Univerzita Karlova v Praze

Přírodovědecká fakulta

Studijní program: Vývojová a buněčná biologie

Mgr. Jana Nejepínská

Účinky exprese dlouhé dvouvláknové RNA v savčích buňkách

The effects of long double-stranded RNA expression in mammalian cells

Disertační práce

Školitel:
Doc. Mgr. Petr Svoboda, Ph.D.

Praha, 2013

Prohlášení

Prohlašuji, že jsem disertační práci zpracovala samostatně a že jsem řádně uvedla a citovala všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 18.4.2013

Jana Nejepínská

Acknowledgement

This work would never be accomplished without the other members of Petr Svoboda's lab, not only for their practical participation in my project but mainly for creating an extraordinary and enviable lab atmosphere.

First of all, I'd like to thank to Petr for giving me the chance to become a member of his lab. He always motivated me to try my best, which has driven me to achieve goals that seemed impossible to me a couple of years ago. In addition, I appreciate the effort and patience he has spent on the training of my writing and presenting skills.

Secondly, thanks to Matyáš, a peer collaborator and competitor at the same time, who has sewed as an exemplary scientist I have wanted to mimic.

Thirdly, many thanks to Radek, the great teacher, who has patiently assisted me with all the troubleshooting without expecting any reward. His perspective has shown how to see science objectively and critically.

Fourthly, I'd like to acknowledge everybody who gave me a hand by providing suggestions, assistance, or material for my experiments.

Additionally, I'd like to thank my family and friends, who were supporting me all the time and were happy for each of my successes, even if they had no idea what I have been doing and what it is good for.

Not to mention, I'd like to thank my colleagues from IMG, who showed me that scientists can have a diverse social life. I will never forget our frequent cycling, cross-country skiing, and canoeing trips, as well as film and gourmet evenings.

Last but not least, many thanks to Vašek for standing by me, for helping me to overcome all difficulties, and for his tolerance.

Table of contents

Abstrakt

Dvouvláknová RNA je cizorodá molekula, která se objevuje v buňce buď jako produkt replikace viru nebo vzniká párováním komplementárních RNA, a to v rámci jedné molekuly nebo mezi dvěma molekulami. Komplementární RNA často pocházejí z repetitivních sekvencí. U savců může dvouvláknová RNA vstupovat do jedné ze tří drah: sekvenčně specifického umlčování RNA (RNA silencing), sekvenčně nezávislé interferonové (IFN) odpovědi nebo editace pomocí adenosin deamin——áz.

Hlavním zaměřením mého PhD projektu bylo detailně zanalyzovat efekty dvouvláknové RNA produkované v savčích buňkách, a to v rámci celého organismu. Za tímto účelem jsme vytvořili konstrukt exprimující dvouvláknovou RNA ve formě mRNA, obsahující dlouhou a perfektně párující vlásenkovitou strukturu.

Transgenní myši, vytvářející dvouvláknovou RNA ve všech buňkách, byly životaschopné a, na rozdíl od výsledků dosavadních studií, neměly aktivovanou IFN odpověď. V somatických buňkách byla dvouvláknová RNA jen málo štěpena na krátké interferující RNA (siRNA), nezpůsobovala transkripční umlčování v *trans* a docházelo k mírné deaminaci adenosinu bez zadržování editované RNA v jádře. Shodné výsledky jsme získali v savčích buňkách, které byly tranzientně transfekovány plazmidem exprimujícím dvouvláknovou RNA. Naproti tomu ve vajíčcích způsobovala dvouvláknová RNA silnou RNA interferenci. Naše výsledky jako první ukazují, ze exprimovaná dvouvláknová RNA, na rozdíl od mnoha jiných forem dvouvláknové RNA, je v savčích buňkách dobře tolerována a že myší vajíčka reprezentují zvláštní buněčný typ, ve kterém je dvouvláknová RNA cílena do RNA interferenční dráhy.

Pokusy založené na tranzientní transfekci dále ukázaly, že plazmidy produkující dvouvláknovou RNA tlumí expresi společně transfekovaných reportérových plazmidů, ale pouze minimálně působí na expresi endogenních genů nebo reportérů stabilně začleněných do genomu. K této inhibici pravděpodobně dochází na úrovni iniciace translace a je zprostředkována místní a přechodnou aktivací protein kinázy R. Naše výsledky naznačují, že buňka by mohla být schopna odlišit endogenní mRNA od exogenní a v přítomnosti dvouvláknové RNA tlumit translaci cizí mRNA.

Analýza transkriptomu buněk transfekovaných běžně používanými plazmidy ukázala, že transkripce plazmidů je komplexní a může způsobit vznik dvouvláknové RNA párováním komplementárních úseků transkriptů. Naše výsledky ukazují, že vznik dvouvláknové RNA nelze předpovídat a že přítomnost dvouvláknové RNA v buňce může vést ke špatné interpretaci dat založených na tranzientních transfekcích.

Abstract

Double stranded RNA (dsRNA) is a foreign molecule that arises in the cell either as a by-product of viral replication or it is produced by the intramolecular or intermolecular pairing of complementary RNAs, often originating from repetitive sequences. In mammals, dsRNA can enter one of three pathways: the sequence-specific RNA silencing, the sequence-independent interferon (IFN) response, or editing by adenosine deaminases.

The main focus of my PhD project was to comprehensively analyze the effects of the expressed dsRNA in mammals in the context of the whole organism. To follow this aim, we generated a construct expressing dsRNA in a form of an mRNA containing a long perfect hairpin structure.

Transgenic mice ubiquitously expressing dsRNA were viable and, in contrast to the previous studies, the IFN response was not activated. In somatic cells, dsRNA was poorly processed into small interfering RNAs, did not cause transcriptional silencing *in trans*, and underwent low adenosine deamination without the nuclear retention. Consistent results were obtained in human cells transiently transfected with a dsRNA-expressing plasmid. On the other hand, dsRNA expression caused robust RNA interference (RNAi) in oocytes. Thus, we show for the first time that expressed dsRNA, in contrast to many other forms of dsRNA, can be well tolerated in mammals and that oocytes represent a privileged cell type in terms of directing dsRNA into the RNAi pathway.

In addition, transient transfection experiments revealed that dsRNA-expressing plasmids inhibit expression of co-transfected reporter plasmids, but they have a minimal impact on the expression of endogenous genes or reporters stably integrated in the genome. The inhibition likely occurs at the level of translation initiation and it is mediated by the local and transient activation of protein kinase R. Accordingly, our results indicate that cells might be able to distinguish between endogenous and exogenous mRNAs and selectively inhibit the translation of foreign mRNAs in response to dsRNA.

Finally, the transcriptome analysis of cells transfected with commonly used plasmids showed that plasmid transcription is complex and may result in the dsRNA formation via basepairing of complementary sense and antisense transcripts. Our results demonstrate that the formation of dsRNA structure cannot be efficiently predicted and can cause misinterpretation of data based on transient transfection experiments.

Abreviations

2-5A	2’,5’-linked oligoadenylates
A	Adenosine
ADAR	Adenosine deaminase acting on RNA
AP-1	Activator protein 1
APAF	Apoptotic protease activating factor
ATF-3	Cyclic AMP-dependent transcription factor 3
A-to-I	Adenosine to inosine
bp	Base pair
C	Cytidine
CMV	Cytomegalovirus
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA
dsRBD	Double-stranded RNA-binding domain
dsRNA	Double-stranded RNA
E2F1	E2F transcription factor 1
EGFP	Enhanced green fluorescent protein
eIF2	Eukaryotic initiation factor 2
endo-siRNA	Endogenous small interfering RNA
ERK	Extracellular signal-regulated kinase
FADD	Fas-associated protein with death domain
FL	Firefly luciferase
G	guanosine
GTP	Guanosine triphosphate
HEK293	Human embryonic kidney 293 cells
Hsp	Heat shock protein
I	Inosine
IFN	Interferon
IL-1	Interleukin 1
IR	Inverted repeat
IRF	Interferon regulatory factor
ISG15	Interferon-stimulated protein of 15 kDa
JNK	c-Jun N-terminal kinase
kan/neo	Kanamycin/neomycin resistance
LGP2	Laboratory of Genetics and Physiology 2
MAPK	Mitogen-activated protein kinase
MDA5	Melanoma differentiation-associated gene 5
MDA7	Melanoma differentiation-associated gene 7
Met-tRNAi	Initiator methionine transfer RNA
miRNA	MicroRNA

Mos	V-mos moloney murine sarcoma viral oncogene homolog
mRNA	Messenger RNA
Mx	Myxovirus resistance
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
nt	Nucleotide
OAS	Oligoadenylate synthetase
OASL	2'-5' oligoadenylate synthetase-like
p53	Protein 53
PACT	Protein activator of interferon-induced protein kinase
piRNA	Piwi-interacting RNA
PKR	Protein kinase R
pre-miRNA	Precursor microRNA
pre-mRNA	Precursor messenger RNA
pri-miRNA	Primary microRNA
RIG-I	Retinoic acid-inducible gene 1
RISC	RNA-induced silencing complex
RL	*Renilla* luciferase
RNA	Ribonucleic acid
RNAi	RNA interference
RNase	Ribonuclease
shRNA	Short hairpin RNA
siRNA	Small interfering RNA
ssDNA	Single-stranded DNA
ssRNA	Single-stranded RNA
STAT	Signal transducers and activators of transcription
TNF-α	Tumor necrosis factor alpha
TRAF	TNF receptor associated factor
TRBP	Trans-activation response RNA-binding protein
Tudor-SN	Tudor staphylococcal nuclease
U	Uridine
UTR	Untranslated region
ZP3	Zona pellucida 3

1. Introduction

Double-stranded RNA (dsRNA) is a double-helical RNA molecule formed by pairing of two antiparallel RNA strands. It has a well-defined structure that is recognized by several proteins acting in different pathways (Saunders & Barber, 2003). In mammals, dsRNA triggers various cellular responses protecting cells against miscellaneous threats (Obbard et al, 2009). Although dsRNA is a quite abundant molecule, it was always at the boundary of interest of molecular biology, considered only as a by-product of important cellular processes (such as transcription or viral replication).

The appreciation of dsRNA significance started fifteen years ago (in 1998), when Andrew Fire and Craig Melo published that injected long dsRNA induces sequence-specific messenger RNA (mRNA) degradation in the nematode *Caenorhabditis elegans* (Fire et al, 1998). The new cellular pathway, where dsRNA acts as a specificity factor, was named RNA interference (RNAi). In following years, RNAi was found to be evolutionary conserved in almost all eukaryotes, including mammals. Its widespread occurrence and the fact that deletions of key factors are lethal in many organisms underline its significance.

The possibility to specifically inhibit gene expression initialized the race to understand RNAi and related pathways in order to establish new ways in research (gene knock-down, functional genomics, genome-wide screenings), biotechnology (genetically modified plants), and medicine (new drug design). Currently, dsRNA is viewed as a molecule playing a variety of roles in different organisms, ranging from antiviral defense to regulation of gene expression.

1.1. Sources of dsRNA

The sources of dsRNA can be either exogenous (from the environment) or endogenous (produced by a cell itself).

First, dsRNA is a common by-product of viral replication. Almost all types of viruses produce dsRNA at some phases of their life cycle. In dsRNA viruses, double-helical RNA constitutes the genome; single-stranded RNA (ssRNA) viruses produce dsRNA as a replication intermediate (e.g. during replication of the second strand). Alternatively, ssRNA transcripts or parts of the genome can anneal to form an

intramolecular or intermolecular dsRNA motif. In DNA viruses, dsRNA often originates from bidirectional overlapping transcription (reviewed in Kumar & Carmichael, 1998).

Second, dsRNA can be produced in cells. Similarly to viral RNA, endogenous dsRNA can originate from pairing of complementary sequences either within one molecule (intramolecular duplexes forming hairpin-like structures) or from two RNA molecules that basepair with each other. The templates for such complementary RNAs generally come from the activity of transposable elements, repetitive sequences, genome inversions, small RNA precursors, or overlapping transcripts (either in *cis* by convergent transcription of the same element or in *trans* by pairing of a transcribed gene and its pseudogene). Accordingly, such endogenous dsRNAs constantly emerge in the cell (Hartig et al, 2009; He et al, 2008; Kumar & Carmichael, 1998; Yelin et al, 2003).

Third, dsRNA molecules can be introduced into cells during experimental procedures (e.g. techniques utilizing siRNAs, shRNAs, long dsRNAs, or generating RNAs that may encounter a complementary partner in the cell).

Depending on an origin, dsRNA can be found either in the nucleus or in the cytoplasm. Cellular localization represents one of the key features that affect dsRNA fate, mainly due to availability of dsRNA-recognizing proteins.

1.2. dsRNA recognition

dsRNA typically adopts a right-handed A-form helix conformation. In contrast to a B-form helix, which is common for DNA duplex, the A-form displays a broader and shallower minor groove, together with a deeper and narrower major groove (Ryter & Schultz, 1998). This conformation is recognized by the dsRNA-binding domain (dsRBD) (St Johnston et al, 1992), an evolutionary conserved and approximately 65-68 amino acid-long motif. In general, dsRBD binds dsRNA without a sequence-specificity, but some types of dsRNA-binding motifs recognize dsRNA structure in a sequence-specific manner, such as in the case of selective adenosine deamination (Stefl et al, 2010). In addition to dsRNA binding, dsRBD has affinity to ssRNA with extensive secondary structures but it ignores dsDNA, ssDNA, or ssRNA (reviewed in Saunders & Barber, 2003).

A presence of dsRBD is a common marker for most proteins interacting with dsRNA. These proteins regulate a variety of cellular functions, namely RNA stability, mRNA elongation, splicing, nuclear export, RNAi, RNA editing, regulation of translation, host defense, or signaling (summarized in Saunders & Barber, 2003). Additionally, there are alternative dsRNA-binding motifs, containing zinc finger motifs, lysine-rich regions, or other special motifs reviewed in (Ben-Asouli et al, 2002; Saunders & Barber, 2003).

1.3. **Responses to dsRNA**

The variability of dsRNA sources is closely associated with the complexity of responses elicited by its recognition. Depending on the organism and cell type, dsRNA can be engaged in one of three well characterized pathways (or their combination): RNA interference (RNAi), interferon (IFN) response, or A-to-I RNA editing (Fig. 1). Although all pathways share a common substrate - long dsRNA, they differ in their evolutionary conservation, cellular localization, and outcome.

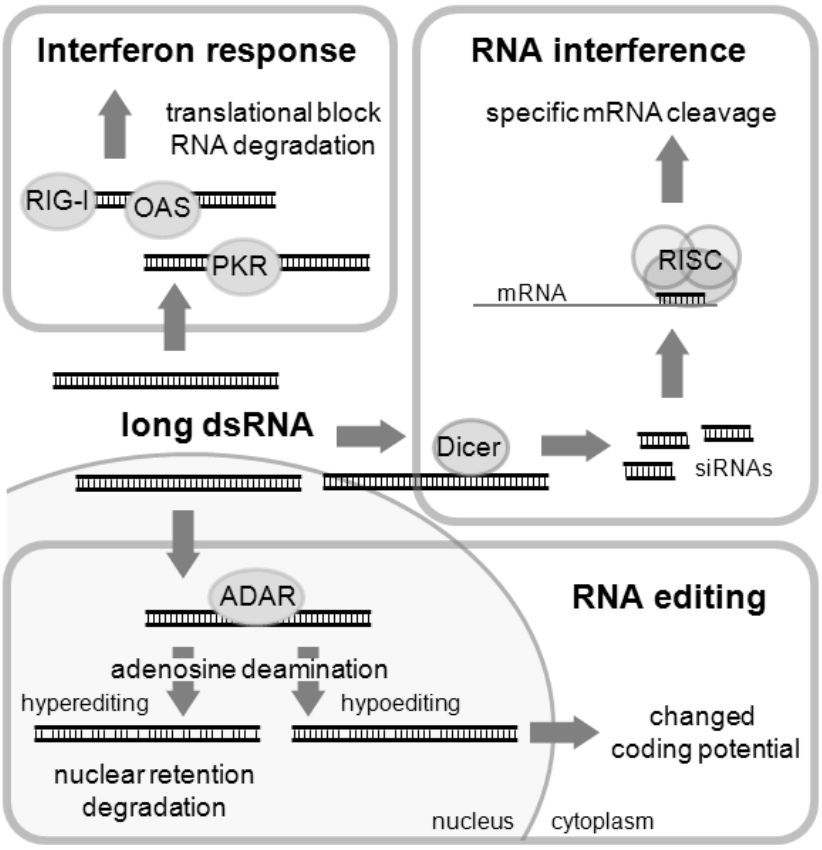


Figure 1: Pathways activated by long dsRNA (adopted from Chalupnikova et al, 2013).

1.3.1. **RNA silencing pathways**

RNA silencing is a term collectively referring to related pathways that share two common features: (1) small RNA molecules that act as specificity factors directing effector proteins to complementary RNAs and (2) proteins from Ago family (Carthew & Sontheimer, 2009).

The mechanism of action of RNA silencing machinery is generally inhibitory, comprising of the regulation of transcription, RNA stability, RNA processing, translation, and chromatin structure (summarized in Carthew & Sontheimer, 2009). Apart from RNAi (which receives the most attention in this work), mammalian RNA silencing pathways include microRNA (miRNA) and PIWI-interacting RNA (piRNA) pathways (reviewed in Kim et al, 2009).

1.3.1.1. **RNA interference**

RNAi is the most ancient RNA silencing pathway, highly conserved among all eukaryotes with a few exceptions, such as *Saccharomyces cerevisiae* (Aravind et al, 2000) or some protozoan parasites (*Leishmania major*, *Trypanosoma cruzi*) (Ullu et al, 2004), which apparently lost the functional RNAi pathway during evolution.

In RNAi, long dsRNA substrate is processed by the ribonuclease (RNase) III enzyme Dicer into 21-22 nucleotide (nt)-long small interfering RNAs (siRNAs); each terminus contains a 5' monophosphate group and a 3' dinucleotide overhang (reviewed in Jinek & Doudna, 2009). One of the siRNA strands is incorporated into the RNA-induced silencing complex (RISC) and acts as a guide for binding to the complementary mRNA. The cognate mRNA is subsequently cleaved in the middle of the basepairing region by a "slicer" (AGO2 protein) and rapidly degraded (reviewed in Ghildiyal & Zamore, 2009). RNAi pathway is sequence-specific; i.e. it cleaves only sequences complementary to long dsRNA that triggered the response. In the case of perfect or nearly perfect complementarity between siRNA strand and target RNA (usually mRNA), target RNA is degraded. Alternatively, in the presence of bulges and mismatches between the two molecules, the response results rather in the translational repression of cognate mRNAs than its cleavage (Doench et al, 2003), typical for the miRNA pathway.

1.3.1.2. **MicroRNA pathway**

miRNAs represent a specific class of regulatory small RNAs originating from a short hairpin precursor. In many organisms, including mammals, miRNA pathway shares components with RNAi machinery. Consequently, both pathways extensively overlap, mainly in biogenesis and mechanism of action (reviewed in Nejepinska et al, 2012a).

miRNAs are generated from endogenous primary transcripts (pri-miRNAs), ~70 nt-long capped and polyadenylated RNAs with local hairpin structures. The imperfect stem-loops are recognized and cleaved by a nuclear RNase III Drosha-containing complex into short hairpin intermediates (pre-miRNAs). Exportin-5 transports pre-miRNAs into the cytoplasm, where Dicer performs another endonucleolytic cleavage to release a short duplex of approximately 20 bp flanked by 2-nt 3' overhangs. The strand that is less stable at the 5' end is loaded into the RISC complex and forms the mature miRNA. A typical mammalian miRNA imperfectly base-pairs with a cognate 3' untranslated region (UTR) and inhibits protein translation (summarized in Umbach & Cullen, 2009). In contrast, miRNAs perfectly basepairing with their targets are engaged into the RNAi pathway (Yekta et al, 2004).

1.3.1.3. **The importance of RNA silencing**

Although the world of small RNAs has been discovered relatively recently, their significance is enormous. There are several lines of evidence highlighting the importance of RNA silencing: (1) the key components of the silencing machinery are conserved from yeasts to humans (Ketting, 2011), (2) genes in the siRNA pathway undergo fast adaptive evolution due to a strong positive selection in response to mutating viruses (Obbard et al, 2006), and (3) the deletion of Dicer or AGO2 in mammals is lethal (Bernstein et al, 2003; Liu et al, 2004).

The most probable evolutional role of RNA silencing is the antiviral defense. The importance of RNAi against viral infection has been demonstrated in plants and many invertebrate taxons by fulfilling these criteria: (1) virus-derived siRNAs are present in infected cells, (2) inhibition of RNA silencing enhances viral replication, and (3) many viruses have evolved so called 'suppressors of RNA silencing', molecules encoded by

many viruses that block different steps of RNA silencing or the IFN response (summarized in Marques & Carthew, 2007; Nejepinska et al, 2012a).

On the other hand, convincing evidence for the antiviral role of RNA silencing is missing in vertebrates, including mammals. Although the complete RNAi machinery is present and experimentally induced RNAi inhibits viral replication in mammals (reviewed in van Rij & Andino, 2006), the active role of RNAi in combating viruses has not been proven. Consistently, the exhaustive search for virus-derived siRNAs has also been unsuccessful. It is believed that the primitive nucleic acid-based immune system (represented by RNAi) was replaced by a more sophisticated protein-based system (mediated by antibodies and components of the IFN response) (reviewed in Cullen, 2006; Marques & Carthew, 2007).

In addition to combating viruses, RNA silencing pathways probably serve as a defense mechanism against invasive nucleic acids. In contrast to the antiviral role of RNAi, the suppression of transposable elements was likely retained in mammals, as endo-siRNA and piRNA pathways help to protect mouse germline (reviewed in Nejepinska et al, 2012a).

Although the significance of RNAi and piRNA pathways in mammals is not completely understood, miRNA pathway plays a crucial role. Since its discovery, it has been found to be involved in the regulation of almost all studied biological processes, including development, cellular proliferation, apoptosis, homeostasis, or carcinogenesis (reviewed in Umbach & Cullen, 2009). It is estimated that more than 60 % of human genes are miRNA-regulated (Friedman et al, 2009). Remarkably, miRNA pathway is suppressed and non-essential in mouse oocytes (Ma et al, 2010; Suh et al, 2010).

1.3.2. **Interferon response**

The origin of dsRNA as a by-product of viral replication makes it a convenient target for the innate immune response machinery. Indeed, dsRNA is recognized by a handful of proteins collectively operating in the complex of pathways whose common aim is to signal infection, restrict viral growth, and limit viral spread (Gantier & Williams, 2007). As the hallmark of this pathway is a production of antiviral cytokines IFNs, the

response was termed the IFN response. In contrast to RNAi, the IFN response is limited to vertebrates and it recognizes dsRNA molecule without any concern to its sequence (with some exceptions mentioned below).

Two branches form the backbone of the IFN response: protein kinase R (PKR) and 2'5'-oligoadenylate synthetase (OAS)/RNase L systems.

1.3.2.1. **PKR pathway**

PKR alias 'IFN-inducible protein kinase' or 'eukaryotic translation initiation factor 2-alpha kinase 2' (EIF2αK2) is a cytoplasmic dsRNA-activated serine/threonine kinase. PKR is expressed ubiquitously at low levels but its expression increases upon exposure to type I and type III IFNs (Ank et al, 2006; Garcia et al, 2006).

PKR contains a dsRBD and a kinase domain. In the inactive state, PKR exists as a monomer, in which the dsRBD sterically blocks the kinase domain. The activation of PKR by dsRNA or other stress signals leads to PKR homodimerization and autophosphorylation. Activated PKR dimers phosphorylate their substrates; the main substrate is alpha subunit of eukaryotic initiation factor 2 (eIF2α) (reviewed in Garcia et al, 2006; Garcia et al, 2007; Sadler & Williams, 2008). eIF2α is responsible for the delivery of initiator methionine transfer RNA (Met-tRNAi) to 40S ribosomal subunit to initiate protein synthesis (summarized in Garcia et al, 2006). Phosphorylated eIF2α has up to 100-fold increased affinity to GTP-exchange factor eIF2B, which is present in a limiting amount compared to eIF2α. As a consequence, PKR activation results in a dramatic drop in translation initiation (Sudhakar et al, 2000).

PKR can be activated by different types of molecules, including dsRNA. As mentioned above, dsRNA binds directly to dsRBD at the N-terminus of PKR to trigger the antiviral response. The activating dsRNA can be of viral, cellular, or synthetic origin (such as polyinosinic : polycytidylic acid). Earlier data indicated that PKR is activated only by dsRNA fragments longer than 33 base pairs (bp) (Manche et al, 1992). This minimal length requirement can be explained by the structure of dsRBD (two identical dsRNA-binding motifs with a random-coiled 20 amino acid linker), which wrap around the dsRNA molecule for optimal protein-RNA interactions (Nanduri et al, 1998).

The PKR activation only by molecules longer than 30 bp is frequently considered as a fact by many scientists working with siRNAs and shRNAs, as it avoids the need to examine the non-specific effects in their experiments. However, the 30-bp size limit is not absolutely valid. Several groups reported the activation of PKR by shorter dsRNA molecules (Hornung et al, 2005; Judge et al, 2005; Schlee et al, 2006; Sledz et al, 2003). Interestingly, an "ss-dsRNA motif" (an imperfect dsRNA stem of at least 16 bp with flanking ss tail(s) of 11-15 nt) efficiently binds to and activates PKR (Zheng & Bevilacqua, 2004) if it contains 5'-triphosphate (Nallagatla et al, 2007). Nevertheless, PKR is optimally activated by dsRNAs longer than 80 bp in the A-conformation without any sequence requirements (Garcia et al, 2006).

Although PKR activation is known to inhibit cellular translation globally, several reports demonstrated the inhibition of plasmid-derived proteosynthesis without affecting the expression of endogenous genes (Ben-Asouli et al, 2002; Kaufman et al, 1989). Moreover, except its prevalent cytoplasmic localization, PKR was also detected in nuclei (Jeffrey et al, 1995) or associated with ribosomes (Minks et al, 1979), suggesting that the role for PKR is not fully elucidated, yet.

Except dsRNA, there are several other activators of PKR, such as proinflammatory stimuli (e.g. toll-like receptors), cytokines, growth factors, or oxidative stress. Proinflammatory cytokines (such as IL-1, TNF-α, IFN-α, IFN-β, or IFN-γ) not only activate PKR, but they are also upregulated upon PKR activation, creating a positive feedback loop in the PKR pathway (reviewed in Beutler, 2004; Garcia et al, 2006).

In addition to well-established role for PKR in translational regulation, it also regulates many signal transduction pathways. The main signaling molecules affected by PKR are NF-κB, signal transducers and activators of transcription (STAT1, 3), interferon-regulatory factors (IRF-1, 3, 7), p53, members of MAPK pathway (JNK, p38, ERK), TRAF proteins (TRAF2, 3, 6), ATF-3, or AP1. These factors control reactions to stress, cell damage, mediate cell cycle arrest, growth inhibition, or apoptosis (reviewed in Garcia et al, 2006; Garcia et al, 2007).

Apoptosis is an essential mechanism to prevent spreading of an infection. In the IFN response, there are several independent triggers of apoptosis and some of them are mediated by PKR. The first mechanism causing the cell death is the phosphorylation of eIF2α. The subsequent absence of proteosynthesis, if it is

persistent, leads ultimately to cell death. The second mediator of apoptosis is NF-κB. Although this factor is generally considered antiapoptotic, under some circumstances, such as viral infection, it can be involved in pathways leading to apoptosis (Donze et al, 2004). The third connection of PKR to apoptosis is mediated by caspases. Both FADD/caspase 8 and APAF/caspase 9 pathways were found to be involved in PKR-mediated apoptosis (Garcia et al, 2006). In addition, apoptosis can be induced by OAS/RNase L pathway (Silverman, 2007a).

The pleiotropic effects of PKR activation need to be tightly regulated. PKR is activated by PACT, MDA7, and E2F1. Conversely, several cellular proteins inhibit PKR, such as TRBP, Hsp70, Hsp90, nucleophosmin, p58, or p67 (summarized in Garcia et al, 2006). Interestingly, high concentrations of long dsRNA also inhibit PKR enzyme (Manche et al, 1992). In addition, plenty of viruses evolved their own inhibitors of PKR pathway, acting at different levels (such as sequestering dsRNA, inhibition of PKR activation or dimerization, synthesis of PKR pseudosubstrates, PKR inactivation or degradation) (Garcia et al, 2006; Garcia et al, 2007). This underlines the importance of PKR as a key defense protein against viral infection.

1.3.2.2. **OAS/RNase L pathway**

Another part of the IFN response is represented by OAS/RNase L system. Similarly to PKR, OAS and RNase L are expressed ubiquitously at low levels and their expression is enhanced upon IFN induction (Sadler & Williams, 2008).

