligation, reverse transcription and PCR amplification, and sequencing were carried out by BGI Genomics. For both total and chromatin-associated RNA sequencing, >200 ng of total RNA was used for libraries with ribosomal RNA removal prior to DNA nanoball synthesis and sequencing on

Table 2.3 Overview of antibodies used for western blotting

Antibody target	Supplier (catalog nr)	Origin	Dilution	MW (kDa)
α -Tubulin	Calbiochem (CP06)	Mouse monoclonal	1:10000	50
DGCR8 N-terminus	Abcam (ab90579)	Rabbit polyclonal	1:1000	~100
DGCR8 C-terminus	Abcam (ab191875)	Rabbit monoclonal	1:1000	~100
Drosha	Novus (NBP1-03349)	Rabbit polyclonal	1:1000	~170
Histone H3	CST (4499)	Rabbit monoclonal	1:2000	17
MDA5	CST (5321)	Rabbit monoclonal	1:1000	140
MAVS (E6)	SCB (sc- 365334)	Mouse monoclonal	1:100- 1:1000	~90
Rabbit IgG (HRP)	CST (7074)	Goat polyclonal	1:2500	
Mouse IgG (HRP)	Bio-Rad	Goat polyclonal	1:2500	

DNBseq PE100 (Drmanac *et al.*, 2010; Li *et al.*, 2018). Small RNA-seq libraries were made using unique molecular identifiers (UMIs) and sequenced using DNBseq (Li *et al.*, 2018). All sequencing reads were pre-filtered by BGI, with adapters, low quality reads and contaminants being removed, and data was returned in compressed FASTQ format. Quality control for all three datasets by FastQC (v0.11.5) (Andrews, 2010) showed that reads were of high quality and no further quality filtering was required.

2.12 ANALYSIS OF TOTAL RNA-SEQ

Paired-end reads of 100 nts were aligned to the human genome (GRCh38.p13) using HISAT2 (v2.1.0) with options `--no-discordant --no-mixed --no-unal` (figure 2.1A; Kim, Paggi, Park, Bennett, & Salzberg, 2019). Transcript counts for each sample were created with featureCounts (Rsubread v2.2.6), using reverse counts and excluding reads that aligned to the genome multiple times (Liao, Smyth and Shi, 2014). Normalisation and statistical analysis was carried out with DESeq2 (v1.32.0) , using apegIrm for fold change shrinkage (Love, Huber and Anders, 2014; Zhu, Ibrahim and Love, 2019). RNAs with an adjusted p-value <0.05 and a shrunk absolute log2 fold change >0.5 were subsequently inputted in EnrichR (v3.0) for enrichment analysis (E. Y. Chen *et al.*, 2013; Kuleshov *et al.*, 2016; Xie *et al.*, 2021). Information about transcript in specific pathways was taken from the KEGG database (v99.0; Kanehisa, Goto, Furumichi, Tanabe, & Hirakawa, 2009).

2.13 ANALYSIS OF SMALL RNA-SEQ

For each sample, identical small RNA-seq reads were collapsed and only reads that were present more than once were included in further analysis (figure 2.1B).

Mature miRNAs were identified and counted using miRDeep2, using the quantification function with human hairpin and mature miRNA sequence files from miRbase (v22.1) and allowing no mismatches (Friedländer *et al.*, 2012; Kozomara and Griffiths-Jones, 2014; Kozomara, Birgaoanu and Griffiths-Jones, 2019). In addition, the small RNA samples were analyzed with blastn (v2.6.0+) using the commands `-num_threads 20 -max_target_seqs 1 -max_hsps 1 1 -outfmt '6 std qseq sseq' -task blastn -word_size 6 -dust no` in order to identify small RNAs (Camacho *et al.*, 2009).

DESeq2 was used for statistical analysis using apegIrm as log fold-change shrinkage model, using the miRDeep2 counts as input (Love, Huber and Anders, 2014; Zhu,

Ibrahim and Love, 2019). Instead of using DESeq2's built-in normalization, the counts were normalized to the amount of tRNAs, relative to WT sample 1, identified in each sample in the blastn analysis using the normalizationFactors option of DESeq2.

Iteration of statistical analysis occurred eliminating the miRNAs with the overall lowest expression until all miRNAs analyzed passed the independent filtering threshold with $\alpha = 0.1$. In doing this, overall lowly expressed miRNAs were filtered up to the point that no miRNAs were removed from statistical analysis due to low read count, resulting in expression data for 700 mature miRNAs. MiRNAs with an FDR lower than 0.05 were considered significantly differentially expressed.

Pathway analysis of predicted targets for differentially expressed miRNAs (adjusted p-value <0.01) was carried out with mirPath (v3) utilizing Tarbase (v7) for KEGG pathways, using Fisher's exact test and a FDR-corrected p-value threshold of 0.05 (Vlachos *et al.*, 2015).

2.14 ANALYSIS OF CHROMATIN-ASSOCIATED RNA-SEQ AND MPI

CALCULATION

Paired-end sequencing data from chromatin-associated RNA samples were aligned to the human genome (GRCh38.p13) using HISAT2 (version 2.1.0) with the options -no-discordant --no-mixed --no-unal (figure 2.1C; Kim *et al.*, 2019).

Human pre-miRNA locations were determined by aligning human precursor sequences, obtained using the mature miRNA and hairpin sequences from miRbase (version 22.1) against the same genome using Bowtie2 with options --very-sensitive --no-unal (Langmead and Salzberg, 2012). The genomic locations and coverage of the pre-miRNAs plus 100 nts on each side were determined by Alisdair Ivens using bedtools getfasta -s.

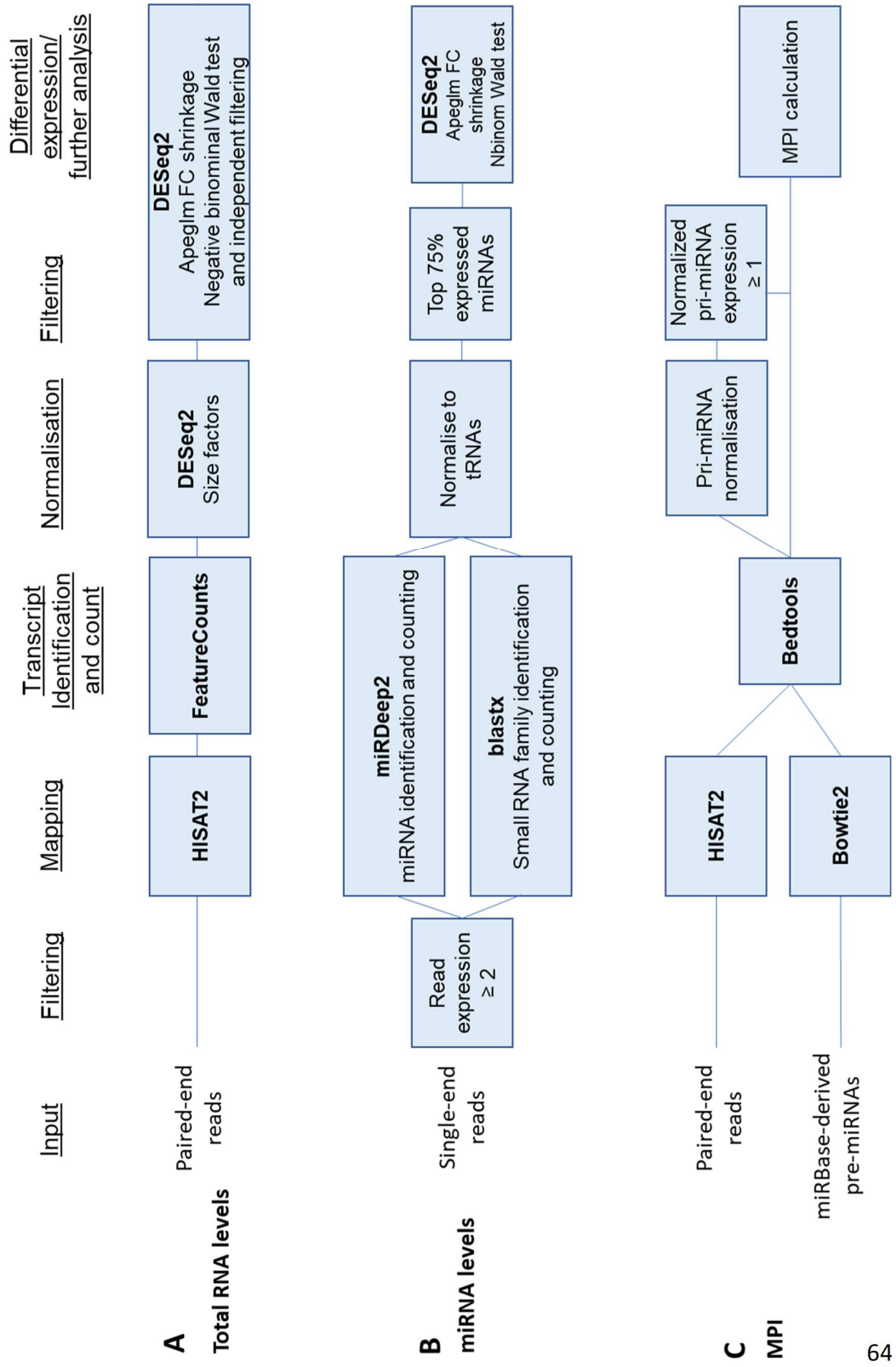


Figure 2.1. Flowchart of bioinformatic analysis. A total-RNA-seq. B small-RNA-seq. C Chromatin-associated RNA-seq.

A 10 nt gap between each pre-miRNA and its flanking regions was created by excluding 5 nts from the flanking regions and both sides of the pre-miRNA, creating leeway for potential alternative DROSHA cleaving. For each region, the average read depth of all included nucleotides was used for MPI calculation. To filter out Microprocessor-independent pri-miRNAs, reduce noise and exclude artifacts, pri-miRNAs were only included if they: were no mirtrons or tailed mirtrons as annotated in USCSC Genome Browser (Fujita et al., 2011; GRCh38/hg38); showed expression of a mature miRNA in the small RNA-seq analysis; had an overall normalised flank read depth < 1 ; if the ratio between the two flanks differed more than 4-fold.

The MPI calculation was considered invalid if either the pre-miRNA sequence or the flank sequences had no reads, making it impossible to calculate the MPI. When there were three either or four valid MPI calculations per cell line for each cell line, the mean MPI was included. Calculating the normalised flank read depth was done by normalising flank counts from each sample by its total flank read depth; this was done independently of and did not affect MPI calculation.

2.15 MIRTARRNASEQ

To identify possible miRNA-mRNA interactions, the R package MirTarRnaSeq (v1.0.0) was utilised. As input, the analysis from the small RNA-seq and total RNA-seq with DESeq2 were used: for target RNAs, the shrunken $\log_2$ fold changes of transcripts with an adjusted p-value < 0.05 were included. For miRNAs, the shrunken $\log_2$ fold changes of all 700 miRNAs were imported. Fold change differences between miRNAs and potential targets were calculated using twoTimePoint and

twoTimePointSamp with 10 rounds was used to make a background distribution. Using this background distribution, miRNA-target RNA correlations with a p-value <0.05 were combined with the Miranda database for human miRNA-target interactions using mirandaIntersectInter. The resulting MirTarRnaSeq score indicates the correlation between miRNA and target transcript over the replicates in case $p < 0.05$.

CHAPTER 3: TRANSCRIPTOMIC
CHARACTERIZATION OF *DGCR8* KNOCKOUT
AND HETEROZYGOTE PA-1 CELL LINES

3.1 INTRODUCTION

DGCR8 is an RNA binding protein and part of the nuclear microprocessor complex, which can bind and cleave endogenous RNAs, including pri-miRNAs (Macias *et al.*, 2012; Macias, Cordiner and Cáceres, 2013). Absence of a functional DGCR8 protein is lethal in mammals, complicating studies of the role of DGCR8 *in vivo*. Both miRNA-deficiency and non-canonical functions of DGCR8 are thought to lead to a disruption of differentiation (Wang *et al.*, 2007; Cirera-Salinas *et al.*, 2017; DeVeale, Swindlehurst-Chan and Blelloch, 2021). While the microprocessor has been studied extensively since its discovery, most studies focused on the biochemical aspects of the microprocessor, including structural models and *in vitro* cleaving of miRNAs (Sohn *et al.*, 2007; Senturia *et al.*, 2010; Partin *et al.*, 2020; Rice *et al.*, 2020). While these experiments bring valuable knowledge about the molecular basis of microprocessor cleaving, they do not elucidate the functional relevance of miRNA biogenesis and its non-canonical functions on a cellular or organismal level.

Due to the lethal phenotype in absence of canonical miRNAs, cell lines or animal models with knockouts for miRNA biogenesis factors are often difficult to obtain (Suh *et al.*, 2010; Nguyen *et al.*, 2018). Therefore, in many animal models, Dicer or *Dgcr8* function is studied by using conditional knockouts using Cre-Lox recombination, making it possible to study the local effects of DGCR8 depletion. These conditional deletions of DGCR8 seriously disrupted kidney, cardiomyocyte, glial cell and neuronal function, showing the organism-wide importance of DGCR8 and miRNAs (Rao *et al.*, 2009; Hsu *et al.*, 2012; Bartram *et al.*, 2015; Lin *et al.*, 2015; Marinaro *et al.*, 2017).

While a mESC line knocked out for *Dgcr8* is available (Wang *et al.*, 2007), creating a stable human *DGCR8* knockout cell line has proven to be difficult so far, probably due to the essential function of DGCR8. Previous attempts to target exon 2 of *DGCR8* in HCT116 cells proved unsuccessful, however a HCT116 cell line with a C-terminal truncated DGC8 protein is available (Nguyen *et al.*, 2018). Deletion of the

CTD of DGCR8 abolished DROSHA interaction, and therefore a functional microprocessor. Furthermore, a *DGCR8* knockout was successfully created in human ESCs by knocking out exon 3, but these cells showed abnormal karyotypes (Deng *et al.*, 2019). Disruption of exon 2 led to hESCs that were unable to maintain pluripotency. Due to these concerns both knockout cell lines were deemed unsuitable for further study.

In our approach to study the role of DGCR8 in immunity in early development, PA-1 cells were used as a model cell line for pluripotency. While these cells have limited differentiation capability, they are genetically more stable and easier to culture than hESCs (Zeuthen, 1981). PA-1 *DGCR8* knockout cell lines were provided by Sara Heras' group (Genyo, Granada). To our best knowledge, this is the first reported human cell line with complete ablation of DGCR8 that does not spontaneously differentiate.

In addition to characterisation of this *DGCR8* knockout cell line, we also aimed to create *DGCR8* heterozygote PA-1 cells to examine the role of *DGCR8* heterozygosity in 22q11.2 DS. The transcriptome of wildtype and mutant cell lines was researched using RNA-seq and subsequent differential expression analysis.

3.2 RESULTS

3.2.1 Creation of PA-1 *DGCR8* heterozygous cell lines

PA-1 *DGCR8* knockout cell lines were created by Sara Heras' group (Genyo, Granada), using CRISPR-Cas with a Cas9 9 D10A nickase mutant and targeting the first coding exon (exon 2), resulting in two cell lines with different mutations (figure 3.1a). Cell line *DGCR8* knockout 1 has insertions in both *DGCR8* alleles: a 7-nucleotide insertion, and a 60-nucleotide insertion. *DGCR8* knockout 2 cells have a

22-nucleotide insertion and an indel in their respective *DGCR8* alleles. The mutations in both cell lines led to a frameshift and an early stop codon.

In addition to using these cell lines, we were interested in *DGCR8* heterozygosity. To create *DGCR8* heterozygous cell lines, two different strategies were carried out: repairing one of the mutated alleles in a knockout cell line; and a novel *DGCR8* targeting strategy in wildtype cells.

Firstly, the creation of *DGCR8* heterozygous cell lines was attempted by repairing one of the mutated alleles in *DGCR8* knockout cell line 1. To achieve this, Cas9 and a guide RNA were used. CRISPR-Cas is dependent of complementarity of a guideRNA to a 20 nt protospacer in the target DNA (Jinek *et al.*, 2012; Cong *et al.*, 2013; Ran *et al.*, 2013). Directly next to the protospacer lies a protospacer adjacent motif (PAM) of sequence NGG. Binding of the guideRNA to the protospacer recruits the endonuclease Cas9 and mediates cleaving of both genomic DNA strands in the protospacer sequence. These double-stranded breaks are subsequently repaired by non-homologous end joining (NHEJ), often resulting in deletions and insertions and premature stop codons when genes are targeted (Jinek *et al.*, 2012; Cong *et al.*, 2013; Ran *et al.*, 2013).

The 7 nt insertion in allele 1 of *DGCR8* knockout cell line 1, with the sequence ACGTTGG, contained a PAM sequence preceded by a genomic sequence that is absent in allele 2, and could therefore be uniquely targeted by an allele-specific guide RNA (figure 3.1A). In addition, a DNA donor repair template was designed, containing the wildtype sequence of the protospacer and surrounding regions. To restore allele 1 to wildtype, *DGCR8* knockout cell line 1 was transfected with Cas9, the guideRNA targeting allele 1 and the repair template (figure 3.1A). Following this treatment, cells were serially diluted and grown into single cell colonies. Restoration of allele 1 to wildtype was examined by sequencing of genomic DNA fragments. This strategy created cell line *DGCR8* heterozygote 1 (figure 3.1A; figure 3.2B).



Figure 3.1 PA-1 DGCR8 wild type and mutant sequences and CRISPR-Cas9 strategy. A. Genomic DNA sequence of wildtype and two DGCR8 knockout cell lines created by Sara Heras' lab. Allele 1 of DGCR8 knockout cell line 1 was targeted for repair using allele-specific CRISPR-RNA, resulting in DGCR8 heterozygote cell line 1. B. Genomic DNA sequence of wildtype and DGCR8 heterozygote cell line 2, which was created by targeting exon 10.

Secondly, novel *DGCR8* mutants were created in wildtype PA-1 cells by targeting exon 10 of *DGCR8*, close to the region coding for dsRBD2 (figure 3.1B). This was done using the same CRISPR-Cas9 approach as described above, but without transfection of a repair template. This approach led to the creation of *DGCR8* heterozygous cell line 2, which has an 8-nucleotide deletion in exon 10 (figure 3.1B; figure 3.2B). All genomic mutations in *DGCR8* were confirmed by sanger sequencing.

While the PA-1 *DGCR8* heterozygotes show no altered phenotype compared to wildtype cells in culture, the *DGCR8* knockouts show worse cell adhesion and grow slower, suggested to be caused by a delay in the G1 phase of mitosis (Sara Heras' lab, unpublished; Deng *et al.*, 2019).

The presence of *DGCR8* protein in the different PA-1 cell lines was determined by western blot. Since cell lines *DGCR8* knockout 1 and 2 and heterozygotes 1.1 and 1.2 have mutations at the N-terminal side of *DGCR8*, there was the possibility that the presence of a downstream start codon could still result in a functional, albeit truncated, variant of *DGCR8*. These shortened alternative proteins would not be detectable by antibodies against *DGCR8* recognizing the N-terminus of the protein, since the immunogen generated for this antibody only contained the first 50 amino acids of *DGCR8*. Therefore, two different antibodies against *DGCR8* were used for western blotting: one recognizing the N-terminus, the other targeting the C-terminus. Neither antibody could detect the presence of *DGCR8* by western blotting, suggesting that these cell lines are complete *DGCR8* knockouts (figure 3.2C). On the other hand, as expected, *DGCR8* could be detected in all heterozygous

cell lines. The presence of DGCR8 in heterozygous cell lines 1.1 and 1.2 indicate that the defective allele 1 in knockout cell line 1 has been successfully restored, which Sanger sequencing confirmed. While DROSHA could be detected in all cell lines, protein levels were strongly reduced in DGCR8 knockout cell lines (figure 3.2C), probably due to destabilization by the lack of direct protein-protein interactions with DGCR8 (Han *et al.*, 2009).

3.2.2 Transcriptomic changes upon *DGCR8* knockout and heterozygosity

RNA-seq from total RNA was carried out by BGI, using 4 biological replicates each from PA-1 wildtype, *DGCR8* knockout 1 and *DGCR8* heterozygote 1.1 cells. For each sample, around 50 million filtered reads were returned (figure 3.3A). Using HISAT2, these reads were aligned to the human genome and a novel transcriptome was assembled (Kim *et al.*, 2019). Subsequently, transcripts from all cell lines were identified and counted with FeatureCounts and normalized using DESeq2 (Cheng *et al.*, 2014; Liao, Smyth and Shi, 2014).

Determination of principle components using rlog-transformed gene expression data from all samples indicated that 85% of all variation between samples can be explained by the first two principal components (figure 3.3B). Samples belonging to the same cell line were plotted close together in the biplot, indicating that the differences between cell lines are much bigger than the differences within the same cell line. For principal component 1, explaining 64% of the variation, wildtype samples and *DGCR8* knockout samples were furthest apart, with heterozygote cells showing an intermediate phenotype. This indicates that complete knockout of *DGCR8* leads to more severe changes than still having one functional gene copy as in the heterozygote. Pairwise differential expression analysis between wildtype and mutant cell lines was carried out using the R package DESeq2, and fold changes between the *DGCR8* mutants and wildtype cells were shrunk with the apegglm algorithm (Love, Huber and Anders, 2014; Zhu, Ibrahim and Love, 2019).

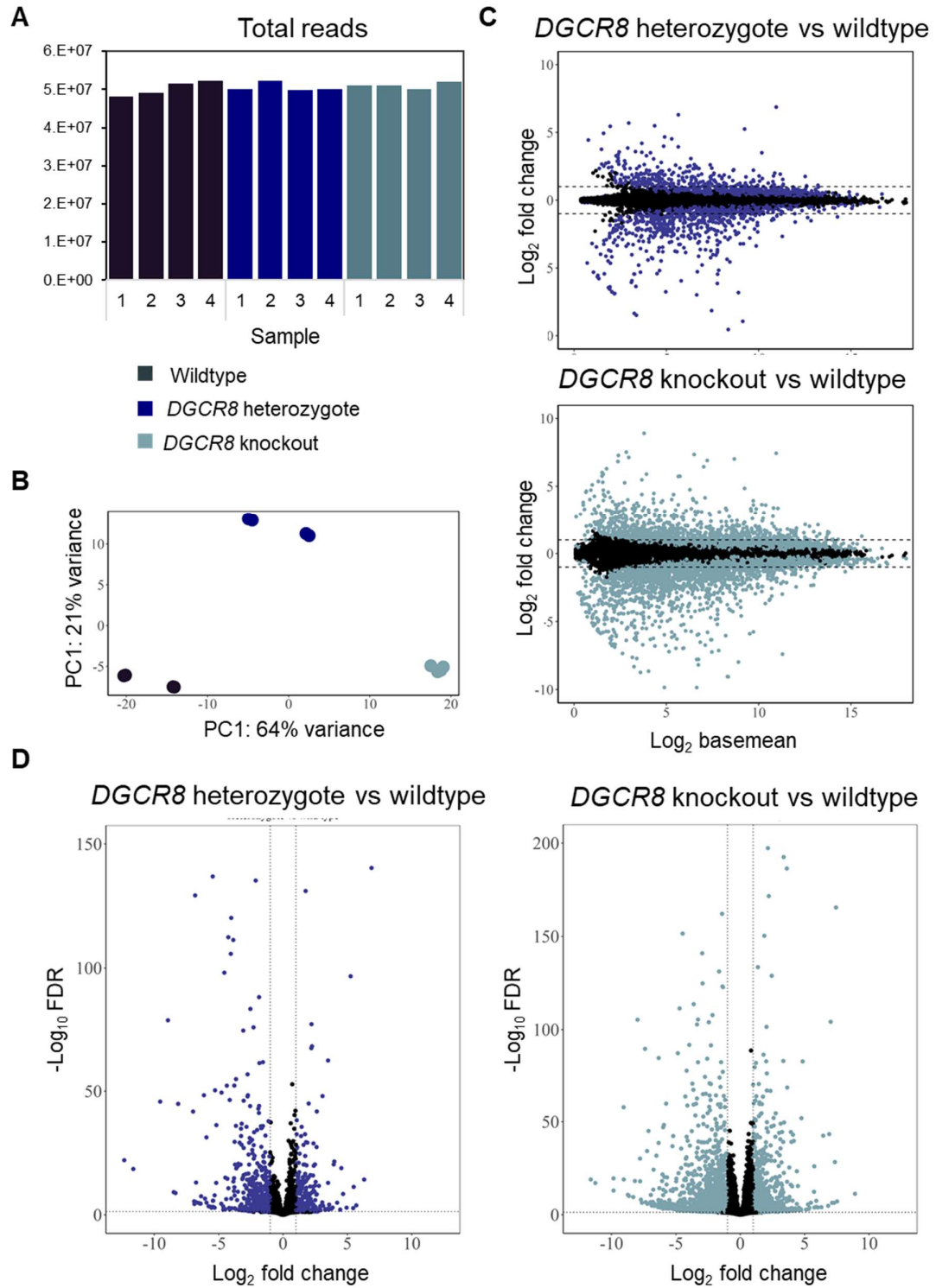


Figure 3.3 Transcriptomic analysis of PA-1 DGCR8 mutants. A. Library size of paired reads obtained after sequencing of PA-1 cell lines by BGI. B. Principal component biplot of rlog-transformed normalised reads from 500 genes with the highest variance between samples. C. Volcano plots of DGCR8 heterozygote and knockout cell lines compared to wildtype. Significant datapoints ($FDR < 0.05$) with an absolute fold change >0.5 are shown in colour, all other datapoints are shown in black. D. Comparison of $\log_2$ fold changes between DGCR8 knockout and heterozygote cells compared to wildtype cells.

When cells heterozygous for DGCR8 were compared with wildtype cells, 24,960 RNAs were analysed after filtering for average expression by DESeq2's independent filtering. Of these, 2210 transcripts were considered to be differentially expressed ($FDR < 0.05$), of which 1,062 upregulated and 1,148 were downregulated (figure 3.3C). DGCR8 knockout cells showed more significant changes: out of 25,624 transcripts analysed, 3,301 transcripts were upregulated, and 4,010 transcripts were found to be downregulated, accumulating to a total of 7,311. *DGCR8* knockout cell lines were expected to have a more severe phenotype, since DGCR8 protein was still detectable in the heterozygous cell lines. Therefore, it could be hypothesized that this cell line has an intermediate phenotype compared to wildtype and knockout cells. To study this, the transcript changes between the two mutant cell lines were compared to each other (figure 3.3D). While several RNAs in the *DGCR8* mutants seemed to behave similarly relative to wildtype cells, no strong overall correlation was found, indicating that the DGCR8 heterozygote cells show a distinct transcriptome phenotype from both wildtype and knockout cells.

3.2.3 Validation of RNA-seq data and concordance between cell lines

To validate the transcriptomics analysis and determine variability between cell lines, the expression of several mRNAs involved in immunity and pluripotency was measured by RT-qPCR (figure 3.4A; individual RNAs RT-qPCRs shown in chapter 5). When the $\log_2$ fold changes of the RNA-seq analysis and the qPCRs for the measure

RNAs were compared, a strong correlation was found between the two methods for both *DGCR8* knockout 1 and *DGCR8* heterozygote 1 samples, with R^2 values of 0.97 and 0.7034, respectively, indicating that our RNA-seq analysis was reliable (figure 3.4A).

Only one *DGCR8* knockout and one *DGCR8* heterozygote were analysed using RNA-seq. In addition, we were interested in how similar the mutant cell lines were to each other. When the quantification of transcripts by RT-qPCR was compared between *DGCR8* knockout 1 and knockout 2, the tested mRNAs showed highly similar fold changes (figure 3.4B; $R^2=0.84$). However, there was no correlation between cell lines *DGCR8* heterozygote 1.1 and heterozygote 2 (figure 3.4B; $R^2=0.0024$).

To test the effect of *DGCR8* mutation on transcript levels, RT-qPCR of *DGCR8* mRNA was carried out with two different primer pairs. The first pair annealed on both sites of the junction between exon 2 and 3, downstream of the mutated site in cell lines knockout 1 and 2 and heterozygote cell lines 1.1 and 1.2. The second primer pair annealed at the 3' untranslated region. The RT-qPCRs and the RNA-seq data both showed that, compared to wildtype cells, the *DGCR8* transcript was upregulated in both knockout cell lines and in heterozygote cell line 1.1 (figure 3.4D). This suggests that the mutated RNA is not degraded by nonsense-mediated mRNA decay, but instead accumulates in cells. However, we did not detect any signs of the presence of a truncated protein in the western blots we ran. Both RT-qPCR primers showed the same trends, indicating that the mutations did not influence the splicing out of intron 2. The accumulation of *DGCR8* mRNA in knockout cell line 1 was higher compared to heterozygote cell line 1.1, indicating that both mutated alleles caused accumulation. Contrastingly, *DGCR8* mRNA from heterozygote cell line 2 did not show any accumulation by RT-qPCR, indicating that turnover of the mutated mRNA could be similar to that of wildtype (figure 3.4C).

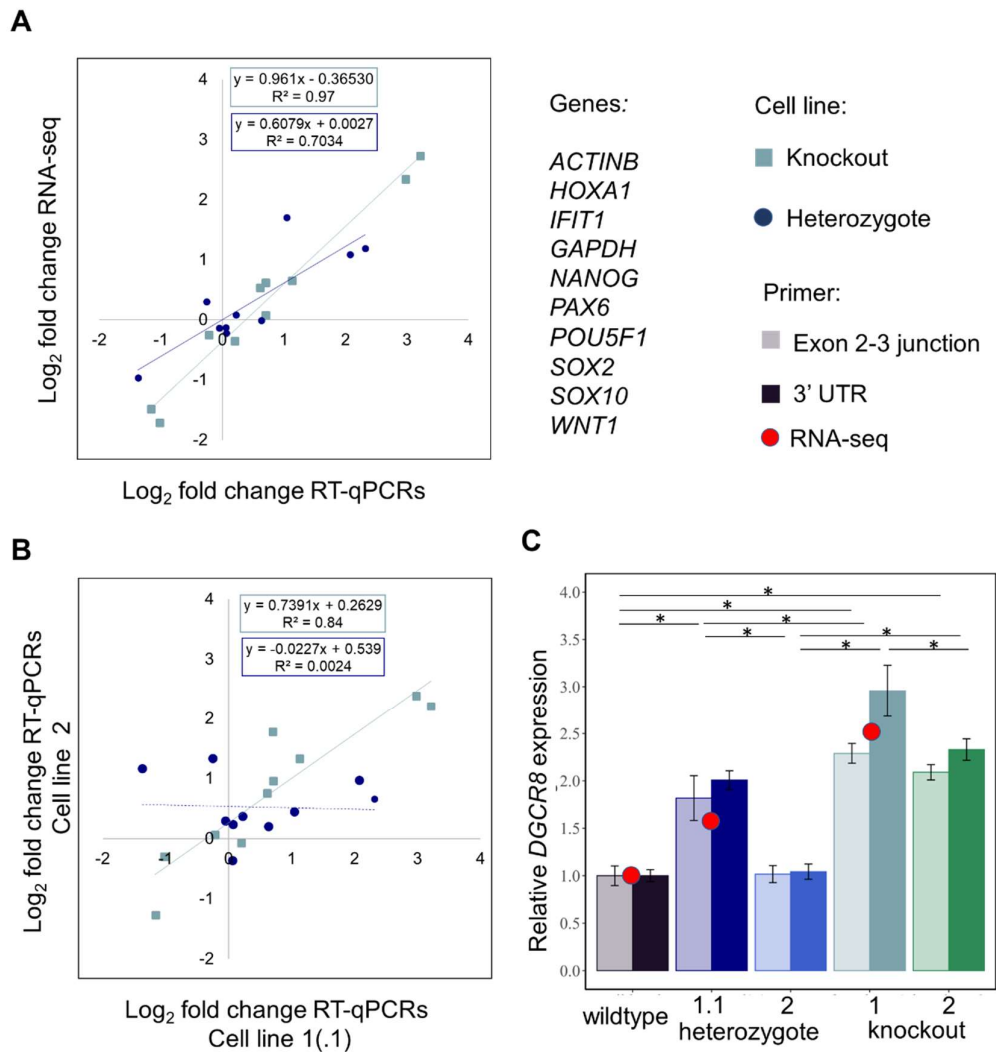


Figure 3.4. Validation of RNA-seq and transcript expression between cell lines. Correlation between changes in transcript expression from DGCR8 mutant cells vs wild type measure by A RNA-seq versus RT-qPCR, B.) RT-qPCRs of DGCR8 knockout 2 versus DGCR8 knockout 2 and DGCR8 heterozygote cell line 2 versus DGCR8 heterozygote cell line 1, D.) RT-qPCRs of DGCR8 heterozygote cell lines versus DGR8 knockout cell lines. Some of RT-qPCR data of the measured transcripts will be shown as bar graphs in chapter 5. D. RT-qPCR comparing DGCR8 transcript levels between cell lines.

3.2.4 Functional analysis of RNA-seq data show changes in pluripotency upon *DGCR8* mutation

To study the consequences of *DGCR8* mutation for molecular processes, overrepresentation analysis with the RNA-seq data was carried out using EnrichR (Xie *et al.*, 2021). As input for EnrichR, a gene list containing all differentially expressed transcripts (FDR < 0.05) with log₂ fold change > 0.5 was used, either for *DGCR8* knockout versus wildtype cells, or for *DGCR8* heterozygotes versus wildtype cells. EnrichR determines enrichment of specific genesets by Fisher's exact test and calculates odds ratios, tests statistics and a combined score of both.

For comparisons between the cell lines, two different approaches were used. Firstly, pathway analysis was carried out to study whether differentially expressed genes are overrepresented in specific biological pathways as annotated by the KEGG database (Kanehisa *et al.*, 2009). Secondly, the list of differentially expressed genes were compared with transcriptome signatures that occur upon perturbation of specific transcription factors. Both methods indicate altered activity of specific biological pathways or mechanisms, the latter may also indicate the cause of the altered activity. Interestingly, when KEGG pathway analysis was carried out to study transcriptome differences upon *DGCR8* heterozygosity, many pathways involving pluripotency and differentiation were shown to contain differentially expressed transcripts, such as 'Signalling pathways regulating pluripotency of stem and HIPPO signalling' (table 3.1). The WNT signalling pathway was ranked 12th on this list by significance. Upon studying the transcripts present in these pathways, it was found that several transcripts of the *WNT* family are differentially expressed in *DGCR8* heterozygote cells. In addition, the two highest scoring pathways are involved in neuropathy and cardiopathy, symptoms relevant for 22q11.2 DS syndrome. When KEGG pathway analysis was carried out to study transcriptome differences upon *DGCR8* knockout, the pathway "pluripotency signalling in stem cells" showed the highest combined score (table 3.2).

Table 3.1 Enriched KEGG pathways upon *DGCR8* heterozygosity as determined by adjusted *p*-value < 0.05

KEGG pathway	Total genes in list	Differentially expressed	Combined Score	Adjusted p-value
Axon guidance	182	30	28	5.08E-03
Dilated cardiomyopathy	96	19	32	6.94E-03
Arrhythmogenic right ventricular cardiomyopathy	77	16	31	8.82E-03
Cushing syndrome	155	25	22	8.82E-03
ECM-receptor interaction	88	17	27	9.31E-03
Hypertrophic cardiomyopathy	90	17	25	1.03E-02
Basal cell carcinoma	63	12	20	4.27E-02
Protein digestion and absorption	103	17	17	4.27E-02
Signalling pathways regulating pluripotency of stem cells	143	21	15	4.27E-02
Hippo signalling pathway	163	23	14	4.27E-02
Focal adhesion	201	27	13	4.27E-02

Furthermore, several pathways involved in cardiopathy were shown to be affected, which was mainly driven by differential expression of several *integrin* genes.

Apart from their presence in KEGG pathways, differentially expressed transcripts were compared to gene expression changes upon overexpression or depletion of specific transcription factors. This analysis showed that the changes seen in upon *DGCR8* heterozygosity are similar to those upon downregulation of *SOX2* and *POU5F1* (coding for OCT4) (table 3.3). Together with *NANOG*, these two genes are the master regulators of pluripotency (Greve, Judson and Blelloch, 2013).

Table 3.2 Enriched KEGG pathways upon DGCR8 knockout as determined by adjusted p-value < 0.05

KEGG pathway	Total genes in list	Differentially expressed	Combined Score	Adjusted p-value
Signalling pathways regulating pluripotency of stem cells	143	55	24	3.92E-03
Dilated cardiomyopathy	96	39	23	8.47E-03
Hypertrophic cardiomyopathy	90	36	20	1.56E-02

Interestingly, these transcription factors themselves are not differentially expressed, suggesting regulation of pluripotency pathways downstream of *SOX2*, *POU5F1* and *NANOG*. Other transcription factors which repression shows similar changes as in *DGCR8* heterozygous cells are *NEUROG1* and *DMRTA2*, that are involved in differentiation into neuronal cell lines (Konno *et al.*, 2012; Han *et al.*, 2018). All in all, these results indicate a potential differentiation defect in *DGCR8* heterozygote cells.

Similar to *DGCR8* heterozygous cells, differentially expressed RNAs upon *DGCR8* knockout were also found to be affected upon downregulation of *SOX2* and *POU5F1* (table 3.4). Since *Dgcr8* mESC knockout cell lines are known to have defective differentiation, aberrant expression of pluripotency factors can also be expected in human *DGCR8* knockout cell lines.

Table 3.3 Top 15 transcription factor perturbation datasets enriched upon DGCR8 heterozygosity as determined by adjusted p-value.

Affected transcription factor	Total genes in list	Differentially expressed	Combined Score	Adjusted P-value
SOX2 KO human gse59234 creedsid gene 2956 up	333	73	149	1.06E-14
<i>Dmrta2</i> OE mouse gse25179 creedsid gene 2204 down	305	57	74	2.50E-08
Otx2 KO mouse gse27630 creedsid gene 78 up	197	40	67	4.29E-07
<i>Cbx2</i> KO mouse gse29485 creedsid gene 381 up	195	39	63	8.72E-07
<i>Cbx2</i> KO mouse gse29485 creedsid gene 380 up	230	44	62	4.35E-07
Sox3 KO mouse gse53760 creedsid gene 1668 up	343	59	60	1.75E-07
<i>Dmrta2</i> OE mouse gse25179 creedsid gene 2203 up	285	51	58	3.38E-07
<i>FLI1</i> KD human gse27524 creedsid gene 1598 up	434	70	57	1.06E-07
<i>NEUROG1</i> OE human gse18296 creedsid gene 1426 down	328	56	55	3.38E-07
<i>POU5F1</i> KD human gse21135 creedsid gene 1328 up	400	64	51	3.38E-07
<i>POU5F1</i> KD human gse21135 creedsid gene 1329 up	418	66	51	3.38E-07
SOX2 KO human gse59234 creedsid gene 2957 down	326	54	49	1.18E-06
Rest shrna C2 human gse90068 pbmpa rnaseq down	118	25	49	6.00E-05
SOX2 KO human gse59234 creedsid gene 2957 up	259	45	48	3.68E-06
<i>Dmrta2</i> OE mouse gse25179 creedsid gene 2204 up	180	34	47	1.94E-05

Table 3.4 Top 15 transcription factor perturbation datasets enriched upon DGCR8 knockout as determined by adjusted p-value.

Affected transcription factor	Total genes in list	Differentially expressed	Combined Score	Adjusted p-value
SOX2 KO human gse59234 creedsid gene 2957 down	326	137	147	2.32E-18
POU5F1 KD human gse21135 creedsid gene 1328 up	400	161	145	3.17E-19
POU5F1 KD human gse21135 creedsid gene 1329 up	418	163	128	2.73E-18
SOX2 KO human gse59234 creedsid gene 2957 up	259	109	118	1.00E-14
SOX2 KO human gse59234 creedsid gene 2956 up	333	133	116	8.55E-16
EOMES OE human gse43805 creedsid gene 92 up	306	120	97	1.38E-13
EBF3 OE sknsh human gse90682 rnaseq down	276	110	95	5.31E-13
NFKB1 activation human gse20736 creedsid gene 2522 up	268	105	85	6.13E-12
Cebpg KO mouse gse47777 creedsid gene 3104 down	197	81	81	1.92E-10
Sox3 KO mouse gse53760 creedsid gene 1668 up	343	126	77	5.70E-12
FLI1 KD human gse27524 creedsid gene 1598 up	434	152	75	1.59E-12
Msgn1 OE mouse gse29848 creedsid gene 1631 up	233	91	73	2.57E-10
NEUROG1 OE human gse18296 creedsid gene 1426 down	328	119	69	6.58E-11
FLI1 KD human gse27524 creedsid gene 1597 up	359	127	66	7.96E-11
FLI1 KD human gse27524 creedsid gene 1601 up	398	137	63	9.43E-11

3.3 DISCUSSION

3.3.1 PA-1 *DGCR8* knockout and heterozygote cells can help further study the role of *DGCR8* and the microprocessor in the human context

PA-1 cells heterozygous for *DGCR8* were created using CRISPR-Cas9, either by repairing a defective allele, or by novel targeting of a coding sequence. By having multiple knockouts and heterozygote cell lines obtained by different approaches, *DGCR8*-specific effects can be detected, independent of the background of the cell line and individual variation of each cell line.

The *DGCR8* knockout cell lines showed no protein levels by western blot when blotted against either the N- or C-terminus of *DGCR8*, suggesting that no truncated variant caused by a methionine downstream of the start codon is present and that the cell line is a complete knockout. Furthermore, lower protein levels of *DROSHA* were detected in the knockout cells. This has been described before in the absence of a functional microprocessor: the loss of *DGCR8* removes its stabilizing effect on *DROSHA*, leading to degradation of the latter (Han *et al.*, 2009; Nguyen *et al.*, 2018).

For *DGCR8* heterozygote cells, no big changes in protein levels of either *DGCR8* or *DROSHA* compared to wildtype cells were detected. Recent work by our collaborators in Sara Heras' lab suggested that *DGCR8* accumulates in the heterozygote cells with increasing passage number. While all sequencing experiments were carried out in cells with a low passage number, extract for western blots may have been taken from cells with a higher passage number, and therefore show accumulated *DGCR8*. Furthermore, while the protein is expressed in heterozygous cells, it is unknown if these proteins still interact with each other to form a functional microprocessor. Immunoprecipitation of *DGCR8* followed by immunoblotting for *DROSHA*, or vice versa, may help clarify this.

Apart from the studies described in this thesis, *DGCR8* heterozygote and knockout cells can be used for future research. To further study *DGCR8* structure and molecular function of specific domains, mutated *DGCR8* can be reintroduced in the knockout cell line. For example, expression of a C-terminal truncated version of *DGCR8* abolishes interaction with DROSHA, while N-truncated *DGCR8* alters heterochromatin structure, making it possible to look at non-canonical functions of *DGCR8* (Nguyen *et al.*, 2018; Deng *et al.*, 2019). Furthermore, the effect of *DGCR8* heterozygosity could be further studied and directly compared with the *DGCR8* knockout cell line to study the role of *DGCR8* protein and gene dosage in the context of 22q11.2 DS.

3.3.2 Transcriptomic analyses show a changed RNA landscape upon *DGCR8* depletion

To find the effect of *DGCR8* depletion on the RNA landscape, RNA-seq was carried out for PA-1 wildtype, *DGCR8* heterozygote and *DGCR8* knockout cells. Statistical analysis showed that 8.9% and 29% of the detected transcripts was differentially expressed in heterozygote and knockout cells, respectively, indicating a substantial shift in the RNA landscape. The observed phenotype is more severe in the knockout cell line, presumably due to the complete absence of *DGCR8*, and therefore of canonical miRNAs.

This is reflected in their reduced proliferation and dependence on rock inhibitor during splitting. Our analysis does not reveal whether the observed transcriptomics changes are the result of altered microprocessing cleaving, miRNA regulation or indirect effects. The effect of *DGCR8* mutations on DROSHA cleaving could be studied by comparing RNAs found with DROSHA HITS-CLIP in the different cell lines, while methods such as CLASH can help understanding the altered miRNA-target interactions (Macias *et al.*, 2012; Helwak *et al.*, 2013). Furthermore, previous research in human ESCs with a truncated variant of *DGCR8* showed a role for *DGCR8*

in maintaining heterochromatin stability, suggesting that changes in transcription could explain of the changes in RNA levels observed in *DGCR8* knockout cells (Deng *et al.*, 2019). Methods such as ATAC-seq could help further elucidate this, while inhibition of RNA degradation could indicate if altered degradation plays a role in the *DGCR8* mutants (Buenrostro *et al.*, 2013; Lai *et al.*, 2019). This is currently a focus in our lab.

Both coding- and noncoding RNAs are affected by the loss of one or two *DGCR8* copies, potentially leading to altered metabolism and signal transduction of several pathways, as suggested by our pathway analysis. Altered expression of mRNAs could affect protein biogenesis and could be further studied by mass spectroscopy, while changes in the noncoding genome could influence genomic stability and alter regulation of transcripts by miRNAs (Macias *et al.*, 2012; Macias, Cordiner and Cáceres, 2013). The effect of *DGCR8* heterozygosity or knockout on the processing of pri-miRNAs and levels of mature miRNAs is the focus of chapter 4.

A high correlation was found between RNA-seq and qPCRs, supporting the validity of the RNA-seq analysis. While the two knockout cell lines used in this thesis behaved very similar in qPCR experiments, there was a big difference between the cell lines *DGCR8* heterozygote 1.1 and heterozygote 2. This might be caused by the different background of the two mutants, since one cell line was created by repairing a *DGCR8* knockout cell line instead of wild type cells. However, for the transcripts measured by qPCR, both heterozygous cell lines showed limited correlation of fold change in expression compared to the knockout. In future experiments, comparison of multiple *DGCR8* mutant cell lines versus wildtype will help distinguish *DGCR8*-dependent effects from variable background noise within cell lines.

The cell lines *DGCR8* knockout 1, 2 and heterozygote 1.1 all showed an accumulation of *DGCR8* transcript. These cell lines all have mutations in the DNA sequence containing *miR-1306*. Normally, the microprocessor can cleave the *DGCR8*

transcript in exon 2 to release *pre-miRNA-1306*, resulting in degradation of the remaining transcript and thereby regulating *DGCR8* expression (Han *et al.*, 2009). In case of a *DGCR8* knockout, a functional microprocessor is absent, and this negative feedback loop is missing. However, accumulation of *DGCR8* is also observed in heterozygous cell line 1.1, even though *DGCR8* expression can be measured by western blot. A possible explanation is that the 60-nucleotide insertion in allele 1 of heterozygote cell line 1 disrupts the hairpin structure of *pre-miR-1306*, and therefore the RNA cannot be recognised and cleaved by the microprocessor. Furthermore, the accumulation also indicates that the mutated transcripts are not targeted by nonsense-mediated RNA decay.

Overrepresentation analysis using EnrichR showed that both *DGCR8* knockouts and heterozygotes showed aberrant expression of transcripts involved in pluripotency, cardiopathies and (neuronal) differentiation. While the importance of the microprocessor for differentiation and neuronal development has been widely reported in conditional *DGCR8* knockouts and knockdown studies in mice and mouse cell lines (Rao *et al.*, 2009; Fénelon *et al.*, 2011; Earls *et al.*, 2012; Knuckles *et al.*, 2012; Ouchi *et al.*, 2013; Marinaro *et al.*, 2017), only limited studies have focussed on human models. The effect of *DGCR8* knockout or heterozygosity on differentiation is further explored in chapter 5.

CHAPTER 4: THE IMPACT OF DGCR8 ON THE MIRNA LANDSCAPE

4.1 INTRODUCTION

DGCR8, DROSHA, EXP5, and DICER are proteins necessary for canonical miRNA biogenesis. DGCR8 and DROSHA together form the microprocessor complex that facilitates the nuclear cleaving of pri-miRNAs, while EXP5 exports the cleaved pre-miRNAs and DICER facilitates the cytoplasmic cleaving of pre-miRNAs, releasing mature miRNAs (Bohnsack, Czapinski and Görlich, 2004; Denli et al., 2004; Gregory et al., 2004; Han et al., 2004; Hutvagner et al., 2001; Ketting et al., 2001; Okada et al., 2009; Zhang et al., 2004).

While processing of pre-miRNAs by DICER is well understood, a comprehensive view of the factors influencing microprocessor precision and efficacy is missing. Cleavage of miRNA precursor molecules is often biochemically examined in *in vitro* cleavage studies (Auyeung *et al.*, 2013; Fang and Bartel, 2015; Kwon *et al.*, 2020; Li *et al.*, 2020; Rice *et al.*, 2020). While these studies have been vital in understanding the effect of hairpin length or presence of sequence motifs in a specific subset of pri-miRNAs, these studies focus on a small amount of well annotated miRNAs. In addition, they do not take the cellular constraints, such as pri-miRNA transcription and presence of cofactors involved in microprocessor cleaving, into account and therefore do not capture the complexity of *in vivo* processing. In 2014, a novel way of measuring endogenous microprocessor efficacy was published, based on measuring the ratio between cleaved and uncleaved pri-miRNAs (Conrad *et al.*, 2014). Since this method does not rely on *in vitro* studies or transfection of miRNA precursors, it can be used to measure biologically relevant processing under specific conditions.

Human knockout cell lines for the miRNA biogenesis factors DROSHA, DICER and EXP5 have been reported before and their effect of miRNA levels has been studied (Kim, Kim and Kim, 2016). While deletion of these proteins led to an overall decrease in miRNAs the levels of downregulation were protein specific, indicating the potential role of non-canonical pathways. A human complete *DGCR8* knockout

has never been reported before, although C-terminal truncation of DGCR8 prevents interaction with DROSHA and leads to an overall decrease in canonical miRNA levels (Nguyen *et al.*, 2018). Depleting miRNA biogenesis factors and comparing the effects can help identify non-canonical miRNAs and miRNA-independent functions. Furthermore, these knockout cell lines can be used to re-introduce specific miRNAs to study their function. Since miRNAs are known to be dysregulated in several diseases, better understanding their biogenesis and function will lead to better intervention options in the future.

In chapter 3, we obtained PA-1 cell lines that were *DGCR8* knockouts or heterozygotes. In this chapter we use these cell lines to study the effect of DGCR8 depletion on mature miRNA levels using small RNA-seq. In addition, using chromatin-associated RNA-seq, we study microprocessor efficacy.

4.2 RESULTS

4.2.1 *DGCR8* mutation changes small RNA levels in PA-1 cells

In order to determine the effect of losing either one or both genomic copies of *DGCR8* on mature miRNA levels, small RNA-seq was carried out. For PA-1 cells that were either wildtype, *DGCR8* knockout or *DGCR8* heterozygous (obtained by reparation of the knockout cell line), RNA was extracted from 4 biological replicates and used for small RNA-seq. For each sequenced RNA sample, over 20 million reads were obtained (figure 4.1A). Reads that were present only once in the sample were removed prior to further analysis.

The identity of the sequenced small RNAs was analysed by blastn and revealed big differences between samples: while both *DGCR8* wild type and heterozygous cell lines showed that over 90% of the identified reads belonged to miRNAs, this was only 50-60% for the *DGCR8* knockout cells lines (figure 4.1A). Instead, DGCR8

knockout cells contained large proportions of other RNA types, particularly tRNAs and rRNAs. Notably, the proportion of all other RNA types other than miRNAs was ~6 times larger in the knockout cell lines, and when excluding miRNAs, the proportions of remaining small RNAs are virtually identical among cell lines. Since the absence of DGCR8 should largely abolish miRNA biogenesis, an overall decrease in small RNAs is expected in DGCR8 knockout cells. Therefore, we suspected that the increase in small RNAs other than miRNA is the result of more cycles of either RNA library preparation or sequencing in order to get similar read depth to the other samples, rather than a biological effect. Consequently, this meant that neither absolute nor relative miRNA counts could be used for comparisons between cell lines. Therefore, we looked at other ways to normalise the small RNA counts.

Since most normalisation techniques assume that the distributions of the expression data have equal means, these would compensate for the lack of miRNAs in the knockout cell lines and therefore artificially inflate their numbers. Instead, we opted for miRNA identification using MirDeep2, and normalisation by the total number of tRNAs identified by blastn, to recreate realistic counts of miRNAs in each sample. Principle component analysis (PCA) showed little variation in miRNA expression within the biological replicates for each cell line, while there was great variation between cell lines (figure 4.1B). Remarkably, principal component 1 explained 90% of the variance and placed the DGCR8 knockout samples opposite both wild type and heterozygote samples. This indicates that the knockout of *DGCR8* accounts by far for most of the differences in miRNA expression between the measured samples. Wildtype and *DGCR8* heterozygote cell lines showed very little variation with respect to PC1, but explained most of the variation in PC2, indicating the *DGCR8* heterozygote does not display an intermediate phenotype between wild type and knockout, but shows different characteristics altogether.

Studying the relative levels of the top 50 miRNAs with the highest overall variance showed that for all these miRNAs, the levels were dramatically decreased in *DGCR8*

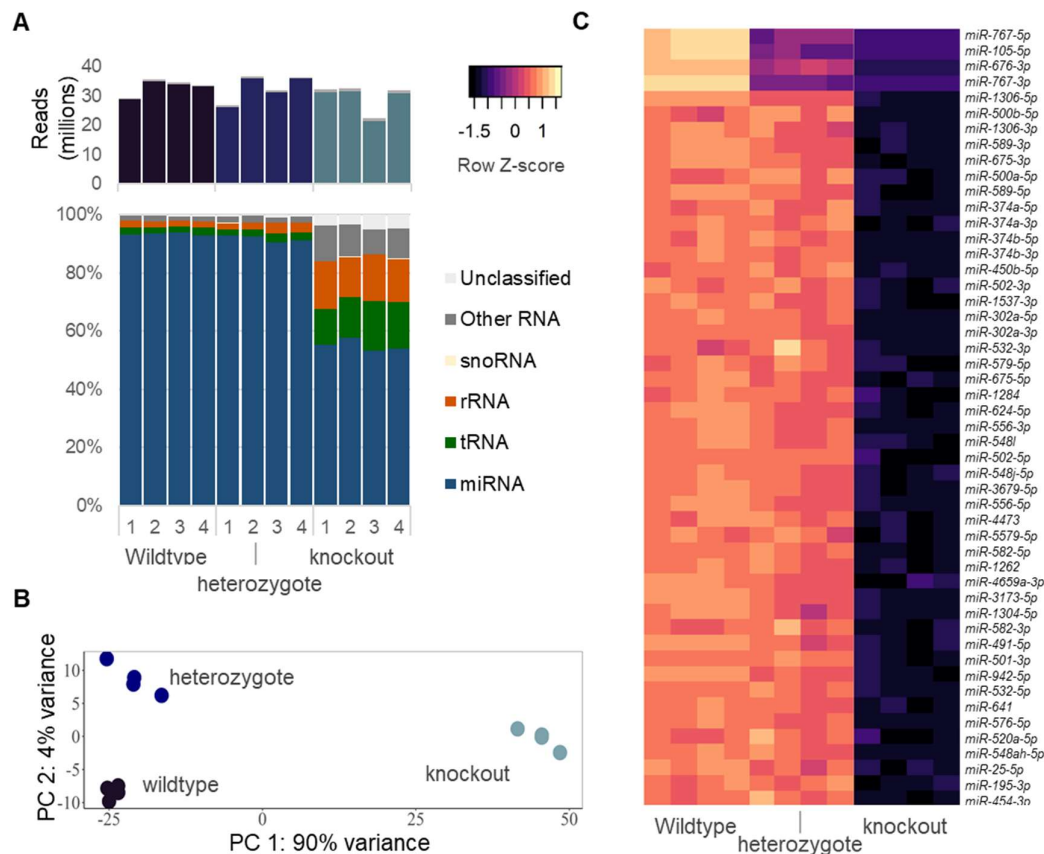


Figure 4.1. Sequencing characteristics and miRNA identification of *DGCR8* wildtype, heterozygote, and knockout cell lines. **A.** Top: total reads identified by small RNA-seq. Reads that were excluded due to low abundance are coloured grey. Bottom: relative levels of small RNAs present in the sequenced samples, as identified by *blastn*. **B.** PCA biplot of Rlog transformed miRNA counts normalized by tRNAs. **C.** Heatmap showing the 50 miRNAs with the highest variance between samples, sorted by decreasing variance. The colour indicates the z-transformed expression levels for each miRNA.

knockout cells compared to both wildtype and heterozygote cells (figure 4.1C).

Apart from one cluster of miRNAs, *DGCR8* heterozygote cell line samples showed miRNA levels very similar to wild type in this selection. This is consistent with the observations from the PCA that indicated few differences within cell lines and major differences between wild type and knockout cell lines.

4.2.2 DGCR8 knockout shifts the relative levels of miRNA subtypes

Apart from the absolute differences in miRNA numbers, there were also relative differences noted between wildtype cells and the *DGCR8* mutants. In figure 4.2A the top 10 most occurring miRNAs in each cell line are shown, together with microprocessor-independent miRNAs. For PA-1 wildtype cells, the top 10 of most detected miRNAs accounted for over 90% of all assigned miRNA reads. Members of the *miR-302* family were by far the most occurring, accounting for over 80% of all miRNAs. High expression of *miR-302* members was expected, since this is also observed in human ESCs and germline tumors (Greve, Judson and Belloch, 2013). *Mir-203a-3p* and *miR-302-3p* each accounted for around 25% of total miRNA reads. The other miRNAs in the top 10 are *miR-21-5p*, *miR-92a-3p*, *miR-26a-5p*, *miR-16-5p* and *miR-151a-5p*.

DGCR8 heterozygote cells showed very similar trends in overall miRNA expression compared to wildtype cells, although miRNAs outside the top 10 form a slightly larger portion of the total reads. *DGCR8* knockout cells, however, showed an altered relative miRNA distribution. Most *miR-302* family members showed decreased relative levels of reads: while *miR-302b-3p* still accounted for over 20% of miRNA reads, the other members each represented less than 5% of total miRNA reads. Instead, the relative levels of other miRNAs is increased, including miRNAs with non-canonical biogenesis.

To study this latter observation further, we looked at the normalized expression data of mirtrons, which are microprocessor-independent miRNAs. Compared to wildtype cells, *DGCR8* heterozygote cells did not show any changes in mirtron or non-mirtron levels (figure 4B). *DGCR8* knockout cells, however, showed an increase in detected mirtrons in our dataset, while non-mirtrons were decreased overall.

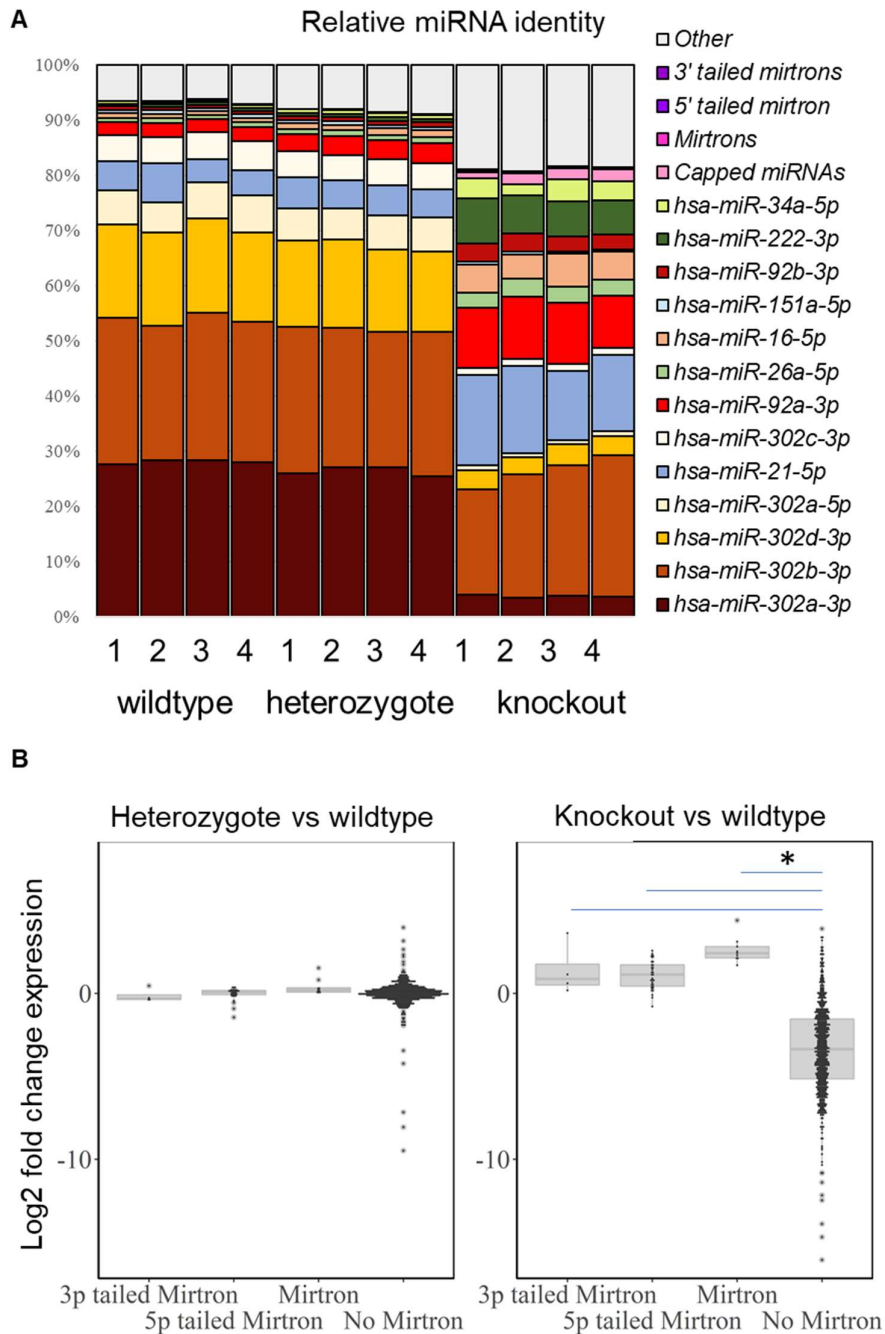


Figure 4.2. The contribution of miRNA family members and subtypes to the miRNA landscape. A. stacked bar graph showing the relative contribution of individual miRNAs and DGCR8-independent miRNA subtypes. Of each cell line, the top 10 most occurring miRNAs in included and sorted by relative presence in wildtype cells. DGCR8-independent miRNAs are represented as their concatenated value.

B. Boxplot showing the apegln shrunk fold changes of mirtrons and non-mirtrons in the DGCR8 mutant cell lines compared to wildtype. Statistical analysis by ANOVA followed by Benjamin-Hoch corrected pairwise -test.

4.2.3 Differential expression of miRNAs upon *DGCR8* mutation

In order to identify differentially expressed miRNAs, statistical analysis of miRNA expression between cell lines was carried out using the DESeq2 package in R (Love, Huber, & Anders, 2014). As indicated by the heatmap, there were a lot of changes in the *DGCR8* knockout cells compared to wildtype, while the *DGCR8* heterozygote cells were very similar to wildtype. Out of the 700 miRNAs analysed, 87 were significantly up or downregulated (FDR <0.05) in the *DGCR8* heterozygote samples (figure 4.3A, top). Of these, 19 miRNAs were more than two-fold upregulated and 19 were more than two-fold downregulated. Contrastingly, in the *DGCR8* knockout samples almost all miRNA were differentially expressed, with a substantial overall decrease in miRNA expression (figure 4.3A, bottom). 565 miRNAs out of 700 were downregulated, of which 545 miRNAs were downregulated more than two-fold. Since knockout of *DGCR8* abolishes canonical miRNAs biogenesis, a big decrease in total miRNA expression was expected. Interestingly, 66 miRNAs were significantly upregulated, with 59 miRNAs having a more than two-fold-upregulation. When comparing the changes in overall miRNA expression of the mutants to baseline miRNA expression in wildtype, some overall trends could be observed (figure 4.3B). The *DGCR8* heterozygote cells did not show big overall changes between miRNAs and compared to wildtype. On the other hand, *DGCR8* knockout cells did not only show a big overall miRNA decrease as indicated before, but also showed more variation in fold change between individual miRNAs. MiRNAs with either a low or high basemean seemed to be even more downregulated, albeit with a higher overall variance.