Humans possess three functional OAS genes (*OAS1, OAS2, OAS3*), resulting in a dozen of alternatively spliced isoforms; mice have eight *Oas1* gene variants, one *Oas2*, and one *Oas3* gene. In addition, both species encode 'oligoadenylate synthetase-like' OASL gene, which binds dsRNA but it does not have the oligoadenylate synthetase activity; hence, it functions as dominant negative regulator (reviewed in Silverman, 2007a). Interestingly, although OAS is activated by dsRNA, it does not contain any defined dsRBD. Therefore, its activation is mediated by dsRNA molecules with special structural requirements (NNWWNNNNNNNNNWGN) forming the A-helix (Ben-Asouli et al, 2002). This indicates that there might be a mechanism distinguishing between viral and endogenous mRNA, though many cellular mRNAs fulfill the criteria. In

addition, many siRNAs also match this structure, suggesting that the IFN response can be triggered by molecules even shorter than 30 bp (Ben-Asouli et al, 2002), despite previous contradictory observations (Gantier & Williams, 2007; Minks et al, 1979).

Upon binding dsRNA, OAS is activated and converts adenosine triphosphates to pyrophosphates and 2'5'-linked oligoadenylates (2-5A) (Silverman, 2007b). 2-5A must contain at least one 5' phosphoryl group, 2' to 5' internucleotide linkage and more than two adenosines (Sadler & Williams, 2008) to activate latent RNase (RNase L). Activated RNase L non-specifically degrades viral and cellular RNA (including mRNA and ribosomal RNA); thereby resulting in decreased proteosynthesis (summarized in Player & Torrence, 1998; Silverman, 2007a).

1.3.2.3. **Other members of the IFN response**

Besides PKR and OAS/RNase L, there are additional players involved in the activation or regulation of the IFN response (reviewed in Gantier & Williams, 2007). The circumstances under which these molecules are activated depend on various factors, such as cell type, dsRNA origin, localization, length, structure, or termini (Gantier & Williams, 2007; Hornung et al, 2005; Judge et al, 2005; Marques et al, 2006; Sioud, 2005; Stein et al, 2005; Yang et al, 2001).

The integral part of the IFN response represent helicases from 'retinoic acid-inducible gene 1' (RIG-I)-like family, namely RIG-I, melanoma differentiation-associated gene-5 (MDA-5), and LGP2. These proteins sense viral RNAs present in the cell and complement each other to evoke an antiviral state. RIG-I (alias DDX58) recognizes short dsRNA or ssRNA with a 5' triphosphate (Hornung et al, 2006; Luo et al, 2011), as features distinguishing RNA generated by viral polymerases from 5'-capped cellular mRNAs. On the other hand, MDA-5 recognizes long dsRNA substrates without the 5' triphosphate (Kato et al, 2008). This collective action of RIG-I and MDA-5, together with their positive regulator LGP2 (alias DHX58), helps to detect different types of viruses (Satoh et al, 2010).

Toll-like receptors (TLR) represent additional molecules recognizing viral infection. The detection of dsRNA by TLR-3 results in NF-κB activation and IFN production

(Alexopoulou et al, 2001). Last but not least, IFN response is induced upon activation of ISG15 or proteins from Mx GTPase family (reviewed in Sadler & Williams, 2008).

1.3.3. **RNA editing**

In addition to RNA silencing and the interferon response, dsRNA serves as a substrate for 'adenosine deaminase acting on RNA' (ADAR) enzymes. One to three ADAR paralogues are present in all animals; in contrast, they are completely absent in protozoa, yeasts, and plants (Jin et al, 2009). Mammals possess three adenosine-to-inosine (A-to-I) editing enzymes (ADAR1-3) and ADAR1 has two isoforms (p110 and p150). All of them contain a deaminase domain plus 2-3 dsRBDs and operate either in the nucleus or in the cytoplasm; yet, ADAR3 is enzymatically inactive (reviewed in Nishikura, 2010).

ADARs recognize largely or completely double-stranded RNAs in the size over 20 bp (two turns of the dsRNA helix) (Nishikura et al, 1991). In these substrates, adenosine (A) is deaminated to inosine (I) (a nucleoside in which hypoxanthine is attached to ribose). Newly-arisen inosine is recognized as guanosine (G) by the translation and splicing machineries; hence, ADAR editing can influence mRNA processing and protein coding. In addition, inosine preferentially pairs with cytidine (C) (instead of original adenosine – uridine (U) pair), which changes RNA structure (reviewed in Hundley & Bass, 2010; Nishikura, 2010).

For short dsRNAs (20-30 bp) or dsRNAs containing bulges, mismatches, or loops, ADAR editing is usually selective (hypoediting). In contrast, long (>100 bp) and perfect substrates are edited randomly and massively, with up to 50 % of all adenosines converted (hyperediting); hyperedited RNAs are often captured in the nucleus and degraded (reviewed in DeCerbo & Carmichael, 2005; Nishikura, 2010).

1.3.3.1. **Roles of ADAR editing**

A significance of ADARs is underlined by its conservation and the fact that ADAR deletion is lethal in vertebrates (Higuchi et al, 2000; Wang et al, 2000). Although a physiological function of ADARs remains enigmatic, several roles for different types and regions of RNAs have been described.

The first discovered ADAR substrates were mRNAs edited in the protein-coding region. In this case, A-to-I editing serves as an additional (post-transcriptional) level in protein diversification by creating or eliminating splice sites or stop codons or by changing amino acid composition of proteins. About 30 receptors and ion channels undergo editing in the brain, resulting for example in a change of ion channel permeability (summarized in Jepson & Reenan, 2008).

The first discovered and the best known example of ADAR editing in the protein-coding region represents glutamate receptor B. The mechanism of this site-selective editing is based on an imperfect RNA hairpin that folds from an exon and a proximal intron site of pre-mRNA. Subsequently, A-to-I editing at RNA level results in a conversion from glutamine to arginine at protein level. Interestingly, the lethal phenotype of ADAR2 knock-out mice can be rescued by the introduction of edited glutamate receptor B (Higuchi et al, 1993), suggesting a highly specific role of ADAR2.

However, the majority of ADAR editing sites are located in non-coding sequences, such as introns or UTRs; a large portion maps to repetitive elements (Athanasiadis et al, 2004; Levanon et al, 2004). Although these conversions do not influence protein sequence, they seem to be important for many cellular processes (mentioned below).

One of the proposed functions of ADARs is to modulate dsRNA stability. As I pairs preferentially with C, a change of AU pairs to IU pairs destabilizes double-stranded structure; conversely, an alteration from AC to IC stabilizes dsRNA (reviewed in Hundley & Bass, 2010). The structure of dsRNA may affect for instance binding of dsRNA-binding proteins, including Dicer or PKR (chapter 1.3.4).

Additionally, A-to-I editing may impair translation. Inosine-containing RNAs were found to be retained in the nucleus (by binding to p54nrb protein) (DeCerbo & Carmichael, 2005; Zhang & Carmichael, 2001), which makes them spatially separated from translation machinery. In addition, mRNAs with inosine-rich 3'UTRs seem to associate with fewer ribosomes compared to unedited mRNAs (Hundley et al, 2008). Although the regulation of mRNA translation via compartmentalization plays an important role in the case of 'CAT2 transcribed nuclear RNA' (Prasanth et al, 2005), nuclear retention of edited RNAs is not a general rule (Hundley et al, 2008; Nejepinska et al, 2012b).

Furthermore, edited RNAs can become substrates for degradation. Tudor staphylococcal nuclease (Tudor-SN) or its interacting factor is probably an effector (Scadden, 2005). As Tudor-SN associates with RISC complex, a role in RNAi is presumable.

Interestingly, a large portion of A-to-I editing was detected in 3'UTRs, common miRNA-binding sites. Conversely, several pri-miRNAs, pre-miRNAs, or mature miRNAs contain inosine conversions (reviewed in Hundley & Bass, 2010). A selective deamination of miRNAs can affect miRNA pathway at different steps. (1) The editing of miRNA precursors may inhibit Drosha or Dicer cleavage or (2) prevent pre-miRNA export into the cytoplasm. (3) Extensively edited miRNA precursors can be targeted for degradation. (4) Editing can also result in the onset of a new set of targets or a selection of the other leading strand. (5) A tissue-specific ADAR editing can arrange a targeting in selected cell types. (6) A-to-I conversion in 3'UTR may create or eliminate miRNA target sites (reviewed in Nishikura, 2010).

1.3.4. Interaction among dsRNA-activated pathways

As mentioned above, dsRNA is a substrate for several enzymes, mediating different cellular functions; hence, the selection of a particular dsRNA-recognizing enzyme depends on various factors. Accordingly, different dsRNA-binding proteins process different types of dsRNA, with respect to dsRNA length, termini, or mismatches. In addition, some pathways are present only in some organisms, tissues, or subcellular localization. For instance, RNAi and the IFN response are prevalently cytoplasmic; on the other hand, ADAR editing takes place mainly in the nucleus. Additionally, the IFN response is conserved only in vertebrates, ADAR enzymes are absent in yeasts or plants, while miRNA pathway is suppressed in fully-grown oocytes and early embryos. In the following sections, I will summarize how dsRNA-responding pathways collaborate or compete with each other for the same substrate and how they are balanced.

1.3.4.1. **Interplay between RNA silencing and ADAR editing**

Both Dicer and ADAR enzymes compete for dsRNA substrate, so they negatively influence each other's activity. Therefore, A-to-I editing system might have evolved to modulate RNAi efficacy (Nishikura, 2010). As described in detail in section 1.3.3.1, edited RNAs become often resistant to Dicer processing. Moreover, ADAR1 expression is enhanced upon injection of a high amount of siRNAs (Hong et al, 2005), suggesting a feedback mechanism that maintains the balance between RNA silencing and RNA editing pathways.

1.3.4.2. **Interplay between RNA silencing and the IFN response**

Possibly due to a common original antiviral role, there is a mechanistical connection between RNA silencing and the IFN response. For example, two dsRNA-binding proteins, TRBP and PACT, interact both with Dicer and PKR; while TRBP mediates the inhibition of PKR (Park et al, 1994), PACT activates it (Patel & Sen, 1998). In addition, there is a significant homology between helicases acting in RNA silencing in *C. elegans* or the IFN response in mammals (e.g. RIG-I, LGP2, MDA5), further indicating that the IFN response could have partly evolved from RNA silencing machinery (reviewed in Nejepinska et al, 2012a). Therefore, there is a variety of proteins recognizing different dsRNA features (such as its length, termini, or structure) to distinguish between dsRNAs that should be recruited into the IFN response or RNAi pathways.

1.3.4.3. **Interplay between the IFN response and ADAR editing**

ADAR editing and the IFN response are closely coupled because ADAR1 isoform p150 has an IFN-inducible promoter; in other words, IFN-α or IFN-γ proteins upregulate ADAR1 expression (Patterson & Samuel, 1995). This suggests that ADAR can play a role in the antiviral response, possibly by editing viral RNA to inhibit viral replication. Indeed, hyperedited viral transcripts were isolated from cells infected by different types of RNA and DNA viruses, including hepatitis C virus or influenza virus (summarized in Nishikura, 2010; Samuel, 2001; Samuel, 2011). Additionally, some viruses were found to produce molecules that inhibit ADAR1 activity (vaccinia virus E3L or adenovirus VAI RNA) (reviewed in Samuel, 2011).

On the other hand, there are numerous examples showing that ADARs can stimulate viral infection. Importantly, ADAR1 inhibits PKR activation; thus, it antagonizes the IFN response (Nie et al, 2007). In addition, IU-rich RNAs inhibit IRF3; thereby, they suppress IFN-stimulated genes and apoptosis (Vitali & Scadden, 2010). Consistently, a proviral effect of ADAR was demonstrated in several viruses, such as measles virus or human immunodeficiency virus (reviewed in Samuel, 2011).

Interestingly, site-selective editing of viral RNA is also documented. In the case of hepatitis D virus (HDV), A-to-I editing is essential for the production of the longer version of delta antigen protein (HDAg-L), which mediates packaging of the viral genome. However, overexpression of ADARs leads to the hyperediting of viral RNA and inhibits HDV replication (Jayan & Casey, 2002). Last but not least, site-selective editing is also present in Epstein-Barr virus and human herpesvirus 8, where editing involves even a virus-encoded miRNA (reviewed in Samuel, 2011).

Taken together, ADARs can play either proviral or antiviral role during infection depending on a particular virus, host, and level of ADAR and other antiviral molecules.

1.3.5. dsRNA-activated pathways in oocytes

Mammalian oocytes respond to dsRNA presence much differently than somatic cells. The IFN response is absent in mouse oocytes and preimplantation embryos and the crucial genes (*Pkr*, *Rnasel*, and active *Oas* isoforms) are not expressed (Stein et al, 2005). Similarly, miRNA pathway is suppressed in oocytes (Ma et al, 2010; Suh et al, 2010). In contrast, RNAi is much more efficient in mouse oocytes compared to somatic cells, probably due to an oocyte-specific isoform of Dicer (Flemr *et al.*, submitted).

Taken together, mouse oocytes represent a special cell type where the sequence-specific effect of long dsRNA can be studied without the risk of non-specific effects caused by the IFN response.

1.4. Background of the project

This project studying dsRNA-activated pathways in mammals was based on preceding results from Richard Schultz group (Stein et al, 2003; Svoboda et al, 2000; Svoboda et

al, 2001). In all these studies, mouse *Mos* gene coding sequence was used to generate dsRNA to study dsRNA-activated pathways in oocytes.

1.4.1. **Mos**

Mos (c-mos, v-mos Moloney murine sarcoma viral oncogene homolog) is a maternal mRNA expressed specifically in oocytes and early embryos (Heikinheimo et al, 1995). MOS is a serine/threonine kinase that activates MAP kinase cascade through direct phosphorylation of the MAPK activator MEK (Gebauer & Richter, 1997). During meiosis, a high level of MOS protein (generating a high level of MAPK) is critical for maintaining the meiotic arrest at metaphase II prior to fertilization (Singh & Arlinghaus, 1997). The insufficient level of MOS promotes parthenogenetic activation of oocytes, which is accompanied by sterility or reduced female fertility. On the other hand, there is no effect of Mos knock-out in somatic cells or males (Gebauer & Richter, 1997; Hashimoto et al, 1994).

There were several reasons for choosing Mos gene as a model studying the effects of long dsRNA expression. First, Mos knock-out has a well-defined phenotype (the parthenogenesis in oocytes). Second, Mos expression is oocyte-specific; hence, Mos dsRNA would not cause a sequence-specific phenotype in somatic cells. Third, Mos is a dormant maternal mRNA that is abundant in oocytes (cca. 10,000 copies) (Keshet et al, 1988), which facilitates expression analysis. Fourth, MOS protein is a serine/threonine kinase with a well-established role in MAPK-pathway, which allows for single-cells assays monitoring the recruitment of Mos mRNA for translation in oocytes.

For the overmentioned reasons, the mouse *Mos* gene coding sequence was originally used to generate dsRNA inducing Mos knock-down by RNAi mechanism (Svoboda et al, 2000). The same sequence was utilized in the subsequent projects investigating the effects of long dsRNA in oocytes and preimplantation embryos (Stein et al, 2003; Stein et al, 2005; Svoboda et al, 2004; Svoboda et al, 2001).

1.4.2. **Overview of the previous findings**

The foundation of the current project dates back to 2000, when *in vitro*-transcribed long Mos dsRNA was microinjected into mouse oocytes to assess mammalian RNAi

activity (Svoboda et al, 2000). The injected dsRNA induced a sequence-specific degradation of Mos mRNA in a time-dependent and concentration-dependent manner, phenocopying the Mos knock-out. Therefore, results showed that dsRNA can be used to study function of maternal mRNAs by the knock-down approach, which is more efficient than previous antisense RNA approach and less laborious than gene knock-out by homologous recombination. In addition, it opened the possibility that RNAi may function in mammals (Svoboda et al, 2000). It should be mentioned that a similar study employing Mos dsRNA in mouse oocytes and early embryos and giving comparable results was published earlier in 2000 (Wianny & Zernicka-Goetz, 2000).

Subsequently, Mos dsRNA was expressed from a plasmid in the form of a hairpin, representing a model system similarly efficient as previously utilized *in vitro*-transcribed dsRNA. The optimization of individual plasmid components resulted in a construct with either ubiquitously expressed cytomegalovirus (CMV) or oocyte-specific zona pellucida 3 (ZP3) promoter, followed by a short intron, enhanced green fluorescent protein (EGFP) coding sequence, the long inverted repeat (IR) (535 bp) with a short (20-50 nt) spacer, and ampicillin resistance (Svoboda et al, 2001). The arisen construct was suitable for both microinjection and transgene-based induction of RNAi in mouse oocytes and preimplantation embryos.

Later, the previously generated construct expressing MosIR from oocyte-specific ZP3 promoter was introduced into mice. The phenotype of mice mimicked Mos knock-out phenotype and represented one of the first RNAi transgenes in mammals (Stein et al, 2003).

Lastly, the MosIR-expressing oocytes from the previously generated transgenic mice (Stein et al, 2003), were also used for microarray and real-time PCR analysis to reveal the presence of potential off-target effects and the IFN response. Results revealed that long dsRNA induces highly specific knock-down of target mRNA without affecting other genes. In addition, it was demonstrated that mouse oocytes do not express genes required for the IFN response (*Oas*, *RNase L*), but there is a high level of the inhibitory *Oas1d* isoform (Stein et al, 2005).

2. Aims of the project

The main objective of this PhD project was to explore how mammals respond to long dsRNA expression. This study provides a comprehensive view on the effects of dsRNA (1) in the context of the whole organism (mice expressing dsRNA) and (2) under a variety of cell culture conditions. In contrast to previously used dsRNA model systems, we utilized dsRNA expressed in the nucleus in the form of an mRNA with a long hairpin structure.

We aimed to address following objectives:

- Develop a model system that generates dsRNA *in vivo*
- Determine if dsRNA expression can be tolerated in mammals
- Analyze which pathways are recruited for MosIR processing
- Discover how dsRNA-responding pathways intersect
- Monitor how the response to dsRNA differs across different cell types
- Examine the behavior of different dsRNA types and origins

The revelation that expressed dsRNA inhibits proteosynthesis of co-transfected reporters without affecting cellular viability established additional aims:

- Investigate the mechanism of the reporter inhibition
- Examine the fate of transfected plasmids
- Reveal how cells distinguish between endogenous and exogenous mRNAs

3. Summary of results and discussion

3.1. Model system for dsRNA expression

There are several strategies for introducing long dsRNA into cells. The easiest way is to produce dsRNA by *in vitro* transcription and deliver it directly into cells (e.g. by transfection or microinjection). Alternatively, dsRNA can be expressed from polymerase II or polymerase III vectors by choosing one of these options. (1) Sense and antisense RNA can be expressed separately from two vectors, or (2) from one vector with two separately transcribed regions. (3) A target sequence can be transcribed from two promoters placed in opposite directions. (4) A single promoter can transcribe an IR; the main benefit of the last technique is a high probability of forming a duplex and the regulation of the dsRNA amount by the strength of a promoter. On the other hand, the cloning of the IR is laborious (summarized in the Supplement 1).

In this project, we used a transgene (pCagEGFP-MosIR, Fig. 2) expressing long dsRNA in the form of a perfect hairpin from a polymerase II promoter. The previously optimized construct design (Stein et al, 2003; Svoboda et al, 2001) was modified by employing the strong chimeric promoter fusing the CMV enhancer and chicken β-actin core promoter to allow for the expression of the transgene in all tissues/cell lines (Kaname & Huxley, 2001). The promoter is followed by an intron that enhances the expression of the transgene (Svoboda et al, 2001). The expression of the plasmid can be monitored via EGFP fluorescence. The 3'UTR of the EGFP contains an insert of the IR expressing dsRNA corresponding to a fragment from the coding region of the *Mos* sequence (535 bp, bases 554-1088) (Svoboda et al, 2000). Upon transcription, the IR presumably folds into a perfect hairpin with a short loop (Svoboda et al, 2001); thereby, the resulting dsRNA is embedded in a spliced, capped, and polyadenylated mRNA. The ability of the transgene to fold into dsRNA and to induce RNAi effect in mouse oocytes was verified in previous studies (Stein et al, 2003; Stein et al, 2005; Svoboda et al, 2001).

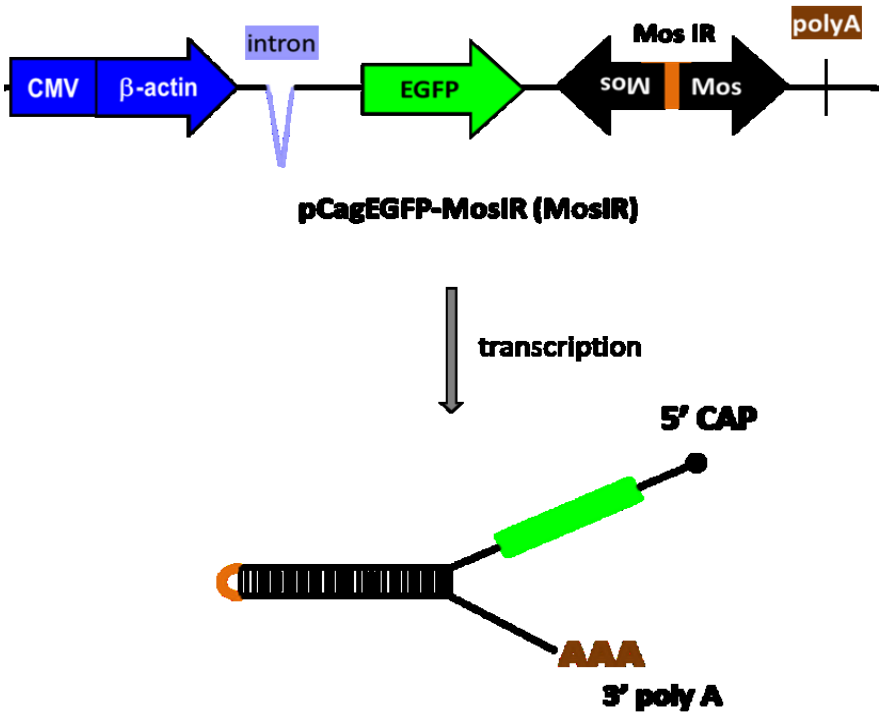


Figure 2: pCagEGFP-MosIR construct and its transcription into the hairpin-containing mRNA.

The pCagEGFP-MosIR construct was employed to generate transgenic mice ubiquitously expressing dsRNA (Supplement 2). In addition, transient or stable transfection of pCagEGFP-MosIR into several cell lines was used to analyze the effects of the expressed dsRNA in more detail (Supplement 2, 3, 4).

The ability of the construct to efficiently fold into dsRNA is supported by several lines of evidence. (1) MosIR expression downregulates *Mos* mRNA and depletes MOS protein in mouse oocytes (evidenced by low MAPK activity and parthenogenesis) (Stein et al, 2003). (2) Sequences mapping to MosIR contain A/G conversion and a peak in the size of 20-23 nt, indicating adenosine deamination and Dicer processing. (3) A fraction of MosIR-derived sequences is resistant to digestion by RNase T1 (enzyme cleaving ssRNA but not dsRNA) (Supplement 2).

3.2. Effects of long dsRNA expression in the mouse

The main part of the project studying long dsRNA expression in mammals investigated the phenotype of transgenic mice ubiquitously expressing Mos hairpin RNA. Previous studies concerning mammals have characterized only individual dsRNA-responding

pathways, while their mutual interactions were poorly understood (reviewed in the Supplement 5). Long dsRNA had been regarded as an efficient activator of the IFN response, which has been limiting its use as RNAi trigger to a few cell types where the IFN response is absent, namely the oocytes, blastomeres of early embryos, and undifferentiated embryonic stem cells. Additionally, some reports demonstrated the specific silencing in neuronal cells (summarized in Supplements 1 and 6). Besides, dsRNA has been a common tool for inducing RNAi in non-vertebrate models.

Mice expressing dsRNA in all cells were viable and exhibited no sign of activation of the IFN response (Supplement 2), which represents one of the key outcomes of this study. Accordingly, we provide the first model system demonstrating that ubiquitous dsRNA expression can be well tolerated in a living mammal.

Although the transgene was expressed in all tested tissues and MosIR efficiently formed dsRNA, its loading into dsRNA-activated pathways differed between somatic cells and oocytes. The response to MosIR in oocytes mimicked the phenotype of Mos knock-out mice (Gebauer & Richter, 1997; Hashimoto et al, 1994) and ZP3-driven MosIR transgene (Stein et al, 2003). Strictly speaking, there was a specific down-regulation of *Mos* mRNA (a 90 % decrease in average), accompanied by the lack of MOS protein (manifested as parthenogenesis and decreased fertility of females). Therefore, our results provide additional evidence that mouse oocytes possess efficient RNAi (Supplement 2).

In contrast, RNAi effects were negligible in somatic cells. To overcome that *Mos* is not expressed in somatic tissues, we produced a transgene carrying the *Mos* sequence in the 3'UTR of an EGFP reporter (Mos3, Fig. 3). The MosIR x Mos3 hybrids did not exhibit an RNAi effect on Mos3 reporter in any tested somatic tissue; on the other hand, a sequence-specific disruption of Mos3 sequence was clearly detectable in oocytes (Supplement 2). The difference in the RNAi efficiency between somatic and germ cells can be at least partly explained by the alternative and more processive isoform of Dicer present in oocytes (Flemr *et al*., submitted). Nevertheless, low but detectable processing long dsRNA substrates is also present in somatic cells, as documented by the cleavage of MosIR sequences into ~22 nt RNAs in the brain and kidney (Supplement 2), suggesting the presence of Dicer activity.

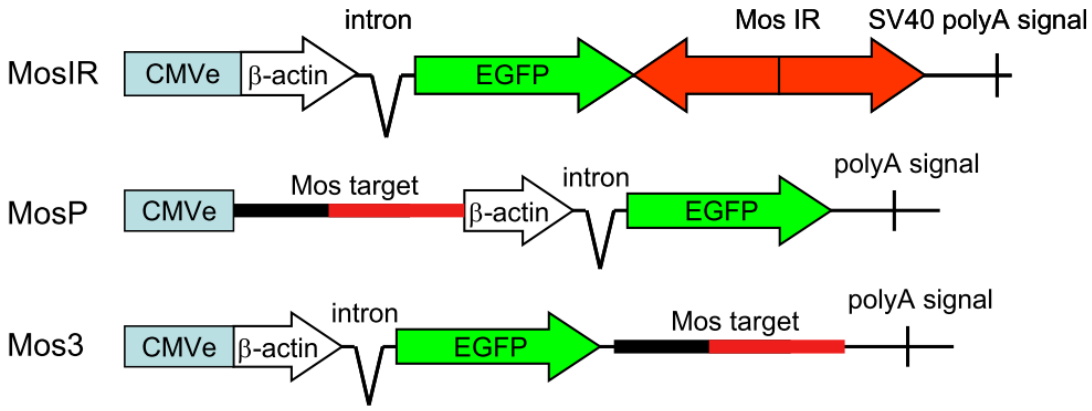


Figure 3: Scheme of the transgenes (adopted from the Supplement 2).

In addition, we examined whether MosIR expression can induce transcriptional silencing. For that, MosIR mice were crossed with transgenic mice in which Mos target sequence was inserted into the promoter (MosP, Fig. 3). Although we observed an occasional spontaneous silencing and DNA methylation of the Mos sequence in the promoter, the transcriptional silencing has never been observed in MosIR x MosP hybrids (Supplement 2). This argues against the functional transcriptional silencing of homologous sequences by RNAi in mammalian somatic cells and expands the previous data demonstrating that MosIR transcript does not induce DNA methylation in mouse oocytes (Svoboda et al, 2004).

In our dsRNA-expressing model system, we did not found any evidence for the activation of the IFN response. This notion is supported by normal appearance of MosIR mice and no increased expression of IFN-stimulated genes (Supplement 2). This is in agreement with analyses of IFN response activation in cultured cells (see section 3.3). The explanation for the lack of the IFN response in our system, which differs from a great deal of other studies (summarized in the Supplement 1), plausibly lies in the nature of dsRNA used in our model system. There are several types of dsRNA (differing in the length, structure, termini, *etc.*) that likely determine which response is activated. In addition, the dose, subcellular localization of dsRNA, its delivery into the cell, and the cell type used for the study creates another set of factors influencing the activation of the IFN response and other dsRNA-activated pathways (reviewed in the Supplement 1). Accordingly, our dsRNA expressed in the nucleus in the form of a capped, polyadenylated, and spliced mRNA seems not to be recognized as deleterious;

at the same time, it preserves the dsRNA structure necessary for its recognition by other dsRNA-activated pathways. The ability of cells to tolerate expressed dsRNA may have an important evolutionary role. As there is a widespread overlapping transcription in the genome (coming from repetitive sequences, activity of transposable elements, or regulatory antisense RNAs), it is vital to tightly regulate whether and which response to dsRNA will be triggered.

Last but not least, dsRNA derived from MosIR underwent low but clearly detectable adenosine deamination. We observed a modest A/G conversion in the region of predicted RNA duplex but not in ssRNA part of the MosIR transcript; similar results were obtained from the brain, kidney, and liver. The editing did not likely interfere with nuclear export and translation, as we detected A-to-I edited MosIR-derived RNA in both cytoplasmic fraction and polysome fraction from MosIR-derived mouse embryonic fibroblasts (Supplement 2). These results are consistent with cell culture data (section 3.3.). The absence of a significant loading in ADAR editing pathways suggests that the expressed dsRNA is not an optimal substrate for ADAR. As one of the proposed roles of ADARs is to counteract RNAi, low A-to-I editing of the long perfect dsRNA could reflect the inefficient Dicer processing in mammalian somatic cells.