To validate the changes in expression observed upon *DGCR8* depletion, RT-qPCRs for mature miRNAs were validated using the miRCURY LNA kit (Qagen). For *DGCR8*

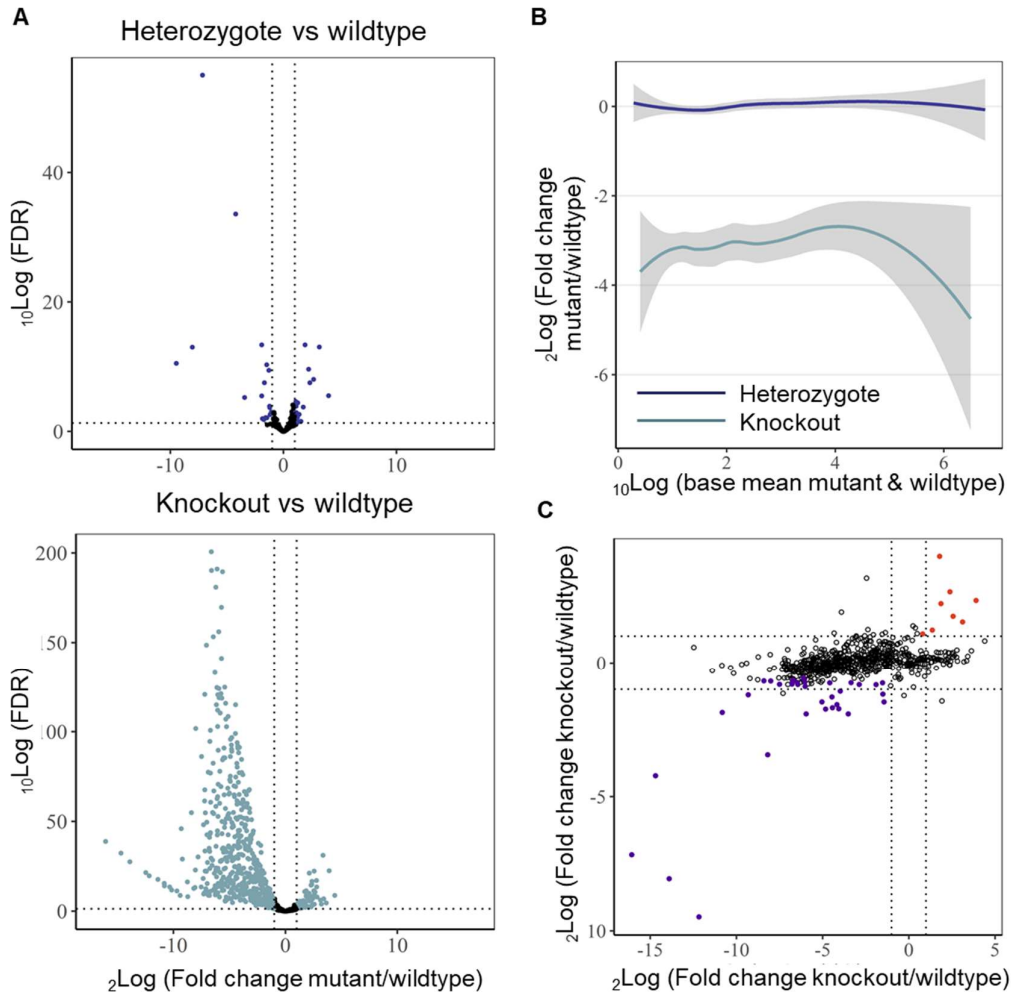


Figure 4.3. Differential expression analysis of miRNAs in wildtype and DGCR8 mutant cells. For fold change, *apeglm* shrunken values are used. **A.** miRNA volcano plots of DGCR8 heterozygote and knockouts cell lines compared to wild type. Significant miRNAs ($FDR < 0.05$) with an absolute fold change >1 are shown in colour, all other datapoints are shown in black. **B.** Adapted miRNA MA plot of DGCR8 heterozygote and knockouts cell lines compared to wildtype. The mean expression of wildtype and respective mutant are shown on the x-axis, while the y-axis represent the fold change of the respective cell line compared to wildtype cells. Overall trend per cell lines are represented by local polynomial regression. **C.** Correlation between fold change expression of both mutants versus wildtype. miRNAs who are differentially expressed in both mutants compared to wildtype ($FDR < 0.05$) are coloured, with downregulated in both is shown in purple and upregulated in both is shown in orange.

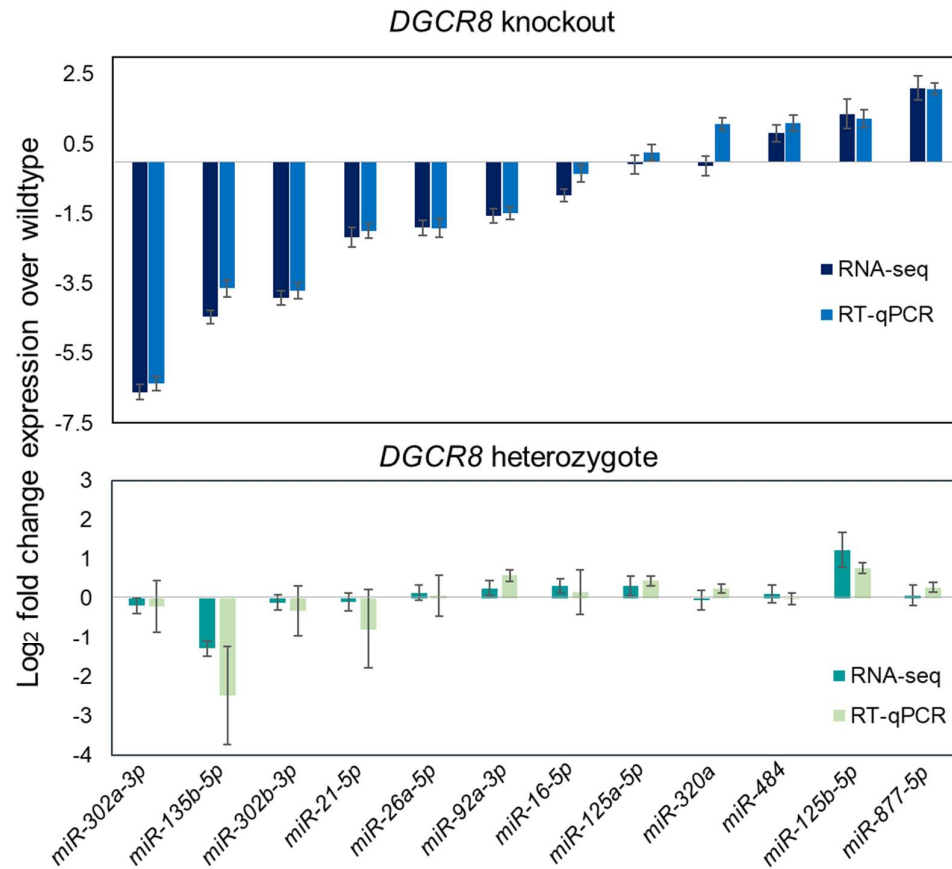


Figure 4.4 Validation of RNA-seq using RT-qPCR. The log₂ fold changes between the mature miRNA expression levels of *DGCR8*-deficient mutants and wildtype cells is shown +/- SE. N=4 for RNA-seq data, n=6 for *DGCR8* knockout qPCRs, n= 4 for *DGCR8* heterozygote qPCRs.

knockout versus wildtype miRNA expression, very similar changes in expression were obtained using both methods (figure 4.4). For *DGCR8* heterozygotes, the RT-qPCR data showed high variability, although the RNA-seq values showed similar trends.

Even though *DGCR8* knockout had a more severe effect compared to *DGCR8* heterozygosity, there were overlaps between significant changes in miRNA levels compared to wildtype (figure 4.3C). Of the 700 identified miRNAs, 8 and 34 miRNAs

were significantly up- and downregulated in both mutant cell lines, respectively (table 4.1). Since a decrease in miRNA levels is expected upon DGCR8 mutation, the upregulated miRNAs are of special interest. For example, *hsa-miR-10a-5p* and *hsa-miR-10a-5p* were upregulated in both *DGCR8* knockout and heterozygote cells. The two miRNAs with the highest wildtype expression in this table are *miR-135b-5p* and *miR-135a-5p*. Since these miRNAs belong to the same family, they are predicted to have similar targets, yet *miR-135a-5p* is upregulated and *miR-135b-5p* is downregulated in both mutant cell lines.

4.2.4 Differentially expressed miRNAs in *DGCR8* heterozygous cells target pathways involved in differentiation and pluripotency

Since in the *DGCR8* knockout cells the majority of miRNAs are differentially expressed, it is expected that this affects the expression of multiple target mRNAs in virtually all biological pathways to some extent. In the heterozygote cells, a smaller subset of miRNAs showed a change in expression, and therefore pathway analysis was carried out on the 55 miRNAs with an adjusted p-value below 0.01. These miRNAs were shown to be overrepresented in 74 KEGG pathways in Mirpath analysis (figure 4.5A). Remarkably, there was an overlap between the KEGG pathways overrepresented in our total transcriptomics analysis and the pathways shown here. Of these, several pathways are involved in pluripotency, differentiation and neuronal function, suggesting that the observed changes in mRNAs can potentially be explained by increased or reduced miRNA targeting. These results are relevant since *DGCR8* is one of the hemizygous genes in patients with 22q11.2 DS, which is characterised with developmental symptoms.



Figure 4.5. miRNA differential expression analysis by mirPath analysis. A. KEGG pathways significantly enriched ($FDR < 0.05$) in differential expression analysis of miRNAs. Pathways labelled in green are involved in pluripotency and development, while pathways labelled in pink are involved in neuronal function. B. Overview of differentially expressed miRNAs in heterozygote cells versus wildtype and involvement in each of the pathways labelled in A. miRNAs are sorted by $2\log$ fold change. The dotted line represents the change from downregulated to upregulated.

When comparing the individual miRNAs associated with the 5 KEGG pathways overlapping between the two analyses, only 36 miRNAs of 55 were involved in at least one of the pathways, including members of the *miR-767-105* cluster, the *miR-135* and *miR-10* family and *miR-125b-5p* that were mentioned previously (figure 4.5B). Of these, 22 miRNAs, both up- and downregulated, were involved in all 5 pathways, including *miR-105-5p*, *miR-125b-5p* and *miR-10a-5p*. This suggests that these miRNAs target either multiple transcripts in multiple pathways, or targets who themselves are involved in multiple pathways, and makes them interesting targets for future study.

4.2.5 The microprocessor index

To determine the efficacy of microprocessor cleaving, we applied an adapted version of the MPI (Conrad *et al.*, 2014). After pri-miRNA transcripts are processed by the microprocessor, the released pre-miRNAs are transported out of the nucleus, while the neighbouring sequences (from now on called the “flanks”) are retained before they are degraded. When processing is absent or ineffective, the unprocessed pri-miRNA will accumulate. This phenomenon forms the basis of the MPI index, which measures microprocessing efficacy by comparing the ratio of pre-miRNA sequence to its flanks interacting with chromatin (Conrad *et al.*, 2014)(figure 4.6A). In this thesis, the MPI is defined as the negative $\log_2$ -transformed ratio between the mean read depth of the pre-miRNA region to the mean read depth of the regions flanking the pre-miRNA.

Table 4.1. miRNAs that are either up- or downregulated in both mutant cell lines compared to wildtype.

miRNA	Wildtype count	Fold change heterozygote vs wildtype	FDR heterozygote vs wildtype	Fold change knockout vs wildtype	FDR knockout vs wildtype
hsa-miR-767-3p	46	-9.5	3.E-11	-12.2	2.E-20
hsa-miR-105-5p	132	-8.1	9.E-14	-13.9	2.E-28
hsa-miR-767-5p	509	-7.2	9.E-56	-16.1	1.E-39
hsa-miR-676-3p	215	-4.2	2.E-34	-14.7	4.E-33
hsa-miR-584-3p	29	-3.4	6.E-06	-8.2	2.E-13
hsa-miR-486-5p	67	-1.9	3.E-06	-3.5	4.E-22
hsa-miR-584-5p	827	-1.9	4.E-14	-6.0	1.E-124
hsa-miR-548s	21	-1.9	1.E-02	-10.8	8.E-15
hsa-miR-3138	21	-1.7	1.E-02	-4.8	1.E-12
hsa-miR-486-3p	31	-1.7	1.E-02	-4.1	4.E-10
hsa-miR-135b-3p	219	-1.7	3.E-08	-4.4	9.E-53
hsa-miR-184	51	-1.6	7.E-03	-4.2	3.E-15
hsa-miR-126-3p	5457	-1.5	5.E-11	-1.4	1.E-11
hsa-miR-548b-3p	30	-1.5	7.E-03	-5.0	1.E-19
hsa-miR-135b-5p	14113	-1.3	3.E-10	-4.5	1.E-115
hsa-miR-1306-5p	172	-1.2	2.E-04	-9.3	1.E-46
hsa-miR-126-5p	95	-1.2	2.E-03	-1.5	4.E-06
hsa-miR-193a-5p	240	-1.1	1.E-03	-4.0	9.E-43
hsa-miR-1301-3p	405	-0.9	1.E-02	-6.0	2.E-71
hsa-miR-29b-1-5p	186	-0.8	1.E-03	-1.9	1.E-18
hsa-miR-193a-3p	43	-0.8	4.E-02	-2.9	5.E-17
hsa-miR-942-5p	1334	-0.8	2.E-03	-6.4	4.E-154
hsa-miR-589-5p	520	-0.8	3.E-02	-7.5	7.E-87
hsa-miR-3144-3p	53	-0.8	5.E-02	-6.8	4.E-34
hsa-miR-1293	69	-0.8	3.E-02	-1.5	3.E-07
hsa-miR-3187-3p	315	-0.8	1.E-02	-4.6	2.E-65
hsa-miR-194-5p	352	-0.8	1.E-02	-3.4	1.E-34
hsa-miR-491-5p	285	-0.8	3.E-02	-6.5	3.E-83
hsa-miR-3173-5p	190	-0.7	9.E-03	-6.8	3.E-95
hsa-miR-192-5p	2923	-0.7	1.E-02	-6.0	3.E-121
hsa-miR-675-3p	307	-0.7	2.E-02	-8.0	2.E-102
hsa-miR-589-3p	147	-0.7	4.E-02	-8.4	1.E-55
hsa-miR-624-5p	359	-0.7	4.E-02	-6.7	2.E-101
hsa-miR-625-5p	6248	-0.6	1.E-02	-6.1	9.E-192
hsa-miR-135a-5p	6318	1.1	2.E-05	0.8	4.E-04
hsa-miR-125b-5p	150	1.2	7.E-03	1.4	5.E-04
hsa-miR-1228-3p	6	1.5	3.E-02	3.1	6.E-07
hsa-miR-10b-5p	62	1.7	2.E-04	2.6	4.E-10
hsa-miR-219a-1-3p	29	2.2	2.E-10	1.9	5.E-09
hsa-miR-196b-5p	160	2.3	3.E-08	3.9	3.E-23
hsa-miR-219a-5p	8	2.7	9.E-09	2.4	8.E-09
hsa-miR-10a-5p	15	4.0	3.E-06	1.8	6.E-03

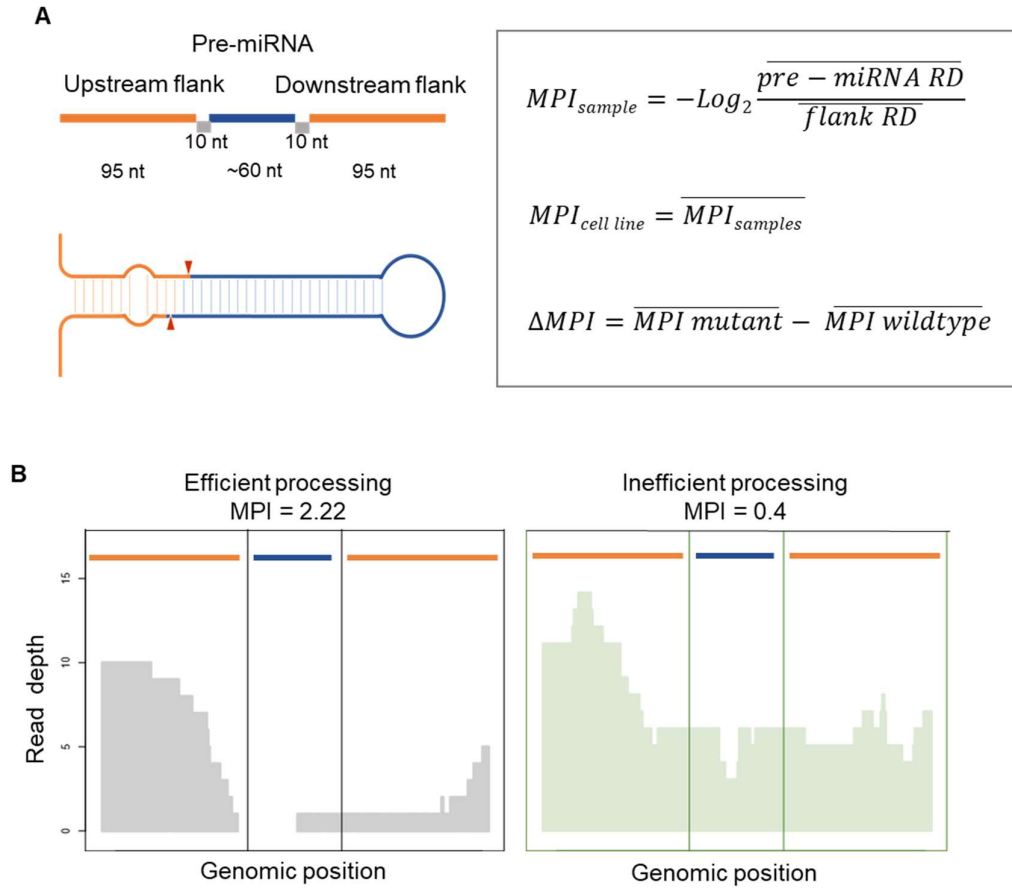


Figure 4.6. MPI as a measurement of pri-miRNA processing. A. Overview of pri-miRNA regions and MPI calculations. RD=read depth. B. Read depth around pre-miRNA regions, for pri-miRNA-34a. Left: wildtype sample 3; right: DGCR8 knockout sample 2.

Therefore, a high MPI indicates effective processing by the microprocessor while a negative value or a value close to zero indicates there is either no processing or processing is very ineffective (figure 4.6B). By comparing the MPIs of pri-miRNAs of our DGCR8 mutant cell lines, the effect of loss of one or two copies of the DGCR8 gene on pri-miRNA processing can be determined.

4.2.6 Purification of chromatin-associated RNA and MPI determination

Since pri-miRNAs are mostly processed co-transcriptionally by the microprocessor in the nucleus, chromatin-associated RNA was purified to see whether this fraction was enriched for pri-miRNAs. Cellular fractions were obtained by subjecting the PA-1 cells to multiple lysis steps and sonication of their chromatin to release the chromatin-associated RNA. This protocol led to a clear separation of fractions on both protein and RNA level, with both Histone 3 and unspliced mRNA being strongly enriched in the chromatin-associated fraction, as expected (figures 4.7A, B). Moreover, the chromatin-associated RNA showed enriched levels of pri-miRNAs (figure 4.7B) and therefore this fraction was used for RNA-seq and subsequent analysis. The RNA-seq returned at least 24 million high quality reads per sample (figure 4.7C). These reads were mapped to the human genome and the average read depth for pre-miRNAs and the surrounding flanks was determined per sample in order to calculate the MPI.

In order to be included in our dataset, the pri-miRNAs had to pass several requirements:

- The overall mean expression of the flanks should exceed 1, meaning that on average at least 1 read is covering each nucleotide in the pri-miRNA flanking regions
- The overall mean ratio of upstream and downstream flanks and vice versa should not exceed 4, to prevent the presence of exons in the flank region to confound pri-miRNA expression.
- The miRNA is no mirtron, tailed mirtron or capped miRNA, since these are microprocessor independent.

This approach led to MPI values for 101 pri-miRNAs. Although there was a clear correlation between the MPIs of individual samples of the same cell line, the overall read depth for individual samples was quite low. Therefore, in order to reduce noise, it was decided to batch the data per cell line and to look at overall trends in

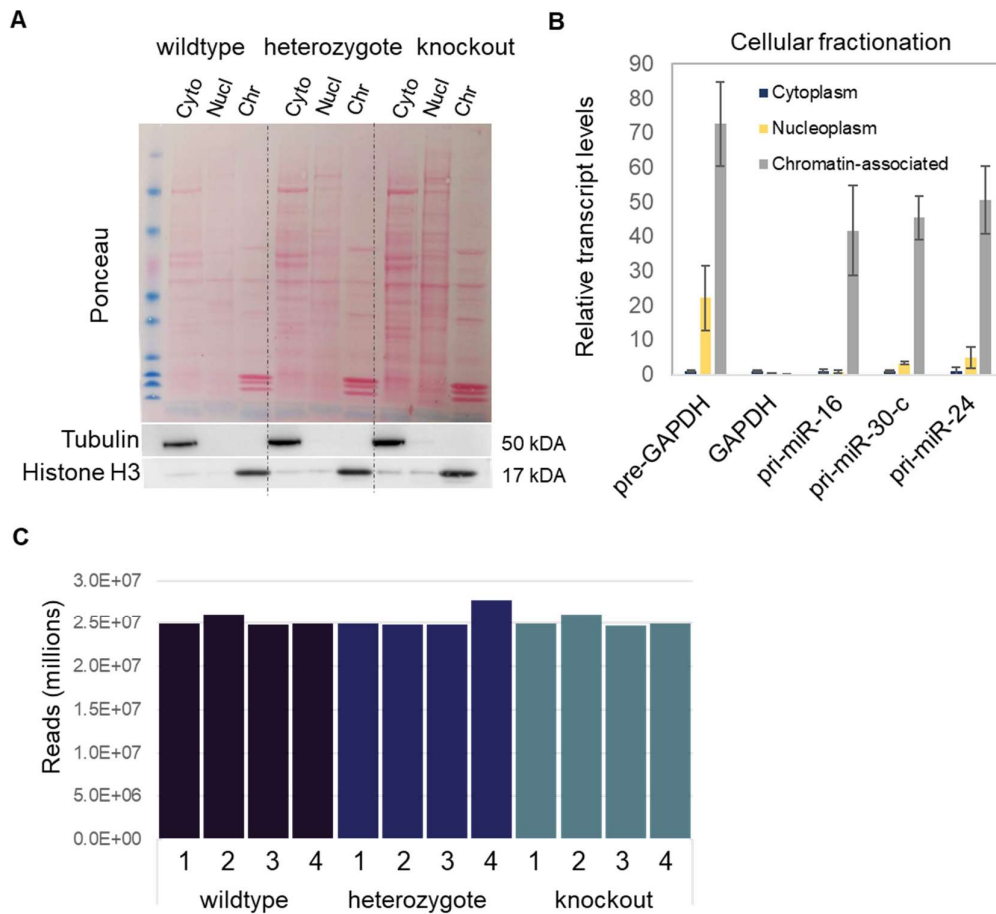


Figure 4.7. Cellular fractionations and chromatin-associated RNA-seq. A. Ponceau staining and Western blot from cellular fractions. B. RT-qPCR analysis from cellular fractions in PA-1 wildtype cells. Results are shown as the average transcript normalised to cytoplasm +/- SE with n=3. C. Library size of paired reads obtained after chromatin-associated RNA-seq.

MPI within and between cell lines instead of focussing on individual pri-miRNAs. A range of MPI values was observed for the three cell lines (figure 4.8).

An MPI value >1 indicated that there the sequencing read depth for the flanking sequences is at least twice as high as the pre-miRNA sequence and was chosen as a cut-off for effective processing. 29 pri-miRNAs showed effective processing in wildtype samples, compared to 19 in *DGCR8* heterozygotes and 9 in knockouts.

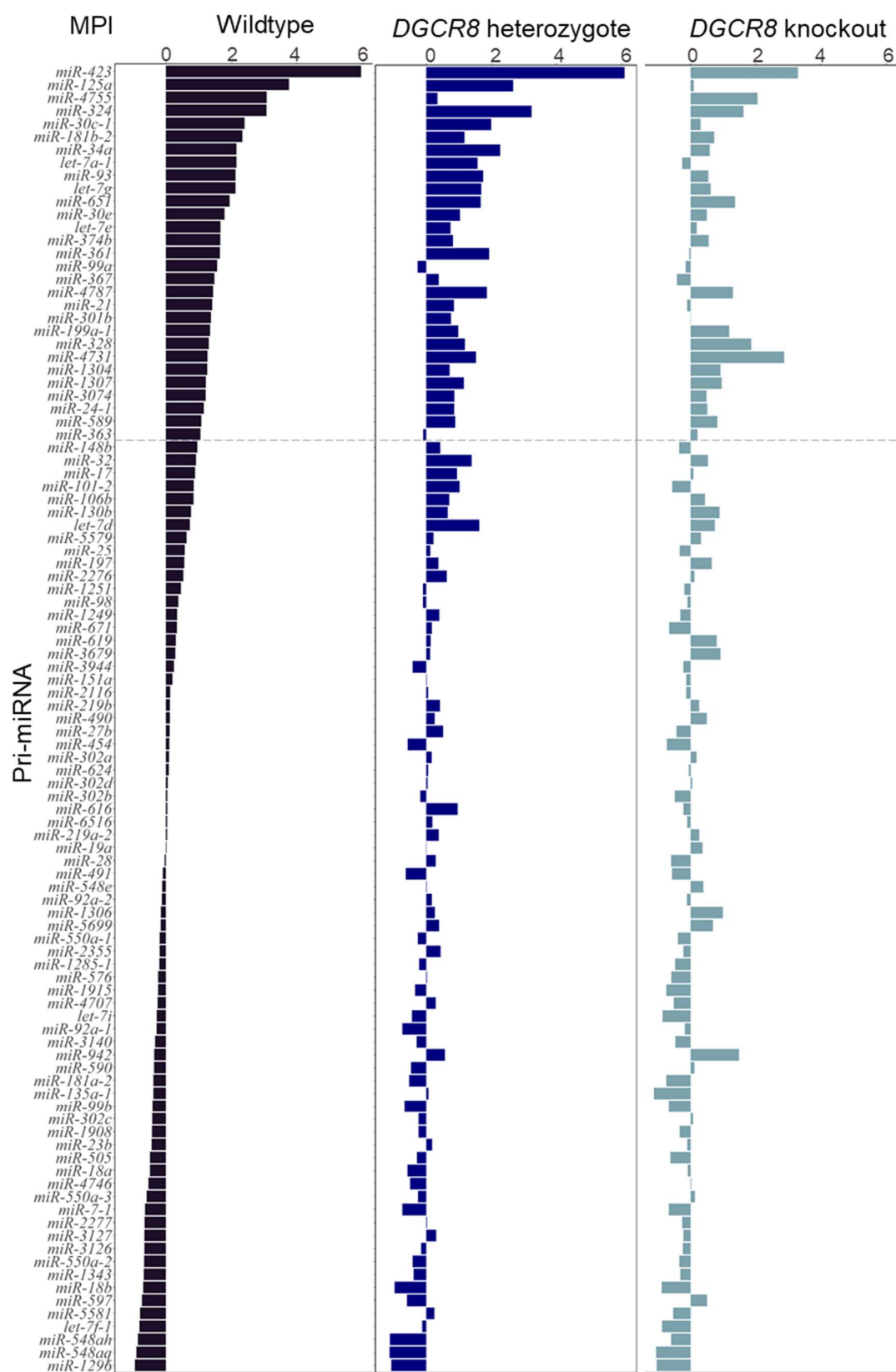


Figure 4.8. Overview of detected pri-miRNAs and their MPI values in different cell lines. Pri-miRNAs are sorted by MPI in wildtype samples, the horizontal dashed line indicates an MPI of 1.

4.2.7 Microprocessor efficacy is affected in *DGCR8* knockout cells

When the MPI values for *DGCR8* wildtype, heterozygote and knockout cell lines were compared to each other, the MPI in the knockout was significantly reduced compared to wildtype (figure 4.9A). *DGCR8* heterozygotes did not show a statistically significant change but seemed to have an intermediate distribution. Next, we studied the change in MPI of *DGCR8* mutants versus wildtype cells, dependent on efficacy in wildtype cells. For both *DGCR8* heterozygote and knockout cells, the MPI was more affected in pri-miRNAs that were effectively processed in wildtype cells (figure 4.9B).

Next, we studied the relationship between pri-miRNA abundance and the MPI. For this, the normalised mean read depth of the flanking regions was calculated per cell line. Although a weak relationship between pri-miRNA levels and MPI was detected for wildtype cells, there was no correlation between changes in pri-miRNA and MPI between *DGCR8* mutant and wildtype cells detected (figure 4.9C, D).

Previous research has shown that the presence of specific sequencing motifs influences microprocessor efficacy (Auyeung *et al.*, 2013; Conrad *et al.*, 2014; Fang and Bartel, 2015; Nguyen *et al.*, 2015; Dang *et al.*, 2020). We studied the MPI for pri-miRNAs with UG, CNNC or UGU/GUG motifs based on their annotation in MirGeneDB 2.0 (Fromm *et al.*, 2020). Pri-miRNAs with a UG motif showed higher MPI levels in wildtype cells, albeit not significant (figure 4.10). The pri-miRNAs with this motif were more affected by *DGCR8* knockout. Apart from the UG motif, no correlations between sequence motif and processing was found.

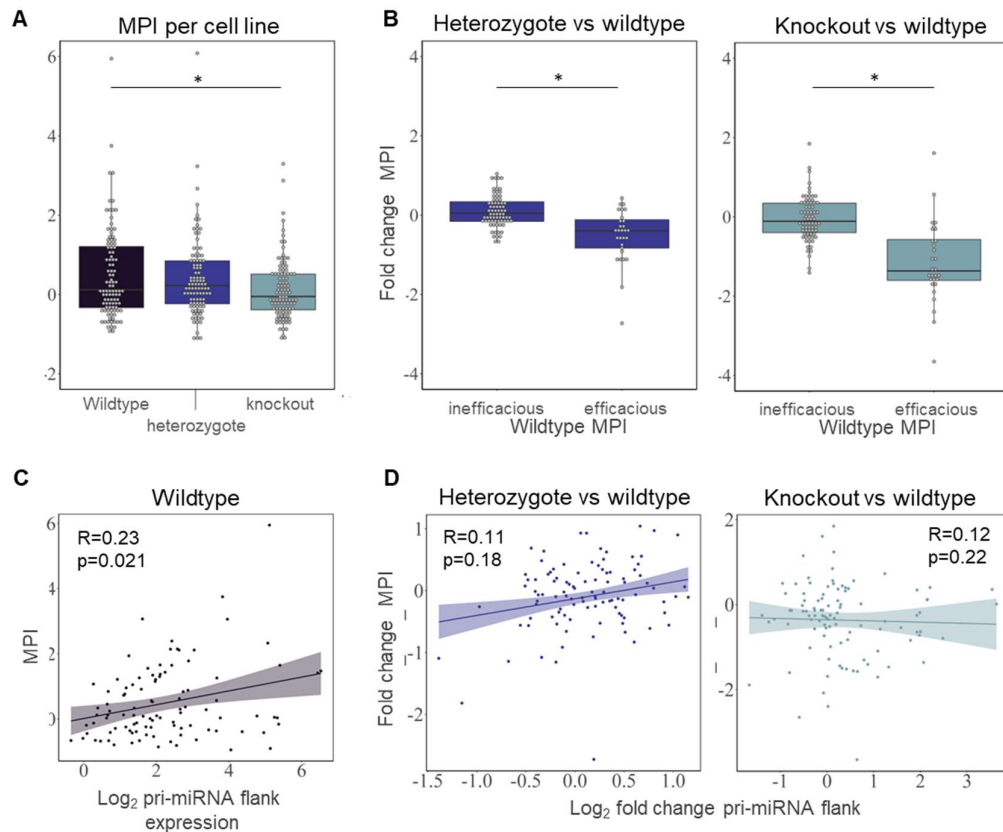


Figure 4.9 The MPI in DGCR8 wildtype and mutant cells. **A.** Boxplots of MPI distribution of DGCR8 wildtype, heterozygote, and knockout cells. * $p < 0.05$ for Tukey's HSD test following Anova, followed. **B.** Boxplots of fold change MPI between DGCR8 mutant and wild type cells, separated by MPI efficacy (wildtype MPI above or below 1). * $p < 0.05$ for Welch's t -test. **C.** Correlation between wildtype pri-miRNA expression and wildtype MPI. Linear regression is shown by line surrounded by 95% confidence interval. Correlation coefficient and statistical analysis was carried out using Spearman's ranking. **D.** Correlation between changes in flank expression and MPI between DGCR8 mutants and wildtype cells. Visualisation and analysis as in C.

Lastly, we were interested in the relationship between pri-miRNA, MPI and mature miRNA levels, since previous studies found correlations between them (Feng *et al.*, 2011; Conrad *et al.*, 2014). Upon comparing the MPI and mature miRNA levels in wildtype cells, we found no difference between effectively and non-effectively processed pri-miRNAs (figure 4.11A). Next, we divided the fold change MPI levels in affected (MPI increased) and not affected (MPI decreased) (figure 4.11B). While

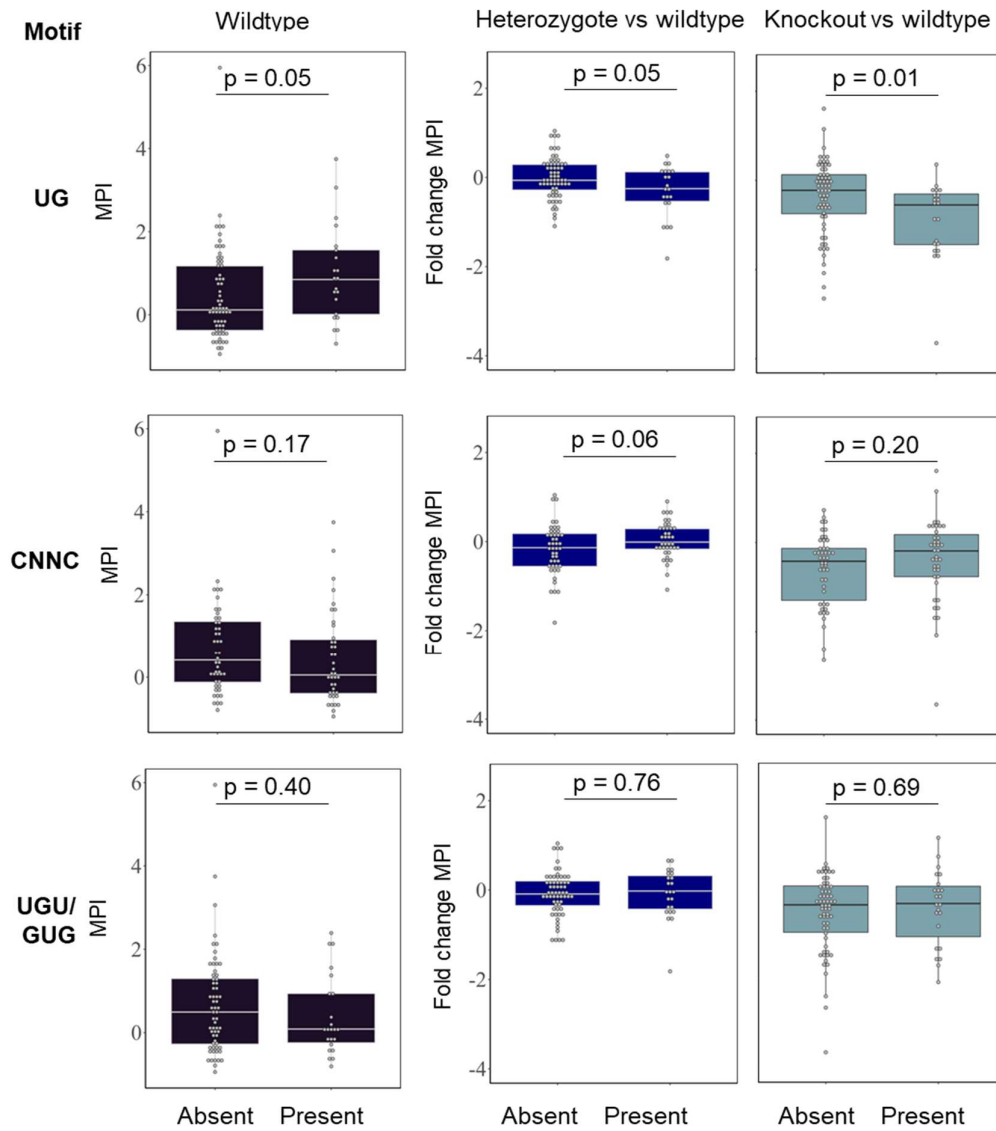


Figure 4.10. Relationship between the presence of previously described sequence motifs related with pri-miRNA processing and measured MPI. To test for significance, the Wilcoxon rank sum test was carried out.

changes in MPI did not change the relative miRNA levels in heterozygotes, surprisingly a decrease in processing correlated with an increase in mature miRNA levels. While MPI was not a good predictor for miRNA expression, there was, however, a weak correlation between pri-miRNA levels and mature miRNAs (figure

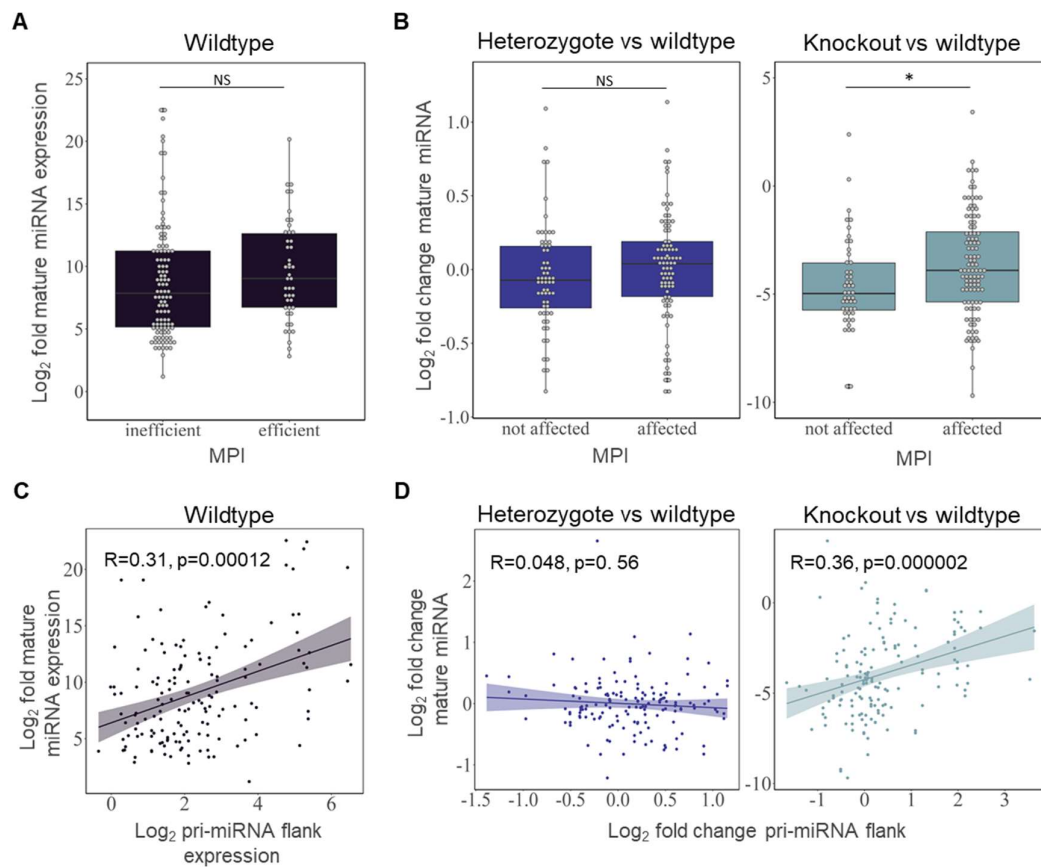


Figure 4.11. Relationship between MPI and pri-miRNA expression and mature miRNA levels. A,B. Boxplots showing the correlation between MPI and mature miRNA levels for wildtype or fold changes between the variables for the DGCR8 mutants and wildtype. * $p < 0.05$ for Wilcoxon signed-rank test. C. Correlation between wildtype pri-miRNA expression and mature miRNA levels. Linear regression is shown by line surrounded by 95% confidence interval. Correlation coefficient and statistical analysis was carried out using Spearman's ranking. D. Correlation between changes in flank expression and mature miRNAs between DGCR8 mutants and wildtype cells. visualisation and analysis as in C.

4.11). While no relationship was found for DGCR8 heterozygotes, changes in pri-miRNA levels weakly correlated with changes in mature miRNAs for DGCR8 knockout cells compared to wildtype cells (figure 4.11C, D).

4.3 DISCUSSION

4.3.1 Knockout of *DGCR8* transforms the miRNA landscape

Upon knockout of *DGCR8*, a dramatic decrease in overall miRNA levels was detected. Since *DGCR8* is part of the canonical miRNA biogenesis machinery, such a decrease was expected, and has been described before for truncated *DGCR8* and knockouts of *DROSHA* and *DICER* in human cell lines (Kim, Kim and Kim, 2016; Nguyen *et al.*, 2018). Given the regulatory effect of miRNAs on the coding- and noncoding transcriptome, it is expected that the observed changes in miRNA levels are at least partly responsible for the transcriptome changes described in chapter 3.

Although the majority of mature miRNAs are downregulated in *DGCR8* knockout cells, they still can be detected, albeit to a lesser extent compared to wildtype, suggesting that the miRNA biogenesis pathway is not completely ablated in absence of *DGCR8*. Apart from the absolute changes in miRNA levels, the relative amounts of miRNAs also showed a difference, potentially reflecting an altered turnover of miRNAs with relative overexpression of essential or non-redundant miRNAs. A few well-established miRNAs even showed upregulation, such as *miR-125b-5p*, *miR-135a-5p* and members of the *miR-10* family. When comparing *DGCR8* wildtype and knockout cells, mirtron levels were overall increased. Since mirtrons are per definition *DGCR8*-independent, their biogenesis should not be impaired by knockout of *DGCR8*. A hypothesis explaining the increased expression could be that there are overall less miRNAs available to associate with AGO, and therefore the expressed miRNAs are stabilized by being part of a RISC.

The presence of canonical miRNAs in absence of microprocessor factors has been reported before (Kim, Kim and Kim, 2016; Nguyen *et al.*, 2018). There could be several reasons for the expression of these miRNAs, both technical and biological: sample contamination, sequencing artifacts, *DGCR8*-independent *DROSHA* cleaving of pri-miRNAs or cleaving by alternative endonucleases.

Preliminary validation of several miRNAs by RT-qPCR in RNA samples from novel RNA extractions detected the presence of canonical miRNAs, at similar levels as detected in this sequencing analysis, making it unlikely that the presence of miRNAs in *DGCR8* knockout cells are caused by contamination of the RNA samples used for sequencing. It is also improbable that sequencing artifacts are responsible for the presence of canonical miRNAs, since there is a clear overall correlation between miRNAs in wildtype and knockout samples. In addition, if a read contains mismatches or deviates too much from the miRBase annotation, it is excluded from our analysis.

MirDeep2's default parameters include reads as miRNAs when they map within 2 nucleotides upstream or 5 nucleotides downstream of the annotated mature miRNA sequence, thereby allowing certain isomiRs to be included (Friedländer *et al.*, 2012). However, in our dataset we found no indications for substantial differences in isomiR ratios between our cell lines, neither by manual inspection of annealing reads nor by varying the inclusion parameters in miRDeep2 (data not shown). Since annotated RNase III proteins have specialized functions and are optimized for a specific substrate (MacRae and Doudna, 2007), it is unlikely that pri-miRNA cleaving by an alternative, non-canonical endonuclease would have the same precision, resulting in isomiRs. Since we could not find evidence of alternative pri-miRNA cleaving in our MPI analysis or in the presence of isomiRs, it is unlikely that the absence of *DGCR8* causes alternative pri-miRNA processing.

Recent research indicates that a large group of miRbase-annotated pri-miRNAs showed little to no processing or microprocessor association in DROSHA cleavage studies, indicating that many pri-miRNAs are possibly microprocessor-independent and that their miRNAs are wrongly annotated as canonical (Kim, Jeong and Kim, 2017; Rice *et al.*, 2020; Kim *et al.*, 2021). This might explain the lowly expressed, but detectable miRNAs in the *DGCR8* knockout cells. An alternative reason explaining the presence of canonical miRNAs would be that DROSHA is able to cleave pri-miRNAs independently of *DGCR8 in vivo*. Earlier research showed that when

DROSHA was knocked down in hESCs a subset of miRNAs was preferentially processed (Sperber *et al.*, 2014). In this case, the absence of DGCR8 could have led to lower DROSHA levels and, together with decreased processing efficacy, have reduced the miRNA biogenesis rate and altered pri-miRNA preference. This hypothesis should be further explored, for example by CLIP-seq of DROSHA in *DGCR8* knockout cells.

4.3.2 *DGCR8* heterozygosity leads to limited changes in miRNA levels compared to wildtype cells

Compared to a complete *DGCR8* knockout, *DGCR8* heterozygote cells showed a miRNA profile more similar to wildtype cells. Previous research in both human and mouse models show either limited changes, or an overall decrease in miRNAs in *DGCR8* heterozygosity (Stark *et al.*, 2008; Marinaro *et al.*, 2017). Apart from altered miRNA biogenesis, the observed changes could be explained by several other transcriptional and post-transcriptional parameters, such as chromatin accessibility, transcription factor availability, alternative splicing of long pri-miRNA precursors and miRNA turnover.

Members of the same miRNA family often showed similar changes in expression when *DGCR8* wildtype and heterozygote cells are compared. In other cases expression of a whole cluster of miRNA, such as with *miR-767*, *miR-676* and *miR-105*. This might be explained by altered pri-miRNA transcription of a specific cluster as opposed to microprocessor cleaving. Methods such as RNA-pol II ChIP-seq could further research the impact of transcription on miRNA biogenesis.

Out of 700 measured miRNAs, 87 were significantly up- or downregulated. Interestingly, the majority of the most differentially expressed miRNAs showed an overlap with *DGCR8* knockout cells. Since the *DGCR8* heterozygote cells used in this chapter stem from the *DGCR8* knockout cells, this might be caused by an altered

background compared to wildtype: validation in the other *DGCR8* mutant cell lines by RT-qPCR could clarify this. Pathway analysis using miRpath showed that many of the differentially expressed miRNAs are involved in the same pathways. The enriched KEGG pathways showed an overlap with the pathways detected in chapter 3 for messenger RNAs that are involved in pluripotency and neuronal function. The effect of *DGCR8* mutation on pluripotency is studied in more detail in chapter 5.

4.3.3 The MPI as a determinant of microprocessor cleaving

To learn more about the effect of *DGCR8* mutation on miRNA biogenesis, we studied pri-miRNA processing in wildtype, *DGCR8* heterozygote and *DGCR8* knockout cells by measuring expression of endogenous chromatin-associated RNA levels. While our fractionation protocol effectively enriched for chromatin, overall read depth across pri-miRNAs was very low and complicated determining the MPI of individual pri-miRNAs. Therefore, the results of individual replicates from the same cell lines were pooled to infer microprocessor function on a cell line level. After filtering of low-read pri-miRNAs and microprocessor-independent miRNAs, our conclusions are based on a limited set of pri-miRNAs and make it difficult to detect trends and draw conclusions. On the other hand, filtering might not be stringent enough, since pre-miRNA regions might overlap with transcription start- and end site and intron/exon junctions. Furthermore, a weak correlation between flank expression and while this suggest a positive correlation between pri-miRNA levels and processing, it could also mean that our filtering was not stringent enough.

Interestingly, pri-miRNAs of the *miR-302* family, although having the most highly expressed mature miRNAs, are not effectively processed, with an MPI of around 0. This suggest that either downstream processes such as low miRNA degradation could be responsible for the high mature miRNA levels, or that, even if ineffective, enough processing is carry out to maintain the high levels of expression. The latter

explanation is likely since the flanking regions of the *miR-302* pre-miRNAs are among the most highly detected in our chromatin-associated RNA-seq dataset. On the other hand, pri-miRNAs might accumulate due to limited processing.

Despite the technical challenges, there was a clear difference in MPI observed between wildtype and *DGCR8* knockout cells, indicating that pri-miRNA processing is indeed negatively affected. Effectively processed pri-miRNAs in wildtype cells were more affected by *DGCR8* mutation than poorly processed pri-miRNAs, corresponding with previous research that showed effectively processed hairpins were found to be more sensitive to mutations (Rice *et al.*, 2020). No significant difference in overall MPI was found between wildtype and *DGCR8* heterozygote cells, indicating that the differentially expressed miRNAs observed are the result of altered turnover regulated elsewhere.

Previous studies with the MPI showed that the presence of UG, UGU/GUG and CNNC motifs all were indicative of better processing, with the CNNC motif showing the strongest correlation (Conrad *et al.*, 2014). In our analysis, only the UG motif was associated with better processing, and pri-miRNAs in the *DGCR8* knockout cells without this motif were more severely affected in processing. The other reported motifs were not associated with better processing or affected in *DGCR8* mutants. Besides using a different cell line, there could be several explanations for this. Firstly, our read depth might be too low to precisely detect the MPI or our filtering of pri-miRNAs might not be stringent enough, thereby introducing noise. On the other hand, only a small subset of pri-miRNAs contains these motifs. Lastly, the interaction between the presence of multiple motifs was not taken into consideration. This could be a focus of future studies.

Since we measured both microprocessor efficacy and mature miRNA levels, we could test if the two measurements were correlated. In our limited dataset, we did not find a significant correlation between effective processing and higher miRNA levels. There could be several technical and biological explanations for this. Firstly,

the number of miRNAs with an MPI score is low, therefore our combined dataset has limited power. Alternatively, contrarily to previous reports, it is possible that the MPI is not a good predictor of miRNA biogenesis. Previous studies reported that AGO loading is a bottleneck in miRNA biogenesis (Reichholf *et al.*, 2019), and that microprocessor cleaving happens so quickly it is not a limiting factor, which might explain the lack of changes observed in the MPI of *DGCR8* heterozygote cells. In our dataset, pri-miRNA levels in wildtype cells correlated with the MPI, suggesting transcription is more important for biogenesis than microprocessor activity. Alternatively, it has been reported that a subset of pri-miRNAs is processed post-transcriptionally (Louloupi *et al.*, 2017). Since we measure the MPI on chromatin-associated RNA, the reported MPI of post-transcriptionally pri-miRNAs might therefore be too low. Measuring the MPI of the nuclear cellular fraction or pulse chase experiments can give more insights in the timing and localisation of microprocessor cleaving.

CHAPTER 5: THE ROLE OF DGCR8 IN INNATE IMMUNITY AND DEVELOPMENT

5.1 INTRODUCTION

DGCR8 plays an important role for in both immunity and development, either through its role in canonical miRNA biogenesis or because of non-canonical mechanisms.

As described in section 1.4, the relationship between miRNAs and the type-I IFN response is intricate: while several miRNAs are known to affect the expression of IFNs and ISGs, IFN activation downregulates overall miRNA biogenesis (Sperber *et al.*, 2014; Witteveldt, Ivens and Macias, 2018). Interestingly, mammalian ESCs are unable to elicit an IFN response upon stimulation with the dsRNA analogue poly I:C, which can be recognized by the RLRs RIG-I and MDA5 (Kato *et al.*, 2006).

Besides their importance in immunity, both the microprocessor and miRNAs are also essential for development. In absence of DGCR8, mouse ESCs fail to silence the pluripotency factors *Oct4* and *Nanog* and to form differentiated teratomas upon injection in mice (Wang *et al.*, 2007). However, the effects of *DGCR8* knockout on differentiation in human ESCs had not been tested so far. Previous studies in mouse models and cell lines showed that knockdown of *DGCR8* causes a lethal phenotype. This has been attributed to both canonical and non-canonical functions of DGCR8. Indeed, the role of individual miRNAs in development is well studied (Giraldez *et al.*, 2005; Barca-Mayo and De Pietri Tonelli, 2014; Parchem *et al.*, 2015; Pons-Espinal *et al.*, 2017). However, several studies have found non-canonical functions of DGCR8 to be the cause of aberrant pluripotency and neurogenesis (Cirera-Salinas *et al.*, 2017; Marinaro *et al.*, 2017).

Interestingly, patients with 22q11.DS are hemizygous for DGCR8, among other genes, and often have developmental delays. The impact of DGCR8 and its role in miRNA biogenesis in 22q11.2 syndrome has been understudied.

In this chapter we investigate the role of DGCR8 in the IFN response in the pluripotent context using both a mouse and a human cell line. Furthermore, we

study the effect of DGCR8 knockout and heterozygosity on the pluripotency status and the differentiation potential in human PA-1 cells.

5.2 RESULTS

5.2.1 *mmu-miR-673* downregulates MAVS and increases viral sensitivity in mESCs

To study activation of the type-I IFN response in mammalian early development, the mouse ES cell line V6.5, and a *Dgcr8* knockout cell line derived from it, were used (Wang *et al.*, 2007). Transfection of these cells with poly I:C did not result in *Ifnb* production (figure 5.1A). In contrast, NIH3T3 murine fibroblast cells showed higher baseline expression of *Ifnb*, while poly I:C stimulation led to a more than 400-fold increase over mock transfection, suggesting that ESCs are incapable of eliciting a type-I IFN response, as has been reported previously (Wang *et al.*, 2013). Interestingly, knockout of *Dgcr8* in these cells led to increased *Ifnb* production upon poly I:C transfection (figure 5.1B). In addition, the same trend was observed for mESCs knockout for Dicer compared to its Dicer-proficient counterpart (Witteveldt, Knol and Macias, 2018). This suggests that the absence of miRNAs, as opposed to miRNA-independent functions of DGCR8, is responsible for the downregulation of the type-I IFN response in mESCs.

In order to understand the functional relevance of this regulation in a more physiological context, we infected mESCs with Theiler's murine encephalomyelitis virus (TMEV). We showed that mESCs were more susceptible to this virus than differentiated murine cells lines (Witteveldt, Knol and Macias, 2018). Next, to study the effect of *Dgcr8* knockout on viral susceptibility, v6.5 wildtype cells, *Dgcr8* ^{-/-} cells and stable *Dgcr8* rescue cell lines of the knockout were infected with different infectious doses of TMEV (figure 5.1C). Compared to wildtype cells, *Dgcr8* ^{-/-} lead to decreased viral susceptibility, as measured by the relative amount of viral genome

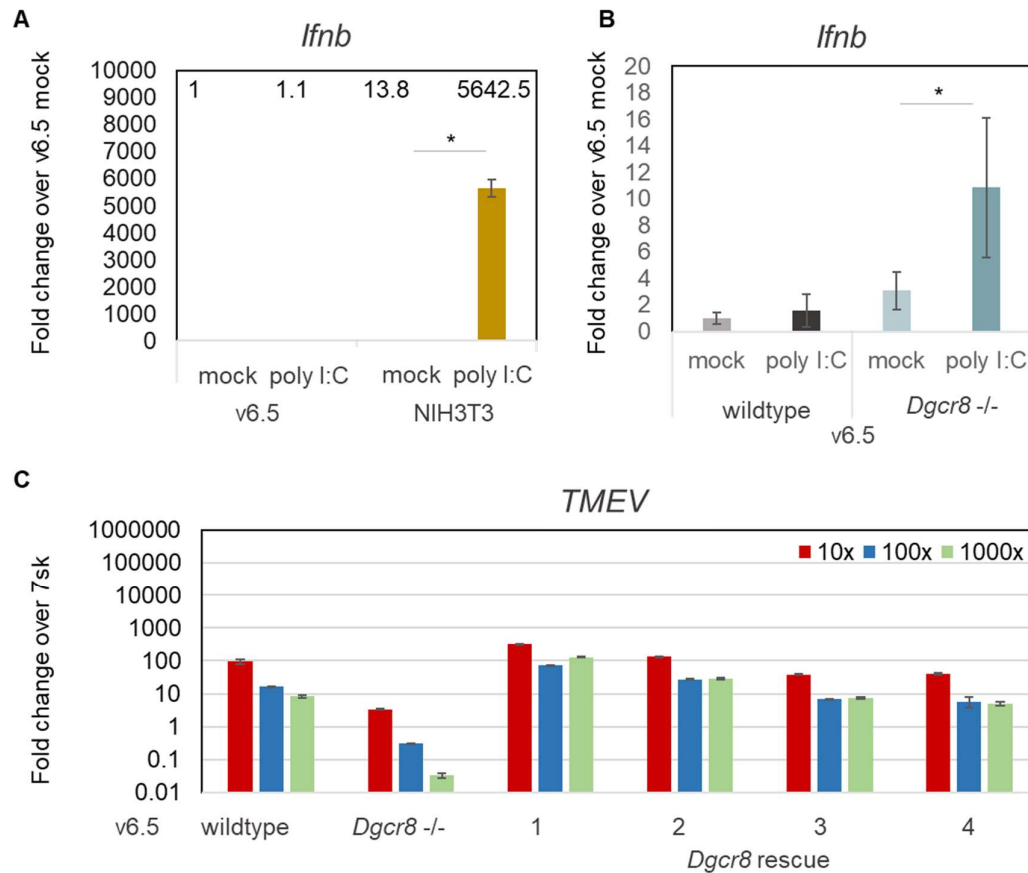


Figure 5.1 miRNAs repress the type-I IFN response and increase viral susceptibility. **A.** *Ifnb* expression of v6.5 and NIH3T3 cells after either mock or poly I:C transfection. Expression was measured by RT-qPCR, normalized to *Gapdh* and represented as fold change over v6.5 wildtype cells mock transfected. Data are represented as mean $\pm$ SD with $n=3$ for each group. Significance level of the difference between mock and poly I:C treatment was determined by two-tailed Welch *t*-test. * $p<0.05$. **B.** *Ifnb* expression of v6.5 wildtype and *Dgcr8*^{-/-} cells as described in A. **C.** Cellular quantities of TMEV relative to expression of endogenous 7sk as measured by RT-qPCR 24 hours after infection of v6.5 wildtype cells and mutants. Serial dilutions of TMEV were used, corresponding to an estimated MOI of 1.5, 0.15 and 0.015 in wildtype cells. Data are represented as mean $\pm$ SD with $n=2$ for each group.

by RT-qPCR. Rescue of the *Dgcr8* knockout by stable overexpression of *Dgcr8* reversed this phenotype, suggesting that the observed effects are *Dgcr8*-specific and not caused by *Dgcr8*-independent alteration in the knockout cell line. This led

us to the hypothesis that specific miRNAs are targeting the type-I IFN pathway in mESCs and are affecting viral susceptibility.

In order to identify the mechanism by which miRNAs silence the IFN response in ESCs, we performed proteomics analysis. Mass spectrometry datasets of total cell extracts derived from *Dgcr8* ^{-/-} cells were compared to *Dgcr8* rescue cells, and showed alterations in expression of mitochondrial and mitochondria-associated proteins, including MAVS (Witteveldt, Knol and Macias, 2018). Therefore, research focused on murine miRNAs that have been predicted to target MAVS or that have been associated with viral immunity. A potential candidate was found in *mmu-miR-673*. This murine-specific miRNA is predominantly expressed in mESCs and is predicted to target *Mavs* (Agarwal *et al.*, 2015; Fromm *et al.*, 2015). Knockdown or transfection of *miR-673* led to and overexpression or reduction of MAVS, respectively (Witteveldt, Knol and Macias, 2018). To study the role of this miRNA in the antiviral response, we aimed to create v6.5 *mmu-miR-673* knockout cell lines using CRISPR-Cas9.

In order to create *miR-673* mutants, two guide RNAs were designed that targeted the *miR-673* hairpin (figure 5.2A). Since disruption of the RNA hairpin structure would already prevent microprocessor recognition and cleaving, the sequence of *miR-673-5p* itself did not necessarily have to be mutated in order to create a null mutant. Transfection of Cas9 and guide RNAs resulted in several clones selected for further analysis. Sanger sequencing of the *miR-673* hairpin region confirmed deletions in the hairpin sequence in both alleles for 3 clones, which were selected for further characterization (figure 5.2B).

First, absence of mature *miR-673-5p* expression was confirmed by miRNA-specific RT-qPCR (figure 5.3A). As opposed to wildtype cells, no *miR-673* could be detected in either *Dgcr8* ^{-/-} or any of the *miR-673* ^{-/-} cells. While *Dgcr8* ^{-/-} cells were significantly delayed in cell growth, *miR-673* ^{-/-} cells showed similar rates of cell division in culture (figure 5.3B). Furthermore, no altered cell morphology or

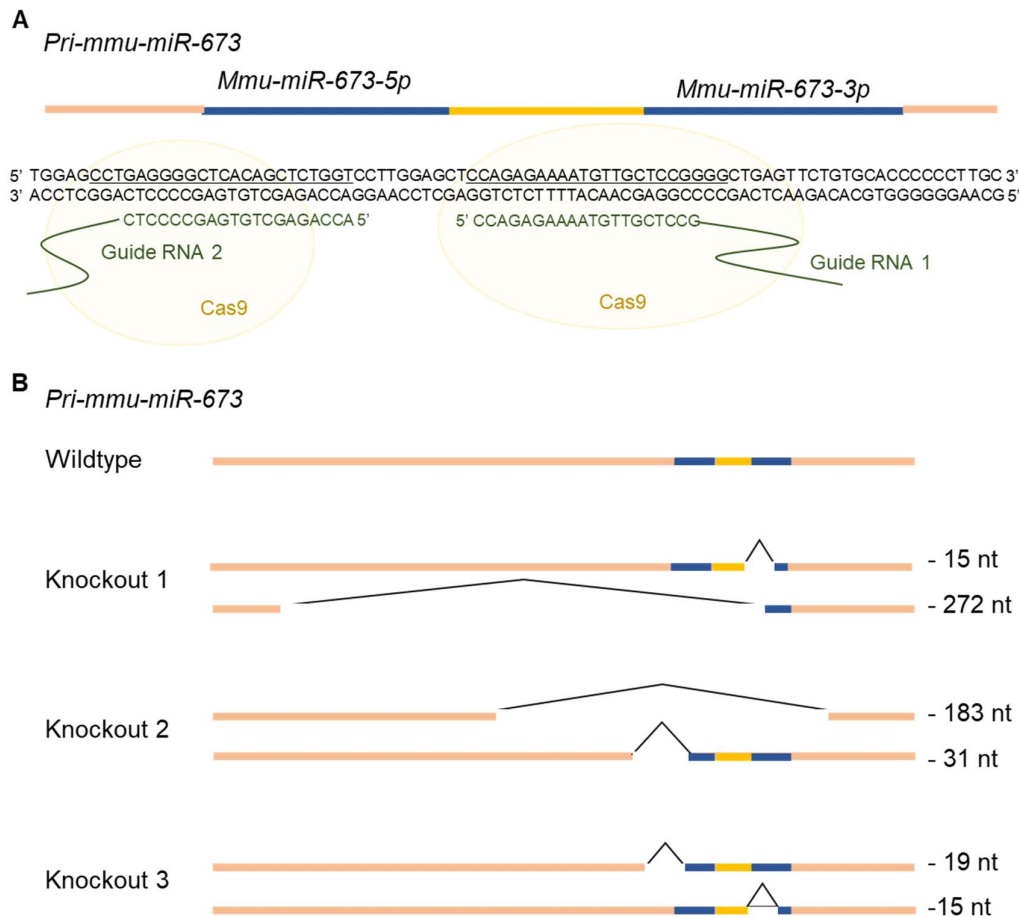


Figure 5.2. Knockout of *mmu-miR-673* using CRISPR-Cas9. A. Two different guide RNA targeting the *miR-673* hairpin were used to induce mutations. B. Three different *miR-673* knockouts were created with deletions in the hairpin sequence.

expression of pluripotency factors was detected in *miR-673* $-/-$ cells compared to wildtype, minimising potential explanatory parameters for altered antiviral immunity in these cells (Witteveldt, Knol and Macias, 2018).