3.3. Effects of long dsRNA expression in the cell culture

dsRNA expression using a transgenic model gives the opportunity to study the response to dsRNA in the whole organism, but it does not allow to manipulate experimental conditions. Therefore, we performed a series of experiments in cultured cells, which enabled us to use different cell types, species (a shift from the mouse to the human), amount of dsRNA-expressing plasmid, or different hairpin sequences. Furthermore, we could have employed different kinds of reporters or compare the effects of transiently transfected plasmids with plasmids integrated into the genome. In addition, the overexpression of the MosIR-expressing plasmid (pCagEGFP-MosIR) opened the possibility to study the effects of dsRNA-expression which were barely detectable with lower dose of dsRNA expressed in the transgenic mouse model.

First, the absence of the activation of the IFN response was confirmed by the microarray analysis of pCagEGFP-MosIR-transfected HeLa and HEK293 cell lines. Interestingly, there was a small number of genes with altered expression and no pattern of upregulation of IFN-stimulated genes (confirmed by real-time polymerase chain reaction analysis). Together with a minor mortality and low PKR-phosphorylation in pCagEGFP-MosIR-transfected cells, we assume that our dsRNA-inducing system does not elicit the IFN response (Supplement 2). These results extend the analysis of dsRNA response in the mouse and suggest that the IFN response is absent even when MosIR is expressed in higher doses.

Second, deep sequencing of pCagEGFP-MosIR-transfected cells yielded more sequences derived from MosIR and EGFP regions of the transgene compared to mouse tissues; thereby, we were able to analyze siRNA production quantitatively. We detected hundreds of putative siRNAs mapping to different parts of MosIR sequence, further supporting the processing of dsRNAs by Dicer in somatic cells. As siRNAs did not cluster to the ends of long dsRNA, we assume that Dicer has no preferential cleavage sites (Supplement 2 and 4). However, the amount of MosIR-derived sequences was approximately hundred-times lower compared to reads mapping to abundant miRNAs (Supplement 2) suggesting that, at least in mammalian somatic cells, Dicer likely plays a more important role in miRNA pathway than in RNAi pathway.

In addition, we observed an abundant editing of MosIR-derived RNAs in transfected HEK293 cells. Remarkably, we also detected numerous putative siRNAs with A/G conversion, indicating that both ADAR and Dicer can process the same substrate (Supplement 2 and 4). However, it remains to investigate, whether there is a required order in which these enzymes process dsRNA.

3.4. Effects of dsRNA-expressing plasmids on co-transfected plasmids

The sequence-specific and sequence-independent effects of dsRNA expression can be easily monitored using co-transfected reporter plasmids, such as luciferase-expressing plasmids. Originally, *Renilla* luciferase (RL) reporter carried a Mos target sequence to study sequence-specific effects and firefly luciferase (FL) reporter (without the homology to pCagEGFP-MosIR plasmid) was used to monitor sequence-independent

effects. Since there was no sequence-specific effect in transfected cells, in later experiments, we used RL and FL reporters without the Mos target sequence (Supplement 2 and 4).

A transient co-transfection of an increasing amount of dsRNA-expressing plasmid together with a constant amount of luciferase-expressing plasmids revealed a concentration-dependent decrease of luciferase activities. This inhibition was independent of a cell type, transfection method, or sequence of dsRNA-expressing plasmid (tested by three different hairpin sequences). Additionally, the effect was independent of luciferase reporters and luminometric assay (tested by a fluorescent reporter). On the other hand, the effect required the presence of an IR and its transcription. Further experiments showed that the inhibition of co-transfected reporters occurs post-transcriptionally, plausibly at the level of translation initiation, and it involves PKR activity (Supplement 4).

In contrast, dsRNA-expressing plasmids did not affect endogenous genes. This statement is supported by the absence of reporter inhibition in stable lines with integrated luciferase reporters as well as no decrease in cellular proteosynthesis (Supplement 2 and 4). In addition, these data are consistent with our observation that dsRNA expression does not block proteosynthesis in MosIR transgenic mice (Supplement 2).

The selective translation inhibition of transiently co-transfected plasmids but not genome-integrated genes mediated by PKR indicates dsRNA can induce two types of PKR responses. (1) the classical IFN response, inducing the global repression of translation and activation of IFN-stimulated genes and (2) a partial response, marked by selective translation repression of specific “unfavorable” transcripts.

3.5. Transcription of transiently transfected plasmids

The analysis of RNA transcriptome of transfected plasmids by massive parallel sequencing revealed that RNA is produced from the entire plasmid, including the plasmid backbone, antisense strand of the coding region, and the intron (Supplement 4). Complex transcription was demonstrated for several different plasmids: pCagEGFP-MosIR, pBluescript (a small cloning plasmid), pEGFP-C1 (a plasmid

for protein expression), phRL-SV40 and pGL4-SV40 (RL- and FL-expressing plasmids, respectively) (Supplement 3 and 4).

Although the annotated eukaryotic transcription units yielded the highest read densities, significant amounts of RNA were generated from regions that are not expected to be transcribed in mammalian cells. The lowest level of the spurious transcription was present in pGL4-SV40 plasmid, which was modified to reduce putative transcription factor-binding sites (Paguio et al, 2005). The same sequences in different plasmids (such as bacterial origin of replication or antibiotic resistance) exhibited similar expression patterns. The transcripts typically generated clusters in either sense or antisense direction with a minimal overlap, except the kanamycin/neomycin (kan/neo) resistance cassette in the pEGFP-C1 plasmid (Supplement 3). Accordingly, we demonstrate that plasmid expression is not restricted to the annotated transcription units, but there are many cryptic promoters responsible for the complex and undesirable transcription.

The spurious transcription can affect results of transient co-transfection experiments as demonstrated by luciferase activities in cells co-transfected with different plasmids. Most plasmids inhibited luciferase activities, with the most prominent effect of pEGFP-C1 and its derivatives. Further experiments revealed that the expression of the kan/neo resistance cassette in pEGFP-C1 is likely responsible for the inhibitory effect (Supplement 3). We propose here the mechanism of the reporter inhibition by the formation of dsRNA from complementary sense and antisense transcript in the kan/neo resistance region. However, the widespread reporter inhibition resulting from the co-transfection with many commonly used plasmids indicates that plasmids significantly influence each other's expression (Supplement 3). The presence of dsRNA originating from plasmid transcription is an important but probably not the only culprit of this non-specific effect. Although the adverse effects of co-transfection are usually neither tested nor studied in detail, during the last decades, it has been described that a combination of some chemicals, type of promoter, transcription factors binding sites, cell line and many other factors may affect reporter expression (Ghazawi et al, 2005; Chang & Stoltzfus, 1985; Montgomery et al, 1998; Osborne & Tonissen, 2002; Shifera & Hardin, 2010; Schule et al, 1988; Sorscher & Cordeiro-Stone, 1994; Vopalensky et al, 2008).

Interestingly, the kan/neo region of pEGFP-C1 plasmid exhibited several characteristics similar to MosIR region of pCagEGFP-MosIR plasmid, namely the concentration-dependent inhibition of reporter plasmid expression, convergent transcription, clearly detectable A/G conversions, and the peak of small RNAs (~22 nt), including reads that were processed to small RNAs and edited at the same time (Supplement 2 and 3). All these features support the hypothesis that the kan/neo resistance cassette generates dsRNA. Accordingly, both pEGFP-C1 and MosIR-expressing plasmids may share the same mechanism of the reporter inhibition indicating that studies of dsRNA expression have broad implications.

As transient co-transfection represents a basic molecular biology method, a detailed understanding of the fate of plasmids in cells is crucial for a wide spectrum of scientists. The analysis of plasmid transcription provided in this thesis does not offer a complete guideline how to avoid false readout in transfection experiments. Rather, our results serve as a warning highlighting the need to include all kinds of controls and to carefully choose an extra DNA that is used to equalize the total amount of transfected DNA.

4. Conclusions

Long dsRNA triggers RNAi in plants and invertebrates, while it is a potent inducer of the antiviral IFN response in vertebrates. However, no integrated study of the effects of the expressed dsRNA in mammals has been conducted, yet. Most reports focus either on sequence-independent inhibitory effects of dsRNA or they utilize dsRNA for RNAi studies in cell types where the sequence-specific RNAi is functional. Moreover, the majority of studies employ *in vitro*-transcribed dsRNA that is more immunostimulatory than expressed dsRNA used in this project. Therefore, this work represents a comprehensive study of dsRNA effects in mammals, demonstrating that dsRNA (expressed either in transgenic mice or in transfected cultured cells) can be well tolerated without provoking the IFN response. At the same time, expressed dsRNA retains features necessary for the processing into siRNAs by Dicer or editing by ADARs. Accordingly, it seems that the limiting factor for RNAi in mammalian somatic cells is inneficient Dicer processing.

In addition, this work demonstrates that transcription of transfected plasmids is not restricted to defined transcription units, but it involves all plasmid regions. This spurious convergent transcription can lead to dsRNA formation and can suppress translation of co-transfected reporter plasmids, leading to misinterpretation of results. Many commonly used plasmids have a potential to generate dsRNA; however, dsRNA formation cannot be efficiently predicted. As transient transfection is a widely used method, our report documenting the inhibition of co-transfected plasmids indicates that further studies of dsRNA are highly desirable.

The follow-up research needs to address which factors are responsible for different RNAi efficiency between somatic cells and oocytes and explore in more detail how cells distinguish between endogenous and exogenous DNA and RNA.

5. References

Alexopoulou L, Holt AC, Medzhitov R, Flavell RA (2001) Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. *Nature* **413:** 732-738

Ank N, West H, Bartholdy C, Eriksson K, Thomsen AR, Paludan SR (2006) Lambda interferon (IFN-lambda), a type III IFN, is induced by viruses and IFNs and displays potent antiviral activity against select virus infections in vivo. *Journal of virology* **80:** 4501-4509

Aravind L, Watanabe H, Lipman DJ, Koonin EV (2000) Lineage-specific loss and divergence of functionally linked genes in eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America* **97:** 11319-11324

Athanasiadis A, Rich A, Maas S (2004) Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. *PLoS biology* **2:** e391

Ben-Asouli Y, Banai Y, Pel-Or Y, Shir A, Kaempfer R (2002) Human interferon-gamma mRNA autoregulates its translation through a pseudoknot that activates the interferon-inducible protein kinase PKR. *Cell* **108:** 221-232

Bernstein E, Kim SY, Carmell MA, Murchison EP, Alcorn H, Li MZ, Mills AA, Elledge SJ, Anderson KV, Hannon GJ (2003) Dicer is essential for mouse development. *Nature genetics* **35:** 215-217

Beutler B (2004) Inferences, questions and possibilities in Toll-like receptor signalling. *Nature* **430:** 257-263

Carthew RW, Sontheimer EJ (2009) Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136:** 642-655

Cullen BR (2006) Is RNA interference involved in intrinsic antiviral immunity in mammals? *Nature immunology* **7:** 563-567

DeCerbo J, Carmichael GG (2005) Retention and repression: fates of hyperedited RNAs in the nucleus. *Current opinion in cell biology* **17:** 302-308

Doench JG, Petersen CP, Sharp PA (2003) siRNAs can function as miRNAs. *Genes & development* **17:** 438-442

Donze O, Deng J, Curran J, Sladek R, Picard D, Sonenberg N (2004) The protein kinase PKR: a molecular clock that sequentially activates survival and death programs. *The EMBO journal* **23:** 564-571

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391:** 806-811

Friedman RC, Farh KK, Burge CB, Bartel DP (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome research* **19:** 92-105

Gantier MP, Williams BR (2007) The response of mammalian cells to double-stranded RNA. *Cytokine & growth factor reviews* **18:** 363-371

Garcia MA, Gil J, Ventoso I, Guerra S, Domingo E, Rivas C, Esteban M (2006) Impact of protein kinase PKR in cell biology: from antiviral to antiproliferative action. *Microbiology and molecular biology reviews : MMBR* **70:** 1032-1060

Garcia MA, Meurs EF, Esteban M (2007) The dsRNA protein kinase PKR: virus and cell control. *Biochimie* **89:** 799-811

Gebauer F, Richter JD (1997) Synthesis and function of Mos: the control switch of vertebrate oocyte meiosis. *BioEssays : news and reviews in molecular, cellular and developmental biology* **19:** 23-28

Ghazawi I, Cutler SJ, Low P, Mellick AS, Ralph SJ (2005) Inhibitory effects associated with use of modified Photinus pyralis and Renilla reniformis luciferase vectors in dual reporter assays and implications for analysis of ISGs. *Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research* **25:** 92-102

Ghildiyal M, Zamore PD (2009) Small silencing RNAs: an expanding universe. *Nature reviews Genetics* **10:** 94-108

Hartig JV, Esslinger S, Bottcher R, Saito K, Forstemann K (2009) Endo-siRNAs depend on a new isoform of loquacious and target artificially introduced, high-copy sequences. *The EMBO journal* **28:** 2932-2944

Hashimoto N, Watanabe N, Furuta Y, Tamemoto H, Sagata N, Yokoyama M, Okazaki K, Nagayoshi M, Takeda N, Ikawa Y, et al. (1994) Parthenogenetic activation of oocytes in c-mos-deficient mice. *Nature* **370:** 68-71

He Y, Vogelstein B, Velculescu VE, Papadopoulos N, Kinzler KW (2008) The antisense transcriptomes of human cells. *Science* **322:** 1855-1857

Heikinheimo O, Lanzendorf SE, Baka SG, Gibbons WE (1995) Cell cycle genes c-mos and cyclin-B1 are expressed in a specific pattern in human oocytes and preimplantation embryos. *Hum Reprod* **10:** 699-707

Higuchi M, Maas S, Single FN, Hartner J, Rozov A, Burnashev N, Feldmeyer D, Sprengel R, Seeburg PH (2000) Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. *Nature* **406:** 78-81

Higuchi M, Single FN, Kohler M, Sommer B, Sprengel R, Seeburg PH (1993) RNA editing of AMPA receptor subunit GluR-B: a base-paired intron-exon structure determines position and efficiency. *Cell* **75:** 1361-1370

Hong J, Qian Z, Shen S, Min T, Tan C, Xu J, Zhao Y, Huang W (2005) High doses of siRNAs induce eri-1 and adar-1 gene expression and reduce the efficiency of RNA interference in the mouse. *The Biochemical journal* **390:** 675-679

Hornung V, Ellegast J, Kim S, Brzozka K, Jung A, Kato H, Poeck H, Akira S, Conzelmann KK, Schlee M, Endres S, Hartmann G (2006) 5'-Triphosphate RNA is the ligand for RIG-I. *Science* **314:** 994-997

Hornung V, Guenthner-Biller M, Bourquin C, Ablasser A, Schlee M, Uematsu S, Noronha A, Manoharan M, Akira S, de Fougerolles A, Endres S, Hartmann G (2005) Sequence-specific potent induction of IFN-alpha by short interfering RNA in plasmacytoid dendritic cells through TLR7. *Nature medicine* **11:** 263-270

Hundley HA, Bass BL (2010) ADAR editing in double-stranded UTRs and other noncoding RNA sequences. *Trends in biochemical sciences* **35:** 377-383

Hundley HA, Krauchuk AA, Bass BL (2008) C. elegans and H. sapiens mRNAs with edited 3' UTRs are present on polysomes. *RNA* **14:** 2050-2060

Chalupnikova K, Nejepinska J, Svoboda P (2013) Production and application of long dsRNA in mammalian cells. *Methods Mol Biol* **942:** 291-314

Chang LJ, Stoltzfus CM (1985) Gene expression from both intronless and intron-containing Rous sarcoma virus clones is specifically inhibited by anti-sense RNA. *Molecular and cellular biology* **5:** 2341-2348

Jayan GC, Casey JL (2002) Increased RNA editing and inhibition of hepatitis delta virus replication by high-level expression of ADAR1 and ADAR2. *Journal of virology* **76:** 3819-3827

Jeffrey IW, Kadereit S, Meurs EF, Metzger T, Bachmann M, Schwemmle M, Hovanessian AG, Clemens MJ (1995) Nuclear localization of the interferon-inducible protein kinase PKR in human cells and transfected mouse cells. *Experimental cell research* **218:** 17-27

Jepson JE, Reenan RA (2008) RNA editing in regulating gene expression in the brain. *Biochimica et biophysica acta* **1779:** 459-470

Jin Y, Zhang W, Li Q (2009) Origins and evolution of ADAR-mediated RNA editing. *IUBMB life* **61:** 572-578

Jinek M, Doudna JA (2009) A three-dimensional view of the molecular machinery of RNA interference. *Nature* **457:** 405-412

Judge AD, Sood V, Shaw JR, Fang D, McClintock K, MacLachlan I (2005) Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA. *Nature biotechnology* **23:** 457-462

Kaname T, Huxley C (2001) Simple and efficient vectors for retrofitting BACs and PACs with mammalian neoR and EGFP marker genes. *Gene* **266:** 147-153

Kato H, Takeuchi O, Mikamo-Satoh E, Hirai R, Kawai T, Matsushita K, Hiiragi A, Dermody TS, Fujita T, Akira S (2008) Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-I and melanoma differentiation-associated gene 5. *The Journal of experimental medicine* **205:** 1601-1610

Kaufman RJ, Davies MV, Pathak VK, Hershey JW (1989) The phosphorylation state of eucaryotic initiation factor 2 alters translational efficiency of specific mRNAs. *Molecular and cellular biology* **9:** 946-958

Keshet E, Rosenberg MP, Mercer JA, Propst F, Vande Woude GF, Jenkins NA, Copeland NG (1988) Developmental regulation of ovarian-specific Mos expression. *Oncogene* **2:** 235-240

Ketting RF (2011) The many faces of RNAi. *Developmental cell* **20:** 148-161

Kim VN, Han J, Siomi MC (2009) Biogenesis of small RNAs in animals. *Nature reviews Molecular cell biology* **10:** 126-139

Kumar M, Carmichael GG (1998) Antisense RNA: function and fate of duplex RNA in cells of higher eukaryotes. *Microbiology and molecular biology reviews : MMBR* **62:** 1415-1434

Levanon EY, Eisenberg E, Yelin R, Nemzer S, Hallegger M, Shemesh R, Fligelman ZY, Shoshan A, Pollock SR, Sztybel D, Olshansky M, Rechavi G, Jantsch MF (2004) Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nature biotechnology* **22:** 1001-1005

Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song JJ, Hammond SM, Joshua-Tor L, Hannon GJ (2004) Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305:** 1437-1441

Luo D, Ding SC, Vela A, Kohlway A, Lindenbach BD, Pyle AM (2011) Structural insights into RNA recognition by RIG-I. *Cell* **147:** 409-422

Ma J, Flemr M, Stein P, Berninger P, Malik R, Zavolan M, Svoboda P, Schultz RM (2010) MicroRNA activity is suppressed in mouse oocytes. *Current biology : CB* **20:** 265-270

Manche L, Green SR, Schmedt C, Mathews MB (1992) Interactions between double-stranded RNA regulators and the protein kinase DAI. *Molecular and cellular biology* **12:** 5238-5248

Marques JT, Carthew RW (2007) A call to arms: coevolution of animal viruses and host innate immune responses. *Trends in genetics : TIG* **23:** 359-364

Marques JT, Devosse T, Wang D, Zamanian-Daryoush M, Serbinowski P, Hartmann R, Fujita T, Behlke MA, Williams BR (2006) A structural basis for discriminating between self and nonself double-stranded RNAs in mammalian cells. *Nature biotechnology* **24:** 559-565

Minks MA, West DK, Benvin S, Baglioni C (1979) Structural requirements of double-stranded RNA for the activation of 2',5'-oligo(A) polymerase and protein kinase of interferon-treated HeLa cells. *The Journal of biological chemistry* **254:** 10180-10183

Montgomery MK, Xu S, Fire A (1998) RNA as a target of double-stranded RNA-mediated genetic interference in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America* **95:** 15502-15507

Nallagatla SR, Hwang J, Toroney R, Zheng X, Cameron CE, Bevilacqua PC (2007) 5'-triphosphate-dependent activation of PKR by RNAs with short stem-loops. *Science* **318:** 1455-1458

Nanduri S, Carpick BW, Yang Y, Williams BR, Qin J (1998) Structure of the double-stranded RNA-binding domain of the protein kinase PKR reveals the molecular basis of its dsRNA-mediated activation. *The EMBO journal* **17:** 5458-5465

Nejepinska J, Flemr M, Svoboda P. (2012a) The Canonical RNA Interference Pathway in Animals. In Mallick B, Ghosh Z (eds.), *Regulatory RNAs: Basics, Methods and Applications*. Springer.

Nejepinska J, Malik R, Filkowski J, Flemr M, Filipowicz W, Svoboda P (2012b) dsRNA expression in the mouse elicits RNAi in oocytes and low adenosine deamination in somatic cells. *Nucleic acids research* **40:** 399-413

Nie Y, Hammond GL, Yang JH (2007) Double-stranded RNA deaminase ADAR1 increases host susceptibility to virus infection. *Journal of virology* **81:** 917-923

Nishikura K (2010) Functions and regulation of RNA editing by ADAR deaminases. *Annual review of biochemistry* **79:** 321-349

Nishikura K, Yoo C, Kim U, Murray JM, Estes PA, Cash FE, Liebhaber SA (1991) Substrate specificity of the dsRNA unwinding/modifying activity. *The EMBO journal* **10:** 3523-3532

Obbard DJ, Gordon KH, Buck AH, Jiggins FM (2009) The evolution of RNAi as a defence against viruses and transposable elements. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* **364:** 99-115

Obbard DJ, Jiggins FM, Halligan DL, Little TJ (2006) Natural selection drives extremely rapid evolution in antiviral RNAi genes. *Current biology : CB* **16:** 580-585

Osborne SA, Tonissen KF (2002) pRL-TK induction can cause misinterpretation of gene promoter activity. *BioTechniques* **33:** 1240-1242

Paguio A, Almond B, Fan F, Stecha P, Garwin D, Wood M, Wood K (2005) pGL4 vectors: A new generation of luciferase reporter vectors. *Promega Notes* **89:** 7-10

Park H, Davies MV, Langland JO, Chang HW, Nam YS, Tartaglia J, Paoletti E, Jacobs BL, Kaufman RJ, Venkatesan S (1994) TAR RNA-binding protein is an inhibitor of the interferon-induced protein kinase PKR. *Proceedings of the National Academy of Sciences of the United States of America* **91:** 4713-4717

Patel RC, Sen GC (1998) PACT, a protein activator of the interferon-induced protein kinase, PKR. *The EMBO journal* **17:** 4379-4390

Patterson JB, Samuel CE (1995) Expression and regulation by interferon of a double-stranded-RNA-specific adenosine deaminase from human cells: evidence for two forms of the deaminase. *Molecular and cellular biology* **15:** 5376-5388

Player MR, Torrence PF (1998) The 2-5A system: modulation of viral and cellular processes through acceleration of RNA degradation. *Pharmacology & therapeutics* **78:** 55-113

Prasanth KV, Prasanth SG, Xuan Z, Hearn S, Freier SM, Bennett CF, Zhang MQ, Spector DL (2005) Regulating gene expression through RNA nuclear retention. *Cell* **123:** 249-263

Ryter JM, Schultz SC (1998) Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *The EMBO journal* **17:** 7505-7513

Sadler AJ, Williams BR (2008) Interferon-inducible antiviral effectors. *Nature reviews Immunology* **8:** 559-568

Samuel CE (2001) Antiviral actions of interferons. *Clinical microbiology reviews* **14:** 778-809, table of contents

Samuel CE (2011) Adenosine deaminases acting on RNA (ADARs) are both antiviral and proviral. *Virology* **411:** 180-193

Satoh T, Kato H, Kumagai Y, Yoneyama M, Sato S, Matsushita K, Tsujimura T, Fujita T, Akira S, Takeuchi O (2010) LGP2 is a positive regulator of RIG-I- and MDA5-mediated antiviral responses. *Proceedings of the National Academy of Sciences of the United States of America* **107:** 1512-1517

Saunders LR, Barber GN (2003) The dsRNA binding protein family: critical roles, diverse cellular functions. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **17:** 961-983

Scadden AD (2005) The RISC subunit Tudor-SN binds to hyper-edited double-stranded RNA and promotes its cleavage. *Nature structural & molecular biology* **12:** 489-496

Shifera AS, Hardin JA (2010) Factors modulating expression of Renilla luciferase from control plasmids used in luciferase reporter gene assays. *Analytical biochemistry* **396:** 167-172

Schlee M, Hornung V, Hartmann G (2006) siRNA and isRNA: two edges of one sword. *Molecular therapy : the journal of the American Society of Gene Therapy* **14:** 463-470

Schule R, Muller M, Kaltschmidt C, Renkawitz R (1988) Many transcription factors interact synergistically with steroid receptors. *Science* **242:** 1418-1420

Silverman RH (2007a) A scientific journey through the 2-5A/RNase L system. *Cytokine & growth factor reviews* **18:** 381-388

Silverman RH (2007b) Viral encounters with 2',5'-oligoadenylate synthetase and RNase L during the interferon antiviral response. *Journal of virology* **81:** 12720-12729

Singh B, Arlinghaus RB (1997) Mos and the cell cycle. *Progress in cell cycle research* **3:** 251-259

Sioud M (2005) Induction of inflammatory cytokines and interferon responses by double-stranded and single-stranded siRNAs is sequence-dependent and requires endosomal localization. *J Mol Biol* **348:** 1079-1090

Sledz CA, Holko M, de Veer MJ, Silverman RH, Williams BR (2003) Activation of the interferon system by short-interfering RNAs. *Nature cell biology* **5:** 834-839

Sorscher DH, Cordeiro-Stone M (1994) Inhibition of reporter gene expression in mammalian cells. Effects of distinct carcinogen lesions in DNA. *Carcinogenesis* **15:** 1093-1096

St Johnston D, Brown NH, Gall JG, Jantsch M (1992) A conserved double-stranded RNA-binding domain. *Proceedings of the National Academy of Sciences of the United States of America* **89:** 10979-10983

Stefl R, Oberstrass FC, Hood JL, Jourdan M, Zimmermann M, Skrisovska L, Maris C, Peng L, Hofr C, Emeson RB, Allain FH (2010) The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. *Cell* **143:** 225-237

Stein P, Svoboda P, Schultz RM (2003) Transgenic RNAi in mouse oocytes: a simple and fast approach to study gene function. *Developmental biology* **256:** 187-193

Stein P, Zeng F, Pan H, Schultz RM (2005) Absence of non-specific effects of RNA interference triggered by long double-stranded RNA in mouse oocytes. *Developmental biology* **286:** 464-471

Sudhakar A, Ramachandran A, Ghosh S, Hasnain SE, Kaufman RJ, Ramaiah KV (2000) Phosphorylation of serine 51 in initiation factor 2 alpha (eIF2 alpha) promotes complex formation between eIF2 alpha(P) and eIF2B and causes inhibition in the guanine nucleotide exchange activity of eIF2B. *Biochemistry* **39:** 12929-12938

Suh N, Baehner L, Moltzahn F, Melton C, Shenoy A, Chen J, Blelloch R (2010) MicroRNA function is globally suppressed in mouse oocytes and early embryos. *Current biology : CB* **20:** 271-277

Svoboda P, Stein P, Filipowicz W, Schultz RM (2004) Lack of homologous sequence-specific DNA methylation in response to stable dsRNA expression in mouse oocytes. *Nucleic acids research* **32:** 3601-3606

Svoboda P, Stein P, Hayashi H, Schultz RM (2000) Selective reduction of dormant maternal mRNAs in mouse oocytes by RNA interference. *Development* **127:** 4147-4156

Svoboda P, Stein P, Schultz RM (2001) RNAi in mouse oocytes and preimplantation embryos: effectiveness of hairpin dsRNA. *Biochemical and biophysical research communications* **287:** 1099-1104

Ullu E, Tschudi C, Chakraborty T (2004) RNA interference in protozoan parasites. *Cellular microbiology* **6:** 509-519

Umbach JL, Cullen BR (2009) The role of RNAi and microRNAs in animal virus replication and antiviral immunity. *Genes & development* **23:** 1151-1164

van Rij RP, Andino R (2006) The silent treatment: RNAi as a defense against virus infection in mammals. *Trends in biotechnology* **24:** 186-193

Vitali P, Scadden AD (2010) Double-stranded RNAs containing multiple IU pairs are sufficient to suppress interferon induction and apoptosis. *Nature structural & molecular biology* **17:** 1043-1050

Vopalensky V, Masek T, Horvath O, Vicenova B, Mokrejs M, Pospisek M (2008) Firefly luciferase gene contains a cryptic promoter. *RNA* **14:** 1720-1729

Wang Q, Khillan J, Gadue P, Nishikura K (2000) Requirement of the RNA editing deaminase ADAR1 gene for embryonic erythropoiesis. *Science* **290:** 1765-1768

Wianny F, Zernicka-Goetz M (2000) Specific interference with gene function by double-stranded RNA in early mouse development. *Nature cell biology* **2:** 70-75

Yang S, Tutton S, Pierce E, Yoon K (2001) Specific double-stranded RNA interference in undifferentiated mouse embryonic stem cells. *Mol Cell Biol* **21:** 7807-7816

Yekta S, Shih IH, Bartel DP (2004) MicroRNA-directed cleavage of HOXB8 mRNA. *Science* **304:** 594-596

Yelin R, Dahary D, Sorek R, Levanon EY, Goldstein O, Shoshan A, Diber A, Biton S, Tamir Y, Khosravi R, Nemzer S, Pinner E, Walach S, Bernstein J, Savitsky K, Rotman G (2003) Widespread occurrence of antisense transcription in the human genome. *Nature biotechnology* **21:** 379-386

Zhang Z, Carmichael GG (2001) The fate of dsRNA in the nucleus: a p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell* **106:** 465-475

Zheng X, Bevilacqua PC (2004) Activation of the protein kinase PKR by short double-stranded RNAs with single-stranded tails. *RNA* **10:** 1934-1945

Supplementary material

Supplement 1

Chalupnikova K, Nejepinska J, Svoboda P.