Since *miR-673-5p* was predicted to target *Mavs*, the presence of MAVS was measured by western blot (figure 5.3C). Absence of *Dgcr8* $-/-$ showed an increase in MAVS protein, as indicated by mass spectrometry. Knockdown of *miR-673* also led to an increase of MAVS compared to wildtype cells, although not to the extent of

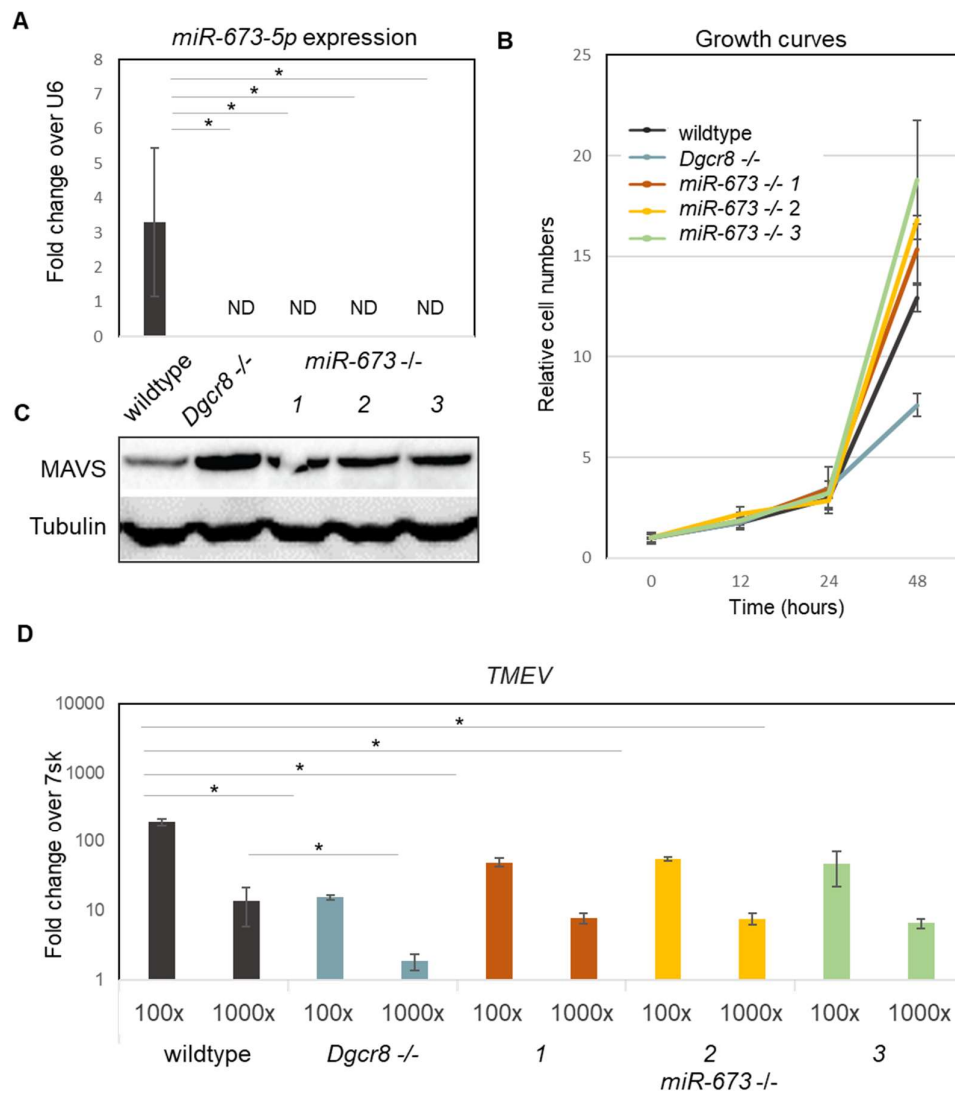


Figure 5.3 *mmu-miR-673* knockout is associated with higher Mavs levels and decreased viral susceptibility. **A.** RT-qPCR of *mmu-miR-673* in wildtype and mutant v6.5 cells, normalized to U6 snRNA. * $p < 0.05$ as calculated by ANOVA followed by Tukey's HSD test. Data are represented as mean $\pm$ SD with $n=4$ for each group. ND = not detected. **B.** Growth curves in wildtype and mutant v6.5 cells, normalized to cell numbers 24 hours after seeding. Cells were kept in T25 flasks up to the point of harvesting and were counted using the CASY system. **C.** Western blot against MAVS in wildtype and mutant v6.5 cells. Tubulin serves as a loading control. **D** Cellular quantities of TMEV relative to expression of endogenous 7sk as measured by RT-qPCR 24 hours after infection of v6.5 wildtype cells and mutants. Serial dilutions of TMEV were used, corresponding to an estimated MOI of 0.15 and 0.015 in wildtype cells. * $p < 0.05$ as calculated by ANOVA followed by Tukey's HSD test. Data are represented as mean $\pm$ SD with $n=3$ for each group.

Dgcr8 ^{-/-}. These results suggest a role for *miR-673* in downregulating *Mavs*, although not being the sole miRNA responsible.

Next, we studied the effect of *miR-673* absence on viral susceptibility, by infecting the different cell lines with TMEV and studying susceptibility by TCID₅₀ (Witteveldt, Knol and Macias, 2018) or RT-qPCR (figure 5.3D). mESCs without *miR-673* showed a decrease in viral material by RT-qPCR compared to wildtype cells at both concentrations tested, and the change was significant for the higher viral concentration. As showed before, *Dgcr8* ^{-/-} cells were more resistant to viral infection compared to wildtype. The viral levels detected by RT-qPCR were lower than *miR-673*, although not significantly, suggesting that *miR-673* is at least partly responsible for the TMEV susceptibility in mESCs. We hypothesised that this effect was due to activation of the type-I IFN response in *miR-673* ^{-/-} cells. To test this further, cells were again infected with TMEV, now in presence or absence of the JAK1/2 inhibitor Ruxolitinib, which prevents expression of ISGs. Ruxolitinib treatment significantly increased TMEV susceptibility of *miR-673* ^{-/-} cells, indicating the observed changes in viral susceptibility are type-I IFN-dependent (Witteveldt, Knol and Macias, 2018). Taken together, these results suggest *mmu-miR-673* is a key player in the (down)regulation of the dsRNA-mediated type-I IFN response in mESCs.

5.2.2 Absence of *DGCR8* increases the type-I IFN response in PA-1 cells

Since human ESCs are challenging to culture due to spontaneous differentiation and genomic instability (Henry *et al.*, 2019), embryonic carcinoma PA-1 cells were used instead as a model to study innate immunity during pluripotency in human cells. While these cells are not derived from a human blastocyst, embryonic markers and pluripotency factors such as SOX2, OCT4 and LIN28 are expressed at both RNA and protein level (Peng, Maihle and Huang, 2010; Yin *et al.*, 2014; Moriwaki *et al.*, 2017). Furthermore, PA-1 cells are able to differentiate and form embryoid bodies when

incubated with retinoic acid (Smith et al., 2001; Jesper Zeuthen, 1981) and can form complex multicellular structures upon plating in a non-adhesive dish upon high passage (Zeuthen *et al.*, 1980).

To study the role of DGCR8 in activation of the IFN response, wildtype, *DGCR8* knockout and *DGCR8* heterozygous PA-1 cells were transfected with poly I:C in a pilot experiment (figure 5.4). Both *MDA5* and *IFNB1* levels in absence of IFN stimulation were low, while *MAVS* and *IFIT1* were constitutively expressed.

Since PA-1 cells do not represent a pre-implantation blastocyst state, wildtype cells were able to express some *IFNb* after stimulation, although their response was attenuated in comparison with other cell types, such as the epithelial lung adenocarcinoma cell line, A549 (data not shown). Furthermore, the ISGs *IFIT1* and *MDA5* were upregulated, indicating that poly I:C infection resulted in successful first and second wave IFN signaling (figure 5.4).

However, *DGCR8* knockout cells showed much higher expression of *IFNB1*, *MDA5* and *IFIT1*, suggesting that miRNAs also downregulate the type-I IFN response in the human pluripotent context. Both *DGCR8* heterozygous cell lines tested showed an upregulation of *IFNB1* and ISGs after poly I:C transfection, but the expression levels did not change markedly from wildtype cells, indicating one copy of *DGCR8* was enough to revert to the wildtype phenotype. Transcript levels of *MAVS*, not an ISG itself, did not show great fluctuations in expression between the cell lines with or without poly I:C transfection, indicating that the higher IFN signaling in *DGCR8* knockout cells is not *MAVS*-dependent, as we showed in mESCs. Preliminary results from our group indicate that steady state levels of *MDA5* and *IFNb* show higher expression in *DGCR8* knockout cells in the absence of infection, suggesting there is some basal stimulation of the IFN response. Therefore, we hypothesize that the mechanism repressing the type-I IFN response in human and mouse pluripotent cells could be different.

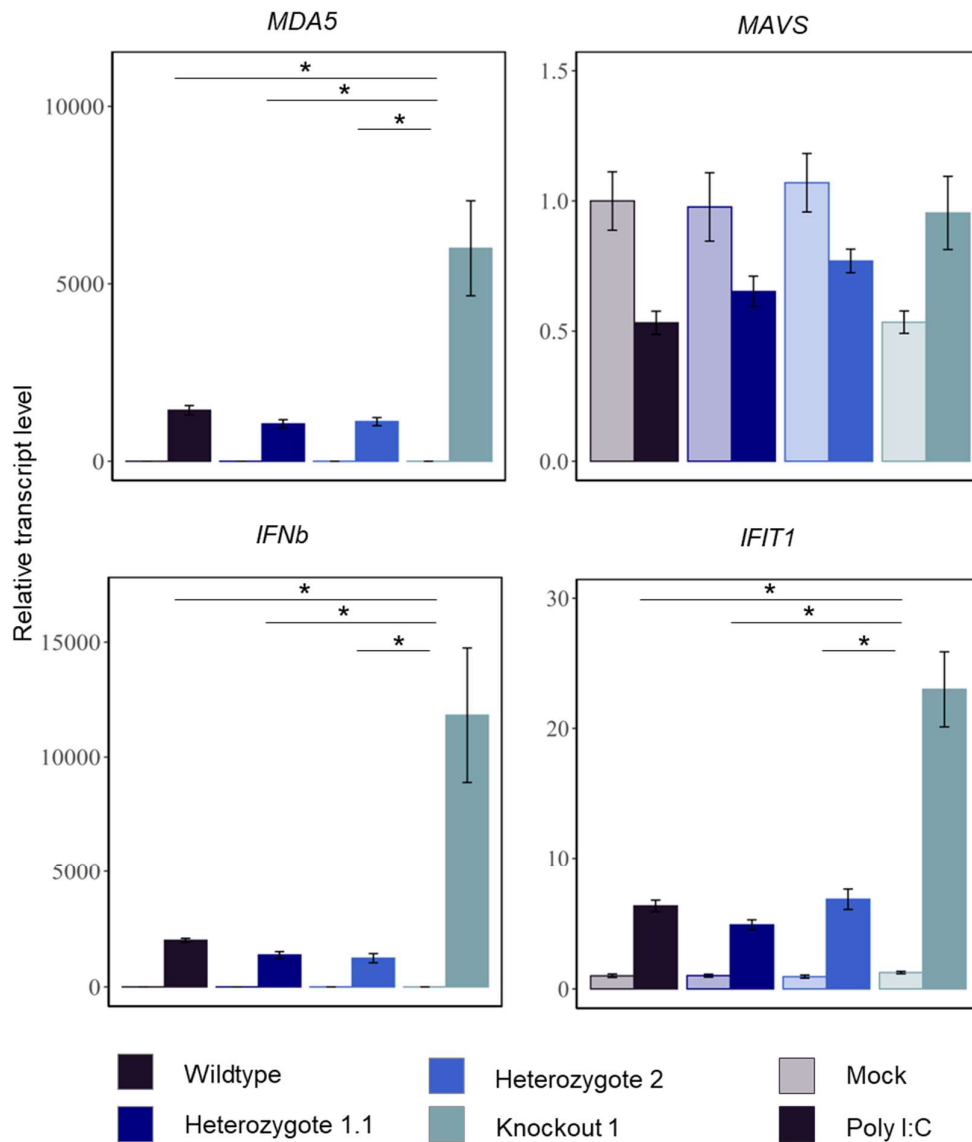


Figure 5.4. DGCR8 knockout cells express higher levels of IFN β and IGS after poly I:C stimulation. Expression of transcripts involved in the type-I IFN response in PA-1 cell lines after either mock (light color, uneven column number) or poly I:C transfection (dark color, even column number). Expression was measured by RT-qPCR, normalized to GAPDH and ACTB and represented as fold change over mock transfected wildtype cells. Data are represented as mean $\pm$ SE with $n=6$ for each group. Statistical analysis was carried out using a two-way ANOVA, with interaction effect for the cell line and the presence of poly I:C. * indicates $p<0.05$ between cell lines by Tukey's HSD test.

5.2.3 Depletion of *DGCR8* leads to changes in steady state levels of pluripotency factors

Overexpression analysis of RNA-seq data indicated that PA-1 *DGCR8* heterozygote and knockout cells showed a change in expression of transcripts involved in regulating pluripotency in stem cells compared to wildtype cells (figure 5.5A). Additionally, the detected transcriptional changes were similar to the transcript signature observed when the transcription factors *POU5F1* (*OCT4*) or *SOX2* are knocked out or depleted (chapter 3). Interestingly, *SOX2*, *POU5F1* and *NANOG* are the central factors in regulating pluripotency.

While *NANOG* and *POU5F1* were respectively differentially down- and upregulated upon *DGCR8* knockout, *SOX2* expression did not change in our RNA-seq data. RT-qPCR analysis showed altered expression of all three genes between wildtype and *DGCR8* knockout cells (figure 5.5B). In *DGCR8* heterozygote cells none of the genes were differentially expressed, and RT-qPCR data only showed moderate changes. This suggests that in *DGCR8* heterozygosity, the dysregulation of pathways regulating pluripotency occurs downstream of the central transcription factors.

Of the transcripts in the KEGG pathway '*regulating pluripotency of stem cells*,' *WNT* and *HOX* transcripts, important in embryonic patterning, were especially affected (figure 5.5A). Further analysis of all *WNT* and *HOX* genes detected in the RNA-seq assay showed that most *WNT* genes were either differentially up- or downregulated in both mutant cell lines, and that there was a general upregulation of *HOX* genes, which are normally not expressed in stem cells (figure 5A & B). Additionally, *DGCR8* heterozygote cell lines showed increased levels of the neural crest marker *SOX10*.

No upregulation of *SOX10* was detected in *DGCR8* knockout cells, however they showed a more severe change in *WNT* and *HOX* transcripts compared to heterozygote cells (figure 5.5B). Overall, these results indicate that both *DGCR8* heterozygote and knockout appear to be in an aberrant state of pluripotency compared to wildtype cells, but that the underlying mechanisms might differ.

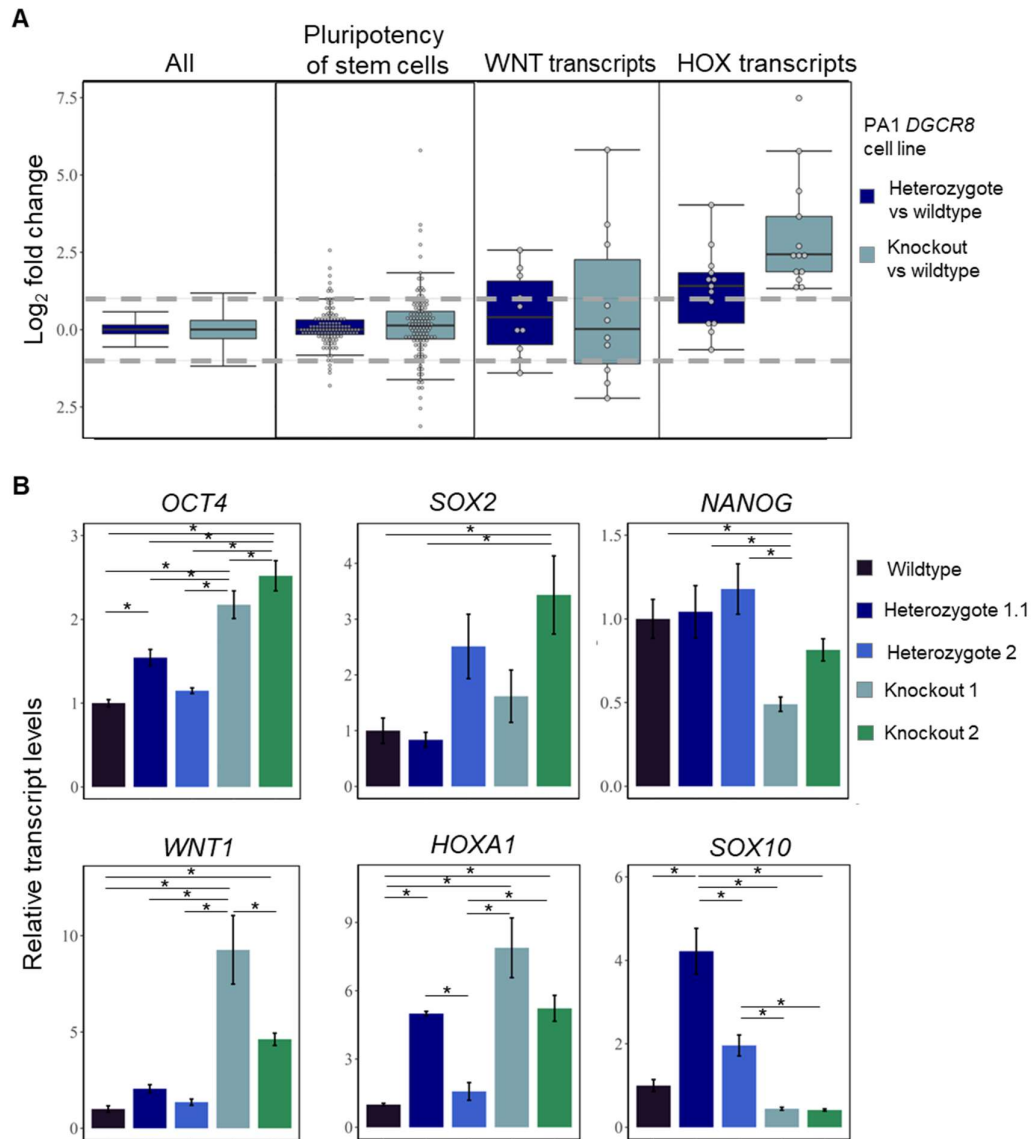


Figure 5.5 Differential gene expression of pluripotency and differentiation factors upon DGCR8 depletion. **A.** Boxplot analysis of total RNA transcriptomics transcript levels, including all measured transcripts; all transcripts in KEGG pathway hsa04550 ('Pluripotency of stem cells'); all WNT transcripts; all HOX transcripts. Only transcripts with a basemean >10 are shown. Dashed horizontal line shows a 2-fold up- or downregulation. **B.** RT-qPCR analysis of pluripotency and differentiation transcript steady-state levels. Data represented are the mean $\pm$ SE of 6 biological replicates relative to wildtype. Statistical analysis was carried out using a one-way ANOVA, followed by Tukey's HSD test. * $p < 0.05$.

The observed changes in the expression of pluripotency and differentiation markers in *DGCR8* knockout and heterozygote cells could be a direct effect of dysregulated miRNA expression, could be caused by secondary effects, or by non-canonical functions of *DGCR8*.

To study potential miRNA-mRNA interactions in *DGCR8* heterozygote cells, mirTarRnaSeq analysis was carried out. This analysis uses the changes in both small and long RNAs between two conditions and utilizes Miranda data to predict whether the observed changes in long RNAs are caused by altered miRNA expression. In our datasets there are several potential miRNA-mRNA interactions that could contribute to gene expression changes observed in *DGCR8* heterozygote cells related to pluripotency (figure 5.6). For instance, *miR-10a-5p*, which is upregulated in *DGCR8* heterozygote cells, is predicted to target several of the downregulated mRNAs involved in regulating pluripotency. On the other hand, the downregulated *miR-105/767* cluster is predicted to regulate transcripts involved in MAPK signalling and ectoderm development (Kanehisa *et al.*, 2009). Furthermore, mirTarRnaSeq analysis predicted that *miR-767-5p* and *miR-767-3p* target several other transcripts involved in differentiation, including multiple *WNT* and *HOX* family members and *SOX10*, emphasizing their potential importance in regulating pluripotency.

5.2.4 *DGCR8* heterozygote and knockout cells show aberrant differentiation

To further investigate the functional impact of *DGCR8* knockout and heterozygosity on embryonic development, PA-1 cell lines were subjected to differentiation. To this end, cells were seeded on poly-l-ornithine and laminin coated plates in medium containing 1 μ M retinoic acid (figure 5.7A). While wildtype and *DGCR8* heterozygous cells both showed high numbers of cell death and severe phenotypical changes, *DGCR8* knockout cells were much less affected. Wildtype and heterozygote cells initially attached well to the coated plate, but from day 2 on the proliferation rate

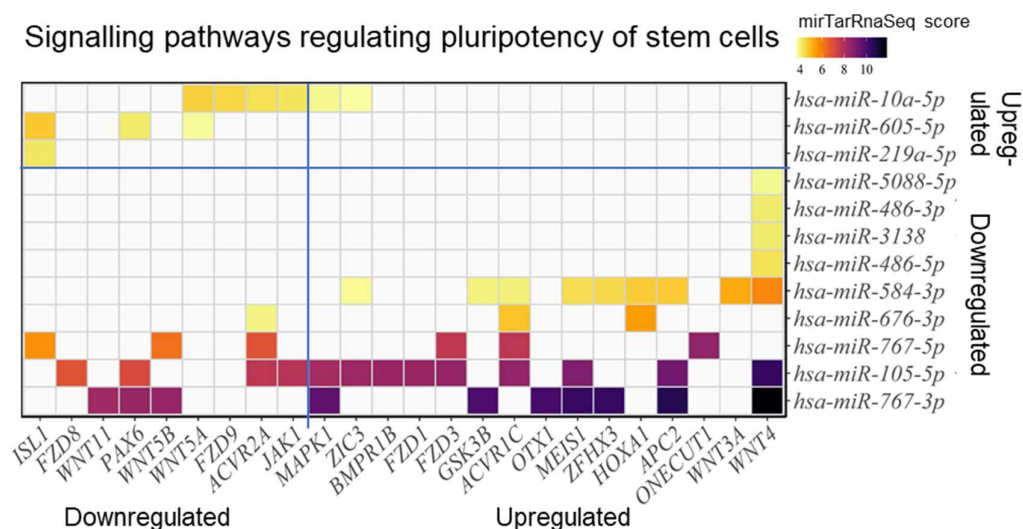


Figure 5.6 Heatmap showing potential miRNA-mRNA interactions. *mirTarRnaSeq* analysis to detect interactions between statistically up- or downregulated mRNAs and miRNAs in PA-1 DGCR8 heterozygote cells compared to wildtype. Only transcripts appearing in KEGG pathway hsa04550 and miRNAs with a possible interaction are shown. The transcripts are sorted by the changes in expression between DGCR8 heterozygote and wildtype cells.

slowed, which is expected when cells are differentiating. Subsequently, their appearance became more star-shaped, granules accumulated around the nucleus and cell death increased (figure 5.7B). While knockout cells attached slower than the other cell lines to the coated plates, overall cell death was low throughout the run of the experiment, and while cells became more elongated few other phenotypical changes were observed.

Changes in transcript levels regulating pluripotency were observed over time and between cell lines (figure 5.7C). Of the three main pluripotency factors, *NANOG*, *POU5F1* and *SOX2*, only *NANOG* showed a decrease in expression over time, resulting in a 10-fold decrease after 6 days, albeit non-significant. Both *SOX2* and *POU5F1* transcript levels were increasing during differentiation, 3.4- and 45-fold, respectively (not significant). *HOXA1* was suggested to steadily increase in expression during differentiation, up to over 400-fold in cells on day 8. These

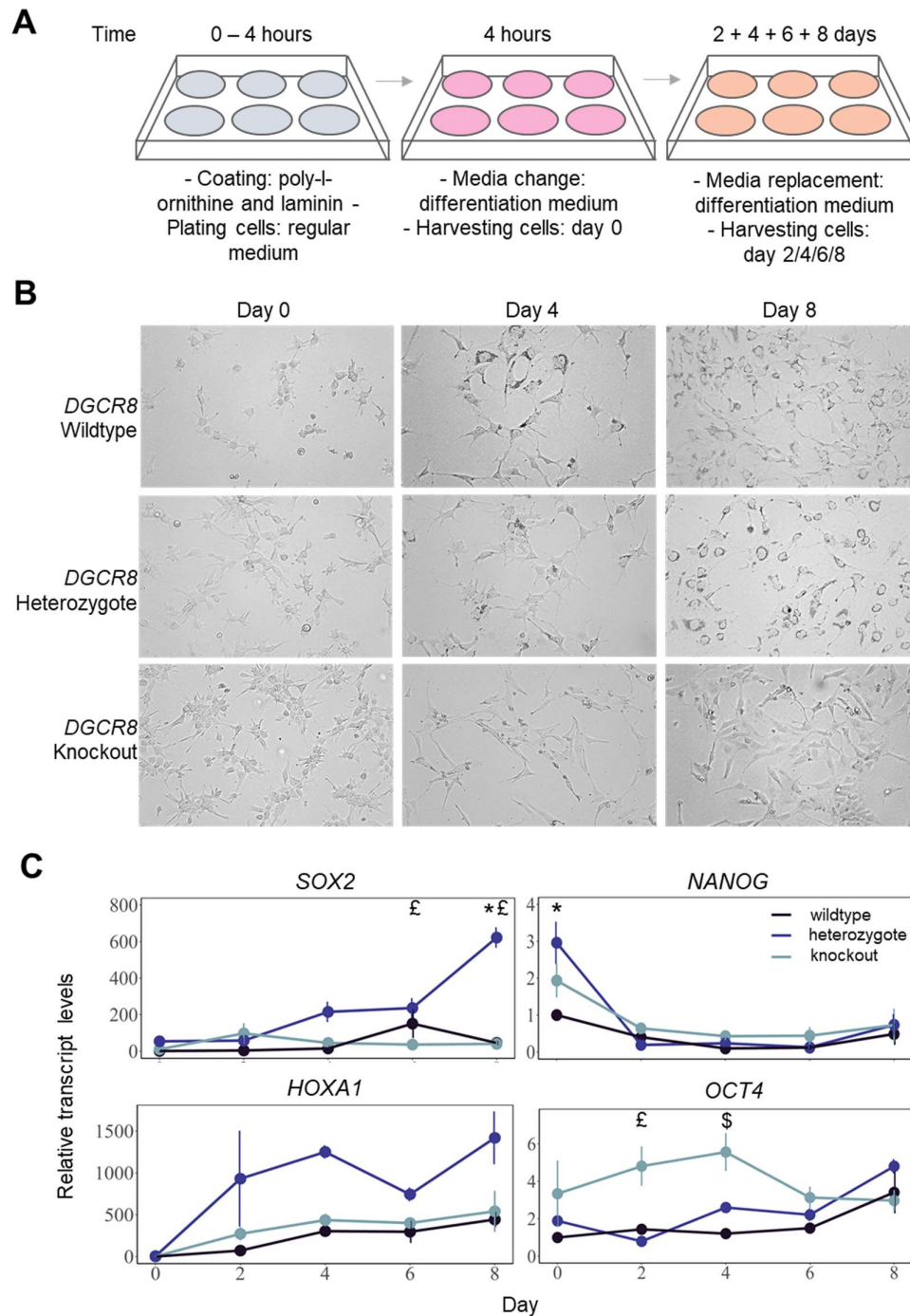


Figure 5.7 Differentiation of PA-1 cell lines. **A.** Differentiation protocol. Cells were plated in 6-well plates coated with poly-l-ornithine and laminin. Four hours after seeding, medium containing serum replacement and retinoic acid was added. This differentiation medium was replaced every two days and cells from different timepoints were harvested every time the medium was changed. **B.** Phenotype of PA-

*1 cell lines during differentiation by phase-contrast microscopy. C. Quantification of pluripotency and differentiation factors by RT-qPCR. Mean $\pm$ SE of 2 or 3 replicates relative to wildtype on day 0 is shown. Statistical analysis was carried out using a two-way ANOVA for an unbalanced design, assuming an interaction between cell line and day of differentiation, followed by Tukey's HSD test. * $p < 0.05$ between wildtype and heterozygote. \$ $p < 0.05$ between wildtype and knockout. £ $p < 0.05$ between heterozygote and knockout.*

results are surprising, since all three central pluripotency factors are expected to decrease over time during differentiation.

While in most cases *DGCR8* knockout and heterozygote cells showed similar trends in the transcripts measured over the course of the differentiation assay, several differences between cell lines could be detected. *DGCR8* wildtype and knockout cells showed similar expression patterns for *SOX2*, *NANOG* and *HOXA1* (figure 5.7C). However, consistent with our sequencing data, *POU5F1* was increased in *DGCR8* knockout cells, and showed elevated levels for the first 4 days. On the other hand, the *DGCR8* heterozygote cells showed very little difference in *POU5F1* expression but showed slightly higher expression of *NANOG* on day 0 compared to wildtype, and consistently higher levels and higher increase of *SOX2* and *HOXA1* over time. Overall, the phenotypical and transcriptional changes in the differentiation assay were more severe for *DGCR8* heterozygote cells compared to *DGCR8* knockout cells.

In addition to PA-1 differentiation using retinoic acid, wildtype and *DGCR8* heterozygote H9 human ESCs were differentiated towards specific cell types (Sara Heras lab; unpublished, data not shown). This more specialised differentiation into the three germ layers showed that *DGCR8* heterozygote cells are defective in mesodermal and endodermal differentiation, indicating that losing only one gene copy of *DGCR8* is already detrimental for development.

5.3 DISCUSSION

5.3.1 *DGCR8* and miRNAs are involved in downregulation of the RLR-pathway in the type-I IFN response

Previous work from our lab showed that activation of the type-I IFN response resulted in a decrease in miRNA biogenesis (Witteveldt, Ivens and Macias, 2018). The results presented in this chapter show that the inverse is also true: miRNAs are responsible for downregulating the IFN response during pluripotency. In mESCs, this occurs via downregulation of MAVS, which depletion of the miRNA *miR-673* restored. Furthermore, *miR-673* knockout cells were less susceptible for TMEV infection. Interestingly, *mir-673* has only been reported in mice and rats, with no human equivalent (Kozomara, Birgaoanu and Griffiths-Jones, 2019). Furthermore, *MAVS* expression in PA-1 cells was not found to be upregulated in PA-1 *DGCR8* knockout cells. This suggests that, while both human and mouse ESCs are unable to produce type-I IFNs, the mechanisms orchestrating this differ from each other.

While miRNA depletion led to the production of IFNs in both the mouse and human cell lines, stimulation of the cGAS/STING DNA-sensing pathway using G3-YSD did not (Witteveldt, Knol, & Macias, 2018; unpublished results). This suggest that factors other than miRNAs are repressing cGAS/STING signaling. Since PA-1 cells are derived from a carcinoma, it is possible that this pathway has been inactivated for reasons other than pluripotency.

Several reasons for the suppression of the IFN response has been suggested, including that expression of IFNs leads to an antiproliferative state that could interfere with embryonic development (Hertzog, Hwang and Kola, 1994; Bekisz *et al.*, 2010; Eggenberger *et al.*, 2019). Since in ESCs both retrotransposons and endogenous retroviruses are highly active (Yin, Zhou and Yuan, 2018), their RNA might be sensed in the cytoplasm and lead to unwanted IFN signalling. Therefore,

silencing the RNA sensing mechanism might protect natural embryonic development. This hypothesis is currently a focus in our lab.

The preliminary results from our lab using poly I:C transfections shown here and done by other lab members show consistently higher expression of *IFNB* in *DGCR8* knockout cells. However, the magnitude of the difference varies between experiments, even if all other conditions are kept the same to our knowledge. This led to the hypothesis that there is a great heterogeneity in response between individual cells. At the moment it is unclear whether the observed upregulation in IFNs and ISGs is caused by a moderate increase in expression of all cells, or by big changes in a small subset of cells. Therefore, single cell RNA-seq would be a useful addition to further study the mechanisms of *DGCR8* and miRNAs in immune activation.

5.3.2 *DGCR8* heterozygosity leads to an aberrant state of pluripotency

While *DGCR8* knockout models in mice are known to have limited differentiation capacity (Wang *et al.*, 2007; Z. Chen *et al.*, 2012), this has not yet been studied in the human context, and the effects of *DGCR8* heterozygosity have not yet been elucidated. Here we show differential expression of pluripotency and differentiation factors for both *DGCR8* knockout and heterozygote cell lines, indicating an altered state of pluripotency.

To further study the effect of human *DGCR8* on differentiation, PA-1 cells were differentiated with retinoic acid. Retinoic acid is well known for its capabilities of inducing differentiation, however the response is cell line, dose and duration-dependent (De Angelis *et al.*, 2018).

Upon culturing of wildtype PA-1 cells, we noticed a decrease in *NANOG* over time, while no change in *SOX2* and a slight increase in *POU5F1* was detected. There could be several explanations for the lack of pluripotency factor silencing. Firstly, they

might not be good markers to quantify differentiation in this context. While high levels of *POU5F1* are essential for maintaining pluripotency, its expression is also needed for germ-layer differentiation (Radziskeuskaya *et al.*, 2013). In mesenchymal stem cells, overexpression of *SOX2* and *POU5F1* even improved differentiation, indicating that levels of pluripotency factors and stemness are not directly correlated (Han *et al.*, 2014). On the other hand, it could also be that the protocol used is suboptimal for differentiation of PA-1 cells, or that PA-1 cells are not the most suitable for studying embryonic differentiation. Since our collaborators have generated H9 hESCs heterozygous for *DGCR8*, these cells will be the focus of further study.

During culturing in differentiation medium, *DGCR8* knockout cells followed the same trends in pluripotency transcript expression as wildtype cells, with the exception of *POU5F1*, which showed higher expression levels in early timepoints. The knockout cells underwent fewer morphological changes and suffered less cell death compared to both wildtype and heterozygous cells, however. A possible explanation could be that, since *DGCR8* knockouts are predicted to be lethal, they stall differentiation early in development. Since the timing of phenotypic changes between PA-1 wildtype and *DGCR8* knockout cells and changes in pluripotency factors do not match, it is unlikely that that aberrant differentiation in knockouts is driven by either *SOX2*, *POU5F1* or *NANOG*. Interestingly, small RNA-seq showed an upregulation of *let-7* genes in *DGCR8* knockout cells, which are known to be important for embryonic development. These miRNAs could be the focus of further study.

Although *DGCR8* heterozygote cells showed similar morphological changes as wildtype cells upon differentiation, RT-qPCRs suggested increased *SOX2* and *HOXA1* levels over the course of differentiation. Differences in *HOXA1* expression between heterozygote and wildtype cells preceded differences in *SOX2*. This suggests none of the central pluripotency factors *POU5F1*, *SOX2* or *NANOG* are causing these changes. Therefore, it is likely that the altered pluripotency state is regulated by

either downstream regulation of signalling transcripts or alternative signalling pathways. Our transcriptomics analysis showed altered expression of many transcripts of the *WNT* and *HOX* families in *DGCR8* heterozygous cells. These transcripts are known to drive germ layer differentiation and pattern formation (Kmita and Duboule, 2003; Neijts *et al.*, 2016; Kishimoto *et al.*, 2020). Such an early deviation of expression of these factors will have serious deleterious effects on embryogenesis and differentiation, since each germ layer will generate many different cell types, enabling and supporting essential functions in an organism.

Besides from the *HOX* and *WNT* family members, the transcription factor *SOX10* was shown to be increased in steady state levels in *DGCR8* heterozygotes. *SOX10* is a neural crest marker and is suggested to be regulated by canonical *WNT* pathway, which itself is known to play a role in neural crest formation (Britsch *et al.*, 2001; Aoki *et al.*, 2003; Wu, Saint-Jeannet and Klein, 2003; Kelsh, 2006; Yang *et al.*, 2020). This is of great clinical significance, since many of the organs affected in 22q11.2 syndrome stem from neural crest cells (Schinke and Izumo, 1999). These defects have previously been attributed to haploinsufficiency of *TBX1* or other genes affected in 22q11.2 syndrome (Motahari *et al.*, 2019). The data here suggests that aberrant *WNT*, *HOX* and *SOX10* signalling as a result of *DGCR8* heterozygosity could contribute to disrupted neural crest formation, and that the contribution of *DGCR8* to symptoms found in 22q11.2 DS patients may be greater than previously estimated. Therefore, *DGCR8* should also be considered as an important factor in future research into the developmental delays associated with 22q11.2 DS.

CHAPTER 6: DISCUSSION

6.1 COMBINING RNA-SEQ, SMALL RNA-SEQ AND MPI DETERMINATION IN THE SAME MODEL

DGCR8 has been extensively studied as an essential protein in canonical biogenesis of miRNAs. However, due to the lethal phenotype associated with a lack of DGCR8, little research of the effect of DGCR8 depletion in human cells has been carried out, with only two papers reporting expression of a truncated version of DGCR8 (Nguyen *et al.*, 2018; Deng *et al.*, 2019). Here we report both *DGCR8* knockout and *DGCR8* heterozygote PA-1 cells, allowing us to study the effect of complete DGCR8 depletion, ablating canonical miRNA biogenesis, or heterozygosity, as observed in 22q11.2 DS. Since the mutant cell lines have been derived from the same parental cell line, the effect of different gene doses of *DGCR8* can be directly compared in this setting. Using the different mutant cell lines, we carried out both total transcriptomics and small RNA transcriptomics. By combining these methods, the relationship between expression of specific miRNAs and their potential targets can be directly inferred, by manual curation or by using predictive databases such as miRDB or MiRTarBase, or predictive algorithms based on expression such as mirTarRnaSeq (Chou *et al.*, 2018; Chen and Wang, 2020). In addition, CLASH on PA-1 wildtype cells was carried out to establish miRNA targets under steady state conditions (work by Pablo Tristan, unpublished). The relevance of specific miRNAs of interest can be further studied by knocking miRNAs down in PA-1 wildtype cells or by reintroduction in DGCR8 knockout cells.

Our research showed that deletion of *DGCR8* leads to a drastically altered phenotype at the transcriptomic level. Furthermore, DROSHA levels were depleted due to the lack of stabilizing DGCR8, and no functional microprocessor could be formed. As expected, this led to an overall decrease of miRNAs in *DGCR8* knockout cells. There were however some notable exceptions: several miRNAs that are annotated as canonical could still be detected or were even increased in expression in knockout cells compared to wildtype cells. This might indicate alternative miRNA

biogenesis pathways, which should be further studied. Furthermore, the changes in miRNA expression could be used to curate whether miRNAs are synthesized by the canonical pathway.

Most previously published research looking into microprocessor function took a biochemical approach and studied microprocessor cleaving *in vitro* (Auyeung *et al.*, 2013; Fang and Bartel, 2015; Kwon *et al.*, 2020; Li *et al.*, 2020; Rice *et al.*, 2020).

While these processing experiments contribute greatly to our knowledge about the RNA requirements necessary for effective processing, they don't reflect *in vivo* cleaving conditions. In this thesis we used the microprocessor index to determine the effects of DGCR8 knockout and heterozygosity on pri-miRNA processing *in vivo*. As expected, pri-miRNA processing was decreased in *DGCR8* knockout cells, while heterozygote cell did not show significant differences. Unexpectedly, we found that only the UG sequencing motif impacted microprocessor cleaving, although this could be due to our limited dataset. Further knowledge of features assisting in efficient processing allows design of de novo pri-miRNAs with specific targets or can help increase or decrease miRNA biogenesis for clinical purposes.

6.2 THE ROLE OF DGCR8 AND SMALL RNAs IN ANTIVIRAL IMMUNITY AND DIFFERENTIATION

Previous research showed that mammalian ESCs are not able to elicit a type-I IFN response. Here we show that miRNAs are responsible for inhibition of this response. In mESCs, this is mediated through downregulation of MAVS, in hESCs this mechanism is not yet known. Given the increasing applications of induced pluripotent stem cells in regenerative medicine, it is imperative that the immune potential of these cells is further elucidated.

While in plants, insects and nematodes antiviral RNAi is a major antiviral mechanism, the presence and relevance of siRNAs in mammals remains

controversial. It has been suggested that the antiviral RNAi response is active in ESCs, but is overruled by the interferon response in somatic cells (Maillard *et al.*, 2016). With the research presented here we show that miRNAs are another type of small RNAs mediating antiviral immunity in mammals, and that previous research focusing on siRNA by looking at DICER may have been confounded by the effect of miRNAs.

One of the hypotheses explaining the absence of type-I IFN signaling is the presence of endogenous PRR substrates during pluripotency. Indeed, ESCs were found to have high levels of retrotransposition (Yin, Zhou and Yuan, 2018). In somatic cells, retrotransposons were found to be cleaved by the microprocessor (Macias *et al.*, 2012; Heras *et al.*, 2013). Now, recent unpublished research in our lab also observes an increase in LINE-I and Alu elements upon *DGCR8* depletion. Whether these transposable elements might contribute to a more primed state of innate immunity is currently being investigated.

Interestingly, our transcriptomics analysis and differentiation studies indicated an altered state of pluripotency and differentiation capacity in *DGCR8* heterozygous cells. This is especially relevant for 22q11.2DS, as so far most of the symptoms observed in patients have been contributed to hemizyosity of *TBX1*. The results presented here indicate that hemizyosity of *DGCR8* could also contribute.

6.3 FINAL REMARKS

This thesis emphasizes the importance of *DGCR8* in both immunity and pluripotency. Apart from research focusing on individual miRNAs, the non-canonical roles of *DGCR8* should be further investigated. Previous studies have shown that the C-terminus of *DGCR8* is essential for DRISHA interaction and the microprocessor function, while the N-terminus of *DGCR8* is essential for mesenchymal stem cell proliferation and heterochromatin stability (Nguyen *et al.*,

2018; Deng *et al.*, 2019). Currently, ATAC-seq analysis of the different PA-1 cell lines is being carried out to determine if gene accessibility plays a role in the observed changes in RNA levels between wildtype and DGCR8 mutant cells. Additionally, the DGCR8 knockout cells enable us to reintroduce truncated DGCR8 proteins to distinguish between miRNA-specific and non-canonical DGCR8 functions, which could be a powerful tool.

Further research into the mechanisms through which DGCR8 modulates innate immunity and differentiation could contribute to clinical interventions in 22q11.2 DS and dysregulated immune activation.

VIII. SUPPLEMENTARY TABLE

Overview of oligonucleotides used for this thesis

Name	Sequence (5' to 3')	Function; product size
Repair template <i>DGCR8</i>	TTAGAGAAGGATCCTTTGGAGAGAAGAGAAG CTCCGTAGAAGTTGAAGGGTCCTCAGCAGG GAGTTCGGACTGTCCATCACCACCAGAGCCA ACGTCCATTACCTCTGCACCACTGGAC	Repair template for PA-1 <i>DGCR8</i> knockout 1 allele 1
<i>DGCR8</i> exon 2 fw	AGGAGAAGCGGTGATGGAG	CRISPR gDNA screen PCR; 650 bp
<i>DGCR8</i> exon 2 rv	CATCCACTCTGTCTCTCTGAAC	
<i>Mmu-miR-673</i> fw	AGCAGTGATGGGTGTGCTAC	CRISPR gDNA screen PCR; 632 bp
<i>Mmu-miR-673</i> rv	TCCATTTCCCATCCCCTTGC	
<i>DGCR8</i> exon 10 fw	ATGTACCTGGGGCCCTCTTA	CRISPR gDNA screen PCR; 437 bp
<i>DGCR8</i> exon 10 rv	AGAATATGAGACGCCAGAGAGG	

Human <i>ACT1Nb</i> fw	ACCGAGCGCGGCTACAG	RT-qPCR; 60 bp
Human <i>ACT1Nb</i> rv	CTTAATGTCACGCACGATTTCC	
Human <i>DGCR8</i> exon 2-3 fw	GGGTTCACGGCTAAAGCAATC	RT-qPCR; 87 bp
Human <i>DGCR8</i> exon 2-3 rv	ATCGTTGTCAAAGTCATCCTCGT	
Human <i>DGCR8</i> exon 10 fw	CTCACCCACCCTCTGTGTTT	RT-qPCR; 128 bp
Human <i>DGCR8</i> exon 10 rv	CTTGATGGTTTGCCTCCAAT	
Human <i>GAPDH</i> fw	GGCCTCCAAGGAGTAAGACC	RT-qPCR; 147 bp
Human <i>GAPDH</i> rv	AGGGGTCTACATGGCAACTG	
Human <i>HOXA1</i> fw	CAGGGTGGGCAAACCTCTCTA	RT-qPCR; 149 bp
Human <i>HOXA1</i> rv	GCCGGGTTTCTGAAGTTTGT	

Human <i>IFIT1</i> fw	ACCATTGGCTGCTGTTTAGC	RT-qPCR; 100 bp
Human <i>IFIT1</i> rv	CAGGCTCCTCTGAGATCTGG	
Human <i>IFNb</i> fw	CCTGAAGGCCAAGGAGTACA	RT-qPCR; 145 bp
Human <i>IFNb</i> rv	AGCAATTGTCCAGTCCCAGA	
Human <i>MAVS</i> fw	AGAAGGAGCCAAGTTACCCC	RT-qPCR; 199 bp
Human <i>MAVS</i> rv	TCTGTGTCCTGCTCCTGATG	
Human <i>IFIH1</i> (<i>MDA5</i>) fw	GGCACCATGGGAAGTGATT	RT-qPCR; 86 bp
Human <i>IFIH1</i> (<i>MDA5</i>) rv	ATTTGGTAAGGCCTGAGCTG	
Human <i>NANOG</i> fw	TGCAGTTCCAGCCAAATTCTC	RT-qPCR; 92 bp
Human <i>NANOG</i> rv	CCTAGTGGTCTGCTGTATTACATTAAGG	

Human <i>PAX6</i> fw	CAGACACAGCCCTCACAAAC	RT-qPCR; 113 bp
Human <i>PAX6</i> rv	GGAGTATGAGGAGGTCTGGC	
Human <i>pre-GAPDH</i> fw	CATGGGTGTGAACCATGAGA	RT-qPCR; 110 bp
Human <i>pre-GAPDH</i> rv	CAGGGGAGCGTGTCCATAG	
Human <i>POU5F1</i> fw	AGTGAGAGGCAACCTGGAGA	RT-qPCR; 107 bp
Human <i>POU5F1</i> rv	ACACTCGGACCACATCCTTC	
Human <i>SOX2</i> fw	GAGCTTTGCAGGAAGTTTGC	RT-qPCR; 190 bp
Human <i>SOX2</i> rv	GCAAGAAGCCTCTCCTTGAA	
Human <i>SOX10</i> fw	CTGAAGGCAGGAAGGAGTTG	RT-qPCR; 189 bp
Human <i>SOX10</i> rv	GAGGGAGGCTCTGTGAATTG	
Human <i>WNT1</i> fw	CGGCGTTTATCTTCGCTATC	RT-qPCR; 77 nt

Human <i>WNT1</i> rv	TTCGATGGAACCTTCTGAGC	
Mouse <i>7sk</i> fw	GACATCTGTCACCCCATTGA	RT-qPCR; 281 bp
Mouse <i>7sk</i> rv	GCCTCATTGGATGTGTCTG	
Mouse <i>Gapdh</i> fw	TGTGTCCGTCGTGGATCTGA	RT-qPCR; 77 bp
Mouse <i>Gapdh</i> rv	CCTGCTTCACCACCTTCTTGA	
Mouse <i>Ifnb</i> fw	AAGAGTTACACTGCCTTTGCCATC	RT-qPCR; 135 bp
Mouse <i>Ifnb</i> rv	CACTGTCTGCTGGTGGAGTTCATC	
<i>TMEV</i> fw	TGTGGACTTGGACGATGACG	RT-qPCR; 95 bp
<i>TMEV</i> rv	CAGTATCGCATACGAGCGGT	

IX. BIBLIOGRAPHY

Agarwal, V. *et al.* (2015) 'Predicting effective microRNA target sites in mammalian mRNAs', *eLife*, 4(AUGUST2015). doi: 10.7554/eLife.05005.

Akamatsu, M. A. *et al.* (2021) 'Off balance: Interferons in COVID-19 lung infections', *EBioMedicine*, 73, p. 103642. doi: 10.1016/j.ebiom.2021.103642.

Alavi Darazam, I. *et al.* (2021) 'Role of interferon therapy in severe COVID-19: the COVIFERON randomized controlled trial', *Scientific Reports*. Nature Publishing Group UK, 11(1), pp. 1–11. doi: 10.1038/s41598-021-86859-y.

Alberti, C. *et al.* (2018) 'Cell-Type specific sequencing of microRNAs from complex animal tissues', *Nature Methods*. Nature Publishing Group, 15(4), pp. 283–289. doi: 10.1038/nmeth.4610.

Alberti, C. and Cochella, L. (2017) 'A framework for understanding the roles of miRNAs in animal development', *Development (Cambridge)*, 144(14), pp. 2548–2559. doi: 10.1242/dev.146613.

Altuvia, Y. *et al.* (2005) 'Clustering and conservation patterns of human microRNAs', *Nucleic Acids Research*, 33(8), pp. 2697–2706. doi: 10.1093/nar/gki567.

Alvarez-Saavedra, E. and Horvitz, H. R. (2010) 'Many Families of *C. elegans* MicroRNAs Are Not Essential for Development or Viability', *Current Biology*. Elsevier Ltd, 20(4), pp. 367–373. doi: 10.1016/j.cub.2009.12.051.

Ameres, S. L. and Zamore, P. D. (2013) 'Diversifying microRNA sequence and function', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 14(8), pp. 475–488. doi: 10.1038/nrm3611.

Andrews, S. (2010) 'FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>'.

De Angelis, M. T. *et al.* (2018) 'Short-term retinoic acid treatment sustains pluripotency and suppresses differentiation of human induced pluripotent stem cells', *Cell Death and Disease*. Springer US, 9(1), pp. 1–13. doi: 10.1038/s41419-017-0028-1.

- Aoki, Y. *et al.* (2003) 'Sox10 regulates the development of neural crest-derived melanocytes in *Xenopus*', *Developmental Biology*, 259(1), pp. 19–33. doi: 10.1016/S0012-1606(03)00161-1.
- Auyeung, V. C. *et al.* (2013) 'Beyond secondary structure: Primary-sequence determinants license Pri-miRNA hairpins for processing', *Cell*. Elsevier Inc., 152(4), pp. 844–858. doi: 10.1016/j.cell.2013.01.031.
- Azlan, A., Dzaki, N. and Azzam, G. (2016) 'Argonaute: The executor of small RNA function', *Journal of Genetics and Genomics*. Elsevier Limited and Science Press, 43(8), pp. 481–494. doi: 10.1016/j.jgg.2016.06.002.
- Babiarz, J. E. *et al.* (2008) 'Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs', *Genes & Development*, 22(20), pp. 2773–2785. doi: 10.1101/gad.1705308.
- Backes, S. *et al.* (2012) 'Degradation of host microRNAs by poxvirus poly(a) polymerase reveals terminal RNA methylation as a protective antiviral mechanism.', *Cell Host and Microbe*, 12(2), pp. 200–210. doi: 10.1016/j.chom.2012.05.019.
- Backes, S. *et al.* (2014) 'The Mammalian Response to Virus Infection Is Independent of Small RNA Silencing', *Cell Reports*. The Authors, 8(1), pp. 114–125. doi: 10.1016/j.celrep.2014.05.038.
- Barca-Mayo, O. and De Pietri Tonelli, D. (2014) 'Convergent microRNA actions coordinate neocortical development', *Cellular and Molecular Life Sciences*, 71(16), pp. 2975–2995. doi: 10.1007/s00018-014-1576-5.
- Barlow, D. P., Randle, B. J. and Burke, D. C. (1984) 'Interferon synthesis in the early post-implantation mouse embryo', *Differentiation*. International Society of Differentiation, 27(1–3), pp. 229–235. doi: 10.1111/j.1432-0436.1984.tb01433.x.
- Barr, I. *et al.* (2012) 'Ferric, not ferrous, heme activates RNA-binding protein DGCR8 for primary microRNA processing', *Proceedings of the National Academy of Sciences of the United States of America*, 109(6), pp. 1919–1924. doi: 10.1073/pnas.1114514109.
- Bartel, D. P. (2009) 'MicroRNAs: Target Recognition and Regulatory Functions', *Cell*, 136(2),

pp. 215–233. doi: 10.1016/j.cell.2009.01.002.

Bartram, M. P. *et al.* (2015) 'Loss of Dgcr8-mediated microRNA expression in the kidney results in hydronephrosis and renal malformation hydronephrosis and renal malformation', *BMC Nephrology*. ???, 16(1), pp. 4–9. doi: 10.1186/s12882-015-0053-1.

Bekisz, J. *et al.* (2010) 'Antiproliferative properties of type I and type II interferon', *Pharmaceuticals*, 3(4), pp. 994–1015. doi: 10.3390/ph3040994.

Bennasser, Y., Yeung, M. L. and Jeang, K. T. (2006) 'HIV-1 TAR RNA subverts RNA interference in transfected cells through sequestration of TAR RNA-binding protein, TRBP', *Journal of Biological Chemistry*. © 2006 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 281(38), pp. 27674–27678. doi: 10.1074/jbc.C600072200.

Berezikov, E. *et al.* (2007) 'Mammalian Mirtron Genes', *Molecular Cell*, 28(2), pp. 328–336. doi: 10.1016/j.molcel.2007.09.028.

Bofill-De Ros, X. *et al.* (2019) 'Structural Differences between Pri-miRNA Paralogs Promote Alternative Drosha Cleavage and Expand Target Repertoires', *Cell Reports*. ElsevierCompany., 26(2), pp. 447–459.e4. doi: 10.1016/j.celrep.2018.12.054.

Bogerd, H. P. *et al.* (2014) 'Replication of Many Human Viruses Is Refractory to Inhibition by Endogenous Cellular MicroRNAs', *Journal of Virology*, 88(14), pp. 8065–8076. doi: 10.1128/jvi.00985-14.

Bohnsack, M. T., Czaplinski, K. and Görlich, D. (2004) 'Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs', *RNA*. doi: 10.1261/rna.5167604.

Bose, M. *et al.* (2020) 'Retrograde trafficking of Argonaute 2 acts as a rate-limiting step for de novo miRNP formation on endoplasmic reticulum–attached polysomes in mammalian cells', *Life Science Alliance*, 3(2), pp. 1–16. doi: 10.26508/lsa.201800161.

Bowie, A. G. and Unterholzner, L. (2008) 'Viral evasion and subversion of pattern-recognition receptor signalling.', *Nature Reviews Immunology*, 8(12), pp. 911–22. doi: 10.1038/nri2436.

Braun, D., Caramalho, I. and Demengeot, J. (2002) 'IFN- α/β enhances BCR-dependent B cell responses', *International Immunology*, 14(4), pp. 411–419. doi: 10.1093/intimm/14.4.411.

Britsch, S. *et al.* (2001) 'The transcription factor Sox10 is a key regulator of peripheral glial development', *Genes and Development*, 15(1), pp. 66–78. doi: 10.1101/gad.186601.

Bronkhorst, A. W. and Van Rij, R. P. (2014) 'The long and short of antiviral defense: Small RNA-based immunity in insects', *Current Opinion in Virology*. Elsevier B.V., 7(1), pp. 19–28. doi: 10.1016/j.coviro.2014.03.010.

Broughton, J. P. *et al.* (2016) 'Pairing beyond the Seed Supports MicroRNA Targeting Specificity', *Molecular Cell*. Elsevier Inc., 64(2), pp. 320–333. doi: 10.1016/j.molcel.2016.09.004.

Bruns, A. M. and Horvath, C. M. (2015) 'LGP2 synergy with MDA5 in RLR-mediated RNA recognition and antiviral signaling', *Cytokine*. Elsevier Ltd, 74(2), pp. 198–206. doi: 10.1016/j.cyto.2015.02.010.

Buenrostro, J. D. *et al.* (2013) 'Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position', *Nature Methods*, 10(12), pp. 1213–1218. doi: 10.1038/nmeth.2688.

Burke, D. C., Graham, C. F. and Lehman, J. M. (1978) 'Appearance of interferon inducibility and sensitivity during differentiation of murine teratocarcinoma cells in vitro', *Cell*, 13(2), pp. 243–248. doi: 10.1016/0092-8674(78)90193-9.

Calmont, A. *et al.* (2009) 'Tbx1 controls cardiac neural crest cell migration during arch artery development by regulating Gbx2 expression in the pharyngeal ectoderm', *Development*, 136(18), pp. 3173–3183. doi: 10.1242/dev.028902.

Camacho, C. *et al.* (2009) 'BLAST+: Architecture and applications', *BMC Bioinformatics*, 10, pp. 1–9. doi: 10.1186/1471-2105-10-421.

Carmell, M. A. *et al.* (2004) 'Argonaute2 Is the Catalytic Engine of Mammalian RNAi', *Science*, 305(5689), pp. 1437–1441. doi: 10.1126/science.1102513.

Chaichian, Y., Wallace, D. J. and Weisman, M. H. (2019) 'A promising approach to targeting type 1 IFN in systemic lupus erythematosus', *Journal of Clinical Investigation*, 129(3), pp.

958–961. doi: 10.1172/JCI127101.

Chen, E. Y. *et al.* (2013) 'Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool', *BMC Bioinformatics*, 14(1), p. 128. doi: 10.1186/1471-2105-14-128.

Chen, L. *et al.* (2012) 'Naturally occurring endo-siRNA silences LINE-1 retrotransposons in human cells through DNA methylation', *Epigenetics*, 7(7), pp. 758–771. doi: 10.4161/epi.20706.

Chen, L. L., Yang, L. and Carmichael, G. G. (2010) 'Molecular basis for an attenuated cytoplasmic dsRNA response in human embryonic stem cells', *Cell Cycle*, 9(17), pp. 3552–3564. doi: 10.4161/cc.9.17.12792.

Chen, Y. *et al.* (2013) 'HCV-Induced miR-21 Contributes to Evasion of Host Immune System by Targeting MyD88 and IRAK1', *PLoS Pathogens*, 9(4). doi: 10.1371/journal.ppat.1003248.

Chen, Y. and Wang, X. (2020) 'MiRDB: An online database for prediction of functional microRNA targets', *Nucleic Acids Research*. Oxford University Press, 48(D1), pp. D127–D131. doi: 10.1093/nar/gkz757.

Chen, Z. *et al.* (2012) 'DiGeorge syndrome critical region 8 (DGCR8) protein-mediated microRNA biogenesis is essential for vascular smooth muscle cell development in mice', *Journal of Biological Chemistry*. © 2012 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 287(23), pp. 19018–19028. doi: 10.1074/jbc.M112.351791.

Chendrimada, T. P. *et al.* (2005) 'TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing', *Nature*, 436(7051), pp. 740–744. doi: 10.1038/nature03868.

Cheng, T. L. *et al.* (2014) 'MeCP2 Suppresses Nuclear MicroRNA Processing and Dendritic Growth by Regulating the DGCR8/Drosha Complex', *Developmental Cell*. Elsevier Inc., 28(5), pp. 547–560. doi: 10.1016/j.devcel.2014.01.032.

Chiang, K., Liu, H. and Rice, A. P. (2013) 'MiR-132 enhances HIV-1 replication', *Virology*. Elsevier, 438(1), pp. 1–4. doi: 10.1016/j.virol.2012.12.016.

Chiappinelli, K. B. *et al.* (2012) 'Reduced DICER1 elicits an interferon response in endometrial cancer cells', *Molecular Cancer Research*, 10(3), pp. 316–325. doi:

10.1158/1541-7786.MCR-11-0520.

Chou, C.-H. *et al.* (2018) 'miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions', *Nucleic Acids Research*, 46(D1), pp. D296–D302. doi: 10.1093/nar/gkx1067.

Chung, H. *et al.* (2018) 'Human ADAR1 Prevents Endogenous RNA from Triggering Translational Shutdown', *Cell*. Elsevier Inc., 172(4), pp. 811–824.e14. doi: 10.1016/j.cell.2017.12.038.

Chung, W.-J. *et al.* (2008) 'Endogenous RNA Interference Provides a Somatic Defense against Drosophila Transposons', *Current Biology*, 18(11), pp. 795–802. doi: 10.1016/j.cub.2008.05.006.

Cirera-Salinas, D. *et al.* (2017) 'Noncanonical function of DGCR8 controls mESC exit from pluripotency', *Journal of Cell Biology*, 216(2), pp. 355–366. doi: 10.1083/jcb.201606073.

Cong, L. *et al.* (2013) 'Multiplex genome engineering using CRISPR/Cas systems', *Science*, 339(6121), pp. 819–823. doi: 10.1126/science.1231143.

Conrad, T. *et al.* (2014) 'Microprocessor activity controls differential miRNA biogenesis in vivo', *Cell Reports*. The Authors, 9(2), pp. 542–554. doi: 10.1016/j.celrep.2014.09.007.

D'Angelo, W. *et al.* (2016) 'Development of Antiviral Innate Immunity During In Vitro Differentiation of Mouse Embryonic Stem Cells', *Stem Cells and Development*, 25(8), pp. 648–659. doi: 10.1089/scd.2015.0377.

D'Angelo, W. *et al.* (2017) 'The Molecular Basis for the Lack of Inflammatory Responses in Mouse Embryonic Stem Cells and Their Differentiated Cells', *The Journal of Immunology*, 198(5), pp. 2147–2155. doi: 10.4049/jimmunol.1601068.

Dalet, A., Gatti, E. and Pierre, P. (2015) 'Integration of PKR-dependent translation inhibition with innate immunity is required for a coordinated anti-viral response', *FEBS Letters*. Federation of European Biochemical Societies, 589(14), pp. 1539–1545. doi: 10.1016/j.febslet.2015.05.006.

Dang, T. L. *et al.* (2020) 'Select amino acids in DGCR8 are essential for the UGU-pri-miRNA interaction and processing', *Communications Biology*. Springer US, 3(1), pp. 1–11. doi:

10.1038/s42003-020-1071-5.

Dang, Y. *et al.* (2011) 'RNA interference in fungi: Pathways, functions, and applications', *Eukaryotic Cell*, 10(9), pp. 1148–1155. doi: 10.1128/EC.05109-11.

Daugaard, I. and Hansen, T. B. (2017) 'Biogenesis and Function of Ago-Associated RNAs', *Trends in Genetics*. Elsevier Ltd, pp. 208–219. doi: 10.1016/j.tig.2017.01.003.

Delorme-Axford, E., Sadovsky, Y. and Coyne, C. B. (2014) 'The placenta as a barrier to viral infections', *Annual Review of Virology*, 1(1), pp. 133–146. doi: 10.1146/annurev-virology-031413-085524.

Deng, L. *et al.* (2019) 'Stabilizing heterochromatin by DGCR8 alleviates senescence and osteoarthritis', *Nature Communications*. Springer US, 10(1), p. 3329. doi: 10.1038/s41467-019-10831-8.

Denli, A. M. *et al.* (2004) 'Processing of primary microRNAs by the microprocessor complex', *Nature*, 432(7014), pp. 231–235. Available at:
<http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=15531879&retmode=ref&cmd=prlinks%0Apapers3://publication/doi/10.1038/nature03049>.