Production and application of long dsRNA in mammalian cells

Methods in Molecular Biology

2013

942:291-314

Chapter 16

Production and Application of Long dsRNA in Mammalian Cells

Katerina Chalupnikova, Jana Nejepinska, and Petr Svoboda

Abstract

Double-stranded RNA (dsRNA) is involved in different biological processes. At least three different pathways can respond to dsRNA in mammals. One of these pathways is RNA interference (RNAi) where long dsRNA induces sequence-specific degradation of transcripts carrying sequences complementary to dsRNA. Long dsRNA is also a potent trigger of the interferon pathway, a sequence-independent response that leads to global suppression of translation and global RNA degradation. In addition, dsRNA can be edited by adenosine deamination, which may result in nuclear retention and degradation of dsRNA or in alteration of RNA coding potential. Here, we provide a technical review summarizing different strategies of long dsRNA usage. While the review is largely focused on long dsRNA-induced RNAi in mammalian cells, it also provides helpful information on both the in vitro production and in vivo expression of dsRNAs. We present an overview of currently available vectors for dsRNA expression and provide the latest update on oocyte-specific transgenic RNAi approaches.

Key words: dsRNA (double-stranded RNA), RNAi (RNA interference), Transgenic RNAi, IFN response, OAS (2′5′-oligoadenylate synthetase), siRNA (short interfering RNA), shRNA (short hairpin RNA), Mammalian somatic cells, Oocytes and embryos, IVT (in vitro transcription)

1. Introduction

Long dsRNA is a unique structure that plays various roles in different organisms, from antiviral defense to regulation of gene expression. Three pathways activated by long dsRNA are well characterized (Fig. 1): sequence-specific gene silencing by RNA interference (RNAi), sequence-independent silencing associated with the interferon (IFN) response, and RNA editing which results either in alteration of the substrate specificity (hypoediting) of short interfering RNAs (siRNAs) or in nuclear retention of dsRNA making it

Katerina Chalupnikova and Jana Nejepinska contributed equally to this article.

Debra J. Taxman (ed.), *siRNA Design: Methods and Protocols*, Methods in Molecular Biology, vol. 942, DOI 10.1007/978-1-62703-119-6_16,

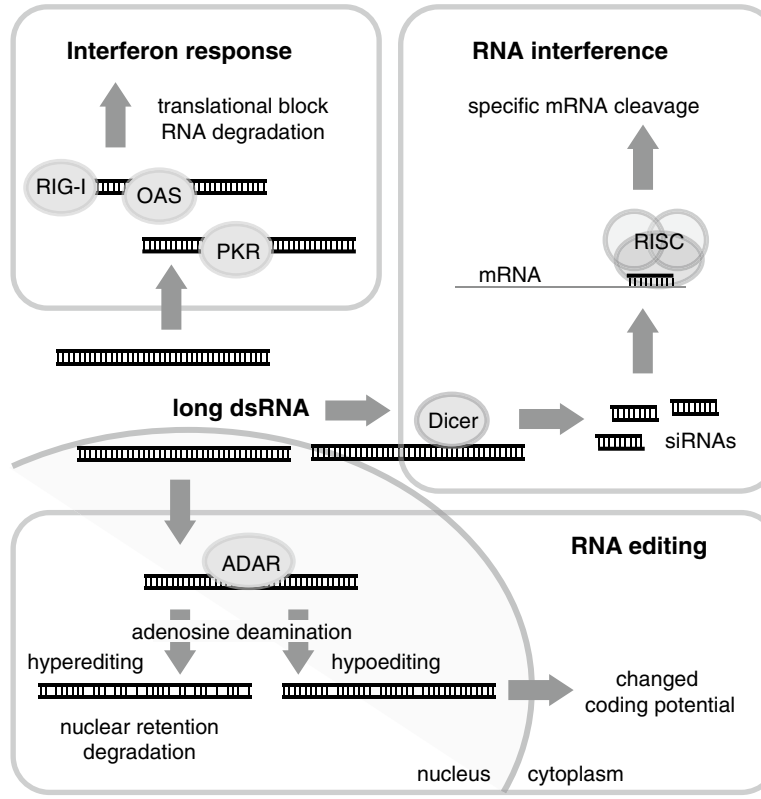


Fig. 1. Pathways activated by long dsRNA. In the RNAi pathway (*upper right frame*), long dsRNA is cleaved by Dicer into siRNAs. One of the strands is incorporated into the RNA-induced silencing complex (RISC) and acts as a guide for binding and cleavage of complementary mRNAs. In the interferon response (*upper left frame*), long dsRNA is recognized by dsRNA-binding proteins (PKR, OAS, RIG-I) and initiates global inhibition of proteosynthesis and RNA degradation. Long dsRNA can also undergo RNA editing (*bottom frame*) by ADARs. Hyperedited dsRNA is degraded in the nucleus, hypoedited RNA is exported into the cytoplasm and can cause alterations during translation.

unavailable for the cytoplasmic silencing machinery (hyperediting). Although all pathways share a common substrate-long dsRNA, they differ in their evolutionary conservation, cellular localization, and outcome. Therefore, the effect of specific long dsRNA depends on the combination of different factors.

The most conserved pathway is RNAi, which is present in almost all eukaryotes. Cytoplasmic long dsRNA is processed by the RNase-III enzyme Dicer into 21-nucleotide (nt)-long siRNAs with two nt 3′ overhangs (reviewed in ref. 1). One of the siRNA strands is incorporated into an RNA-induced silencing complex (RISC) and acts as a guide for binding to the complementary messenger RNA (mRNA). The cognate mRNA is subsequently cleaved in the middle of the basepairing region by a “slicer” (Ago2 protein) and rapidly degraded (reviewed in ref. 2).

Since RNAi discovery (3), dsRNA has been used as an experimental tool to selectively silence gene expression. Specific mRNA degradation mediated by an uptake of exogenous dsRNA is a common way of suppressing gene expression in nematodes (*Caenorhabditis elegans*) (3) and insects (*Drosophila*) (4–7), where RNAi serves as a form of antiviral response (reviewed in ref. 8). However, vertebrates evolved an independent complex response to viruses that includes global inhibition of proteosynthesis, expression of antiviral genes, RNA degradation, and apoptosis. This complex of pathways is known as the "IFN response" and it is activated by long dsRNA (which often serves as a marker of viral infection) (reviewed in ref. 9).

Cytoplasmic dsRNA-binding proteins, such as protein kinase R (PKR) and 2′5′-oligoadenylate synthetase (OAS), are the backbone of the IFN response. Activated PKR phosphorylates the alpha subunit of eukaryotic initiation factor 2 (eIF2α), resulting in a global inhibition of translation (10). The second part of the IFN response is represented by OAS/RNase L system. Upon binding of dsRNA to OAS, this enzyme is activated and converts adenosine triphosphate (ATP) to pyrophosphate and 2′5′-linked oligoadenylates (11). Oligoadenylates activate endoribonuclease RNase L, which nonspecifically degrades viral and cellular RNA (including mRNAs); thereby resulting in decreased proteosynthesis (12). Moreover, there are additional players involved in the activation or regulation of the IFN response (reviewed in ref. 9). The circumstances under which these molecules are activated depend on various factors, such as cell type, dsRNA origin, localization, length, structure, or termini (9, 13–18). This sophisticated system is essential to discriminate foreign dsRNAs from endogenous RNAs with intra or intermolecular complementarity in order to ignore harmless dsRNAs or initiate a deadly IFN response. The antiviral role of RNAi has been replaced by the IFN response in most but not all vertebrate cell types. In oocytes, where the IFN response is not functional (14), RNAi plays an important role by targeting repetitive elements and endogenous genes (reviewed in ref. 19).

In addition to RNAi and the IFN response, dsRNA can be modified by adenosine deaminases acting on RNA (ADARs), nuclear and cytoplasmic enzymes found in Metazoa (20, 21). The conversion of adenosines into inosines in dsRNA substrates can have a number of consequences. Hyperediting of dsRNA results in nuclear retention of the dsRNA and plays an important role in the defense against viruses producing dsRNA (22). In contrast, dsRNAs with a low level of editing are exported into the cytoplasm but modifications in individual codons may affect splicing or cause alterations in translation. Specific modifications by ADARs increase diversity in protein sequences, which is essential for substrate specificity of neuronal channels (reviewed in ref. 23).

In organisms where more dsRNA-induced pathways are functional, individual pathways compete for the common substrate and interfere with each other. In the case of crosstalk between RNA editing and RNAi, ADAR can modify a dsRNA sequence, inducing a change in siRNA specificity and subsequent downregulation of different targets. Alternatively, edited siRNA is not perfectly complementary to its original target, which elicits a translational repression rather than a cleavage of the target mRNA. In the case of hyperediting, long dsRNA acquires excessive I–U pairs that make the duplex unstable; subsequently, edited dsRNA is degraded in the nucleus and unavailable for cytoplasmic RNAi machinery (24).

Opposing roles are played by RNAi and the IFN response. Although both pathways probably evolved as host defense mechanisms, RNAi induces silencing only of the sequences homologous to the original RNAi trigger, while the IFN response induces global RNA degradation and inhibition of translation. The interactions between these two pathways are still not well understood. In fact, while some groups described RNAi effects in mammals, others claim that RNAi is inefficient and masked by nonspecific inhibition of gene expression.

2. Long dsRNA Application in Mammals

In mammals, where dsRNA is part of three different pathways, diverse experimental setups are used to study various effects of dsRNA. Here, we focus primarily on the link between long dsRNA and the RNAi pathway. Sequence-specific mRNA degradation induced by exogenous dsRNA is a common strategy to suppress gene expression in invertebrates and plants. In mammals, however, RNAi induction by long dsRNA is rather uncommon (reviewed in ref. 25). The endogenous RNAi was clearly documented in mouse oocytes and embryonic stem cells (ESCs), where deep sequencing identified endogenous siRNAs (endo-siRNAs) produced from retrotransposons and processed pseudogenes (26–28). In these cell types, dsRNA is preferentially routed to the RNAi pathway while the IFN response is not readily induced (14). In addition, deep sequencing of hippocampal RNA revealed that small RNAs having features of endogenous siRNAs are expressed in the brain and possibly regulate synaptic plasticity of neuronal cells (29). In other cell types, endogenous siRNAs are either not detected or their levels are too low to be physiologically significant. However, experimental induction of RNAi with long dsRNA in different somatic cells indicates that the canonical RNAi pathway is latent or masked rather than absent in these cells.

3. Preparation of Long dsRNA for RNAi Experiments

The great advantage of using long dsRNA to induce RNAi is the simplicity, which, can be appreciated only in permissive model systems. When siRNA and short hairpin RNA (shRNA) are used, their design and use have to follow several essential rules and experimental designs have to take off-target effects into consideration (30–32). In mammalian cells, off-target effects stem from the ability of siRNAs to suppress many other genes via their binding to mRNAs with partial complementarity (typically involving a perfect basepairing at the 5′ end of an siRNA). When siRNA is transfected, it causes detectable off-target effects because of the high abundance of siRNA molecules carrying the same 5′ end sequence, which can basepair with and repress a discrete set of cellular mRNAs (33). On the other hand, long dsRNA causes minimal off-targeting because it is processed into a pool of siRNAs with varying sequences, which binds to different mRNA targets. Thus, this effectively dilutes off-target effects. One should pay attention to select a nonredundant target sequence that would not show sequence identity in stretches longer than 20 nt with other transcripts. This can be easily resolved with a BLAST sequence search.

The position of dsRNA within the mRNA sequence is flexible. Often, dsRNAs derived from the coding sequence of the cognate transcript are used (15, 34–43). However, the 3′ untranslated region (3′UTR) of the transcript can be targeted by dsRNA as well (38). Using dsRNA against the 3′UTR, one can perform a rescue experiment by expressing the coding sequence of the cognate transcript without the 3′UTR or with another nontargeted 3′UTR. Notably, when designing long dsRNA against the 3′UTR, one should also test (e.g., by analyzing available ESTs) whether the dsRNA would target all possible alternative 3′UTRs.

While the relative position of dsRNA in the cognate transcript does not play a crucial role, dsRNA length may influence efficient silencing. dsRNA shorter than 500 base pair (bp) may not be effectively processed by Dicer as suggested by in vitro experiments (44). Nevertheless, it should also be noted that efficient RNAi was induced with 290 bp-long dsRNA (41). It is not clear if there is an upper limit for dsRNA length. However, for practical reasons, dsRNA length typically ranges from 500 to 800 bp and rarely exceeds 1,000 bp. If one cannot use a single unique sequence, two shorter sequences can be fused together. A continuous region of the cognate transcript is not critical for siRNA target recognition.

RNAi silencing can be triggered either by transfection of in vitro transcribed long dsRNA or by long dsRNA expression from vectors transfected in the cell. The next two sections summarize practical considerations for both strategies.

3.1. dsRNA Produced by In Vitro Transcription

A common in vitro dsRNA preparation involves in vitro transcription (IVT) of sense and antisense RNAs, which are annealed and purified. Thus, dsRNA preparation is usually accomplished in <1 week. However, the exact time depends on the availability of suitable sequences, primers, and other reagents. The starting step of dsRNA production, IVT, can be easily done in 1 day. Templates for IVT can either be directly produced from a target cDNA by PCR using primers carrying SP6, T3, or T7 promoter sequences at their 5′ ends or can be cloned into an appropriate vector. Before IVT, a template must be purified to remove primers and nucleotides. PCR-based preparation of templates for IVT is simple and fast but when a large number of different templates needs to be prepared, necessary primers may be costly. Moreover, we have observed that RNA synthesis from different PCR-generated templates is more variable than IVT from a plasmid (45). Therefore, we recommend preparing a template for IVT by cloning PCR product into an appropriate vector. We routinely amplify the desired sequence by reverse transcriptase PCR (RT-PCR) and insert it using the TOPO-TA Cloning Kit (Invitrogen) into the pCRII plasmid (Invitrogen), which carries T7 and SP6 promoters flanking a multiple cloning site. Alternatively, dsRNA can be prepared directly using a template carrying an inverted repeat (IR) of the target sequence (15, 46). However, it should be noted that cloning of IR may not be a trivial task (this issue is discussed in the next section).

A plasmid that is used for IVT needs to be linearized downstream of the transcribed sequence and purified (either by a standard phenol/chloroform extraction or using a DNA purification kit). The final DNA template concentration should be above 500 ng/μl, and 100–500 ng of the template DNA is generally required per IVT reaction. IVT protocols have been already described in details elsewhere, including standard laboratory protocol manuals, RNA polymerase manufacturer's manuals, or protocols available on the Internet. Some laboratories have their own protocols for IVTs while others rely on IVT kits (e.g., MEGAscript kit from Ambion), which contain all necessary reagents and are optimized for a high yield. After IVT, which takes 2–4 h at 37°C, the template is removed by DNase treatment (e.g., by RQ1 DNase (Promega)). Then, one can purify sense and antisense strands, roughly estimate RNA yield by agarose electrophoresis and proceed to annealing. For annealing, approximately equimolar ratio of sense and antisense strands is mixed in up to 200 μl (the concentration of RNA should range between 200 and 500 ng/ml). We have successfully performed annealing of RNA dissolved in MilliQ pure water or in IVT reaction buffer. For efficient annealing, it is important to cool down the reaction slowly. Annealing can be performed in a boiling water bath, in a dry heatblock, or in a PCR machine. For the boiling water method, the 1.5 ml tube must be properly

closed (we use safe-lock tubes) to prevent accidental opening when pressure builds inside the tube upon heating. The tube with the annealing mixture is placed into a large beaker (1 l) with boiling water for 5 min. The beaker is then removed from the heat source and it is left to cool down slowly to room temperature. For even a slower cooling, one can just turn off the heating plate. Annealing in the heatblock is essentially the same: the annealing mixture is incubated for 5 min at 95°C before turning off the heatblock and letting the mixture cool down to room temperature. Alternatively, RNA can be annealed in a PCR machine programmed to cool down from 95 to 25°C for 30–60 min. When dsRNA is annealed from sense and antisense RNA strands, one can either remove the excess of single-stranded RNA and single-stranded overhangs by RNase T1 treatment prior to the purification step or proceed directly to purification. Annealed dsRNA is typically purified by phenol/chloroform extraction and dissolved in pure sterile water (do not use DEPC-treated water if dsRNA will be microinjected into oocytes). Produced dsRNA is fairly stable and can be stored at −20°C for a week or at −80°C for months. In vitro transcribed dsRNA can be microinjected into cells or transfected using Lipofectamine or Oligofectamine reagents.

3.1.1. dsRNA Produced from a Vector or a Transgene

There are several different strategies to express long dsRNA (Fig. 2). Sense and antisense RNA can be expressed separately, either from two different vectors (Fig. 2a), from one vector with two separately transcribed regions for the sense and antisense strands (Fig. 2b), or from one vector with one bidirectionally transcribed region (Fig. 2c). Alternatively, one can generate dsRNA as an intramolecular duplex—an RNA hairpin, which is formed upon transcription of an inserted IR (Fig. 2d). Even if plasmids expressing sense and antisense RNAs separately are easier to produce than plasmids with IR, hybridization of separately expressed sense and antisense RNAs in the cytoplasm may be inefficient. Moreover, the dose of dsRNA cannot be controlled. In contrast, IR forms dsRNA with a high probability and the amount of dsRNA is proportional to the strength of the promoter. Thus, the RNA hairpin expression from IR is a better choice than using the dual promoter plasmid. It should be noted that a dual promoter system (Fig. 2b) worked well in NIH 3T3 and HEK 293 cells (41). However, no systematic comparison of dual promoter and IR systems in mammals are currently available.

Expression of dsRNA can be controlled by various promoters. Polymerase II (pol II) promoters allow for tissue-specific expression or ubiquitous expression with diverse strength. When pol II transcribes IR, the product is typically spliced, capped, and polyadenylated. Such noncoding hairpin RNA is synthesized in the nucleus and presumably transported to the cytoplasm. On the other hand, polymerase III (pol III) induces nuclear production

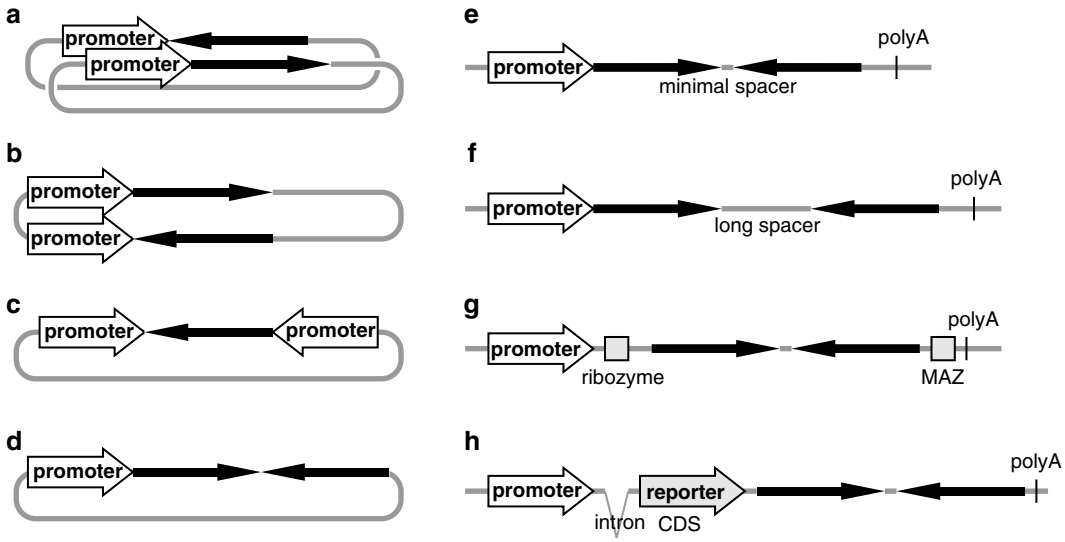


Fig. 2. General plasmid design for long dsRNA expression. (**a**–**d**) Long dsRNA can be produced by sense and antisense expression originating: (**a**) from two different plasmids, (**b**) from a single plasmid expressing independently sense and antisense RNA, (**c**) from a single plasmid where two promoters transcribe the same sequence in opposite directions, or (**d**) by transcription of an inverted repeat (IR). (**e**–**h**) Published hairpin-based expression systems include: (**e**) plasmids consisting of a promoter, an IR with a short spacer, and a polyA signal, (**f**) plasmids consisting of a promoter, an IR with a long spacer, and a polyA signal, (**g**) pDECAP derivates that produce decapped and non-polyadenylated dsRNA because of a presence of a ribozyme and MAZ attenuation site, (**h**) plasmids producing a spliced translatable mRNA carrying a long dsRNA stem in their 3′UTRs.

of non-spliced, non-capped, and non-polyadenylated RNA. Because pol III transcription prematurely terminates at TTTT, target long dsRNA sequence should be designed accordingly. Finally, it is possible to control dsRNA expression by ectopically expressed phage polymerase (15, 46). In this case, dsRNA synthesis can be targeted to the cytoplasm or the nucleus depending on the presence or absence of the nuclear localization signal in the phage polymerase.

Most of the published RNA hairpin expressing vectors fall into one of the four categories (Fig. 2e–h). The simplest two vector types are composed of the promoter, IR, and a polyA signal (Fig. 2e, f). The IR has either a minimal (tens of nucleotides) or a long (hundreds to kilobases) spacer and is followed by a polyA site (39, 46). The plasmid introduced by Shinagawa and Ishii produces decapped and non-polyadenylated dsRNA (Fig. 2g) that should be retained in the nucleus (40). The vector also does not contain an intron because splicing generally enhances the nuclear export of RNA (47). A vector expressing a translatable mRNA carrying a long dsRNA stem in its 3′UTR (Fig. 2h) was developed for transgenic RNAi experiments in mouse oocytes where it was used on numerous occasions (Table 3) (48).

3.2. Cloning of the Inverted Repeat

Several strategies for IR cloning are available (Fig. 3). One strategy uses PCR products with appropriate restriction sites that are ligated in vitro and then inserted into a plasmid in opposite directions forming the IR (Fig. 3a). It is important to avoid cryptic polyadenylation sites in PCR products in any direction, especially when a pol II promoter is used for expression. One of the PCR products can be longer at the end to form the center of the IR (Fig. 3a); this additional sequence will create the spacer, which enhances efficiency of IR cloning. Based on our experience, the short (20–50 bp long) spacer is sufficient for successful cloning without any negative impact on RNAi efficiency. The spacer can be perhaps even longer as successful RNAi was reported with a spacer as long as 700 bp (37). Another strategy uses sequential insertion of IR arms into the plasmid (Fig. 3b). One can also clone an IR with a long spacer first (1–2 kb) and then remove the spacer (Fig. 3c). Finally, one can clone a head to tail tandem array of two fragments, where one is flanked with loxP sites allowing for inverting it by Cre recombinase (37). This strategy (Fig. 3d) helps to minimize difficulties with IR insertion into the plasmid, which may become a frustrating task as transformation of ligated fragments frequently results in a large number of false positives. This is presumably because the presence of IR may interfere with plasmid's replication, thus creating a strong selection against the IR.

4. Effects of Long dsRNAs in Somatic Cells and Their Analysis

Since the 1970s, it has been known that long dsRNA is a potent trigger of the sequence-independent IFN response (49). Therefore, there was a strong skepticism about the existence of RNAi in mammals when the pathway was first identified in *C. elegans* and *Drosophila*. Mammalian RNAi was first demonstrated by microinjecting dsRNA into oocytes (46, 50, 51), which do not possess the IFN response. The problem of the IFN response was eventually minimized by introducing siRNAs and their derivatives (differently modified siRNAs) directly into somatic cells (34). Thus, siRNAs and shRNAs became the main RNAi tools in mammalian cells. It is generally accepted that in mammalian somatic cell types, long dsRNA readily triggers the IFN response causing sequence-independent silencing effects that mask a potential impact of RNAi-induced downregulation on cognate mRNAs (9, 37, 43, 52). The IFN pathway appears to become fully functional during differentiation as suggested by experiments in ESCs. While the specific RNAi effect was masked by sequence-independent silencing in differentiated ESCs, long dsRNA-induced RNAi was present in undifferentiated ESCs. This observation was interpreted as a manifestation of IFN effects although the activation of the IFN pathway was not studied

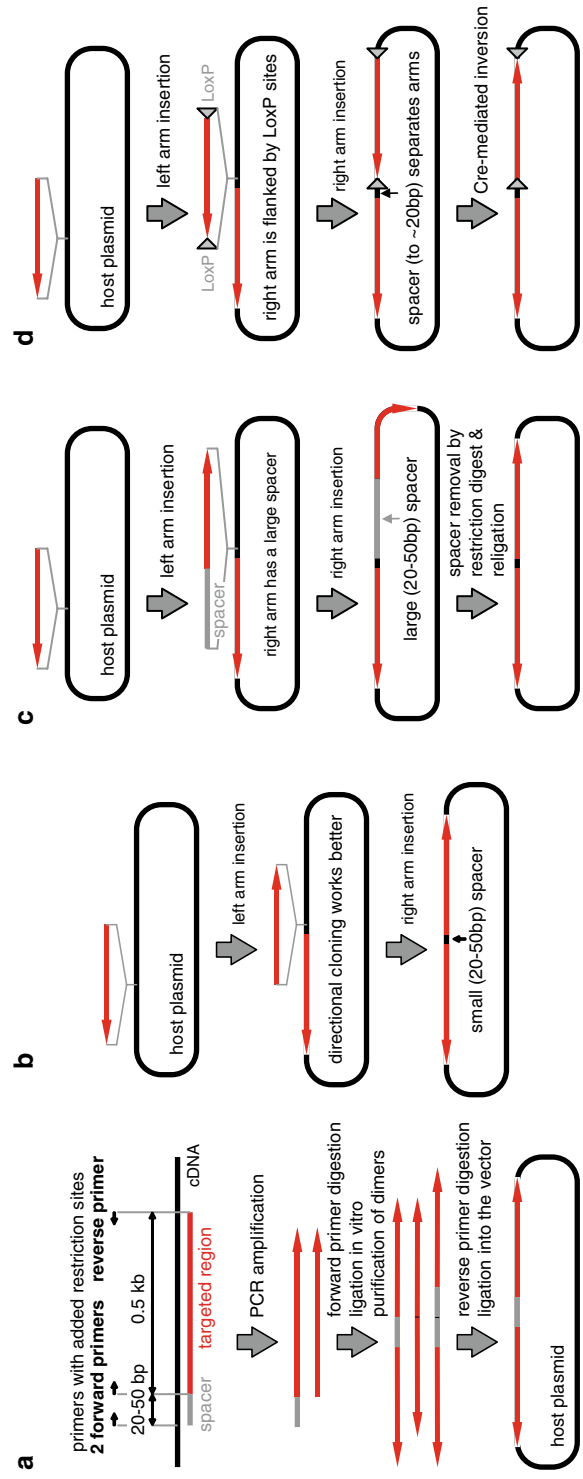


Fig. 3. Strategies for cloning an inverted repeat (IR) into a plasmid. (**a**) In vitro assembly of an IR, which is then inserted into a plasmid in a single cloning step. PCR products with appropriate restriction sites are digested at one end, ligated in vitro, digested at the other end, and the IR is then inserted into a plasmid. (**b**) A sequential insertion of arms (*red arrow*) of the IR. Arms can be prepared as PCR products carrying suitable restriction sites

in detail (15). A recent study of ubiquitous dsRNA expression in transgenic mice suggested a more complex routing of dsRNA into dsRNA-responding pathways in mouse somatic cells. It was shown that expression of long dsRNA embedded in a translatable mRNA (Fig. 2h) was a poor trigger of the IFN pathway; at the same time, it was also poorly processed into siRNAs (53). This suggests that somatic cells recognize different types of long dsRNA. In the following text, we will discuss contradictory results concerning long dsRNA-induced sequence-specific mRNA degradation in mammalian somatic cells.