Denzler, R. *et al.* (2014) 'Assessing the ceRNA Hypothesis with Quantitative Measurements of miRNA and Target Abundance', *Molecular Cell*, 54(5), pp. 766–776. doi: 10.1016/j.molcel.2014.03.045.

DeVeale, B., Swindlehurst-Chan, J. and Blelloch, R. (2021) 'The roles of microRNAs in mouse development', *Nature Reviews Genetics*. Springer US, 22(5), pp. 307–323. doi: 10.1038/s41576-020-00309-5.

Dexheimer, P. J. and Cochella, L. (2020) 'MicroRNAs: From Mechanism to Organism', *Frontiers in Cell and Developmental Biology*, 8(June), pp. 1–18. doi: 10.3389/fcell.2020.00409.

Dhir, A. *et al.* (2015) 'Microprocessor mediates transcriptional termination of long noncoding RNA transcripts hosting microRNAs', *Nature Structural and Molecular Biology*, 22(4), pp. 319–327. doi: 10.1038/nsmb.2982.

Diederichs, S. and Haber, D. A. (2007) 'Dual Role for Argonautes in MicroRNA Processing

- and Posttranscriptional Regulation of MicroRNA Expression', *Cell*, 131(6), pp. 1097–1108. doi: 10.1016/j.cell.2007.10.032.
- Ding, S. W. and Voinnet, O. (2007) 'Antiviral Immunity Directed by Small RNAs', *Cell*, 130(3), pp. 413–426. doi: 10.1016/j.cell.2007.07.039.
- Drmanac, R. *et al.* (2010) 'Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays', *Science*, 327(5961), pp. 78–81. doi: 10.1126/science.1181498.
- Du, Q., de la Morena, M. T. and van Oers, N. S. C. (2020) 'The Genetics and Epigenetics of 22q11.2 Deletion Syndrome', *Frontiers in Genetics*, 10(February), pp. 1–16. doi: 10.3389/fgene.2019.01365.
- Dueck, A. *et al.* (2012) 'MicroRNAs associated with the different human Argonaute proteins', *Nucleic Acids Research*, 40(19), pp. 9850–9862. doi: 10.1093/nar/gks705.
- Duguet, F. *et al.* (2021) 'S95021, a novel selective and pan-neutralizing anti interferon alpha (IFN- α) monoclonal antibody as a candidate treatment for selected autoimmune rheumatic diseases', *Journal of Translational Autoimmunity*, 4(November 2020). doi: 10.1016/j.jtauto.2021.100093.
- Earls, L. R. *et al.* (2012) 'Age-dependent microRNA control of synaptic plasticity in 22q11 deletion syndrome and schizophrenia', *Journal of Neuroscience*, 32(41), pp. 14132–14144. doi: 10.1523/JNEUROSCI.1312-12.2012.
- Eggenberger, J. *et al.* (2019) 'Type I interferon response impairs differentiation potential of pluripotent stem cells', *Proceedings of the National Academy of Sciences*, 116(4), pp. 1384–1393. doi: 10.1073/pnas.1812449116.
- Eichhorn, S. W. *et al.* (2014) 'MRNA Destabilization Is the dominant effect of mammalian microRNAs by the time substantial repression ensues', *Molecular Cell*, 56(1), pp. 104–115. doi: 10.1016/j.molcel.2014.08.028.
- Eisen, T. J. *et al.* (2020) 'MicroRNAs Cause Accelerated Decay of Short-Tailed Target mRNAs', *Molecular Cell*. Elsevier Inc., 77(4), pp. 775–785.e8. doi: 10.1016/j.molcel.2019.12.004.
- Ender, C. *et al.* (2008) 'A Human snoRNA with MicroRNA-Like Functions', *Molecular Cell*,

32(4), pp. 519–528. doi: 10.1016/j.molcel.2008.10.017.

Fang, W. and Bartel, D. P. (2015) 'The Menu of Features that Define Primary MicroRNAs and Enable De Novo Design of MicroRNA Genes', *Molecular Cell*. Elsevier Inc., 60(1), pp. 131–145. doi: 10.1016/j.molcel.2015.08.015.

Fang, W. and Bartel, D. P. (2020) 'MicroRNA Clustering Assists Processing of Suboptimal MicroRNA Hairpins through the Action of the ERH Protein', *Molecular Cell*. Elsevier Inc., 78(2), pp. 289–302.e6. doi: 10.1016/j.molcel.2020.01.026.

Farazi, T. A., Juranek, S. A. and Tuschl, T. (2008) 'The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members', *Development*, 135(7), pp. 1201–1214. doi: 10.1242/dev.005629.

Fénelon, K. *et al.* (2011) 'Deficiency of Dgcr8, a gene disrupted by the 22q11.2 microdeletion, results in altered short-term plasticity in the prefrontal cortex', *Proceedings of the National Academy of Sciences of the United States of America*, 108(11), pp. 4447–4452. doi: 10.1073/pnas.1101219108.

Feng, Y. *et al.* (2011) 'Drosha processing controls the specificity and efficiency of global microRNA expression', *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1809(11–12), pp. 700–707. doi: 10.1016/j.bbagr.2011.05.015.

Flemr, M. *et al.* (2013) 'A Retrotransposon-Driven Dicer Isoform Directs Endogenous Small Interfering RNA Production in Mouse Oocytes', *Cell*. Elsevier Inc., 155(4), pp. 807–816. doi: 10.1016/j.cell.2013.10.001.

Flores, O. *et al.* (2014) 'Differential RISC association of endogenous human microRNAs predicts their inhibitory potential', *Nucleic Acids Research*, 42(7), pp. 4629–4639. doi: 10.1093/nar/gkt1393.

Forster, S. C., Tate, M. D. and Hertzog, P. J. (2015) 'MicroRNA as type I interferon-regulated transcripts and modulators of the innate immune response', *Frontiers in Immunology*, 6(JUN), pp. 1–9. doi: 10.3389/fimmu.2015.00334.

Friedländer, M. R. *et al.* (2012) 'MiRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades', *Nucleic Acids Research*, 40(1), pp. 37–52.

doi: 10.1093/nar/gkr688.

Friedländer, M. R. *et al.* (2014) 'Evidence for the biogenesis of more than 1,000 novel human microRNAs', *Genome Biology*, 15(4), pp. 1–17. doi: 10.1186/gb-2014-15-4-r57.

Fromm, B. *et al.* (2015) 'A Uniform System for the Annotation of Vertebrate microRNA Genes and the Evolution of the Human microRNAome', *Annual Review of Genetics*, 49, pp. 213–242. doi: 10.1146/annurev-genet-120213-092023.

Fromm, B. *et al.* (2020) 'MirGeneDB 2.0: the metazoan microRNA complement', *Nucleic Acids Research*, 48(D1), pp. D132–D141. doi: 10.1093/nar/gkz885.

Fujita, P. A. *et al.* (2011) 'The UCSC genome browser database: Update 2011', *Nucleic Acids Research*, 39(SUPPL. 1), pp. 876–882. doi: 10.1093/nar/gkq963.

Fukunaga, R. *et al.* (2012) 'Dicer partner proteins tune the length of mature miRNAs in flies and mammals', *Cell*. Elsevier Inc., 151(3), pp. 533–546. doi: 10.1016/j.cell.2012.09.027.

Funabiki, M. *et al.* (2014) 'Autoimmune Disorders Associated with Gain of Function of the Intracellular Sensor MDAS', *Immunity*. Elsevier Inc., 40(2), pp. 199–212. doi: 10.1016/j.immuni.2013.12.014.

Gao, Z. *et al.* (2014) 'Phylogenetic analysis of the endoribonuclease Dicer family', *PLoS ONE*, 9(4), pp. 1–7. doi: 10.1371/journal.pone.0095350.

Giraldez, A. J. *et al.* (2005) 'MicroRNAs regulate brain morphogenesis in zebrafish', *Science*, 308(5723), pp. 833–838. doi: 10.1126/science.1109020.

Girard, A. *et al.* (2006) 'A germline-specific class of small RNAs binds mammalian Piwi proteins', *Nature*, 442(7099), pp. 199–202. doi: 10.1038/nature04917.

Gorman, J. A. *et al.* (2017) 'The A946T variant of the RNA sensor IFIH1 mediates an interferon program that limits viral infection but increases the risk for autoimmunity', *Nature Immunology*, 18(7), pp. 744–752. doi: 10.1038/ni.3766.

Goubau, D., Deddouche, S. and Reis e Sousa, C. (2013) 'Cytosolic Sensing of Viruses', *Immunity*. Elsevier Inc., 38(5), pp. 855–869. doi: 10.1016/j.immuni.2013.05.007.

Gregory, R. I. *et al.* (2004) 'The Microprocessor complex mediates the genesis of

- microRNAs', *Nature*, 432(7014), pp. 235–240. doi: 10.1038/nature03120.
- Greve, T. S., Judson, R. L. and Blelloch, R. (2013) 'MicroRNA control of mouse and human pluripotent stem cell behavior', *Annual Review of Cell and Developmental Biology*, 29, pp. 213–239. doi: 10.1146/annurev-cellbio-101512-122343.
- Gromak, N. *et al.* (2013) 'Drosha regulates gene expression independently of RNA cleavage function', *Cell Reports*, 5(6), pp. 1499–1510. doi: 10.1016/j.celrep.2013.11.032.
- Grow, E. J. *et al.* (2015) 'Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells', *Nature*, 522(7555), pp. 221–246. doi: 10.1038/nature14308.
- Gruber, A. J. *et al.* (2014) 'Embryonic stem cell-specific microRNAs contribute to pluripotency by inhibiting regulators of multiple differentiation pathways', *Nucleic Acids Research*, 42(14), pp. 9313–9326. doi: 10.1093/nar/gku544.
- Haase, A. D. *et al.* (2005) 'TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing', *EMBO Reports*, 6(10), pp. 961–967. doi: 10.1038/sj.embor.7400509.
- Hadjadj, J. *et al.* (2020) 'Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients', *Science*, 369(6504), pp. 718–724. doi: 10.1126/science.abc6027.
- Hall, J. C. and Rosen, A. (2010) 'Type I interferons: Crucial participants in disease amplification in autoimmunity', *Nature Reviews Rheumatology*. Nature Publishing Group, 6(1), pp. 40–49. doi: 10.1038/nrrheum.2009.237.
- Hamilton, A. J. and Baulcombe, D. C. (1999) 'A species of small antisense RNA in posttranscriptional gene silencing in plants', *Science*, 286(5441), pp. 950–952. doi: 10.1126/science.286.5441.950.
- Hammond, S. M. (2015) 'An overview of microRNAs', *Advanced Drug Delivery Reviews*. Elsevier B.V., 87, pp. 3–14. doi: 10.1016/j.addr.2015.05.001.
- Han, J. *et al.* (2004) 'The Drosha-DGCR8 complex in primary microRNA processing', *Genes and Development*, 18(24), pp. 3016–3027. doi: 10.1101/gad.1262504.
- Han, J. *et al.* (2006) 'Molecular Basis for the Recognition of Primary microRNAs by the Drosha-DGCR8 Complex', *Cell*, 125(5), pp. 887–901. doi: 10.1016/j.cell.2006.03.043.

- Han, J. *et al.* (2009) 'Posttranscriptional Crossregulation between Drosha and DGCR8', *Cell*. Elsevier Inc., 136(1), pp. 75–84. doi: 10.1016/j.cell.2008.10.053.
- Han, S. *et al.* (2018) 'A non-canonical role for the proneural gene Neurog1 as a negative regulator of neocortical neurogenesis', *Development*, 145(19). doi: 10.1242/dev.157719.
- Han, S. M. *et al.* (2014) 'Enhanced proliferation and differentiation of Oct4- And Sox2-overexpressing human adipose tissue mesenchymal stem cells', *Experimental and Molecular Medicine*. Nature Publishing Group, 46(6). doi: 10.1038/emm.2014.28.
- Han, Y. *et al.* (2012) 'Innate Immune Messenger 2-5A Tethers Human RNase L into Active High-Order Complexes', *Cell Reports*. The Authors, 2(4), pp. 902–913. doi: 10.1016/j.celrep.2012.09.004.
- Hansen, T. B. *et al.* (2016) 'Argonaute-associated short introns are a novel class of gene regulators', *Nature Communications*, 7(May), pp. 1–10. doi: 10.1038/ncomms11538.
- Helwak, A. *et al.* (2013) 'Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding', *Cell*. Elsevier Inc., 153(3), pp. 654–665. doi: 10.1016/j.cell.2013.03.043.
- Henry, M. P. *et al.* (2019) 'The genomic health of human pluripotent stem cells: Genomic instability and the consequences on nuclear organization', *Frontiers in Genetics*, 10(JAN), pp. 1–20. doi: 10.3389/fgene.2018.00623.
- Heo, I. *et al.* (2012) 'Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs', *Cell*. Elsevier Inc., 151(3), pp. 521–532. doi: 10.1016/j.cell.2012.09.022.
- Heras, S. R. *et al.* (2013) 'The Microprocessor controls the activity of mammalian retrotransposons', *Nature Structural and Molecular Biology*. Nature Publishing Group, 20(10), pp. 1173–1183. doi: 10.1038/nsmb.2658.
- Hertzog, P. J., Hwang, S. Y. and Kola, I. (1994) 'Role of interferons in the regulation of cell proliferation, differentiation, and development', *Molecular Reproduction and Development*, 39(2), pp. 226–232. doi: 10.1002/mrd.1080390216.
- Hoffmann, N. *et al.* (2018) 'DGCR8 promotes neural progenitor expansion and represses

neurogenesis in the mouse embryonic neocortex', *Frontiers in Neuroscience*, 12(APR), pp. 1–16. doi: 10.3389/fnins.2018.00281.

Hong, X. X. and Carmichael, G. G. (2013) 'Innate immunity in pluripotent human cells: Attenuated response to interferon- β ', *Journal of Biological Chemistry*. Â© 2013 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 288(22), pp. 16196–16205. doi: 10.1074/jbc.M112.435461.

Hsu, A. C. Y. *et al.* (2017) 'MicroRNA-125a and -b inhibit A20 and MAVS to promote inflammation and impair antiviral response in COPD', *JCI Insight*, 2(7). doi: 10.1172/jci.insight.90443.

Hsu, R. *et al.* (2012) 'Loss of microRNAs in pyramidal neurons leads to specific changes in inhibitory synaptic transmission in the prefrontal cortex', *Molecular and Cellular Neuroscience*, 50(3–4), pp. 283–292. doi: 10.1016/j.mcn.2012.06.002.

Huntzinger, E. and Izaurralde, E. (2011) 'Gene silencing by microRNAs: Contributions of translational repression and mRNA decay', *Nature Reviews Genetics*. Nature Publishing Group, 12(2), pp. 99–110. doi: 10.1038/nrg2936.

van den Hurk, J. A. J. M. *et al.* (2007) 'L1 retrotransposition can occur early in human embryonic development', *Human Molecular Genetics*, 16(13), pp. 1587–1592. doi: 10.1093/hmg/ddm108.

Hutvagner, G. *et al.* (2001) 'A Cellular Function for the RNA-Interference Enzyme Dicer in the Maturation of the let-7 Small Temporal RNA', *Science*, 293, pp. 834–839. doi: 10.1126/science.1062961.

Ivashkiv, L. B. and Donlin, L. T. (2014) 'Regulation of type I interferon responses', *Nature Reviews Immunology*. Nature Publishing Group, 14(1), pp. 36–49. doi: 10.1038/nri3581.

Janas, M. M. *et al.* (2012) 'Alternative RISC assembly: Binding and repression of microRNA-mRNA duplexes by human Ago proteins (RNA (2012) 18, (2041-2055))', *Rna*, 18(12), p. 2345. doi: 10.1261/rna.035675.112.and.

Jin, W. *et al.* (2020) 'Structural Basis for pri-miRNA Recognition by Drosha', *Molecular Cell*.

Elsevier Inc., 78(3), pp. 423–433.e5. doi: 10.1016/j.molcel.2020.02.024.

Jinek, M. *et al.* (2012) 'A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity', *Science*, 337(6096), pp. 816–821. doi: 10.1126/science.1225829.

Jonas, S. and Izaurralde, E. (2015) 'Towards a molecular understanding of microRNA-mediated gene silencing', *Nature Reviews Genetics*. Nature Publishing Group, 16(7), pp. 421–433. doi: 10.1038/nrg3965.

Jopling, C. L. *et al.* (2005) 'Molecular biology: Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA', *Science*, 309(5740), pp. 1577–1581. doi: 10.1126/science.1113329.

Kalil, A. C. *et al.* (2021) 'Efficacy of interferon beta-1a plus remdesivir compared with remdesivir alone in hospitalised adults with COVID-19: a double-blind, randomised, placebo-controlled, phase 3 trial', *The Lancet Respiratory Medicine*, 9(12), pp. 1365–1376. doi: 10.1016/S2213-2600(21)00384-2.

Kanehisa, M. *et al.* (2009) 'KEGG for representation and analysis of molecular networks involving diseases and drugs', *Nucleic Acids Research*, 38(SUPPL.1), pp. 355–360. doi: 10.1093/nar/gkp896.

Kanellopoulou, C. *et al.* (2005) 'Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing', *Genes and Development*, 19(4), pp. 489–501. doi: 10.1101/gad.1248505.

Kang, W. *et al.* (2021) 'MapToCleave: High-throughput profiling of microRNA biogenesis in living cells', *Cell Reports*, 37(7). doi: 10.1016/j.celrep.2021.110015.

Kano, H. *et al.* (2009) 'L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism', *Genes and Development*, 23(11), pp. 1303–1312. doi: 10.1101/gad.1803909.

Karginov, F. V. *et al.* (2010) 'Diverse Endonucleolytic Cleavage Sites in the Mammalian Transcriptome Depend upon MicroRNAs, Drosha, and Additional Nucleases', *Molecular Cell*. Elsevier Ltd, 38(6), pp. 781–788. doi: 10.1016/j.molcel.2010.06.001.

Kataruka, S. *et al.* (2020) 'MicroRNA dilution during oocyte growth disables the microRNA

- pathway in mammalian oocytes', *Nucleic Acids Research*, 48(14), pp. 8050–8062. doi: 10.1093/nar/gkaa543.
- Kato, H. *et al.* (2006) 'Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses.', *Nature*, 441(7089), pp. 101–105. doi: 10.1038/nature04734.
- Kawame, H. *et al.* (2001) 'Graves' disease in patients with 22q11.2 deletion', *Journal of Pediatrics*, 139(6), pp. 892–895. doi: 10.1067/mpd.2001.119448.
- Kelsh, R. N. (2006) 'Sorting out Sox10 functions in neural crest development', *BioEssays*, 28(8), pp. 788–798. doi: 10.1002/bies.20445.
- Ketting, R. F. *et al.* (2001) 'Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*.', *Genes and Development*, 15(516), pp. 2654–2659. doi: 10.1101/gad.927801.2654.
- Kim, B., Jeong, K. and Kim, V. N. (2017) 'Genome-wide Mapping of DROSHA Cleavage Sites on Primary MicroRNAs and Noncanonical Substrates', *Molecular Cell*. Elsevier Inc., 66(2), pp. 258–269.e5. doi: 10.1016/j.molcel.2017.03.013.
- Kim, D. *et al.* (2019) 'Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype', *Nature Biotechnology*. Springer US, 37(8), pp. 907–915. doi: 10.1038/s41587-019-0201-4.
- Kim, K. *et al.* (2018) 'SRSF3 recruits DROSHA to the basal junction of primary microRNAs', *RNA*, 24(7), pp. 892–898. doi: 10.1261/rna.065862.118.
- Kim, K. *et al.* (2021) 'A quantitative map of human primary microRNA processing sites', *Molecular Cell*. Elsevier Inc., 81(16), pp. 3422–3439.e11. doi: 10.1016/j.molcel.2021.07.002.
- Kim, Y.-K., Kim, B. and Kim, V. N. (2016) 'Re-evaluation of the roles of DROSHA , Exportin 5 , and DICER in microRNA biogenesis', *Proceedings of the National Academy of Sciences*, 113(13), pp. E1881–E1889. doi: 10.1073/pnas.1602532113.
- Kim, Y. K. and Kim, V. N. (2007) 'Processing of intronic microRNAs', *EMBO Journal*, 26(3), pp. 775–783. doi: 10.1038/sj.emboj.7601512.
- Kirou, K. A. and Gkrouzman, E. (2013) 'Anti-interferon alpha treatment in SLE', *Clinical Immunology*. Elsevier Inc., 148(3), pp. 303–312. doi: 10.1016/j.clim.2013.02.013.

- Kishimoto, K. *et al.* (2020) 'Bidirectional Wnt signaling between endoderm and mesoderm confers tracheal identity in mouse and human cells', *Nature Communications*. Springer US, 11(1), pp. 1–12. doi: 10.1038/s41467-020-17969-w.
- Kmita, M. and Duboule, D. (2003) 'Organizing axes in time and space; 25 years of colinear tinkering', *Science*, 301(5631), pp. 331–333. doi: 10.1126/science.1085753.
- Knuckles, P. *et al.* (2012) 'Drosha regulates neurogenesis by controlling Neurogenin 2 expression independent of microRNAs', *Nature Neuroscience*. Nature Publishing Group, 15(7), pp. 962–969. doi: 10.1038/nn.3139.
- Konno, D. *et al.* (2012) 'The Mammalian DM Domain Transcription Factor Dmrta2 Is Required for Early Embryonic Development of the Cerebral Cortex', *PLoS ONE*, 7(10), pp. 1–13. doi: 10.1371/journal.pone.0046577.
- Kornfeld, S. J. *et al.* (2000) 'DiGeorge anomaly: A comparative study of the clinical and immunologic characteristics of patients positive and negative by fluorescence in situ hybridization', *Journal of Allergy and Clinical Immunology*, 105(5), pp. 983–987. doi: 10.1067/mai.2000.105527.
- Kozomara, A., Birgaoanu, M. and Griffiths-Jones, S. (2019) 'MiRBase: From microRNA sequences to function', *Nucleic Acids Research*, 47(D1). doi: 10.1093/nar/gky1141.
- Kozomara, A. and Griffiths-Jones, S. (2014) 'MiRBase: Annotating high confidence microRNAs using deep sequencing data', *Nucleic Acids Research*, 42(D1), pp. 68–73. doi: 10.1093/nar/gkt1181.
- Kuleshov, M. V. *et al.* (2016) 'Enrichr: a comprehensive gene set enrichment analysis web server 2016 update', *Nucleic acids research*, 44(W1), pp. W90–W97. doi: 10.1093/nar/gkw377.
- Kwak, P. B. and Tomari, Y. (2012) 'The N domain of Argonaute drives duplex unwinding during RISC assembly', *Nature Structural and Molecular Biology*, 19(2), pp. 145–151. doi: 10.1038/nsmb.2232.
- Kwon, S. C. *et al.* (2016) 'Structure of Human DROSHA', *Cell*. Elsevier Inc., 164(1–2), pp. 81–90. doi: 10.1016/j.cell.2015.12.019.

- Kwon, S. C. *et al.* (2018) 'Molecular Basis for the Single-Nucleotide Precision of Primary microRNA Processing', *Molecular Cell*, (73). doi: 10.1016/j.molcel.2018.11.005.
- Kwon, S. C. *et al.* (2019) 'Molecular Basis for the Single-Nucleotide Precision of Primary microRNA Processing', *Molecular Cell*. Elsevier Inc., 73(3), pp. 505-518.e5. doi: 10.1016/j.molcel.2018.11.005.
- Kwon, S. C. *et al.* (2020) 'ERH facilitates microRNA maturation through the interaction with the N-terminus of DGCR8', *Nucleic Acids Research*. Oxford University Press, 48(19), pp. 11097–11112. doi: 10.1093/nar/gkaa827.
- De la Morena, M. T. *et al.* (2013) 'Signature MicroRNA expression patterns identified in humans with 22q11.2 deletion/DiGeorge syndrome', *Clinical Immunology*. Elsevier Inc., 147(1), pp. 11–22. doi: 10.1016/j.clim.2013.01.011.
- Lai, W. S. *et al.* (2019) 'Inhibiting transcription in cultured metazoan cells with actinomycin D to monitor mRNA turnover', *Methods*. Elsevier, 155(January), pp. 77–87. doi: 10.1016/j.ymeth.2019.01.003.
- Langmead, B. and Salzberg, S. L. (2012) 'Fast gapped-read alignment with Bowtie 2', *Nature Methods*, 9(4), pp. 357–359. doi: 10.1038/nmeth.1923.
- Lee, Y. *et al.* (2002) 'MicroRNA maturation: stepwise processing and subcellular localization', *The EMBO Journal*, 21(17), pp. 4663–4670. doi: 10.1093/EMBOJ/CDF476.
- Lee, Y. *et al.* (2003) 'The nuclear RNase III Drosha initiates microRNA processing', *Nature*, 425(6956), pp. 415–419. doi: 10.1038/nature01957.
- Lee, Y. *et al.* (2004) 'MicroRNA genes are transcribed by RNA polymerase II', *EMBO Journal*, 23(20), pp. 4051–4060. doi: 10.1038/sj.emboj.7600385.
- Lee, Y. S. *et al.* (2004) 'Distinct Roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA Silencing Pathways', *Cell*, 117(1), pp. 69–81. doi: 10.1016/S0092-8674(04)00261-2.
- Lewis, B. P. *et al.* (2003) 'Prediction of Mammalian MicroRNA Targets', 115, pp. 787–798.
- Li, Q. *et al.* (2018) 'Reliable Multiplex Sequencing with Rare Index Mis-Assignment on DNB-Based NGS Platform', *bioRxiv*, pp. 1–13. doi: 10.1101/343137.

- Li, S. *et al.* (2020) 'Mismatched and wobble base pairs govern primary microRNA processing by human Microprocessor', *Nature Communications*. Springer US, 11(1). doi: 10.1038/s41467-020-15674-2.
- Li, Y. *et al.* (2013) 'RNA interference functions as an antiviral immunity mechanism in mammals.', *Science*, 342(6155), pp. 231–4. doi: 10.1126/science.1241911.
- Li, Y. *et al.* (2016) 'Induction and suppression of antiviral RNA interference by influenza A virus in mammalian cells', *Nature Microbiology*, 2(December 2016), pp. 1–9. doi: 10.1038/nmicrobiol.2016.250.
- Li, Y. and Shi, X. (2013) 'MicroRNAs in the regulation of TLR and RIG-I pathways', *Cellular and Molecular Immunology*. Nature Publishing Group, 10(1), pp. 65–71. doi: 10.1038/cmi.2012.55.
- Li, Y., Zhang, Y. and Liu, M. (2021) 'Knockout Gene-Based Evidence for PIWI-Interacting RNA Pathway in Mammals', 9(July). doi: 10.3389/fcell.2021.681188.
- Liao, Y., Smyth, G. K. and Shi, W. (2014) 'FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features', *Bioinformatics*, 30(7), pp. 923–930. doi: 10.1093/bioinformatics/btt656.
- Lin, H. *et al.* (2015) 'Microprocessor Complex Subunit DiGeorge Syndrome Critical Region Gene 8 (Dgcr8) Is Required for Schwann Cell Myelination and Myelin Maintenance * □', *Journal of Biological Chemistry*. Â© 2015 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 290(40), pp. 24294–24307. doi: 10.1074/jbc.M115.636407.
- Lin, S. and Gregory, R. I. (2015) 'MicroRNA biogenesis pathways in cancer', *Nature Reviews Cancer*. Nature Publishing Group, 15(6), pp. 321–333. doi: 10.1038/nrc3932.
- Lindsay, E. A. *et al.* (2001) 'Tbx1 haploinsufficiency in the DiGeorge syndrome region causes aortic arch defects in mice', *Nature*, 410(6824), pp. 97–101. doi: 10.1038/35065105.
- Louloupi, A. *et al.* (2017) 'Microprocessor dynamics shows co- and post-transcriptional processing of pri-miRNAs', *Rna*, 23(6), pp. 892–898. doi: 10.1261/rna.060715.117.
- Love, M. I., Huber, W. and Anders, S. (2014) 'Moderated estimation of fold change and

- dispersion for RNA-seq data with DESeq2', *Genome Biology*, 15(12), pp. 1–21. doi: 10.1186/s13059-014-0550-8.
- Lu, J. *et al.* (2005) 'MicroRNA expression profiles classify human cancers', *Nature*, 435(7043), pp. 834–838. doi: 10.1038/nature03702.
- Luhur, A. *et al.* (2014) 'Drosha-independent DGCR8/Pasha pathway regulates neuronal morphogenesis', *Proceedings of the National Academy of Sciences of the United States of America*, 111(4), pp. 1421–1426. doi: 10.1073/pnas.1318445111.
- Luo, J. *et al.* (2018) 'MicroRNA duplication accelerates the recruitment of new targets during vertebrate evolution', *RNA*, 24(6), pp. 787–802. doi: 10.1261/rna.062752.117.
- Luteijn, M. J. and Ketting, R. F. (2013) 'PIWI-interacting RNAs: From generation to transgenerational epigenetics', *Nature Reviews Genetics*. Nature Publishing Group, 14(8), pp. 523–534. doi: 10.1038/nrg3495.
- Ma, E. *et al.* (2008) 'Autoinhibition of Human Dicer by Its Internal Helicase Domain', *Journal of Molecular Biology*, 380(1), pp. 237–243. doi: 10.1016/j.jmb.2008.05.005.
- Macias, S. *et al.* (2012) 'DGCR8 HITS-CLIP reveals novel functions for the Microprocessor', *Nature Structural and Molecular Biology*. Nature Publishing Group, 19(8), pp. 760–766. doi: 10.1038/nsmb.2344.
- Macias, S. *et al.* (2015) 'DGCR8 Acts as an Adaptor for the Exosome Complex to Degrade Double-Stranded Structured RNAs', *Molecular Cell*, 60(6), pp. 873–885. doi: 10.1016/j.molcel.2015.11.011.
- Macias, S., Cordiner, R. A. and Cáceres, J. F. (2013) 'Cellular functions of the microprocessor', *Biochemical Society Transactions*, 41(4), pp. 838–843. doi: 10.1042/BST20130011.
- MacRae, I. J. *et al.* (2006) 'Structural Basis for Double-Stranded RNA Processing by Dicer', *Science*, 311(5758), pp. 195–198. doi: 10.1126/science.1121638.
- MacRae, I. J. and Doudna, J. A. (2007) 'Ribonuclease revisited: structural insights into ribonuclease III family enzymes', *Current Opinion in Structural Biology*, pp. 138–145. doi: 10.1016/j.sbi.2006.12.002.

- MacRae, I. J., Zhou, K. and Doudna, J. A. (2007) 'Structural determinants of RNA recognition and cleavage by Dicer', *Nature Structural and Molecular Biology*, 14(10), pp. 934–940. doi: 10.1038/nsmb1293.
- Maillard, P. V. *et al.* (2013) 'Antiviral RNA interference in mammalian cells', *Science*, 342(6155), pp. 235–238. doi: 10.1126/science.1241930.
- Maillard, P. V *et al.* (2016) 'Inactivation of the type I interferon pathway reveals long double-stranded RNA-mediated RNA interference in mammalian cells', *The EMBO Journal*, pp. 1–14. doi: 10.15252/emboj.201695086.
- Malathi, K. *et al.* (2007) 'Small self-RNA generated by RNase L amplifies antiviral innate immunity', *Nature*, 448(7155), pp. 816–819. doi: 10.1038/nature06042.
- Malur, M., Gale, M. and Krug, R. M. (2012) 'LGP2 Downregulates Interferon Production during Infection with Seasonal Human Influenza A Viruses That Activate Interferon', *Journal of Virology*, 86(19), pp. 10733–10738. doi: 10.1128/JVI.00510-12.
- Marinaro, F. *et al.* (2017) 'MicroRNA-independent functions of DGCR8 are essential for neocortical development and TBR1 expression', *EMBO reports*, 18(4), pp. 603–618. doi: 10.15252/embr.201642800.
- Marrack, P., Kappler, J. and Mitchell, T. (1999) 'Type I interferons keep activated T cells alive', *Journal of Experimental Medicine*, 189(3), pp. 521–529. doi: 10.1084/jem.189.3.521.
- McCall, M. N. *et al.* (2017) 'Toward the human cellular microRNAome', *Genome Research*, 27(10), pp. 1769–1781. doi: 10.1101/gr.222067.117.
- McDonald-McGinn, D. M. *et al.* (2015) '22Q11.2 Deletion Syndrome', *Nature Reviews Disease Primers*, 1(November). doi: 10.1038/nrdp.2015.71.
- McNab, F. *et al.* (2015) 'Type I interferons in infectious disease', *Nature Reviews Immunology*. Nature Publishing Group, 15(2), pp. 87–103. doi: 10.1038/nri3787.
- Meister, G. *et al.* (2004) 'Human Argonaute2 Mediates RNA Cleavage Targeted by miRNAs and siRNAs', *Molecular Cell*, 15(2), pp. 185–197. doi: 10.1016/j.molcel.2004.07.007.
- Meister, G. (2013) 'Argonaute proteins: Functional insights and emerging roles', *Nature Reviews Genetics*. Nature Publishing Group, pp. 447–459. doi: 10.1038/nrg3462.