4.1. dsRNA Transfection

Even if long dsRNA is of limited use for RNAi in mammalian cells, experiments employing long dsRNA are still a relevant tool to unravel how animal cells respond to foreign dsRNA. The ability of in vitro produced dsRNA to induce RNAi was addressed by several studies (Table 1). An initial study of long dsRNA as a potential RNAi trigger in mammalian somatic cells showed that, unlike siRNAs, long dsRNAs strongly and nonspecifically reduced reporter gene expression in NIH 3T3, COS-7, and HeLa S3 cells (34). The long dsRNAs, which ranged from 50 to 500 bp in length, were in vitro transcribed using T7 polymerase, annealed, and transfected into the cells. Long dsRNA targeting firefly (FF) or renilla (RL) luciferases were co-transfected together with luciferase reporter plasmids. Long dsRNA against GFP was used as an additional control for nonspecific repression of luciferase expression. RNAi efficiency was monitored by dual luciferase assay. The results showed that long dsRNA induced robust sequence-independent repression of luciferase reporters. The nature of the sequence-independent repression and the IFN pathway activation were not analyzed in further detail. A similar result was reported by Yang et al., who suggested that in mammalian cells long dsRNA may induce RNAi, but nonspecific activation of the IFN response may mask the specific effect (15). To repress EGFP reporter expression, they transfected either in vitro produced dsRNA (0.5 kb) or a plasmid harboring an IR (1.1 kb) dowstream of a T7 polymerase promoter, which was co-transfected with the plasmid expressing T7 polymerase (Tables 1 and 2). Delivery of dsRNA into undifferentiated ESCs resulted in a sequence-specific silencing of EGFP suggesting that undifferentiated ESCs may lack the IFN response to

Fig. 3. (continued) allowing for oriented insertion into the plasmid. (**c**) Spacer-in-and-out sequential insertion of IR arms into a plasmid. A large spacer is created either in an insert as shown here or by choosing an insertion site distant from the insertion of the first arm. In the final step, the spacer is removed and the plasmid is self-ligated. This strategy essentially eliminates all head-to-tail tandem insertions, which may cause problems when strategies **a** or **b** are used. (**d**) Production of an IR by Cre-mediated recombination. This approach also helps to minimize the previously mentioned difficulties sometimes experienced with strategies **a** and **b.**

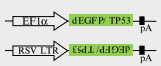
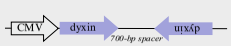
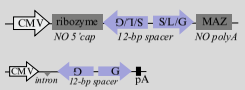
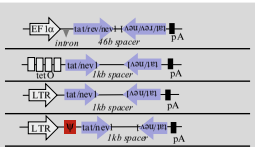
Table 1
Summary of experiments employing in vitro-transcribed long dsRNA in mammalian somatic cells

dsRNA length (bp)	dsRNA production	RNAi effect	Nonspecific effect	Cell line	Reference
50–500	IVT T7 promoter	Weak	+	HeLa S3 NIH 3T3 COS-7	Elbashir et al. (34)
547	IVT T7 promoter from vector	– +	+ –	Diff. ESCs Drosophila S2 Undiff. ESCs	Yang et al. (15)
123–771	IVT T7 promoter	+	– IFN NS	P19 F9 E14 ESCs	Billy et al. (35)
500	IVT T7 promoter	–	+	HeLa C2C12	Paddison et al. (37)
600–800	IVT T3 or T7 promoter	+	NS	Neuroblastama N2a (AGYNB010)	Gan et al. (38)
550–1,000	IVT T7 promoter	+	NS	Hypothalamic cells	Bhargava et al. (42)
472	IVT T7 promoter	+	+	KB-3-1 IMR-32 SK-N-MC	Kabilova et al. (43)

The dsRNA length, promoter used for dsRNA production, and the presence of sequence-specific (RNAi) effects or nonspecific effects as related to the cell lines used are discussed for each experiment. *Abbreviations*: *IVT* in vitro transcription, *bp* base pair, *NS* not studied, *IFN* interferon pathway

dsRNA. However, differentiated ESCs and mammalian cell lines HEK 293 and NIH 3T3 respond to dsRNA in a sequence-independent manner. In vitro assays showed that lysates of differentiated cells are capable of generating siRNAs from long dsRNA suggesting that RNAi effects may be masked by nonspecific effects (15). The ability of in vitro produced dsRNA to induce RNAi in undifferentiated pluripotent cells was also reported by Billy et al., who showed sequence-specific downregulation of exogenous (EGFP) and endogenous (integrin a3 and b1) genes upon electroporation of dsRNA (30 μg of ~700 bp-long dsRNA per 1×10^7 cells) in undifferentiated ESCs and P19 embryonic carcinoma cells (35). Similarly, Paddison et al. showed that in vitro transcribed long dsRNA (~0.5 kb) induced sequence-specific silencing in pluripotent P19 cells but not in HeLa cells or C2C12 myoblasts, where general repression of reporter expression was observed (37). Sequence-independent effects upon dsRNA transfection were also observed in cancer cell lines KB-3-1 (epidermal oral carcinoma) and SK-N-MC (neuroblastoma) (43). In this study, the effects of Myc long dsRNA (~500 bp, sequence from

Table 2
Summary of different constructs used to express long dsRNA in mammalian somatic cells

dsRNA schema design	dsRNA length (bp)	dsRNA production	RNAi effect	Nonspecific effect	Cell line	Reference
	547	Transfection linearized dsEGFP vector + T7 pol. expressing vector	– +	+ –	Diff. ESCs Drosophila S2 Undiff. ESCs	Yang et al. (15)
	1,083 1,195	Stable transfection episomal vectors	+	–	EcR293 HCT116	Tran et al. (36)
	500	Stable transfection	+	–	P19	Paddison et al. (37)
	1,090	Eletroporation Transfection	+	– IFN NS	C2C12	Yi et al. (39)
	540 518 502	Transfection Injection of fer egg Transfection	+ weak	– IFN NS +	MEF Transgenic mice MEF	Shinagawa and Ishii (40)
	289	Transfection	+	NS	NIH 3T3	Wang et al. (41)
	300	Transfection	+ Weak Weak +	–/NS + –/NS +	HEK 293 T C33A Vero	Konstantinova et al. (46)
	520	Transfection Transgenic mice	– +	– –	Somatic cells HEK 293 Mouse oocytes	Nejepinska et al. (53)

The construct design, dsRNA length, method for dsRNA production, presence of sequence-specific (RNAi) effects, and nonspecific effects of long dsRNAs are shown for the cell lines used. *Drawing*: *light gray arrow* polymerase II or III promoter, *CMV* cytomegalovirus enhancer, *green arrow or box* enhance green fluorescent protein (EGFP) coding sequence, *violet arrow* targeting sequence, *light orange box* tetracycline responsive element (TRE), *dark gray box* ribozyme and binding site of the zinc finger protein MAZ, *yellow box* zeocin resistance (Zeo^r), *red triangle* LoxP site for Cre recombination, *dark red box* HIV-1 leader sequence (ψ). *Abbreviations*: *pA* polyadenylation signal, *bp* base pair, *NS* not studied, *IFN* interferon pathway, *ESC* embryonic stem cell, *MEF* mouse embryonic fibroblast, *NS* not studied

exon 3 of c-myc gene) was compared with those of siRNAs and polyinosinic:polycytidylic acid (poly(I:C)). Long dsRNAs and poly(I:C) induced nonspecific suppression of *c-MYC* production and a transient increase in *PKR* and *OAS1* mRNA levels.

There are several observations suggesting that, in addition to ESCs and oocytes, long dsRNA may induce specific silencing in neuronal cells (38, 42). Using in vitro transcribed dsRNAs (~700 bp) in mouse neuronal cell, Gan et al. observed efficient silencing (50–60%) of exogenous GFP or endogenous proteins involved in apoptosis, signaling, and metabolism (38). In another report, injection of ~550 bp-long dsRNA into adult rat hypothalamic cells resulted in a successful sequence-specific knockdown of targeted genes in vivo (42). The observed effect of long dsRNA on two model neuroproteins, corticotrophin-releasing factor (CRF) and arginin vasopressin (VAP), was specific and restricted to injected hypothalamic cells and did not spread to other brain regions. In another set of experiments, 10–15 μg of long dsRNA (~0.5–1 kb) targeting CRF and VAP was mixed with 1 μl of Lipofectamine and injected into the brain. Sequence-specific effects were subsequently present for up to 1 week suggesting that long dsRNA can be applied for sequence-specific silencing in neural cells and for in vivo set studies. The notion that RNAi in neuronal cells may be more prominent than the IFN response is further supported by detection of endogenous siRNAs in the mouse hippocampus (29).

Experiments with delivering in vitro produced dsRNA into mammalian cells showed that its potential to induce RNAi is largely restricted to pluripotent cells. Differentiated cells (perhaps with some exceptions among neuronal cells) generally respond to dsRNA in a sequence-independent manner. This is attributed to the IFN pathway, in agreement with studies of poly(I:C) effects (54). Accordingly, in vitro produced long dsRNA is routinely used for inducing RNAi in invertebrate models, such as *C. elegans*, *Tribolium*, or *Drosophila*. However, it virtually disappeared from RNAi experiments in mammals except oocytes and early embryos. Mammalian RNAi is typically induced by siRNAs and shRNAs (reviewed in ref. 55), while long dsRNA remains a useful tool to study the IFN response and crosstalks between dsRNA-induced pathways.

4.2. dsRNA Expression

In contrast to in vitro-produced long dsRNA, several lines of evidence suggest that expressed long dsRNA more frequently induces sequence-specific silencing without provoking the IFN response in somatic cells (Table 2). For example, Tran et al. used episomal plasmids encoding long complementary RNAs to specifically suppress exogenous and endogenous gene expression in mammalian cells (36). In this system, co-expression of sense and antisense RNAs led to a specific inhibition of GFP or endogenous p53

transcriptional activator. Under these conditions, the IFN response was not triggered as evidenced by the lack of PKR phosphorylation. However, the mechanism of sequence-specific silencing may be different from RNAi as Dicer knockdown did not influence the ability to reduce GFP expression. Furthermore, the gene silencing effect was transferable to other cells not previously exposed to the sense and antisense RNA via culture media, although mammalian cells do not exhibit transitive RNAi (reviewed in ref. 25).

Another system for long dsRNA expression was introduced by Paddison et al., who employed a plasmid expressing a self-complementary stem-loop against the first 500 nt of a GFP coding sequence (37). To produce long dsRNA hairpin, they used a FLIP cassette containing two sense-orientated fragments flanking the zeocin-resistance coding sequence. The second fragment was flanked by LoxP sites allowing its inversion into the antisense orientation by the Cre recombinase (Fig. 3d). Recombination resulted in an IR, which had the potential to form a long hairpin with a zeocin-resistance coding sequence as a loop. This system was used in the P19 embryonal carcinoma cell line and induced sequence-specific silencing of GFP accompanied with processing of GFP dsRNA into siRNAs (37). Whether this system could induce RNAi in differentiated somatic cell lines was not reported.

In contrast to the previously mentioned nonspecific effects of in vitro-produced dsRNA in C2C12 cells, Yi et al. demonstrated that long dsRNA could also trigger sequence-specific repression of Dyxin in differentiated C2C12 myoblast cells (39). To express dsRNA, they used a CMV-driven IR with a 700 bp-long spacer. In undifferentiated C2C12 cells, they observed sequence-specific downregulation of the ectopically expressed Dyxin. Moreover, the dsRNA also functioned in fully differentiated myotubes, where endogenous Dyxin (but not eIF2α, vinculin, or tubulin) was knocked down efficiently. Dyxin dsRNA-induced gene silencing of protein levels was accompanied by reduction of Dyxin mRNA suggesting that the silencing effect was caused by RNAi. However, sequence-independent effects including the activation of the IFN pathway was not studied in detail (39).

An interesting dsRNA expression system designed to prevent the IFN response by nuclear retention of dsRNA was developed by Shinagawa and Ishii (40). The pDECAP (deletion of Cap structure and polyA) vector contained a CMV promoter that controls expression of dsRNA hairpins lacking both the 5′-cap structure and the 3′-poly(A) tail (Fig. 2g). To remove the 5′-cap structure, a ribozyme cassette was added at a site downstream of the RNA start site. To eliminate the poly(A) tail, they inserted a MAZ site downstream of the IR. The zinc finger protein MAZ binding causes polymerase II to pause and fall off. As splicing generally enhances the nuclear export of RNA (47), the vector did not contain any intron. The pDECAP vector harbored an approximately 1 kb long IR containing

the firefly luciferase, β-galactosidase, or the Ski gene sequence that was separated by a short sequence forming a small loop in expressed hairpin transcripts. This design should lead to the retention of long dsRNA in the nucleus where it could be processed to siRNAs while preventing induction of the IFN response by dsRNA in the cytoplasm. Northern blot analysis of the nuclear and cytoplasmic fractions revealed that long dsRNA was not detected in either fraction suggesting that long dsRNA was rapidly processed to siRNAs. While it remains unclear whether dsRNA from pDECAP was exclusively processed in the nucleus, this system successfully induced sequence-specific silencing in HEK 293 cells. Transgenic mice carrying pDECAP vector expressing long dsRNA against the Ski gene showed similar developmental abnormalities as Ski knockout mice. Thus, specifically modified expressed long dsRNA can be routed into the RNAi pathway without provoking the IFN response. However, while the pDECAP vector appeared to be a suitable tool to efficiently generate tissue-specific knockdown mice, it did not get wider attention and was only sporadically used with variable success (56–60).

Another approach to overcome the IFN response was tested in Chock's laboratory (41). To express dsRNA, they used two opposite tetracycline-inducible CMV promoters driving expression of 290 nt sense and antisense sequences of glutaredoxin cDNA. After induction of sense and antisense RNA expression with doxycycline for 6 days in NIH 3T3 cells, glutaredoxin protein reduction was observed. Notably, the antisense RNA alone did not induce significant decrease of glutaredoxin.

Several different vectors containing expressed long hairpin RNAs of approximately 300 bp in length were used to inhibit human immunodeficiency virus type 1 (HIV-1) in infected mammalian cells (46). It was demonstrated that long dsRNAs are capable of inhibiting HIV-1 production in a sequence-specific manner without inducing expression of the class I IFN genes. The most effective constructs contained a viral RNA sequence Ψ cloned downstream of a promoter, or had a T7 promoter producing dsRNA directly in the cytoplasm. The viral Ψ sequence may provide the transcript with a nonself signature and thereby boost RNAi. The T7 promoter bypasses a possible problem with the nuclear export of expressed long dsRNAs, thus increasing HIV-1 repression. Interestingly, the study showed that the length of the loop between the IRs did not play a role in the efficiency of inhibiting HIV-1 production. As long dsRNA can target a large sequence of the viral genome, a major advantage of long dsRNA over shRNA use is the reduced chance of viral escape when mutations occur.

A comprehensive study of the effects of long dsRNA both in vivo and in vitro was done by Nejepinska et al. (53) who adapted a transgene design previously developed for mouse oocytes (48). An IR of ~500 bp was placed in the 3′UTR of spliced EGFP-

encoding mRNA under the control of a chimeric CMV/β-actin promoter. This expression vector was used to generate transgenic mice and for the transfection of human cell lines. In transgenic mice, the IR was ubiquitously expressed and the formation of dsRNA was confirmed by detecting RNase T1-resistant RNA in cells. Interestingly, this long dsRNA did not activate the IFN response in somatic cells as evidenced by the lack of phosphorylation of PKR, analysis of several IFN-stimulated genes, and a microarray analysis in HEK 293 and HeLa cells transfected with the plasmid. There was a low level of RNA editing in mouse organs but a substantial editing in transfected HEK 293 cells, suggesting cell type-dependent activity of the ADAR enzyme. Moreover, edited dsRNA was detected in the cytoplasmic fraction, indicating that RNA editing did not prevent cytoplasmic localization of the dsRNA. Importantly, strong RNAi response was observed in oocytes derived from the transgenic mice but not in any other cell types. In addition, transfected HEK 293 or HeLa cells did not show strong RNAi. High-throughput sequencing showed that long dsRNA was poorly processed into siRNAs in the transgenic liver or brain, or in transfected HEK 293 cells. These data suggest that some long dsRNAs can be tolerated in mammalian somatic cells without induction of the IFN response and that siRNA biogenesis in somatic cells may be limited.

5. Effects of Long dsRNA in Mammalian Oocytes and Early Embryos

Mammalian oocytes and blastomeres of early embryos are the only mammalian cell types where long dsRNA is more frequently employed to induce RNAi. One of the reasons is that the IFN response is not functional in these cells (14). In addition, long dsRNA readily enters RNAi in these cells, presumably because of the presence of a robust endogenous RNAi (26, 27). In fact, mouse oocytes were the first cells where the existence of the mammalian RNAi pathway was demonstrated upon injection of long dsRNA (50, 51). Microinjection of in vitro-produced long dsRNA (typically ~10^6 molecules) offers a simple and robust method to induce RNAi. This method is still being used, although microinjection of siRNAs may be more prevalent nowadays because of the reduced price and improved efficiency of siRNAs. However, the induction of RNAi by long dsRNA microinjection is usually performed only in fully grown oocytes and 1-cell embryos, which are the stages most accessible to the injection. If one wants to target genes during oocyte growth and/or during oocyte-to-embryo transition, expression of dsRNA from a transgene is the method of choice.

5.1. Transgenic RNA

Transgenes for long dsRNA expression in oocytes mostly follow a previously developed design (Fig. 2h), where dsRNA is expressed as a part of EGFP-encoding mRNA (48). The RNA construct expression is specifically induced in oocytes at the post-meiotic stage with the ZP3 promoter or at the pre-meiotic stage with the GDF9 promoter. Detailed protocols for designing a transgene for RNAi in oocytes have been described elsewhere (19, 61). Here, we focus on additional considerations that have emerged recently. To date, 11 published transgenic RNAi experiments have been performed in oocytes (53, 62–71) (Table 3). None of the experiments encountered serious problems with cloning IRs and preparing transgenic constructs. Occasional delays in cloning IRs with a short spacer were solved by growing bacterial clones at room temperature in special bacterial strains (e.g., Stbl4 (Invitrogen) or Sure (Stratagene)). All transgenes yielded a strong mRNA knockdown in the best transgenic lines. In addition, we successfully used transgenic RNAi to silence two genes simultaneously (Flemr and Svoboda, unpublished). The average time from the transgene cloning to the first phenotype analysis typically ranged from 6 to 9 months. Phenotypes of knockdown mice were typically robust, except for the *Wee1B* knockdown, where only approximately 25% of the oocytes exhibited a phenotype and where *Wee1B* mRNA was not knocked down as efficiently as in other transgenic experiments (63).

Despite the successes, there are also several drawbacks that should be considered before deciding to perform a transgenic RNAi experiment. First, one has to select transgenic line(s) with a strong knockdown effect, which may be a problem when a small number of transgenic lines carrying randomly integrated transgenes is obtained. Second, randomly integrated transgenes may be subjected to epigenetic silencing. In our recent experience, some vectors are more prone to epigenetic silencing than others (72), and transgene silencing is more frequently observed if an inbred strain embryo is used to produce transgenic founder animals. Although, our vector for transgenic RNAi (65) is typically well expressed, it was found to be silenced in several C57Bl/6 animals (unpublished results). A hybrid cross of C57Bl/6 and BALB/c partially solved this problem. However, some RNAi transgenes still showed variable expression as judged by variable EGFP fluorescence in oocytes. Notably, RNAi transgenes targeting dormant maternal mRNAs (mRNAs, not translated until ovulation) have typically a uniform EGFP expression while RNAi transgenes targeting translated mRNAs, which may be of importance for oocyte growth, have a more variable EGFP expression. This suggests that some growing oocytes may escape from the silencing effect and fully mature with milder phenotype (Svoboda et al., unpublished observation). Selection against knockdown effects could also explain insufficient knockdown in the *Wee1B* experiment (63).

Table 3
Summary of experiments with transgenic RNAi mice expressing long dsRNA in oocytes

Gene	dsRNA length (bp)	Knockdown efficiency	Phenotype	Reference
Mos	535	90% mRNA reduction	Null phenotype parthenogenetic activation	Stein et al. (65)
Ctcf	700	70–99% mRNA reduction	Abnormal methylation decreased developmental competence	Fedoriw et al. (64)
Msy2	850	60% Protein reduction	Pleotropic effect	Yu et al. (66)
WeelB	566	50% mRNA reduction	Meiotic arrest	Han et al. (63)
Bnc (basonuclin)	~800	>90% mRNA reduction	Not identified	Ma et al. (62)
CPEB	612	~90% Reduction	Abnormal MII spindle parthenogenetic activation sterility	Racki et al. (67)
PLCb1	533	~75% mRNA reduction	Significant decrease in Ca^{2+} transient amplitude	Igarashi et al. (68)
SLBP	654	80% mRNA and >90% protein reduction	DNA replication defect at the 2-cell stage and developmental arrest	Arnold et al. (69)
CTCF	333	50–99% mRNA reduction	Meiotic defect in the egg and mitotic defect in the embryo	Wan et al. (70)
ATRX	450	80–90% Reduction	Abnormal MII spindle and increased aneuploidy	Baumann et al. (71)
Mos	520	90% mRNA reduction	Sterile or subfertile parthogenetic activation	Nejepinska et al. (53)

The table contains the name of the target genes, the length of the dsRNA stem, the knockdown efficiency of the transgene, and the phenotype observed in transgenic mice

6. Outlook

While long dsRNA may not be a perfect tool to induce RNAi in mammalian cells, it remains an important method to investigate mammalian responses to foreign nucleic acids. From this perspective, different strategies to produce and deliver diverse types of dsRNA molecules into mammalian cells are still highly relevant. Furthermore, long dsRNA has a great potential for nonmammalian model systems and many technical aspects discussed in this review are of general importance to anyone working with long dsRNA in any model systems.

Acknowledgements

Authors thank Radek Malik for help with preparation and Camille Du Roure for text revision of this manuscript. This work was supported by the GACR 204/09/0085, EMBO SDIG program #1488, and the Purkynje Fellowship to PS. JN is supported in part by Faculty of Science, Charles University in Prague.

References

1. Carthew RW, Sontheimer EJ (2009) Origins and mechanisms of miRNAs and siRNAs. Cell 136:642–655
2. Ghildiyal M, Zamore PD (2009) Small silencing RNAs: an expanding universe. Nat Rev Genet 10:94–108
3. Fire A, Xu S, Montgomery MK et al (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 391:806–811
4. Caplen NJ, Fleenor J, Fire A et al (2000) dsRNA-mediated gene silencing in cultured Drosophila cells: a tissue culture model for the analysis of RNA interference. Gene 252: 95–105
5. Clemens JC, Worby CA, Simonson-Leff N et al (2000) Use of double-stranded RNA interference in Drosophila cell lines to dissect signal transduction pathways. Proc Natl Acad Sci U S A 97:6499–6503
6. Hammond SM, Bernstein E, Beach D et al (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature 404:293–296
7. Ui-Tei K, Zenno S, Miyata Y et al (2000) Sensitive assay of RNA interference in Drosophila and Chinese hamster cultured cells using firefly luciferase gene as target. FEBS Lett 479:79–82
8. Cullen BR (2006) Is RNA interference involved in intrinsic antiviral immunity in mammals? Nat Immunol 7:563–567
9. Gantier MP, Williams BR (2007) The response of mammalian cells to double-stranded RNA. Cytokine Growth Factor Rev 18:363–371
10. Garcia MA, Gil J, Ventoso I et al (2006) Impact of protein kinase PKR in cell biology: from antiviral to antiproliferative action. Microbiol Mol Biol Rev 70:1032–1060
11. Silverman RH (2007) A scientific journey through the 2–5A/RNase L system. Cytokine Growth Factor Rev 18:381–388
12. Silverman RH (2007) Viral encounters with 2′,5′-oligoadenylate synthetase and RNase L during the interferon antiviral response. J Virol 81:12720–12729
13. Judge AD, Sood V, Shaw JR et al (2005) Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA. Nat Biotechnol 23:457–462
14. Stein P, Zeng F, Pan H et al (2005) Absence of non-specific effects of RNA interference triggered by long double-stranded RNA in mouse oocytes. Dev Biol 286:464–471

15. Yang S, Tutton S, Pierce E et al (2001) Specific double-stranded RNA interference in undifferentiated mouse embryonic stem cells. Mol Cell Biol 21:7807–7816
16. Hornung V, Guenthner-Biller M, Bourquin C et al (2005) Sequence-specific potent induction of IFN-alpha by short interfering RNA in plasmacytoid dendritic cells through TLR7. Nat Med 11:263–270
17. Sioud M (2005) Induction of inflammatory cytokines and interferon responses by double-stranded and single-stranded siRNAs is sequence-dependent and requires endosomal localization. J Mol Biol 348:1079–1090
18. Marques JT, Devosse T, Wang D et al (2006) A structural basis for discriminating between self and nonself double-stranded RNAs in mammalian cells. Nat Biotechnol 24:559–565
19. Svoboda P, Stein P (2009) RNAi experiments in mouse oocytes and early embryos. Cold Spring Harb Protoc (pdb top56)
20. Keegan LP, Leroy A, Sproul D et al (2004) Adenosine deaminases acting on RNA (ADARs): RNA-editing enzymes. Genome Biol 5:209
21. Bass BL (2002) RNA editing by adenosine deaminases that act on RNA. Annu Rev Biochem 71:817–846
22. Kumar A, Crawford K, Close L et al (1997) Rescue of cardiac alpha-actin-deficient mice by enteric smooth muscle gamma-actin. Proc Natl Acad Sci U S A 94:4406–4411
23. Nishikura K (2010) Functions and regulation of RNA editing by ADAR deaminases. Annu Rev Biochem 79:321–349
24. DeCerbo J, Carmichael GG (2005) Retention and repression: fates of hyperedited RNAs in the nucleus. Curr Opin Cell Biol 17:302–308
25. Nejepinska J, Flemr M, Svoboda P (2012) The canonical RNA interference pathway in animals. In: Mallick B, Ghosh Z (eds) Regulatory RNAs. Springer, Heidelberg, p 623
26. Tam OH, Aravin AA, Stein P et al (2008) Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature 453:534–538
27. Watanabe T, Totoki Y, Toyoda A et al (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature 453:539–543
28. Babiarz JE, Ruby JG, Wang Y et al (2008) Mouse ES cells express endogenous shRNAs, siRNAs, and other microprocessor-independent, Dicer-dependent small RNAs. Genes Dev 22:2773–2785
29. Smalheiser NR, Lugli G, Thimmapuram J et al (2011) Endogenous siRNAs and noncoding RNA-derived small RNAs are expressed in adult mouse hippocampus and are up-regulated in olfactory discrimination training. RNA 17: 166–181
30. Fellmann C, Zuber J, McJunkin K et al (2011) Functional identification of optimized RNAi triggers using a massively parallel sensor assay. Mol Cell 41:733–746
31. Taxman DJ, Livingstone LR, Zhang J et al (2006) Criteria for effective design, construction, and gene knockdown by shRNA vectors. BMC Biotechnol 6:7
32. Olson A, Sheth N, Lee JS et al (2006) RNAi Codex: a portal/database for short-hairpin RNA (shRNA) gene-silencing constructs. Nucleic Acids Res 34:D153–D157
33. Birmingham A, Anderson EM, Reynolds A et al (2006) 3′ UTR seed matches, but not overall identity, are associated with RNAi off-targets. Nat Methods 3:199–204
34. Elbashir SM, Harborth J, Lendeckel W et al (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature 411:494–498
35. Billy E, Brondani V, Zhang H et al (2001) Specific interference with gene expression induced by long, double-stranded RNA in mouse embryonal teratocarcinoma cell lines. Proc Natl Acad Sci U S A 98:14428–14433
36. Tran N, Raponi M, Dawes IW et al (2004) Control of specific gene expression in mammalian cells by co-expression of long complementary RNAs. FEBS Lett 573:127–134
37. Paddison PJ, Caudy AA, Hannon GJ (2002) Stable suppression of gene expression by RNAi in mammalian cells. Proc Natl Acad Sci U S A 99:1443–1448
38. Gan L, Anton KE, Masterson BA et al (2002) Specific interference with gene expression and gene function mediated by long dsRNA in neural cells. J Neurosci Methods 121:151–157
39. Yi CE, Bekker JM, Miller G et al (2003) Specific and potent RNA interference in terminally differentiated myotubes. J Biol Chem 278:934–939
40. Shinagawa T, Ishii S (2003) Generation of Ski-knockdown mice by expressing a long double-strand RNA from an RNA polymerase II promoter. Genes Dev 17:1340–1345
41. Wang J, Tekle E, Oubrahim H et al (2003) Stable and controllable RNA interference: investigating the physiological function of glutathionylated actin. Proc Natl Acad Sci U S A 100:5103–5106
42. Bhargava A, Dallman MF, Pearce D et al (2004) Long double-stranded RNA-mediated RNA interference as a tool to achieve site-specific

silencing of hypothalamic neuropeptides. Brain Res Brain Res Protoc 13:115–125

43. Kabilova TO, Vladimirova AV, Chernolovskaya EL et al (2006) Arrest of cancer cell proliferation by dsRNAs. Ann N Y Acad Sci 1091:425–436
44. Bernstein E, Caudy AA, Hammond SM et al (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409:363–366
45. Stein P, Svoboda P, Anger M et al (2003) RNAi: mammalian oocytes do it without RNA-dependent RNA polymerase. RNA 9: 187–192
46. Konstantinova P, de Vries W, Haasnoot J et al (2006) Inhibition of human immunodeficiency virus type 1 by RNA interference using long-hairpin RNA. Gene Ther 13:1403–1413
47. Valencia P, Dias AP, Reed R (2008) Splicing promotes rapid and efficient mRNA export in mammalian cells. Proc Natl Acad Sci U S A 105:3386–3391
48. Svoboda P, Stein P, Schultz RM (2001) RNAi in mouse oocytes and preimplantation embryos: effectiveness of hairpin dsRNA. Biochem Biophys Res Commun 287: 1099–1104
49. Hunter T, Hunt T, Jackson RJ et al (1975) The characteristics of inhibition of protein synthesis by double-stranded ribonucleic acid in reticulocyte lysates. J Biol Chem 250: 409–417
50. Svoboda P, Stein P, Hayashi H et al (2000) Selective reduction of dormant maternal mRNAs in mouse oocytes by RNA interference. Development 127:4147–4156
51. Wianny F, Zernicka-Goetz M (2000) Specific interference with gene function by double-stranded RNA in early mouse development. Nat Cell Biol 2:70–75
52. Akimov IA, Kabilova TO, Vlassov VV et al (2009) Inhibition of human cancer-cell proliferation by long double-stranded RNAs. Oligonucleotides 19:31–40
53. Nejepinska J, Malik R, Filkowski J et al (2012) dsRNA expression in the mouse elicits RNAi in oocytes and low adenosine deamination in somatic cells. Nucleic Acids Res 40(1):399–413
54. Geiss G, Jin G, Guo J et al (2001) A comprehensive view of regulation of gene expression by double-stranded RNA-mediated cell signaling. J Biol Chem 276:30178–30182
55. Bantounas I, Phylactou LA, Uney JB (2004) RNA interference and the use of small interfering RNA to study gene function in mammalian systems. J Mol Endocrinol 33:545–557
56. Hou X, Omi M, Harada H et al (2011) Conditional knockdown of target gene expression by tetracycline regulated transcription of double strand RNA. Dev Growth Differ 53:69–75
57. Dai P, Nakagami T, Tanaka H et al (2007) Cx43 mediates TGF-beta signaling through competitive Smads binding to microtubules. Mol Biol Cell 18:2264–2273
58. Koster MI, Dai D, Marinari B et al (2007) p63 induces key target genes required for epidermal morphogenesis. Proc Natl Acad Sci U S A 104:3255–3260
59. Maekawa T, Shinagawa T, Sano Y et al (2007) Reduced levels of ATF-2 predispose mice to mammary tumors. Mol Cell Biol 27:1730–1744
60. Zipperlen PBaP (2005) Comparison of a range of approaches for RNAi in human cells. QIAGEN News, QIAGEN, Institute of Molecular Biology, University of Zurich, Zurich, Switzerland
61. Svoboda P (2009) Cloning a transgene for transgenic RNAi in mouse oocytes. Cold Spring Harb Protoc (pdb prot5134)
62. Ma J, Zeng F, Schultz RM et al (2006) Basonuclin: a novel mammalian maternal-effect gene. Development 133:2053–2062
63. Han SJ, Chen R, Paronetto MP et al (2005) Wee1B is an oocyte-specific kinase involved in the control of meiotic arrest in the mouse. Curr Biol 15:1670–1676
64. Fedoriw AM, Stein P, Svoboda P et al (2004) Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. Science 303: 238–240
65. Stein P, Svoboda P, Schultz RM (2003) Transgenic RNAi in mouse oocytes: a simple and fast approach to study gene function. Dev Biol 256:187–193
66. Yu J, Deng M, Medvedev S et al (2004) Transgenic RNAi-mediated reduction of MSY2 in mouse oocytes results in reduced fertility. Dev Biol 268:195–206
67. Racki WJ, Richter JD (2006) CPEB controls oocyte growth and follicle development in the mouse. Development 133:4527–4537
68. Igarashi H, Knott JG, Schultz RM et al (2007) Alterations of PLCbeta1 in mouse eggs change calcium oscillatory behavior following fertilization. Dev Biol 312:321–330
69. Arnold DR, Francon P, Zhang J et al (2008) Stem-loop binding protein expressed in growing oocytes is required for accumulation of mRNAs encoding histones H3 and H4 and for early embryonic development in the mouse. Dev Biol 313:347–358

70. Wan LB, Pan H, Hannenhalli S et al (2008) Maternal depletion of CTCF reveals multiple functions during oocyte and preimplantation embryo development. Development 135: 2729–2738
71. Baumann C, Viveiros MM, De La Fuente R (2010) Loss of maternal ATRX results in centromere instability and aneuploidy in the mammalian oocyte and pre-implantation embryo. PLoS Genet 6:e1001137
72. Sarnova L, Malik R, Sedlacek R et al (2010) Shortcomings of short hairpin RNA-based transgenic RNA interference in mouse oocytes. J Negat Results Biomed 9:8

Supplement 2

Nejepinska J, Malik R, Filkowski J, Flemr M, Filipowicz W, Svoboda P.

dsRNA expression in the mouse elicits RNAi in oocytes and low adenosine deamination in somatic cells.

Nucleic Acids Research

2012

40(1):399-413

Published online 8 September 2011 *Nucleic Acids Research, 2012, Vol. 40, No. 1* ***399–413***
doi:10.1093/nar/gkr702

dsRNA expression in the mouse elicits RNAi in oocytes and low adenosine deamination in somatic cells

Jana Nejepinska[1], Radek Malik[1], Jody Filkowski[2], Matyas Flemr[1], Witold Filipowicz[2] and Petr Svoboda[1,*]

[1]Institute of Molecular Genetics AS CR, Videnska 1083, 14220 Prague 4, Czech Republic and [2]Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland

Received August 8, 2011; Accepted August 13, 2011

ABSTRACT

Double-stranded RNA (dsRNA) can enter different pathways in mammalian cells, including sequence-specific RNA interference (RNAi), sequence-independent interferon (IFN) response and editing by adenosine deaminases. To study the routing of dsRNA to these pathways *in vivo*, we used transgenic mice ubiquitously expressing from a strong promoter, an mRNA with a long hairpin in its 3′-UTR. The expressed dsRNA neither caused any developmental defects nor activated the IFN response, which was inducible only at high expression levels in cultured cells. The dsRNA was poorly processed into siRNAs in somatic cells, whereas, robust RNAi effects were found in oocytes, suggesting that somatic cells lack some factor(s) facilitating siRNA biogenesis. Expressed dsRNA did not cause transcriptional silencing *in trans*. Analysis of RNA editing revealed that a small fraction of long dsRNA is edited. RNA editing neither prevented the cytoplasmic localization nor processing into siRNAs. Thus, a long dsRNA structure is well tolerated in mammalian cells and is mainly causing a robust RNAi response in oocytes.

INTRODUCTION

Double-stranded RNA (dsRNA), a double helix formed by two antiparallel RNA strands, is a unique structure whose recognition is important in host defense and regulation of gene expression. The recognition and effects of dsRNA are mediated by a diverse set of proteins harboring dsRNA binding domains (dsRBD), [reviewed in Ref. (1)].

One of the evolutionarily conserved effects of dsRNA is represented by RNA interference (RNAi), sequence-specific degradation of RNAs complementary to the sequence of the dsRNA [reviewed in Ref, e.g. (2,3)]. RNAi is initiated by the RNase III enzyme Dicer, which is cleaving dsRNA at ~21 nt intervals, generating short interfering RNA (siRNA) duplexes with two nucleotide 3′ overhangs. One of the siRNA strands is loaded onto the RNA-induced silencing complex (RISC), where it serves as a guide for cleaving perfectly complementary mRNAs. In mammalian cells, RNAi effects can also be induced experimentally by siRNAs (4) or by microRNAs (5). Endogenous microRNAs (miRNAs) are genome-encoded small RNAs produced by cleavage of pre-processed short hairpin precursors by Dicer and are also loaded on a RISC-like complex [reviewed in Ref. (6)]. For the purpose of this study, we use the term RNAi for the pathway that is induced by long dsRNA, i.e. as it was originally described by Fire *et al.* (7). RNAi operates in most eukaryotes and often serves as a defense mechanism against viruses and repetitive sequences. Hence, it is often viewed as a form of innate immunity (8). However, RNAi in mammals does not appear to play an antiviral role (9) and its endogenous function has only been clearly documented for oocytes and embryonic stem cells where it targets repetitive elements and regulates endogenous genes [reviewed in Ref (10)]. Nevertheless, experimental induction of RNAi with ectopically expressed long dsRNA in different somatic cells (11–14) suggests capacity for siRNA generation from dsRNA also in other cell types.

Mammalian somatic cells can respond to dsRNA in a sequence-independent manner. A pioneering work by Hunter *et al.* (15) showed that different types of dsRNA,

*To whom correspondence should be addressed. Tel: +4161697428; Fax: +41616973976; Email: svobodap@img.cas.cz
Present address:
Jody Filkowski, Department of Biological Sciences, University of Lethbridge, 4401 University Drive, Lethbridge, Alberta, T1K 3M4, Canada.

including the poly I:C duplex, can block translation in reticulocyte lysates. Analysis of this phenomenon identified protein kinase R (PKR), an enzyme activated upon binding to dsRNA that blocks translation by phosphorylating the α-subunit of eukaryotic initiation factor 2 (eIF-2α) (16). Activation of PKR represents part of a complex response to foreign molecules known as the interferon (IFN) response [reviewed in Ref (17)], which includes activation of the NFκB transcription factor and a large number of IFN-stimulated genes (ISGs) (18). In addition to PKR, several other proteins recognizing dsRNA are integrated into the IFN response. DDX58 (RIG-I) and MDA5 sense cytoplasmic dsRNA and activate IFN expression. The 2′,5′-oligoadenylate synthetase produces 2′,5′-linked oligoadenylates that induce a general degradation of RNAs by activating latent RNase L [reviewed in Ref, e.g. (19,20)].

Another dsRNA-associated mammalian pathway is represented by RNA editing, which is mediated by adenosine deaminases acting on RNA (ADARs). These dsRNA recognizing enzymes convert adenosines to inosines, thus affect stability and coding potential of modified RNAs [reviewed in Ref (21,22)]. The effects of dsRNA editing are complex and the degree of editing may affect the subsequent fate of the edited RNA, leading to its nuclear retention (23) or degradation (24). Nevertheless, mRNAs carrying edited dsRNA hairpins in their 3′-UTR can be transported to the cytoplasm and translated, as evidenced by the presence of such mRNAs on polysomes (25). Editing of long dsRNA in *Caenorhabditis elegans* antagonizes the transgene-induced RNAi in somatic cells by retaining edited dsRNA in the nucleus (26). However, the reported role of the cytoplasmic RISC component Tudor-SN in the degradation of hyperedited RNAs (>50% conversion) indicates more complex interplay between editing and RNAi (24). Mammalian ADARs can reduce efficiency of RNAi in two ways. First, they can erode dsRNA to the point where it is either no longer a suitable substrate for Dicer processing or, in the case of successful Dicer cleavage, the resulting siRNAs have changed specificity to base pair with target mRNAs. Second, they can reduce availability of siRNAs by directly binding to them, an effect that seems to be independent of ADAR editing activity (27,28).

While the three main mammalian dsRNA-responding pathways mentioned above have been individually characterized in substantial detail, the interactions between them are still poorly understood. Co-existence of these pathways certainly involves recognition of different types of dsRNA substrates and their possible sequestration in different cellular compartments or cell types [reviewed in Ref (19,29)]. The latter phenomenon underlies the common simplistic view that cytoplasmic dsRNA is toxic to somatic cells because it activates the IFN response, while nuclear dsRNA is edited and thus prevented to enter the cytoplasm. However, such interpretation is challenged by the growing list of reports showing induction of RNAi by intracellular expression of long dsRNA in transformed and primary somatic cells (4,11–13,30,31).

To obtain new insights into the effects of dsRNA in various types of somatic cells, we produced a transgenic mouse model ubiquitously expressing long dsRNA. We have previously developed a transgene that generates dsRNA within the 3′-UTR of a protein-coding transcript. This dsRNA takes the form of a long hairpin with a perfect ~0.5 kb stem, which is flanked by long single-stranded 5′ and 3′ overhangs. Using a transgene with the *Mos* gene sequence in the hairpin and oocyte-specific ZP3 promoter, we induced an efficient and highly specific RNAi effect in mouse oocytes (32,33). Physiologically, the *Mos* gene encodes for a dormant maternal mRNA, which is stored in the oocyte until the resumption of meiosis (34). Elimination of the *Mos* maternal mRNA by transgenic RNAi phenocopies the null mutation (32), which manifests as parthenogenetic activation of ovulated eggs and ovarian cysts. Otherwise, $Mos^{-/-}$ animals appear normal (35,36).

Here, we report an adaptation of the *Mos* hairpin transgene (for simplicity referred to as MosIR) for ubiquitous, constitutive expression of dsRNA in transgenic mice. We show that in somatic cells of transgenic animals, dsRNA does not induce the IFN response, is inefficiently processed by Dicer and its editing is barely detectable. This suggests that a long dsRNA structure embedded in a transcript produced by RNA polymerase II in the nucleus of somatic cells is not a potent trigger of any of the three common pathways responding to dsRNA. When MosIR RNA levels were increased in cell culture experiments, we observed more frequent editing while IFN pathway activation and RNAi effects were still negligible. The IFN response was induced only with high levels of expressed dsRNA in somatic cells. In contrast to somatic cells, the MosIR induced a robust RNAi effect in oocytes suggesting that female germ cells represent a tissue adapted to directing dsRNA into the RNAi pathway.

MATERIALS AND METHODS

Plasmids and transgenes

Schematic structures of the relevant parts of plasmid constructs used in the project are shown in the Supplementary Figure S1. pCAGEGFP-MosIR (for simplicity, referred to as MosIR) was produced by transferring the EagI fragment carrying the *Mos* inverted repeat inserted in the pCR II plasmid (37) into the SspBI site downstream of the enhanced green fluorescent protein (EGFP) coding sequence in the pCAGEGFP plasmid (38). pCAGEGFP-Mos3 (for simplicity, referred to as Mos3) was produced by inserting a PCR-amplified 973 bp fragment of the Mos transcript (corresponding to nucleotides 114–1089 of the Mos cDNA sequence NM020021) into the SspBI site downstream of the EGFP coding sequence in the pCAGEGFP plasmid. pCAGEGFP-MosP (for simplicity, referred to as MosP) was produced by inserting the same PCR-amplified 973 bp fragment of the *Mos* transcript into the SnaBI site between the cytomegalovirus (CMV) enhancer and β-actin promoter. Insertion into the pCAGEGFP was verified by restriction digest and sequencing. The *Mos* sequence in Mos3 and MosP

fragments was inserted in the sense orientation relative to the EGFP transcription and a KpnI site in the *Mos* was eliminated by blunt-ending and re-ligation, allowing for distinguishable reporter sequences from the endogenous *Mos* sequences. EGFP expression from MosP and Mos3 plasmids were compared with the original EGFP plasmid by flow cytometry (FACS) analysis of HEK293 cells (for simplicity, referred to as 293) transfected with equimolar amounts of each plasmid. This analysis showed that, at similar transfection efficiency, Mos3 and MosP EGFP expression reached ~25 and 80% of pCAGEGFP, respectively (data not shown). The lower Mos3 EGFP expression was probably caused by lower stability or reduced translatability of the EGFP mRNA that was expanded in its 3′-UTR by an additional 1 kb of sequence. For each MosIR, Mos3 and MosP plasmid, SalI and HindIII were used to release the transgenic cassette for producing transgenic animals. phRL-TK-Mos3 (for simplicity, referred to as RL-Mos3) production was described previously (39). Firefly luciferase (for simplicity, referred to as FF) served as a non-targeted control.

Transgenic mice

Transgenic mice were generated at the transgenic facility of the Friedrich Miescher Institute by injecting linearized DNA into male pronuclei of C57BL/6 × BALB/c 1-cell embryos. Transgene-positive mice were identified by PCR (primer sequences are listed in the Supplementary Table SIII). For the MosIR transgene, four different founder animals (one female and three males) were obtained and examined further. One female founder animal was sterile and showed the *Mos* null phenotype [parthenogenetic activation of unfertilized oocytes (35,36)]. F1 progeny of founder males was examined for the expression of EGFP and down-regulation of *Mos* mRNA in oocytes. The line with the strongest *Mos* null phenotype in female animals was expanded and used for further analysis. For Mos3 and MosP reporter lines, founder animals were crossed with C57BL/6 × BALB/c animals and the line with the strongest EGFP expression for each reporter was used further.

Oocyte isolation and culture

Adult transgenic or wild-type female siblings were superovulated using 5 U of pregnant mare serum gonadotropin (PMSG, Intervet). Oocytes were isolated as described previously (40). Images of oocytes and organ samples were obtained using a digital camera mounted to a stereomicroscope SZX16 equipped with a 100 W mercury lamp and GFPA filter (Olympus).

Cell culture and transfection

Human 293 cells were maintained in Dulbecco's Modified Eagle Medium (DMEM, Invitrogen) supplemented with 10% FCS (Invitrogen), penicillin (100 U/ml, Invitrogen) and streptomycin (100 μg/ml, Invitrogen) at 37°C and 5% CO_2 atmosphere.

For transfection, cells were typically plated on a 24-well plate, grown to 70% density and transfected using Turbofect *in vitro* Tranfection Reagent (Fermentas) according to the manufacturer's protocol. Cells were transfected with 100 ng of both untargeted (FF) and targeted (RL-Mos3) reporter plasmids and various amount of MosIR plasmid (50–500 ng per well). The total amount of transfected DNA was kept constant by adding pCAGEGFP. After 48 h, cells were washed with PBS and lysed with the Passive Lysis Buffer (Promega). Luciferase reporter expression was assessed using the Dual-Luciferase Reporter Assay (Promega) and luminiscence intensity was measured by Modulus Microplate Multimode Reader (Turner Biosystems).

The 293 cells stably expressing both RL-Mos3 and FF were established by co-transfecting reporter plasmids with the pPuro selection plasmid carrying puromycin resistance (500 ng of each plasmid per well in a 6-well plate). To select positive clones, cells were passaged in the presence of 1.5 μg/ml of puromycin. The positively selected individual clones were tested for *Renilla* and firefly luciferase expression using the Dual-Luciferase Reporter Assay. Clones yielding luciferase expression similar to that of the transiently transfected cells were used for further studies.

Isolation of mouse embryonic fibroblasts

Mouse embryos were isolated at embryonic day 12.5. The head and internal organs were removed, the torso was homogenized using a needle under sterile conditions and subsequently cultured in DMEM supplemented with 10% FCS, penicillin (100 U/ml, Invitrogen), and streptomycin (100 μg/ml, Invitrogen) at 37°C and 5% CO_2 atmosphere. Mouse embryonic fibroblasts (MEFs) within the first three passages were used for experiments.

Detection of RNase T1-resistant MosIR RNA

The fraction of MosIR RNA, which folded into dsRNA, was approximated from the comparison of non-treated RNA (input) and RNase T1-treated (T1) samples. MosIR-positive MEFs or 293 cells transfected with 2 μg of MosIR per well in a 6-well plate were harvested 48 h post-transfection. The cells were washed twice with PBS, liberated with a rubber scraper into a 1.5 ml tube, and hypotonically lysed with three volumes of water [experiments where cells were permeabilized with 0.1% Triton X-100 yielded the same results (data not shown)]. Cell lysates were incubated with RNase T1 (5000 U) for 30 min at 37°C. Negative-control samples were denaturated at 85°C for 5 min and immediately placed on ice to prevent perfect pairing of complementary strands prior to addition of RNase T1. RNase T1 was inactivated by phenol–chloroform extraction using the RNA Blue reagent (Top Bio). DNA was removed by Turbo DNase (Ambion) treatment. All RNA samples were produced in the same volume (30 μl). If needed, concentration of RNA recovered after treatment was equalized using *Escherichia coli* rRNA (Roche). From each sample, 1 μg of RNA was then reverse-transcribed using RevertAid M-MuLV Reverse Transcriptase (Fermentas) and a MosIR-specific primer or using Superscript II (Invitrogen) and random hexamer primers with similar results. Reverse transcriptase was omitted in '–RT' samples. cDNA was amplified

using Maxima SYBR Green/ROX qPCR Master Mix (Fermentas) and PCR products were resolved on 1.5% agarose gel stained with ethidium bromide (Sigma Aldrich).

Microarray analysis

The 293 and HeLa cell lines were analyzed in duplicates 24 h after transfection with 500 ng/well of the MosIR plasmid in a 6-well plate. Controls were transfected with equal amounts of the parental pCAGEGFP plasmid. Microarray analysis of 5 μg of total RNA per sample was performed as described previously (41). Principal component analysis and hierarchical clustering of the GeneChip Robust Multiarray Averaging (GC-RMA) normalized data were performed by Partek software (Partek Inc., USA). The bioconductor Limma package (42) was used to identify differentially expressed genes and generate MA plots. The heatmap for IFN-related gene expression was created using TM4 microarray software suite (43). Microarray data were deposited in the NCBI GEO database (GSE27316).

Next generation sequencing

Brain and kidney tissues were obtained from a freshly sacrificed 17-week old transgenic female. HeLa and 293 cells were transfected with 500 ng/well of MosIR plasmid in a 6-well plate and cells were collected 48 h post-transfection. Small RNA fractions were isolated using mirPremier microRNA Isolation Kit (Sigma) according to the manufacturer's instructions. RNA concentration was determined using a Nanodrop 1000 spectrophotometer, and the quality of RNA was checked by 8% polyacrylamide gel electrophoresis.

Library construction and deep sequencing of small RNAs were performed by Seqomics (Szeged, Hungary) using SOLiD (version 3.0) sequencing platform. For the bioinformatic analysis, the bar code and 3′ adaptor sequences were removed from raw sequence reads using an algorithm requiring at least 3-nt exact matches between 35-nt reads and the adaptor sequence. Sequencing quality and depth of all samples was comparable (Supplementary Figure S3, the complete Seqomics Sequencing Report is available on request). Reads were mapped onto the following sequences: miRNA [MirBase release 15; http://www.mirbase.org/, (44)], tRNA [Genomic tRNA database; http://lowelab.ucsc.edu/GtRNAdb/, (45)], rRNA [http://www.arb-silva.de/, (46)] and MosIR plasmid. Only perfect matches were considered for the first round of analysis. The annotation order was miRNA, tRNA, rRNA and plasmid sequence. To avoid mapping of some reads to multiple categories, we removed each read from the dataset once it was annotated to a particular sequence category. The remaining reads not matching any RNA sequences were marked as 'other'. For RNA (A > I) editing analysis, 'other' reads were remapped to the above mentioned selected sequences allowing for one or more A/G mismatches. All mapping software used in the analysis were programmed using the Visual Basic 2010 platform (Microsoft). Lists of generated 35-nt reads in color-coded format were deposited in the GEO database (GSE26577).

RNA isolation and quantitative real-time RT–PCR

Oocytes for real-time PCR analysis were washed 3-times in PBS, placed individually in pure water with 10 U of RiboLock RNase inhibitor (Fermentas) while 0.14 pg/μl of rabbit globin mRNA (Sigma Aldrich) and 0.2 μg/μl of *E. coli* rRNA (Roche) was added to act as an external standard and to inhibit sample RNA degradation by RNases. Samples were stored at −80°C prior to reverse transcription. Total RNA from organs and cultured cells was isolated by phenol–chloroform extraction using the RNA Blue reagent (Top-Bio) according to the manufacturer's instructions. Expression of specific mRNA was analyzed by quantitative real-time PCR (qPCR) as described previously (39). Briefly, 1 μg of total RNA was reverse transcribed using RevertAid M-MuLV reverse transcriptase (Fermentas) and random hexanucleotides (Fermentas). Reverse transcriptase was omitted in control (−RT) samples. The resulting cDNA was diluted three times with water and a 3 μl aliquot was used as a template for a 10 μl qPCR reaction. qPCR was performed on the Mx3000P (Stratagene) or LC480 (Roche) machines using Maxima SYBR Green qPCR Master Mix (Fermentas). Hypoxanthine–guanine phosphoribosyltransferase (*Hprt1*) was used as an internal housekeeping gene control (all primers are listed in the Supplementary Table SIII). Values of crossing points (CPs) were evaluated and corrected according to PCR efficiency for each reaction. The statistical significance of relative expression changes of target mRNA levels normalized to a housekeeping gene was analyzed by the pair-wise fixed reallocation randomization test using the REST 2008 software (47).

RESULTS AND DISCUSSION

Transgenic mice ubiquitously expressing long dsRNA are viable

To achieve ubiquitous long dsRNA expression in the mouse, we employed the same general design (Figure 1A) that was used for transgenic RNAi in mouse oocytes (32). To control dsRNA expression, we chose a strong chimeric promoter composed of the CMV enhancer and β-actin core promoter (CAG) that provided efficient, ubiquitous expression in transgenic mice (48). The MosIR transgene produced an mRNA that carried an EGFP coding sequence and 3′-UTR containing a long dsRNA hairpin of the *Mos* sequence (Figure 1B).

The MosIR transgene offered several advantages. First, the EGFP coding sequence upstream of the *Mos* hairpin yields a sufficient amount of green fluorescence to monitor transgene expression. Second, this RNA efficiently folds into dsRNA upon transcription *in vitro* (37) and it induces an efficient and highly specific RNAi effect in the oocyte (32,33). Third, the *Mos* null phenotype manifests as a parthenogenetic activation of unfertilized eggs (35,36), providing an easily assayed positive control for RNAi effects. Fourth, the loss of *Mos* has no phenotype in somatic cells (35,36) and *Mos* mRNA levels are negligible

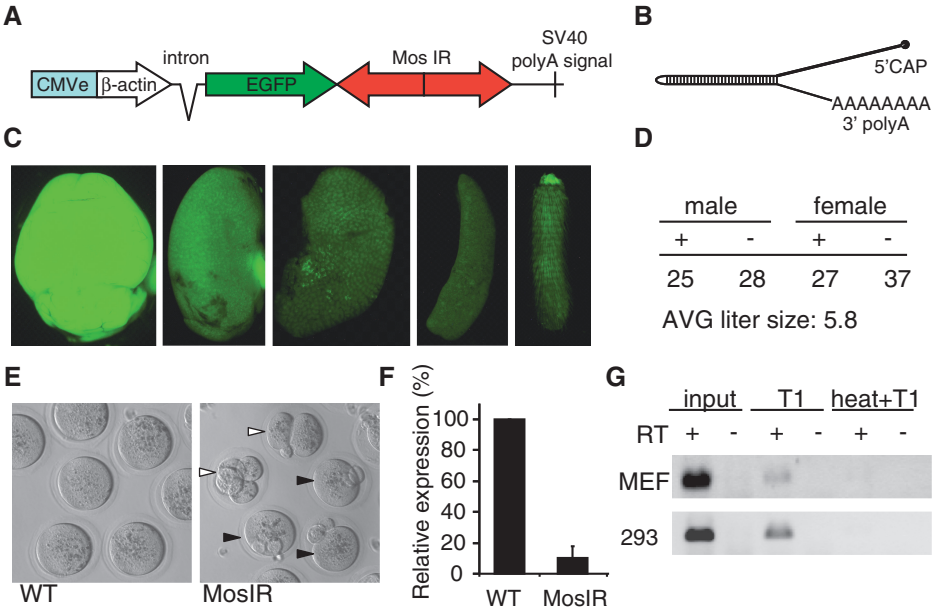


Figure 1. Analysis of MosIR expression and phenotype. (**A**) Schematic composition of the MosIR transgene. (**B**) A schematic structure of the MosIR transcript folded into a long dsRNA hairpin. The length of the stem is 520 bp. (**C**) MosIR transgene produces EGFP in somatic tissues. From left to right: brain, kidney, liver, spleen, tail. All images were taken with the same settings. No autofluorescence was visible under these conditions in internal organs of wild-type mice. (**D**) MosIR transgene (+) shows normal segregation into male and female progeny in crosses of MosIR males and wild-type female mice. (**E**) MosIR positive females phenocopy the *Mos* null phenotype: parthenogenetic activation of ovulated eggs. The left panel shows normal metaphase II-arrested eggs. The right panel shows transgenic parthenogenotes with two extruded polar bodies (black arrowheads) and cleaving parthenogenotes (white arrowheads). (**F**) qPCR of *Mos* mRNA in fully-grown oocytes obtained from three transgenic animals reveals strong reduction of Mos mRNA level compared to wild-type oocytes. Error bars = SEM. (**G**) RNase T1-resistant MosIR RNA can be detected in MEFs isolated from MosIR embryos and in 293 cells transfected with the MosIR plasmid. Hypotonically lysed cells were treated with RNase T1 for 30 min at 37°C before RNA extraction (T1). The efficiency of RNase T1 digestion was tested by disrupting secondary RNA structures by heat before RNase T1 digestion (heat+T1). RT indicates reverse-transcribed RNA (+) and controls without reverse transcriptase (−). cDNA was amplified with MosIR-specific primers for 31 cycles (MEF) and 30 cycles (293) of PCR. No amplification was observed with primers located in the single-stranded EGFP coding sequence (data not shown).

in somatic cells (49). Therefore, while RNAi effects can be conveniently monitored with an appropriate reporter target, the MosIR is also suitable for studying sequence-independent effects of dsRNA expression in either the presence or absence of the target mRNA.