- Melton, C., Judson, R. L. and Blelloch, R. (2010) 'Opposing microRNA families regulate self-renewal in mouse embryonic stem cells', *Nature*. Nature Publishing Group, 463(7281), pp. 621–626. doi: 10.1038/nature08725.
- Mendell, J. T. and Olson, E. N. (2012) 'MicroRNAs in stress signaling and human disease', *Cell*. Elsevier Inc., 148(6), pp. 1172–1187. doi: 10.1016/j.cell.2012.02.005.
- Merscher, S. *et al.* (2001) 'TBX1 is responsible for cardiovascular defects in velo-cardio-facial/DiGeorge syndrome', *Cell*, 104(4), pp. 619–629. doi: 10.1016/S0092-8674(01)00247-1.
- Meyer, O. (2009) 'Interferons and autoimmune disorders', *Joint Bone Spine*, 76(5), pp. 464–473. doi: 10.1016/j.jbspin.2009.03.012.
- Michlewski, G. *et al.* (2008) 'Posttranscriptional Regulation of miRNAs Harboring Conserved Terminal Loops', *Molecular Cell*, 32(3), pp. 383–393. doi: 10.1016/j.molcel.2008.10.013.
- Michlewski, G. and Cáceres, J. F. (2010) 'Antagonistic role of hnRNP A1 and KSRP in the regulation of let-7a biogenesis', *Nature Structural and Molecular Biology*, 17(8), pp. 1011–1018. doi: 10.1038/nsmb.1874.
- Michlewski, G. and Cáceres, J. F. (2019) 'Post-transcriptional control of miRNA biogenesis', *RNA*, 25(1), pp. 1–16. doi: 10.1261/rna.068692.118.
- Moore, M. J. *et al.* (2015) 'MiRNA-target chimeras reveal miRNA 3'-end pairing as a major determinant of Argonaute target specificity', *Nature Communications*. Nature Publishing Group, 6(May). doi: 10.1038/ncomms9864.
- Moriwaki, M. *et al.* (2017) 'POLR2C mutations are associated with primary ovarian insufficiency in women', *Journal of the Endocrine Society*, 1(3), pp. 162–173. doi: 10.1210/js.2016-1014.
- Morlando, M. *et al.* (2008) 'Primary microRNA transcripts are processed co-transcriptionally', *Nature Structural and Molecular Biology*, 15(9), pp. 902–909. doi: 10.1038/nsmb.1475.
- Motahari, Z. *et al.* (2019) 'In the line-up: Deleted genes associated with DiGeorge/22q11.2 deletion syndrome: Are they all suspects?', *Journal of Neurodevelopmental Disorders*.

Journal of Neurodevelopmental Disorders, 11(1), pp. 1–28. doi: 10.1186/s11689-019-9267-z.

Neijts, R. *et al.* (2016) 'Polarized regulatory landscape and Wnt responsiveness underlie Hox activation in embryos', *Genes and Development*, 30(17), pp. 1937–1942. doi: 10.1101/gad.285767.116.

Nejad, C., Stunden, H. J. and Gantier, M. P. (2018) 'A guide to miRNAs in inflammation and innate immune responses', *FEBS Journal*, 285(20), pp. 3695–3716. doi: 10.1111/febs.14482.

Nguyen, T. A. *et al.* (2015) 'Functional anatomy of the human microprocessor', *Cell*. Elsevier Inc., 161(6), pp. 1374–1387. doi: 10.1016/j.cell.2015.05.010.

Nguyen, T. A. *et al.* (2018) 'Microprocessor depends on hemin to recognize the apical loop of primary microRNA', *Nucleic Acids Research*. Oxford University Press, 46(11), pp. 5726–5736. doi: 10.1093/nar/gky248.

Nguyen, T. L. *et al.* (2020) 'The internal loops in the lower stem of primary microRNA transcripts facilitate single cleavage of human Microprocessor', *Nucleic Acids Research*. Oxford University Press, pp. 1–15. doi: 10.1093/nar/gkaa018.

O'Connell, R. M. *et al.* (2007) 'MicroRNA-155 is induced during the macrophage inflammatory response', *Proceedings of the National Academy of Sciences of the United States of America*, 104(5), pp. 1604–1609. doi: 10.1073/pnas.0610731104.

O'Connor, C. M., Vanicek, J. and Murphy, E. A. (2014) 'Host MicroRNA Regulation of Human Cytomegalovirus Immediate Early Protein Translation Promotes Viral Latency', *Journal of Virology*, 88(10), pp. 5524–5532. doi: 10.1128/jvi.00481-14.

Obbard, D. J. *et al.* (2009) 'The evolution of RNAi as a defence against viruses and transposable elements', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1513), pp. 99–115. doi: 10.1098/rstb.2008.0168.

Oda, H. *et al.* (2014) 'Aicardi-goutières syndrome is caused by IFIH1 mutations', *American Journal of Human Genetics*. The American Society of Human Genetics, 95(1), pp. 121–125. doi: 10.1016/j.ajhg.2014.06.007.

Okada, C. *et al.* (2009) 'A high-Resolution structure of the pre-microRNA nuclear export

- machinery', *Science*, 326(5957), pp. 1275–1279. doi: 10.1126/science.1178705.
- Okamura, K. *et al.* (2007) 'The Mirtron Pathway Generates microRNA-Class Regulatory RNAs in *Drosophila*', *Cell*, 130(1), pp. 89–100. doi: 10.1016/j.cell.2007.06.028.
- Okamura, K. *et al.* (2008) 'The *Drosophila* hairpin RNA pathway generates endogenous short interfering RNAs', *Nature*, 453(7196), pp. 803–806. doi: 10.1038/nature07015.
- Olena, A. F. and Patton, J. G. (2010) 'Genomic organization of microRNAs', *Journal of Cellular Physiology*, 222(3), pp. 540–545. doi: 10.1002/jcp.21993.
- Ostermann, E. *et al.* (2012) 'Deregulation of type I IFN-dependent genes correlates with increased susceptibility to cytomegalovirus acute infection of dicer mutant mice', *PLoS ONE*, 7(8). doi: 10.1371/journal.pone.0043744.
- Ouchi, Y. *et al.* (2013) 'Reduced adult hippocampal neurogenesis and working memory deficits in the *Dgcr8*-deficient mouse model of 22q11.2 deletion-associated schizophrenia can be rescued by IGF2', *Journal of Neuroscience*, 33(22), pp. 9408–9419. doi: 10.1523/JNEUROSCI.2700-12.2013.
- Paganini, I. *et al.* (2017) 'A novel PAX1 null homozygous mutation in autosomal recessive otofaciocervical syndrome associated with severe combined immunodeficiency', *Clinical Genetics*, 92(6), pp. 664–668. doi: 10.1111/cge.13085.
- Parchem, R. J. *et al.* (2015) 'miR-302 Is Required for Timing of Neural Differentiation, Neural Tube Closure, and Embryonic Viability', *Cell Reports*. Elsevier, 12(5), pp. 760–773. doi: 10.1016/j.celrep.2015.06.074.
- Partin, A. C. *et al.* (2017) 'Heme enables proper positioning of Drosha and DGCR8 on primary microRNAs', *Nature Communications*. Springer US, 8(1), pp. 1–10. doi: 10.1038/s41467-017-01713-y.
- Partin, A. C. *et al.* (2020) 'Cryo-EM Structures of Human Drosha and DGCR8 in Complex with Primary MicroRNA', *Molecular Cell*. Elsevier Inc., pp. 1–12. doi: 10.1016/j.molcel.2020.02.016.
- Pasquinelli, A. E. *et al.* (2000) 'Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA', *Nature*, 408(6808), pp. 86–89. doi: 10.1038/35040556.

- Patel, V. D. and Capra, J. A. (2017) 'Ancient human miRNAs are more likely to have broad functions and disease associations than young miRNAs', *BMC Genomics*. BMC Genomics, 18(1), pp. 1–13. doi: 10.1186/s12864-017-4073-z.
- Peaston, A. E. *et al.* (2004) 'Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos', *Developmental Cell*, 7(4), pp. 597–606. doi: 10.1016/j.devcel.2004.09.004.
- Peisley, A. *et al.* (2012) 'Kinetic mechanism for viral dsRNA length discrimination by MDA5 filaments', *Proceedings of the National Academy of Sciences of the United States of America*, 109(49). doi: 10.1073/pnas.1208618109.
- Peng, S., Maihle, N. J. and Huang, Y. (2010) 'Pluripotency factors Lin28 and Oct4 identify a sub-population of stem cell-like cells in ovarian cancer', *Oncogene*. Nature Publishing Group, 29(14), pp. 2153–2159. doi: 10.1038/onc.2009.500.
- Pestka, S., Krause, C. D. and Walter, M. R. (2004) 'Interferons, interferon-like cytokines, and their receptors'.
- Piatek, M. J. and Werner, A. (2014) 'Endogenous siRNAs: regulators of internal affairs', *Biochemical Society Transactions*, 42(4), pp. 1174–1179. doi: 10.1042/BST20140068.
- Piliero, L. M. *et al.* (2004) 'T-cell homeostasis in humans with thymic hypoplasia due to chromosome 22q11.2 deletion syndrome', *Blood*, 103(3), pp. 1020–1025. doi: 10.1182/blood-2003-08-2824.
- Pohl, E. *et al.* (2013) 'A hypofunctional PAX1 mutation causes autosomal recessively inherited otofaciocervical syndrome', *Human Genetics*, 132(11), pp. 1311–1320. doi: 10.1007/s00439-013-1337-9.
- Poirier, E. Z. *et al.* (2021) 'An isoform of Dicer protects mammalian stem cells against multiple RNA viruses.', *Science (New York, N.Y.)*, 373(6551), pp. 231–236. doi: 10.1126/science.abg2264.
- Pons-Espinal, M. *et al.* (2017) 'Synergic Functions of miRNAs Determine Neuronal Fate of Adult Neural Stem Cells', *Stem Cell Reports*. ElsevierCompany., 8(4), pp. 1046–1061. doi: 10.1016/j.stemcr.2017.02.012.

- Quicke, K. M., Diamond, M. S. and Suthar, M. S. (2017) 'Negative regulators of the RIG-I-like receptor signaling', *European Journal of Immunology*, 47, pp. 615–628. doi: 10.1002/eji.201646484.
- Radziskeuskaya, A. *et al.* (2013) 'A defined Oct4 level governs cell state transitions of pluripotency entry and differentiation into all embryonic lineages', *Nature Cell Biology*. Nature Publishing Group, 15(6), pp. 579–590. doi: 10.1038/ncb2742.
- Ragab, D. *et al.* (2020) 'The COVID-19 Cytokine Storm; What We Know So Far', *Frontiers in Immunology*, 11(June), pp. 1–4. doi: 10.3389/fimmu.2020.01446.
- Ran, F. A. *et al.* (2013) 'Genome engineering using the CRISPR-Cas9 system', *Nature Protocols*, 8(11), pp. 2281–2308. doi: 10.1038/nprot.2013.143.
- Rao, P. K. *et al.* (2009) 'Loss of cardiac microRNA-mediated regulation leads to dilated cardiomyopathy and heart failure', *Circulation Research*, 105(6), pp. 585–594. doi: 10.1161/CIRCRESAHA.109.200451.
- Reichholf, B. *et al.* (2019) 'Time-Resolved Small RNA Sequencing Unravels the Molecular Principles of MicroRNA Homeostasis', *Molecular Cell*. Elsevier Inc., 75(4), pp. 756–768.e7. doi: 10.1016/j.molcel.2019.06.018.
- Reikine, S., Nguyen, J. B. and Modis, Y. (2014) 'Pattern recognition and signaling mechanisms of RIG-I and MDA5', *Frontiers in Immunology*, 5(JUL), pp. 1–7. doi: 10.3389/fimmu.2014.00342.
- Reinhart, B. J. *et al.* (2000) 'The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*', *Nature*, 403(6772), pp. 901–906. doi: 10.1038/35002607.
- Repetto, G. M. *et al.* (2014) 'Case fatality rate and associated factors in patients with 22q11 microdeletion syndrome: A retrospective cohort study', *BMJ Open*, 4(11), pp. 1–5. doi: 10.1136/bmjopen-2014-005041.
- Ressel, S. *et al.* (2019) 'Extracellular RNA in viral–host interactions: Thinking outside the cell', *Wiley Interdisciplinary Reviews: RNA*, 10(4), pp. 1–23. doi: 10.1002/wrna.1535.
- Rice, G. I. *et al.* (2014) 'Gain-of-function mutations in IFIH1 cause a spectrum of human disease phenotypes associated with upregulated type I interferon signaling', *Nature*

- Genetics*. Nature Publishing Group, 46(5), pp. 503–509. doi: 10.1038/ng.2933.
- Rice, G. M. *et al.* (2020) 'Functional Atlas of Primary miRNA Maturation by the Microprocessor', *Molecular Cell*. Elsevier Inc., 80(5), pp. 892–902.e4. doi: 10.1016/j.molcel.2020.10.028.
- De Rie, D. *et al.* (2017) 'An integrated expression atlas of miRNAs and their promoters in human and mouse', *Nature Biotechnology*, 35(9), pp. 872–878. doi: 10.1038/nbt.3947.
- Roden, C. *et al.* (2017) 'Novel determinants of mammalian primary microRNA processing revealed by systematic evaluation of hairpin-containing transcripts and human genetic variation', *Genome Research*, 27(3), pp. 374–384. doi: 10.1101/gr.208900.116.
- Rodero, M. P. and Crow, Y. J. (2016) 'Type I interferon-mediated monogenic autoinflammation: The type I interferonopathies, a conceptual overview', *The Journal of Experimental Medicine*, 213(12), pp. 2527–2538. doi: 10.1084/jem.20161596.
- Rozas, M. F. *et al.* (2019) 'Association between phenotype and deletion size in 22q11.2 microdeletion syndrome: systematic review and meta-analysis', *Orphanet Journal of Rare Diseases*. Orphanet Journal of Rare Diseases, 14(1), pp. 1–9. doi: 10.1186/s13023-019-1170-x.
- Sabin, L. R. *et al.* (2013) 'Dicer-2 processes diverse viral RNA species.', *PloS one*, 8(2), p. e55458. doi: 10.1371/journal.pone.0055458.
- Santhakumar, D. *et al.* (2010) 'Combined agonist-antagonist genome-wide functional screening identifies broadly active antiviral microRNAs', *Proceedings of the National Academy of Sciences of the United States of America*, 107(31), pp. 13830–13835. doi: 10.1073/pnas.1008861107.
- Satoh, T. *et al.* (2010) 'LGP2 is a positive regulator of RIG-I- and MDA5-mediated antiviral responses', *Proceedings of the National Academy of Sciences*, 107(4), pp. 1512–1517. doi: 10.1073/pnas.0912986107.
- Schinke, M. and Izumo, S. (1999) 'Getting to the heart of DiGeorge syndrome', *Nature Medicine*, 5(10), pp. 1120–1121. doi: 10.1038/13438.
- Schirle, N. T., Sheu-Gruttadauria, J. and MacRae, I. J. (2014) 'Structural basis for microRNA

targeting', *Science*, 346(6209), pp. 608–613. doi: 10.1126/science.1258040.

Schoggins, J. W. *et al.* (2011) 'A diverse range of gene products are effectors of the type I interferon antiviral response', *Nature*. Nature Publishing Group, 472(7344), pp. 481–485. doi: 10.1038/nature09907.

Schoggins, J. W. (2014) 'Interferon-stimulated genes: Roles in viral pathogenesis', *Current Opinion in Virology*. Elsevier B.V., 6(1), pp. 40–46. doi: 10.1016/j.coviro.2014.03.006.

Schuster, S. *et al.* (2019) 'No evidence for viral small RNA production and antiviral function of Argonaute 2 in human cells', *Scientific Reports*, 9(1), pp. 1–11. doi: 10.1038/s41598-019-50287-w.

Sedger, L. M. (2013) 'microRNA control of interferons and interferon induced anti-viral activity', *Molecular Immunology*, 56(4), pp. 781–793. doi: 10.1016/j.molimm.2013.07.009.

Senturia, R. *et al.* (2010) 'Structure of the dimerization domain of DiGeorge Critical Region 8', *Protein Science*, 19(7), pp. 1354–1365. doi: 10.1002/pro.414.

Shabalina, S. A. and Koonin, E. V. (2008) 'Origins and evolution of eukaryotic RNA interference', *Trends in Ecology and Evolution*, 23(10), pp. 578–587. doi: 10.1016/j.tree.2008.06.005.

Shang, R. *et al.* (2020) 'Genomic Clustering Facilitates Nuclear Processing of Suboptimal Pri-miRNA Loci', *Molecular Cell*. Elsevier Inc., 78(2), pp. 303–316.e4. doi: 10.1016/j.molcel.2020.02.009.

Shapiro, J. S. *et al.* (2012) 'Evidence for a cytoplasmic microprocessor of pri-miRNAs', *Rna*, 18(7), pp. 1338–1346. doi: 10.1261/rna.032268.112.

Shapiro, J. S. *et al.* (2014) 'Drosha as an interferon-independent antiviral factor.', *Proceedings of the National Academy of Sciences of the United States of America*, 111(19), pp. 7108–13. doi: 10.1073/pnas.1319635111.

Sheu-Gruttadauria, J. and MacRae, I. J. (2017) 'Structural Foundations of RNA Silencing by Argonaute', *Journal of Molecular Biology*. doi: 10.1016/j.jmb.2017.07.018.

Siomi, M. C. *et al.* (2011) 'PIWI-interacting small RNAs: The vanguard of genome defence', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 12(4), pp. 246–258. doi:

10.1038/nrm3089.

Sohn, S. Y. *et al.* (2007) 'Crystal structure of human DGCR8 core', *Nature Structural and Molecular Biology*, 14(9), pp. 847–853. doi: 10.1038/nsmb1294.

Sood, P. *et al.* (2006) 'Cell-type-specific signatures of microRNAs on target mRNA expression', *Proceedings of the National Academy of Sciences of the United States of America*, 103(8), pp. 2746–2751. doi: 10.1073/pnas.0511045103.

Sperber, H. *et al.* (2014) 'MiRNA sensitivity to Drosha levels correlates with pre-miRNA secondary structure', *Rna*, 20(5), pp. 621–631. doi: 10.1261/rna.043943.113.

Staple, L. *et al.* (2005) 'Allergies in patients with chromosome 22q11.2 deletion syndrome (DiGeorge syndrome/velocardiofacial syndrome) and patients with chronic granulomatous disease', *Pediatric Allergy and Immunology*, 16(3), pp. 226–230. doi: 10.1111/j.1399-3038.2005.00259.x.

Stark, K. L. *et al.* (2008) 'Altered brain microRNA biogenesis contributes to phenotypic deficits in a 22q11-deletion mouse model', *Nature Genetics*, 40(6), pp. 751–760. doi: 10.1038/ng.138.

Su, H. *et al.* (2009) 'Essential and overlapping functions for mammalian argonautes in microRNA silencing', *Genes and Development*, 23(3), pp. 304–317. doi: 10.1101/gad.1749809.

Suh, N. *et al.* (2010) 'MicroRNA Function Is Globally Suppressed in Mouse Oocytes and Early Embryos', *Current Biology*. Elsevier Ltd, 20(3), pp. 271–277. doi: 10.1016/j.cub.2009.12.044.

Sullivan, K. E. *et al.* (1997) 'Juvenile rheumatoid arthritis-like polyarthritis in chromosome 22q11.2 deletion syndrome (DiGeorge anomaly/velocardiofacial syndrome/conotruncal anomaly face syndrome)', *Arthritis and Rheumatism*, 40(3), pp. 430–436. doi: 10.1002/art.1780400307.

Sullivan, K. E. *et al.* (1998) 'Lack of correlation between impaired T cell production, immunodeficiency, and other phenotypic features in chromosome 22q11.2 deletion syndromes (DiGeorge syndrome/velocardiofacial syndrome)', *Clinical Immunology and Immunopathology*, 86(2), pp. 141–146. doi: 10.1006/clin.1997.4463.

- Svoboda, P. and Flemr, M. (2010) 'The role of miRNAs and endogenous siRNAs in maternal-to-zygotic reprogramming and the establishment of pluripotency', *EMBO Reports*, 11(8), pp. 590–597. doi: 10.1038/embor.2010.102.
- Swiecki, M. and Colonna, M. (2011) 'Type I interferons: Diversity of sources, production pathways and effects on immune responses', *Current Opinion in Virology*. Elsevier B.V., 1(6), pp. 463–475. doi: 10.1016/j.coviro.2011.10.026.
- Taganov, K. D. *et al.* (2006) 'NF- κ B-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses', *Proceedings of the National Academy of Sciences of the United States of America*, 103(33), pp. 12481–12486. doi: 10.1073/pnas.0605298103.
- Tam, O. H. *et al.* (2008) 'Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes', *Nature*, 453(7194), pp. 534–538. doi: 10.1038/nature06904.
- Tang, F. *et al.* (2007) 'Maternal microRNAs are essential for mouse zygotic development', *Genes and Development*, 21(6), pp. 644–648. doi: 10.1101/gad.418707.
- Thompson, M. R. *et al.* (2011) 'Pattern recognition receptors and the innate immune response to viral infection', *Viruses*, 3(6), pp. 920–940. doi: 10.3390/v3060920.
- Toyoshima, M. *et al.* (2016) 'Analysis of induced pluripotent stem cells carrying 22q11.2 deletion', *Translational Psychiatry*. Nature Publishing Group, 6(11), pp. e934–10. doi: 10.1038/tp.2016.206.
- Trainor, P. A. (2005) 'Specification and patterning of neural crest cells during craniofacial development', *Brain, Behavior and Evolution*, 66(4), pp. 266–280. doi: 10.1159/000088130.
- Vitelli, F. *et al.* (2002) 'Tbx1 mutation causes multiple cardiovascular defects and disrupts neural crest and cranial nerve migratory pathways', *Human Molecular Genetics*, 11(8), pp. 915–922. doi: 10.1093/hmg/11.8.915.
- Vlachos, I. S. *et al.* (2015) 'DIANA-miRPath v3.0: Deciphering microRNA function with experimental support', *Nucleic Acids Research*, 43(W1), pp. W460–W466. doi: 10.1093/nar/gkv403.
- Wagschal, A. *et al.* (2012) 'Microprocessor, Setx, Xrn2, and Rrp6 co-operate to induce

- premature termination of transcription by RNAPII', *Cell*, 150(6), pp. 1147–1157. doi: 10.1016/j.cell.2012.08.004.
- Wang, R. *et al.* (2013) 'Mouse embryonic stem cells are deficient in type I interferon expression in response to viral infections and double-stranded RNA', *Journal of Biological Chemistry*, 288(22), pp. 15926–15936. doi: 10.1074/jbc.M112.421438.
- Wang, R. *et al.* (2014) 'Antiviral Responses in Mouse Embryonic Stem Cells', *Journal of Biological Chemistry*, 289(36), pp. 25186–25198. doi: 10.1074/jbc.M113.537746.
- Wang, Y. *et al.* (2007) 'DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal.', *Nature genetics*, 39(3), pp. 380–5. doi: 10.1038/ng1969.
- Watanabe, T. *et al.* (2008) 'Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes', *Nature*, 453(7194), pp. 539–543. doi: 10.1007/s12601-017-0045-0.
- Weitz, S. H. *et al.* (2014) 'Processing of microRNA primary transcripts requires heme in mammalian cells', *Proceedings of the National Academy of Sciences of the United States of America*, 111(5), pp. 1861–1866. doi: 10.1073/pnas.1309915111.
- Westholm, J. O. and Lai, E. C. (2011) 'Mirtrons: MicroRNA biogenesis via splicing', *Biochimie. Elsevier Masson SAS*, 93(11), pp. 1897–1904. doi: 10.1016/j.biochi.2011.06.017.
- WHO Solidarity Trial Consortium (2021) 'Repurposed Antiviral Drugs for Covid-19 — Interim WHO Solidarity Trial Results', *New England Journal of Medicine*, 384(6), pp. 497–511. doi: 10.1056/NEJMoa2023184.
- Whyatt, L. M. *et al.* (1993) 'The responsiveness of embryonic stem cells to alpha and beta interferons provides the basis of an inducible expression system for analysis of developmental control genes', *Molecular and Cellular Biology*, 13(12), pp. 7971–7976. doi: 10.1128/mcb.13.12.7971-7976.1993.
- Wilson, R. C. *et al.* (2015) 'Dicer-TRBP complex formation ensures accurate mammalian MicroRNA biogenesis', *Molecular Cell*. Elsevier Inc., 57(3), pp. 397–407. doi: 10.1016/j.molcel.2014.11.030.
- Witteveldt, J., Ivens, A. and Macias, S. (2018) 'Inhibition of Microprocessor Function during

- the Activation of the Type I Interferon Response', *Cell Reports*. Elsevier Company., 23(11), pp. 3275–3285. doi: 10.1016/j.celrep.2018.05.049.
- Witteveldt, J., Knol, L. I. and Macias, S. (2018) 'MicroRNA-deficient embryonic stem cells acquire a functional Interferon response', *eLife*, pp. 1–19. doi: <https://doi.org/10.7554/eLife.44171.001>.
- Witteveldt, J., Knol, L. I. and Macias, S. (2019) 'MicroRNA-deficient mouse embryonic stem cells acquire a functional interferon response', *eLife*, 8, pp. 1–19. doi: 10.7554/eLife.44171.
- Wu, C. W. *et al.* (2018) 'A Comprehensive Approach to Sequence-oriented IsomiR annotation (CASMiR): Demonstration with IsomiR profiling in colorectal neoplasia', *BMC Genomics*. BMC Genomics, 19(1), pp. 1–12. doi: 10.1186/s12864-018-4794-7.
- Wu, J., Saint-Jeannet, J. P. and Klein, P. S. (2003) 'Wnt-frizzled signaling in neural crest formation', *Trends in Neurosciences*, 26(1), pp. 40–45. doi: 10.1016/S0166-2236(02)00011-5.
- Wu, X. *et al.* (2018) 'Intrinsic Immunity Shapes Viral Resistance of Stem Cells', *Cell*. Elsevier Inc., 172(3), pp. 423–438.e25. doi: 10.1016/j.cell.2017.11.018.
- Xie, M. *et al.* (2013) 'Mammalian 5'-capped microRNA precursors that generate a single microRNA', *Cell*, 155(7), pp. 1568–1580. doi: 10.1016/j.cell.2013.11.027.
- Xie, Z. *et al.* (2021) 'Gene Set Knowledge Discovery with Enrichr', *Current Protocols*, 1(3), pp. 1–51. doi: 10.1002/cpz1.90.
- Yang, L. N. *et al.* (2020) 'Sox10 Is a Specific Biomarker for Neural Crest Stem Cells in Immunohistochemical Staining in Wistar Rats', *Disease Markers*, 2020. doi: 10.1155/2020/8893703.
- Yeom, K. H. *et al.* (2006) 'Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing', *Nucleic Acids Research*, 34(16), pp. 4622–4629. doi: 10.1093/nar/gkl458.
- Yin, F. *et al.* (2014) 'LSD1 Regulates Pluripotency of Embryonic Stem/Carcinoma Cells through Histone Deacetylase 1-Mediated Deacetylation of Histone H4 at Lysine 16', *Molecular and Cellular Biology*, 34(2), pp. 158–179. doi: 10.1128/mcb.00631-13.

- Yin, Y., Zhou, L. and Yuan, S. (2018) 'Enigma of Retrotransposon Biology in Mammalian Early Embryos and Embryonic Stem Cells', *Stem Cells International*, 2018. doi: 10.1155/2018/6239245.
- Yoneyama, M. *et al.* (2005) 'Shared and Unique Functions of the DExD/H-Box Helicases RIG-I, MDA5, and LGP2 in Antiviral Innate Immunity', *The Journal of Immunology*, 175(5), pp. 2851–2858. doi: 10.4049/jimmunol.175.5.2851.
- Zeng, Y. and Cullen, B. R. (2003) 'Sequence requirements for micro RNA processing and function in human cells', *Rna*, 9(1), pp. 112–123. doi: 10.1261/rna.2780503.
- Zeng, Y., Yi, R. and Cullen, B. R. (2005) 'Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha', *EMBO Journal*, 24(1), pp. 138–148. doi: 10.1038/sj.emboj.7600491.
- Zeuthen, J. *et al.* (1980) 'Characterization of a human ovarian teratocarcinoma-derived cell line', *International Journal of Cancer*, 25(1), pp. 19–32. doi: 10.1002/ijc.2910250104.
- Zeuthen, J. (1981) 'Human teratocarcinoma cell lines. A review', *International Journal of Andrology*, 4, pp. 61–77. doi: 10.1111/j.1365-2605.1981.tb00654.x.
- Zhang, F. *et al.* (2001) 'Binding of Double-stranded RNA to Protein Kinase PKR is Required for Dimerization and Promotes Critical Autophosphorylation Events in the Activation Loop', *Journal of Biological Chemistry*. © 2001 ASBMB. Currently published by Elsevier Inc; originally published by American Society for Biochemistry and Molecular Biology., 276(27), pp. 24946–24958. doi: 10.1074/jbc.M102108200.
- Zhang, H. *et al.* (2004) 'Single processing center models for human Dicer and bacterial RNase III', *Cell*, 118(1), pp. 57–68. doi: 10.1016/j.cell.2004.06.017.
- Zhang, K. *et al.* (2018) 'A novel class of microRNA-recognition elements that function only within open reading frames', *Nature Structural and Molecular Biology*. Springer US, 25(11), pp. 1019–1027. doi: 10.1038/s41594-018-0136-3.
- Zhang, Q. *et al.* (2020) 'Inborn errors of type I IFN immunity in patients with life-threatening COVID-19', *Science*, 370(6515). doi: 10.1126/science.abd4570.
- Zhang, Z. *et al.* (2005) 'Tbx1 expression in pharyngeal epithelia is necessary for pharyngeal

arch artery development', *Development*, 132(23), pp. 5307–5315. doi: 10.1242/dev.02086.

Zhang, Z., Huynh, T. and Baldini, A. (2006) 'Mesodermal expression of Tbx1 is necessary and sufficient for pharyngeal arch and cardiac outflow tract development', *Development*, 133(18), pp. 3587–3595. doi: 10.1242/dev.02539.

Zhao, D. *et al.* (2015) 'MicroRNA profiling of neurons generated using induced pluripotent stem cells derived from patients with schizophrenia and schizoaffective disorder, and 22q11.2 del', *PLoS ONE*, 10(7), pp. 1–24. doi: 10.1371/journal.pone.0132387.

Zhu, A., Ibrahim, J. G. and Love, M. I. (2019) 'Heavy-Tailed prior distributions for sequence count data: Removing the noise and preserving large differences', *Bioinformatics*, 35(12), pp. 2084–2092. doi: 10.1093/bioinformatics/bty895.

Ziv, O. *et al.* (2018) 'COMRADES determines in vivo RNA structures and interactions', *Nature Methods*. Springer US, p. 1. doi: 10.1038/s41592-018-0121-0.