We obtained four MosIR transgenic lines upon pro-nuclear injection of the MosIR transgene. All of them showed green fluorescence in all tissues examined, documenting that the transgene was ubiquitously expressed (Figure 1C; data not shown). The line derived from the male founder 317.3 showed the strongest *Mos* phenotype in the female progeny and was selected for detailed analysis of effects caused by MosIR expression. Transgenic females had reduced fertility. Mating of five different females for several weeks resulted in only one litter of two animals. This was an expected consequence of parthenogenetic activation caused by RNAi-mediated down-regulation of *Mos* mRNA in oocytes (Figure 1E and F). This effect indirectly demonstrated formation of MosIR-derived dsRNA in oocytes and was identical to the results obtained with the oocyte-specific version of the MosIR transgene (32,33).

Transmission of the MosIR transgene via the male germline showed normal Mendelian distribution and an average litter size of 5.8 (Figure 1D), which did not significantly differ from litters of this strain in our mouse facility. This indicated that the MosIR did not cause embryonic lethality. Apart from the reduced female fertility, transgenic mice showed no other apparent phenotype suggesting the MosIR transcript has no effect on cell growth or viability. At the same time, observable green fluorescence in different organs indicated that the MosIR expression was ubiquitous. The amount of EGFP fluorescence varied across different organs (Figure 1C). The brightest signal was found in the brain and testis and the lowest in the liver and spleen. A similar pattern of EGFP signal was also observed in organs of transgenic mice with other EGFP-carrying transgenes controlled by the same promoter but devoid of the inverted repeat (38) (Figure 2B). Notably, MosIR transgene generated less EGFP than other transgenes driven by the CAG promoter (Supplementary Figure S1C). Transfection of equal amounts of pCAGEGFP and MosIR plasmids into 293 cells showed discrepancy between mRNA and EGFP fluorescence in MosIR transfected cells (Supplementary Figure S2A and B) and thus, suggested that lower EGFP fluorescence in MosIR mice was caused by lower translatability of the MosIR transcript. The MosIR transcript was readily detectable by qPCR in somatic tissues and its relative abundance correlated with relative amounts of EGFP fluorescence among tissues (Supplementary Figure S2C).

Although the *Mos* null phenotype indicated expression of dsRNA in the female germline, further evidence that

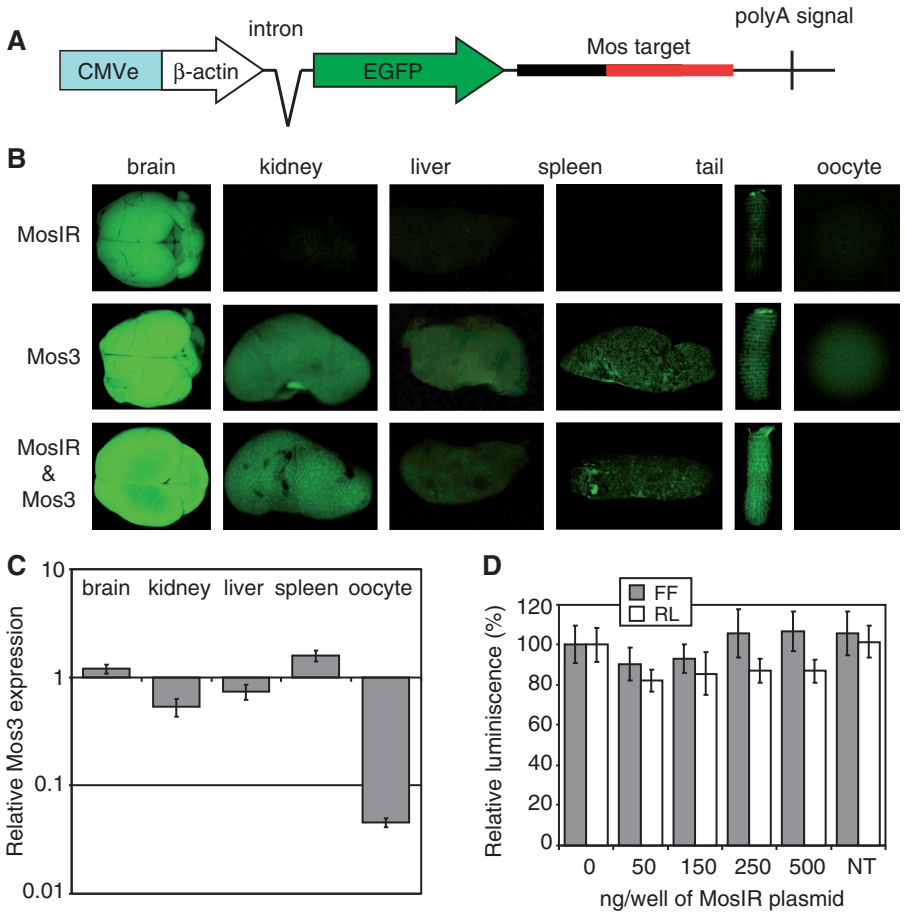


Figure 2. Absence of RNAi in somatic cells. (**A**) A schematic composition of the Mos3 reporter transgene. (**B**) EGFP signal in different organs of MosIR, Mos3 and Mos3 & MosIR mice. Organs were collected from siblings of one litter. All images were taken with the same settings. The exposure of was shorter than in Figure 1C in order to obtain non-saturated brain EGFP signal. (**C**) qPCR analysis of different tissues shows little, if any, RNAi in somatic tissues and a strong RNAi effect in oocytes. Mos3 reporter expression in Mos3 & MosIR mice is shown relative to its expression in Mos3 mice. Samples were collected from two Mos3 and two Mos3 & MosIR animals and analyzed by qPCR in triplicates. *Hprt* was used as an internal normalization standard. Tissue heterogeneity could contribute to the lower level of Mos3 in the kidney. Error bars = SEM. (**D**) MosIR does not induce RNAi when transfected to somatic cells. The 293 cells stably transfected with a non-targeted firefly luciferase (FF) reporter and a targeted *Renilla* luciferase (RL) reporter (carrying *Mos* sequence in its 3′-UTR, Supplementary Figure S1B) were grown in 24-well plates and transfected with increasing amounts of the MosIR plasmid. pCAGEGFP was added to transfection mixtures to balance different amounts of the MosIR plasmid and to maintain the total amount of transfected DNA constant. Cells were harvested 48 h after transfection and luciferase activity was analyzed. Both luciferase activities are shown relative to cells transfected with 0 ng of the MosIR plasmid. NT = non-transfected cells. Transfection efficiency estimated by microscopic examination of EGFP fluorescence was >90%. Error bars = SEM.

the MosIR transcripts formed dsRNA *in vivo* in somatic cells was necessary. To assess formation of MosIR dsRNA in intact cells, we used RNase T1 to degrade single-stranded RNA prior to RNA extraction from MEFs derived from the MosIR mice since the extraction can cause formation of dsRNA that does not exist *in vivo* (50). RNase T1-resistant MosIR RNA was analyzed by qPCR. We successfully amplified the *Mos* fragment from RNase T1-treated samples but not from samples where secondary RNA structure was disrupted by heat before RNase T1 digestion (Figure 1G). These results provide qualitative evidence for the presence of MosIR dsRNA in intact cells. It remains unclear whether the low yield RNase T1-resistant RNA (~5–10% of the input when estimated by qPCR) is caused by instability of dsRNA during T1 treatment or whether it reflects reduced formation of dsRNA from the MosIR transcript. The latter would suggest that somatic cells actively reduce levels of dsRNA originating from endogenous transcripts, either at the level of RNA binding proteins, which prevent dsRNA formation and/or by a duplex unwinding activity. In any case, our data show that the MosIR transgene is ubiquitously transcribed, triggers RNAi in oocytes and generates dsRNA in MEFs derived from transgenic mice. This conclusion is also supported by evidence provided by experiments presented further below, which show that the MosIR RNA is edited and processed into siRNAs in somatic cells.

MosIR transcript efficiently enters the RNAi pathway only in oocytes

The MosIR transcript efficiently entered the RNAi pathway in oocytes as evidenced by the reduction of *Mos* mRNA and presence of the *Mos* null phenotype in

MosIR mice (Figure 1E and F). As *Mos* is not expressed in somatic tissues (49), we produced a transgenic EGFP reporter carrying the *Mos* sequence in the 3′-UTR (Figure 2A), allowing us to monitor RNAi effects in mouse tissues by microscopy. The Mos3 transgene produced bright EGFP fluorescence, which was stronger than that of the MosIR (Figure 2B), allowing for rapid identification of RNAi effects. In any case, the fact that both transgenes produced EGFP was not an obstacle for studying RNAi effects since the expression of MosIR and Mos3 transgenes could be distinguished by qPCR.

With the exception of oocytes, where EGFP fluorescence was reduced, examination of EGFP fluorescence in whole organs or in organ cryosections did not reveal any obvious RNAi effect when samples from mice carrying both MosIR and Mos3 transgenes were compared to samples from mice carrying the Mos3 transgene alone (Figure 2B; data not shown). Likewise, analysis of the Mos3 transcript level by qPCR did not suggest the occurrence of robust somatic RNAi (Figure 2C). Although we did observe relative reduction of the Mos3 RNA in kidney and liver samples, this was much less profound compared with oocytes and may have been a result of tissue heterogeneity.

To further address the potential of the MosIR dsRNA to induce RNAi in somatic cells, we performed experiments in cultured cells. 293 cells carrying stably integrated non-targeted FF and targeted *Renilla* luciferase (RL-Mos3) reporters were transfected with increasing doses of the MosIR transgene that lead to proportionally increasing levels of MosIR RNA (Supplementary Figure S2A). However, even in cells with the highest expression of MosIR, no statistically significant RNAi effect was found (Figure 2D). Furthermore, we did not observe significant RNAi effects upon transient transfection of the MosIR transgene and both luciferase reporters into 293, HeLa, or MCF-7 cells (data not shown).

To directly address whether the MosIR RNA enters the RNAi pathway, we analyzed kidney and brain for the presence of MosIR-derived siRNAs using high throughput sequencing by the SOLiD technology (sequencing depth $\sim 4 \times 10^6$ reads). Kidney and brain samples were selected because both tissues showed high MosIR expression and kidney samples also showed somewhat reduced levels of the Mos3 reporter. SOLiD sequencing of tissue samples produced a distinct peak of 21–23 nt small RNAs, the size expected for Dicer products (Supplementary Figure S3A). The number of putative *Mos* siRNAs was minimal (20 in the brain and 4 in the kidney). This contrasted with hundreds of thousands of potential Dicer cleavage products, including ~70 000 sequences derived from the three most abundant miRNAs (Figure 3A). Notably, only one putative siRNA sequence derived from the single-stranded EGFP region of the MosIR transcript was found in each organ sample. Since virtually all *Mos* sequences in the brain and kidney samples were in the range of Dicer products (Figure 3A), we concluded that a small fraction of MosIR RNA was processed by Dicer into siRNAs in somatic tissues but their abundance was too low to cause a robust RNAi effect.

To further address MosIR dsRNA conversion into siRNA, we performed SOLiD sequencing of small RNA from 293 cells transiently transfected with high dose of MosIR plasmid. SOLiD sequencing yielded 343 putative *Mos* siRNAs (20–23 nt long), which localized along the predicted MosIR dsRNA and were enriched among small RNAs of 17–25 nt in length (Figure 3B and C). Interestingly, mapping of small RNAs onto the transfected plasmid revealed a complex picture of RNA expression, including RNA fragments derived from both strands of EGFP, from the antisense strand of the intron and the plasmid backbone (Supplementary Figure S3B). This spurious transcription likely accounts for higher level of EGFP-derived sequences (434) when compared with the number of MosIR-derived sequences (343) (Figure 3B). However, RNAs of the size expected for Dicer products were not enriched among small RNAs derived from the EGFP coding sequence (Supplementary Figure S3C) arguing that the peak of MosIR-derived 21–23 nt sequences includes Dicer products. Taking into account the background of other *Mos* small RNAs, it appears that *Mos* siRNAs could be up to 10 times more abundant in MosIR-transfected 293 cells than in transgenic tissues, which likely reflects MosIR overexpression in transfected cell. However, the amount of *Mos* siRNAs estimated by SOLiD sequencing in transfected 293 cells was two orders of magnitude below abundant endogenous miRNAs (Figure 3B).

Detailed examination of SOLiD results from 293 cells revealed several thousand EGFP and MosIR sequences longer than 30 nt, whereas, only one was found in the brain (Supplementary Figure S3C). Since fragments of similar length originating from endogenous mRNAs were not detected (for example, we found only one fragment from *Hprt*, a transcript used as an internal standard for qPCR), the MosIR fragments longer than 30 nt likely represent a snapshot of degradation intermediates of overexpressed MosIR RNA, rather than artifacts of sample preparation. Although RNA degradation intermediates would be expected to be much less stable than siRNAs, we found that putative *Mos* siRNAs in transfected 293 cells were an order of magnitude less abundant than the degradation intermediates (Supplementary Figure S3C). This result is consistent with the hypothesis that Dicer does not efficiently process MosIR dsRNA into siRNAs in somatic cells.

Since the MosIR transcript caused robust RNAi effects in oocytes but not in somatic cells, we propose that oocytes are a privileged tissue for endogenous RNAi in which dsRNA is efficiently presented to Dicer. This model is consistent with studies reporting abundant endogenous siRNAs in oocytes (51,52) and few, if any, endogenous siRNAs in somatic cells (53–56). Nonetheless, our results do not rule out that other types of long dsRNA molecules can be processed by Dicer in somatic cells. There are several reports of RNAi induction with long dsRNA in somatic cells where long dsRNA was formed by transcripts, which differ from the MosIR RNA (11–14). The MosIR transcript is a capped and polyadenylated translatable mRNA, which contains long 3′-UTR dsRNA stem and long single-stranded overhangs. These features may

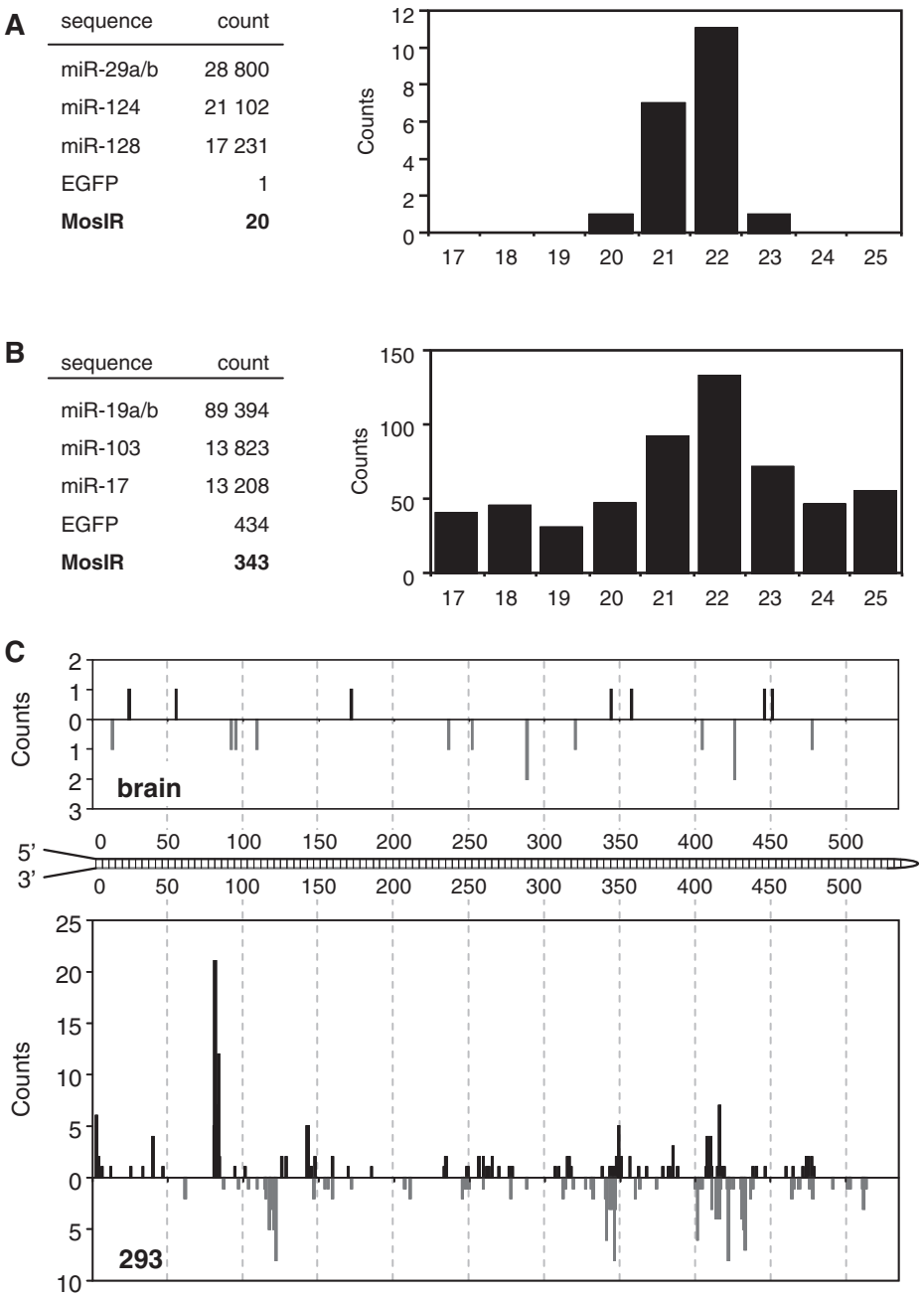


Figure 3. Deep sequencing of MosIR-derived small RNAs. (**A**) Deep sequencing of small RNAs from brain using SOLiD technology. Absolute counts of the most abundant miRNAs and 20–23 nt RNAs derived from the EGFP coding sequence (EGFP) and from the *Mos* inverted repeat (MosIR) are shown on the left. The graph shows distribution of 17–25 nt small RNAs derived from the *Mos* inverted repeat. (**B**) Deep sequencing of small RNAs from 293 cells transfected with MosIR plasmid using SOLiD technology. Data are organized as in panel A. (**C**) Positions of putative MosIR siRNAs along the MosIR hairpin. Graph depicts cumulative counts of 5′-ends of putative MosIR-derived siRNAs (20–24 nt long reads) found in the transgenic brain and 293 cells transiently transfected with the MosIR transgene.

help the dsRNA to be tolerated in somatic cells and negatively affect processing by Dicer. It was shown that long single-stranded overhangs can contribute to the reduced efficiency of Dicer processing *in vitro* (57). Nuclear retention of edited MosIR dsRNA probably does not significantly affect MosIR processing by Dicer. As it will be discussed further below, RNA editing observed in transgenic tissues was below the reported threshold for nuclear retention. In addition, we have found adenosine-deaminated MosIR RNA in the polysomal fraction in transgenic MEFs.

The MosIR expression does not induce transcriptional silencing

Short RNAs produced by Dicer in some model species also cause transcriptional silencing by inducing formation of transcriptionally repressive chromatin [reviewed in Ref

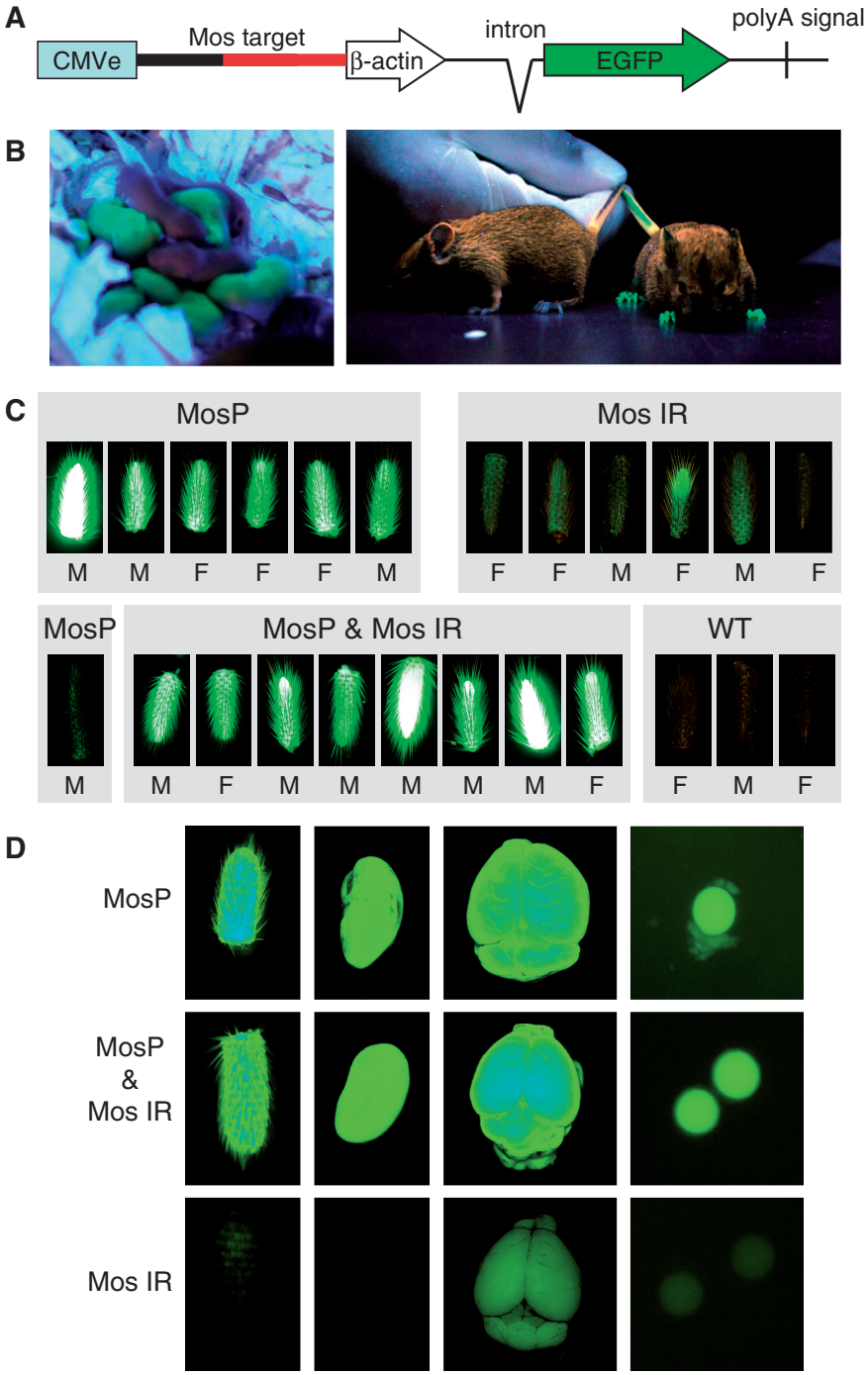


Figure 4. MosIR has no effect on transcriptional silencing of cognate sequences. (**A**) A schematic composition of the MosP reporter transgene. (**B**) EGFP signal in MosP-positive and WT newborn and adult mice. (**C**) EGFP fluorescence in tails of mice carrying MosP and MosIR transgenes is not reduced. Bright signal in some MosP-positive tails (e.g. first on the left in the top row) is caused by the lack of pigmentation of the tails. The MosP tail on the left in the lower row belongs to an F1 animal where we observed a spontaneous transcriptional silencing accompanied by DNA methylation of the MosP promoter (Supplementary Figure S4). All images were obtained using the same settings. (**D**) EGFP signal in tail, kidney, brain and oocytes of MosP, MosIR and MosP & MosIR mice. Organs were collected from siblings of one litter and their images were taken with the same camera settings.

(58,59)]. Although induction of transcriptional silencing by siRNAs has been reported for mammals (e.g. 60–62), the nature of the mechanism is still debated. We have previously shown that the MosIR transcript does not induce DNA methylation in the oocyte (63). However, that experiment was based on oocyte-specific expression of dsRNA targeting the endogenous *Mos* gene coding sequence and not a promoter, which would be a better suited target for studying transcriptional silencing. To address this caveat, we generated another reporter, named MosP, where the cognate sequence was inserted into the chimeric promoter, between the CMV enhancer

and the β-actin core promoter (Figure 4A). The inserted sequence had a minimal effect on the activity of the promoter as estimated by FACS analysis of cultured cells transfected by equimolar amounts of MosP and the parental pCAGEGFP plasmid devoid of *Mos* insertion (data not shown).

Transgenic mice carrying the MosP transgene displayed strong EGFP expression (Figure 4B). In the F1 progeny of the MosP founder male and a wild-type female, we observed one case of spontaneous silencing of the reporter, which was accompanied by DNA methylation of the *Mos* sequence in the promoter (Figure 4C and Supplementary Figure S4). This event suggested that the transgene is prone to silencing involving epigenetic changes to the *Mos* sequence in the promoter. However, when MosP animals were crossed with MosIR animals, the MosP reporter was never silenced in the presence of the MosIR transgene (Figure 4C and D). These results showed that the MosIR transgene did not induce transcriptional silencing of homologous sequences in somatic cells. Likewise, siRNAs generated in oocytes and early embryos did not induce transcriptionally repressive chromatin at the MosP promoter, which would be propagated into the soma. These results expand the list of observations that question the nature of a putative mammalian RNA silencing pathway, which would employ siRNAs to induce transcriptionally silent chromatin *in trans*, a process commonly observed in plants (64–66).

MosIR expression does not induce the IFN response in somatic cells

It is generally believed that long dsRNA causes sequence-independent effects in somatic cells due to the activation of a complex set of pathways generally referred to as the IFN response. However, it is unlikely that the MosIR transgene induced the typical IFN response because the MosIR mice developed normally and showed no obvious phenotype. We used qPCR to analyze transcriptional activation of several genes (*Ifit1*, *Oas1*, *Pkr*, *Ddx58*) associated with the IFN response in the brain, kidney, liver and spleen (Figure 5A). Except for *Ifit1* in the brain (3.6-fold up) and *Oas1* in the liver (increased in one of the three animals), we did not find increased expression of ISGs. Together, these data suggested that the MosIR transcript did not elicit a general IFN response in somatic cells of the MosIR mice.

To further examine IFN activation by the MosIR transgene, we performed a series of transfections of the MosIR plasmid into 293 cells and analyzed several common markers of the IFN response. When 293 cells were examined 48 h post-transfection (this time point was chosen to emulate the situation in transgenic mice where MosIR was continuously expressed), there was little, if any, induction of ISGs even at very high doses of MosIR. In control experiments, we could detect up-regulation of several ISGs in response to poly I:C, a commonly used ISG inducer (Figure 5B). Likewise, phosphorylation of PKR, a common marker of response to dsRNA, was seen only at very high doses of the transfected MosIR vector (Figure 5C).

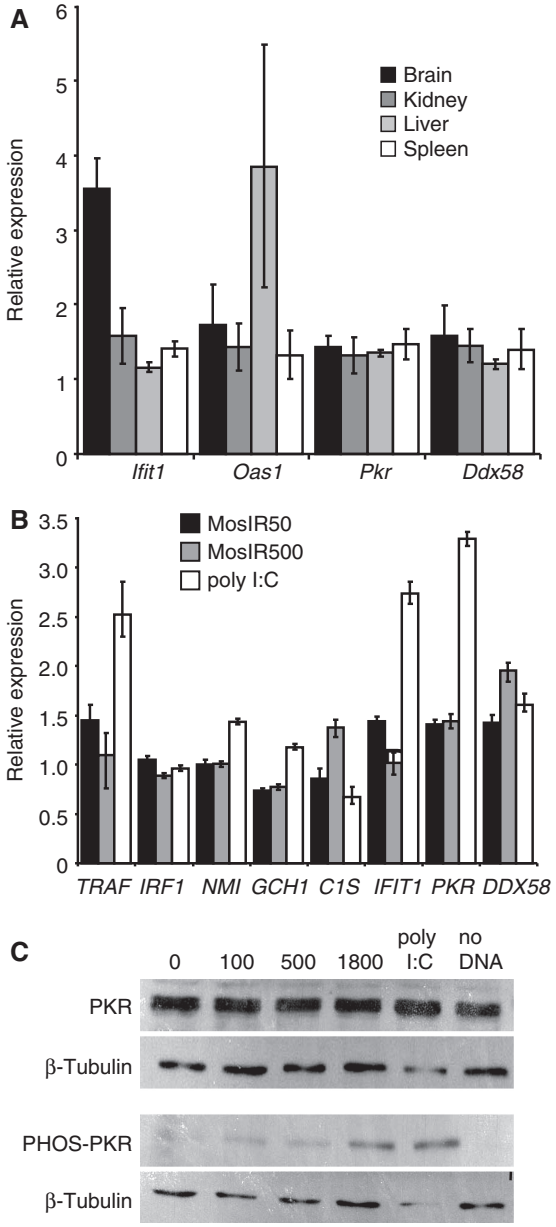


Figure 5. Analysis of the IFN response genes in mice and cell lines. (**A**) Analysis of expression of IFN stimulated genes (ISGs) *Ifit1*, *Oas1*, *Pkr* and *Ddx58* in different tissues of MosIR mice. The expression was estimated by qPCR and is shown relative to wild-type siblings. For each gene, samples from three MosIR animals were analyzed in duplicates. *Hprt* was used as an internal standard. Error bars = SEM. (**B**) Analysis of expression of ISGs in 293 cells transfected with 50 and 500 ng (per well in a 24-well plate, the total amount of the transfected DNA was kept constant by adding pCAGEGFP DNA) of the MosIR plasmid 48 h after transfection. The expression was estimated by qPCR and is shown relative to 293 cells transfected with 100 ng/well of the pCAGEGFP plasmid. For comparison, ISGs were stimulated by adding poly I:C to the media. *HPRT1* was used as an internal standard. The experiment was performed in triplicates. Error bars = SEM. (**C**) Western blot analysis of induction of PKR phosphorylation by the MosIR plasmid. Cells in 6-well plates were transfected with 100, 500 or 1800 ng of the MosIR plasmid and harvested for western blot analysis 24 h after transfection. The total amount of the transfected DNA was adjusted to 2 μg in all samples by adding corresponding amount of pCAGEGFP DNA. The majority of transfected cells showed green fluorescence (data not shown), toxicity of the MosIR to the cells was not apparent. Poly I:C was added to the media to a final concentration of 1 μg/ml and it was apparently toxic to the cells.

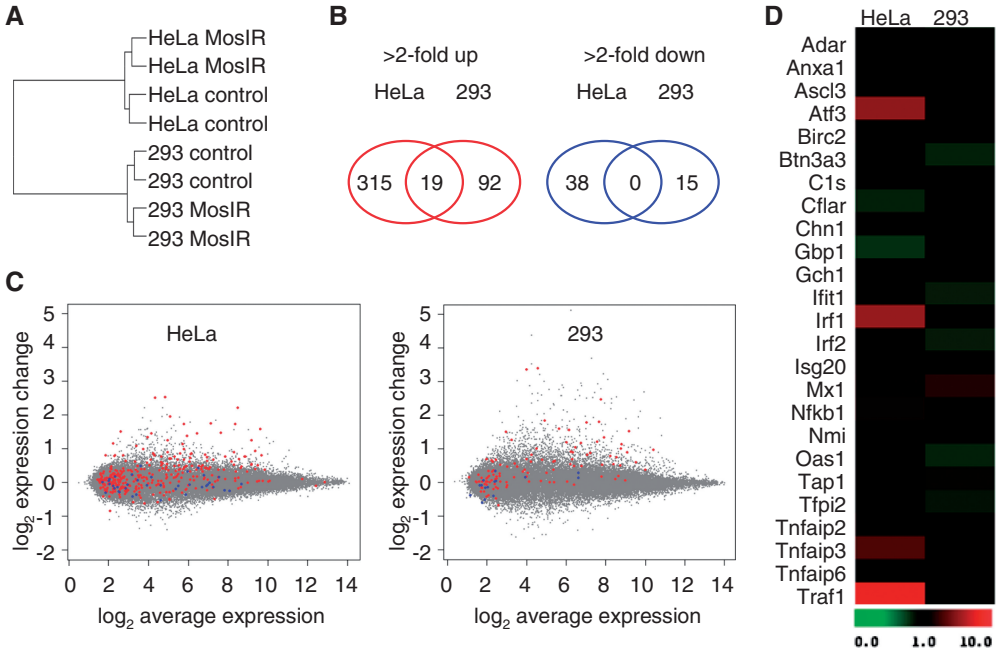


Figure 6. Microarray analysis of HeLa and 293 cells expressing MosIR. Microarray analysis of HeLa and 293 cells transfected with MosIR plasmid (MosIR) or pCAGEGFP plasmid (control). (**A**) A hierarchical clustering of microarray samples using probe sets differentially expressed (>2-fold change) between MosIR and control samples in both lines shows good concordance of duplicates and clustering according to the cell type. This suggests that there is not a common set of genes induced by MosIR in these two cell lines. (**B**) A minimal overlap of differentially expressed genes in HeLa and 293 cells. Genes corresponding to commonly up-regulated probe sets are listed in Supplementary Table SI. (**C**) MA plots of HeLa and 293 transcriptome change upon MosIR expression document small and distinct changes in hybridization intensities of probe sets. The X axis shows Affymetrix probe hybridization signal intensities (log_2 average expression); the Y axis, shows relative changes in probe intensity between MosIR and control samples. Each point represents one probe set. Red and blue color highlight probe sets significantly up-regulated and down-regulated in the other cell line. (**D**) Absence of dsRNA-induced transcriptome signature in HeLa and 293 cells expressing long dsRNA. Genes previously identified as genes induced by dsRNA treatment are shown (18).

To assess in more detail the potential of the MosIR transgene to induce ISGs, we performed microarray analysis of 293 and HeLa cells transfected with the same amount of the MosIR plasmid as in the earlier described SOLiD analysis (Figure 6). Clustering analysis based on differentially expressed genes suggested that there was no common transcriptome signature in cells expressing dsRNA (Figure 6A). Overall, the number of genes with altered expression upon transfection of the MosIR plasmid was rather small and only 19 probe sets (corresponding to 17 genes) were changed more than 2-fold in both cell lines (Figure 6B and C and Supplementary Table SI). Furthermore, detailed examination of differentially expressed genes in 293 and HeLa cells revealed little, if any, evidence for the activation of the IFN pathway (Figure 6D and Supplementary Table SII). This contrasted with our earlier results where siRNA transfection into 293 cells caused a strong up-regulation of many ISGs (67). It should be noted that we have observed up-regulation of a limited set of genes related to IFN activation in HeLa cells (*IRF1*, *IRF7*, *NFKB1*, cytokines *IL1A*, *CCL20*, *CXCL3* and TNF/related genes *TRAF1* and *TNFAIP3*, Supplementary Table SII). However, the transcriptome changes in HeLa cells were small and did not resemble typical transcriptome changes induced either by dsRNA (18) or caused by siRNA transfection in 293 cells (67) (Figure 6C and D).

Taken together, our data show that somatic cells tolerated MosIR expression unexpectedly well. It appears that the IFN response is only elicited when the amount of expressed dsRNA reaches a certain threshold. The tolerance to expressed dsRNA seems to be significant, considering the amounts of MosIR plasmid used in our experiments.

Adenosine deamination of MosIR RNA

Adenosine deamination has been proposed to target nuclear dsRNA and to oppose the RNAi pathway [reviewed in Ref (68,69)]. To estimate levels of adenosine deamination of MosIR transcripts in different cell types, we analyzed the 3′-end of the MosIR hairpin region and neighboring 3′ single-stranded overhang by cloning and sequencing RT–PCR products (Figure 7). This procedure identified edited adenosines as A/G conversions in the sequence. We found a minimal number of A/G conversions in MosIR RNA obtained from kidney, liver and brain of transgenic mice (Figure 7; data not shown). These A/G conversions were mostly found in the region of the predicted RNA duplex indicating that the MosIR dsRNA was likely to be edited (Figure 7).

It has been reported that editing induces nuclear retention. However, the editing frequency observed in our experiments was much lower than editing frequencies implicated in nuclear retention of edited dsRNA in

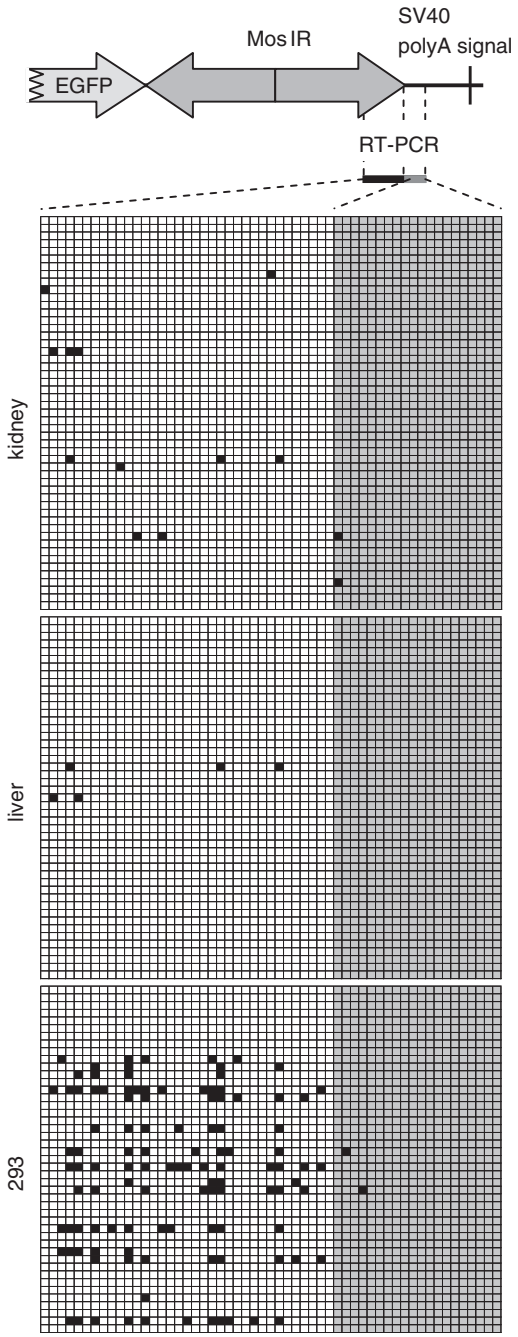


Figure 7. Analysis of adenosine deamination of MosIR RNA. Adenosine deamination was analyzed by cloning RT–PCR products from the MosIR region indicated on the top. Results of editing analysis of MosIR transcripts from transgenic kidney and liver and MosIR-transfected 293 cells are shown below. Each row of squares represents one sequenced clone. Each square represents an individual adenosine in the analyzed sequence. A part of the PCR product is derived from the stem (dsRNA) sequence (white squares) and a part from single-stranded sequence (gray squares). A/G conversions are indicated by black squares.

Xenopus oocytes and mammalian cells (70). In addition, editing analysis in nuclear and cytoplasmic fractions from MosIR MEFs did not reveal a strong evidence for nuclear retention. We have found only 3 of 24 clones and 1 of 23 clones containing A/G conversions in nuclear and cytoplasmic fractions, respectively (Supplementary Figure S5A). Although the low frequency of editing does not allow for precise estimation of enrichment of MosIR in the nuclear fraction, these data suggest that nuclear retention is not strong. We also examined the degree of editing in MosIR transcripts found in the polysomal fraction of MEF cells derived from the MosIR mice. We found that 5 of 23 clones from the polysomal fraction contained A/G conversions (Supplementary Figure S5B–D). Edited clones contained on average 14% of edited adenosines (the most edited sequence had converted 24% of adenosines). These data suggested that editing did not prevent nuclear export and translation. Furthermore, consistent with our results, endogenous human mRNAs with edited 3′-UTRs were found previously to associate with polysomes (25).

Relatively abundant editing of MosIR RNA was observed in 293 cells transfected with the MosIR plasmid. It was evident in about one-third of sequenced clones where, on average, 21% of adenosines were edited (Figure 7). Consistent with previous studies on adenosine deamination, A/G conversion was not randomly distributed along the sequence, but showed specific sequence preferences (71,72). These data further strengthen our conclusion that MosIR transcript forms dsRNA *in vivo*. However, editing apparently affects only a fraction of MosIR dsRNA formed in somatic cells. First, there was no enrichment of editing in RNase T1-resistant MosIR RNA from transgenic MEFs (data not shown). Second, the frequency of A/G conversions in putative MosIR siRNAs was comparable with the editing observed in longer MosIR transcripts (Figures 7 and 8). Analysis of SOLiD data from 293 cells showed that 15% of MosIR-derived small RNAs of 21–23 nt in length contained A/G conversions, while the conversion rate was negligible in small RNAs of other lengths or EGFP sequence-derived ones (Figure 8).

The low A/G conversion rate observed in MosIR-derived small RNAs of other lengths than putative Dicer products is presumably due to increased stability of edited, Dicer-produced, siRNAs. Although we cannot exclude that editing of siRNAs takes place after Dicer processing, this scenario seems less likely than editing of the Dicer substrate. Dicer-mediated cleavage is coupled with RISC loading (73) and there is no biochemical evidence suggesting the association of ADAR with the RISC (74).

In summary, our data suggest that the MosIR forms dsRNA *in vivo* and that a fraction of dsRNA is edited. Since editing frequency is similar in MosIR transcripts and siRNAs derived from them, we conclude that Dicer processes a small fraction of the MosIR dsRNA hairpin regardless of whether they are edited or not. Therefore, the low amount of MosIR siRNAs in somatic cells is likely to be caused by inefficient processing of dsRNA by Dicer with editing having, at best, a secondary role as a factor responsible for the ineffective formation of endogenous siRNAs. While editing of MosIR dsRNA apparently does not appear to interfere with Dicer processing, it could still negatively influence efficiency of RNAi, since editing changes the siRNA substrate specificity.

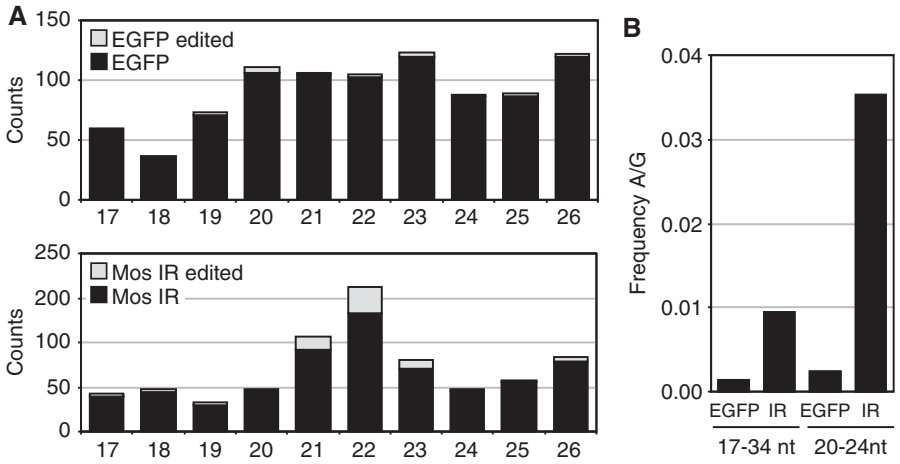


Figure 8. Analysis of adenosine deamination of MosIR-derived siRNAs. (**A**) Distribution of putatively edited SOLiD-identified small RNAs originating from EGFP and *Mos* inverted repeat (MosIR) sequences in 293 cells. Small RNAs are sorted along the *X*-axis according to their length (17–26 nt).The *Y*-axis shows the number of clones with EGFP- (upper graph) or MosIR-derived sequences (lower graph). The gray portion of each column indicates the fraction of siRNAs carrying up to four A/G sequence changes. (**B**) Putative editing is most pronounced in SOLiD-identified siRNA-like RNAs derived from the MosIR sequence. The graph shows A/G mismatch frequency found in the total population of small RNAs and 20–24 nt long RNAs derived from EGFP or MosIR sequences in 293 cells.

CONCLUSIONS

We have established a transgenic mouse model that ubiquitously expresses a long RNA hairpin. There are several independent lines of evidence suggesting that expression of the MosIR transgene in mammalian cells produces dsRNA *in vivo*. At the same time, it appears that somatic cells are able to reduce dsRNA formation, either by preventing dsRNA duplex formation or by unwinding it. In contrast to the common view that dsRNA in mammalian somatic cells imparts a detrimental effect, we did not observe any typical sequence-independent response associated with activation of the IFN pathway. Our study suggests that long dsRNA structures in cellular mRNAs are well tolerated and activate the IFN system only at very high concentrations. Under normal conditions, such dsRNA-containing mRNAs are partially edited and poorly processed by Dicer in somatic cells, perhaps because somatic cells lack a factor facilitating siRNA biogenesis. Effective RNAi has been observed only in oocytes, providing yet more evidence that the female germline is a privileged tissue in terms of directing dsRNA into the RNAi pathway. Altogether, our data support a model that the effects of long dsRNA in mammalian cells are influenced by additional factors, which determine how cells respond to such dsRNA. These factors certainly involve the cellular history of dsRNA and structural features, such as 5′-end modifications and length of sequences flanking dsRNA stem.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Jean-Francois Spetz for production of transgenic mice, Edward Oakeley for microarray analysis, Michael Jantsch for antibodies, Susan Wagner for help with polysome profiling and Radek Sindelka for help with single-cell qPCR.

FUNDING

Czech Science Foundation (grant GACR 204/09/0085); EMBO SDIG (program #1488, partial); Purkynje Fellowship (to P.S., partial); European Community Framework Seven Program Integrated Project "Sirocco" (Research in the laboratory to W.F.); Novartis Research Foundation (to Friedrich Miescher Institute). Funding for the open access charge: GACR 204/09/0085.

Conflict of interest statement. None declared.

REFERENCES

1. Tian,B., Bevilacqua,P.C., Diegelman-Parente,A. and Mathews,M.B. (2004) The double-stranded-RNA-binding motif: interference and much more. *Nat. Rev. Mol. Cell Biol.*, **5**, 1013–1023.
2. Carthew,R.W. and Sontheimer,E.J. (2009) Origins and Mechanisms of miRNAs and siRNAs. *Cell*, **136**, 642–655.
3. Ghildiyal,M. and Zamore,P.D. (2009) Small silencing RNAs: an expanding universe. *Nat. Rev. Genet.*, **10**, 94–108.
4. Elbashir,S.M., Harborth,J., Lendeckel,W., Yalcin,A., Weber,K. and Tuschl,T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, **411**, 494–498.
5. Hutvagner,G. and Zamore,P.D. (2002) A microRNA in a multiple-turnover RNAi enzyme complex. *Science*, **297**, 2056–2060.
6. Kim,V.N., Han,J. and Siomi,M.C. (2009) Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.*, **10**, 126–139.
7. Fire,A., Xu,S., Montgomery,M.K., Kostas,S.A., Driver,S.E. and Mello,C.C. (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, **391**, 806–811.
8. Obbard,D.J., Gordon,K.H., Buck,A.H. and Jiggins,F.M. (2009) The evolution of RNAi as a defence against viruses and

transposable elements. *Philos. Trans. R Soc. Lond. B Biol. Sci.*, **364**, 99–115.
9. Umbach,J.L. and Cullen,B.R. (2009) The role of RNAi and microRNAs in animal virus replication and antiviral immunity. *Genes Dev.*, **23**, 1151–1164.
10. Svoboda,P. and Flemr,M. (2010) The role of miRNAs and endogenous siRNAs in maternal-to-zygotic reprogramming and the establishment of pluripotency. *EMBO Rep.*, **11**, 590–597.
11. Shinagawa,T. and Ishii,S. (2003) Generation of Ski-knockdown mice by expressing a long double-strand RNA from an RNA polymerase II promoter. *Genes Dev.*, **17**, 1340–1345.
12. Diallo,M., Arenz,C., Schmitz,K., Sandhoff,K. and Schepers,U. (2003) Long endogenous dsRNAs can induce complete gene silencing in mammalian cells and primary cultures. *Oligonucleotides*, **13**, 381–392.
13. Yi,C.E., Bekker,J.M., Miller,G., Hill,K.L. and Crosbie,R.H. (2003) Specific and potent RNA interference in terminally differentiated myotubes. *J. Biol. Chem.*, **278**, 934–939.
14. Bhargava,A., Dallman,M.F., Pearce,D. and Choi,S. (2004) Long double-stranded RNA-mediated RNA interference as a tool to achieve site-specific silencing of hypothalamic neuropeptides. *Brain Res. Brain Res. Protoc.*, **13**, 115–125.
15. Hunter,T., Hunt,T., Jackson,R.J. and Robertson,H.D. (1975) The characteristics of inhibition of protein synthesis by double-stranded ribonucleic acid in reticulocyte lysates. *J. Biol. Chem.*, **250**, 409–417.
16. Meurs,E., Chong,K., Galabru,J., Thomas,N.S., Kerr,I.M., Williams,B.R. and Hovanessian,A.G. (1990) Molecular cloning and characterization of the human double-stranded RNA-activated protein kinase induced by interferon. *Cell*, **62**, 379–390.
17. Sadler,A.J. and Williams,B.R. (2007) Structure and function of the protein kinase R. *Curr. Top. Microbiol. Immunol.*, **316**, 253–292.
18. Geiss,G., Jin,G., Guo,J., Bumgarner,R., Katze,M.G. and Sen,G.C. (2001) A comprehensive view of regulation of gene expression by double-stranded RNA-mediated cell signaling. *J. Biol. Chem.*, **276**, 30178–30182.
19. Gantier,M.P. and Williams,B.R. (2007) The response of mammalian cells to double-stranded RNA. *Cytokine Growth Factor Rev.*, **18**, 363–371.
20. Sadler,A.J. and Williams,B.R. (2008) Interferon-inducible antiviral effectors. *Nat. Rev. Immunol.*, **8**, 559–568.
21. DeCerbo,J. and Carmichael,G.G. (2005) Retention and repression: fates of hyperedited RNAs in the nucleus. *Curr. Opin. Cell Biol.*, **17**, 302–308.
22. Bass,B.L. (2006) How does RNA editing affect dsRNA-mediated gene silencing? *Cold Spring Harb. Symp. Quant. Biol.*, **71**, 285–292.
23. Zhang,Z. and Carmichael,G.G. (2001) The fate of dsRNA in the nucleus: a p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell*, **106**, 465–475.
24. Scadden,A.D. (2005) The RISC subunit Tudor-SN binds to hyper-edited double-stranded RNA and promotes its cleavage. *Nat. Struct. Mol. Biol.*, **12**, 489–496.
25. Hundley,H.A., Krauchuk,A.A. and Bass,B.L. (2008) C. elegans and H. sapiens mRNAs with edited 3' UTRs are present on polysomes. *RNA*, **14**, 2050–2060.
26. Knight,S.W. and Bass,B.L. (2002) The role of RNA editing by ADARs in RNAi. *Mol. Cell*, **10**, 809–817.
27. Yang,W., Wang,Q., Howell,K.L., Lee,J.T., Cho,D.S., Murray,J.M. and Nishikura,K. (2005) ADAR1 RNA deaminase limits short interfering RNA efficacy in mammalian cells. *J. Biol. Chem.*, **280**, 3946–3953.
28. Heale,B.S., Keegan,L.P., McGurk,L., Michlewski,G., Brindle,J., Stanton,C.M., Caceres,J.F. and O'Connell,M.A. (2009) Editing independent effects of ADARs on the miRNA/siRNA pathways. *EMBO J.*, **28**, 3145–3156.
29. Wang,Q. and Carmichael,G.G. (2004) Effects of length and location on the cellular response to double-stranded RNA. *Microbiol. Mol. Biol. Rev.*, **68**, 432–452.
30. Gan,L., Anton,K.E., Masterson,B.A., Vincent,V.A., Ye,S. and Gonzalez-Zulueta,M. (2002) Specific interference with gene expression and gene function mediated by long dsRNA in neural cells. *J. Neurosci. Methods*, **121**, 151–157.
31. Tran,N., Raponi,M., Dawes,I.W. and Arndt,G.M. (2004) Control of specific gene expression in mammalian cells by co-expression of long complementary RNAs. *FEBS Lett.*, **573**, 127–134.
32. Stein,P., Svoboda,P. and Schultz,R.M. (2003) Transgenic RNAi in mouse oocytes: a simple and fast approach to study gene function. *Dev. Biol.*, **256**, 187–193.
33. Stein,P., Zeng,F., Pan,H. and Schultz,R.M. (2005) Absence of non-specific effects of RNA interference triggered by long double-stranded RNA in mouse oocytes. *Dev. Biol.*, **286**, 464–471.
34. Gebauer,F. and Richter,J.D. (1997) Synthesis and function of Mos: the control switch of vertebrate oocyte meiosis. *Bioessays*, **19**, 23–28.
35. Colledge,W.H., Carlton,M.B., Udy,G.B. and Evans,M.J. (1994) Disruption of c-mos causes parthenogenetic development of unfertilized mouse eggs. *Nature*, **370**, 65–68.
36. Hashimoto,N., Watanabe,N., Furuta,Y., Tamemoto,H., Sagata,N., Yokoyama,M., Okazaki,K., Nagayoshi,M., Takeda,N., Ikawa,Y. *et al.* (1994) Parthenogenetic activation of oocytes in c-mos-deficient mice (published erratum appears in Nature 1994, **370**, 391). *Nature*, **370**, 68–71.
37. Svoboda,P., Stein,P. and Schultz,R.M. (2001) RNAi in mouse oocytes and preimplantation embryos: effectiveness of hairpin dsRNA. *Biochem. Biophys. Res. Commun.*, **287**, 1099–1104.
38. Kaname,T. and Huxley,C. (2001) Simple and efficient vectors for retrofitting BACs and PACs with mammalian neoR and EGFP marker genes. *Gene*, **266**, 147–153.
39. Sarnova,L., Malik,R., Sedlacek,R. and Svoboda,P. (2010) Shortcomings of short hairpin RNA-based transgenic RNA interference in mouse oocytes. *J. Negat. Results Biomed.*, **9**, 8.
40. Svoboda,P., Stein,P., Hayashi,H. and Schultz,R.M. (2000) Selective reduction of dormant maternal mRNAs in mouse oocytes by RNA interference. *Development*, **127**, 4147–4156.
41. Schmitter,D., Filkowski,J., Sewer,A., Pillai,R.S., Oakeley,E.J., Zavolan,M., Svoboda,P. and Filipowicz,W. (2006) Effects of Dicer and Argonaute down-regulation on mRNA levels in human HEK293 cells. *Nucleic Acids Res.*, **34**, 4801–4815.
42. Smyth,G.K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.*, **3**, Article3.
43. Saeed,A.I., Sharov,V., White,J., Li,J., Liang,W., Bhagabati,N., Braisted,J., Klapa,M., Currier,T., Thiagarajan,M. *et al.* (2003) TM4: a free, open-source system for microarray data management and analysis. *Biotechniques*, **34**, 374–378.
44. Griffiths-Jones,S., Saini,H.K., van Dongen,S. and Enright,A.J. (2008) miRBase: tools for microRNA genomics. *Nucleic Acids Res.*, **36**, D154–D158.
45. Chan,P.P. and Lowe,T.M. (2009) GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res.*, **37**, D93–D97.
46. Pruesse,E., Quast,C., Knittel,K., Fuchs,B.M., Ludwig,W., Peplies,J. and Glockner,F.O. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.*, **35**, 7188–7196.
47. Pfaffl,M.W., Horgan,G.W. and Dempfle,L. (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res.*, **30**, e36.
48. Okabe,M., Ikawa,M., Kominami,K., Nakanishi,T. and Nishimune,Y. (1997) 'Green mice' as a source of ubiquitous green cells. *FEBS Lett.*, **407**, 313–319.
49. Wu,C., Orozco,C., Boyer,J., Leglise,M., Goodale,J., Batalov,S., Hodge,C.L., Haase,J., Janes,J., Huss,J.W. 3rd *et al.* (2009) BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol.*, **10**, R130.
50. Feix,G., Slor,H. and Weissmann,C. (1967) Replication of viral RNA. 13. The early product of phage RNA synthesis in vitro. *Proc. Natl Acad. Sci. USA*, **57**, 1401–1408.
51. Tam,O.H., Aravin,A.A., Stein,P., Girard,A., Murchison,E.P., Cheloufi,S., Hodges,E., Anger,M., Sachidanandam,R., Schultz,R.M. *et al.* (2008) Pseudogene-derived small interfering

RNAs regulate gene expression in mouse oocytes. *Nature*, **453**, 534–538.
52. Watanabe,T., Totoki,Y., Toyoda,A., Kaneda,M., Kuramochi-Miyagawa,S., Obata,Y., Chiba,H., Kohara,Y., Kono,T., Nakano,T. *et al*. (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature*, **453**, 539–543.
53. Mineno,J., Okamoto,S., Ando,T., Sato,M., Chono,H., Izu,H., Takayama,M., Asada,K., Mirochnitchenko,O., Inouye,M. *et al*. (2006) The expression profile of microRNAs in mouse embryos. *Nucleic Acids Res*., **34**, 1765–1771.
54. Cummins,J.M., He,Y., Leary,R.J., Pagliarini,R., Diaz,L.A. Jr., Sjoblom,T., Barad,O., Bentwich,Z., Szafranska,A.E., Labourier,E. *et al*. (2006) The colorectal microRNAome. *Proc. Natl Acad. Sci. USA*, **103**, 3687–3692.
55. Yi,R., Pasolli,H.A., Landthaler,M., Hafner,M., Ojo,T., Sheridan,R., Sander,C., O'Carroll,D., Stoffel,M., Tuschl,T. *et al*. (2009) DGCR8-dependent microRNA biogenesis is essential for skin development. *Proc. Natl Acad. Sci. USA*, **106**, 498–502.
56. Smalheiser,N.R., Lugli,G., Thimmapuram,J., Cook,E.H. and Larson,J. (2011) Endogenous siRNAs and noncoding RNA-derived small RNAs are expressed in adult mouse hippocampus and are up-regulated in olfactory discrimination training. *RNA*, **17**, 166–181.
57. Bernstein,E., Caudy,A.A., Hammond,S.M. and Hannon,G.J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, **409**, 363–366.
58. Grewal,S.I. and Elgin,S.C. (2007) Transcription and RNA interference in the formation of heterochromatin. *Nature*, **447**, 399–406.
59. Moazed,D. (2009) Small RNAs in transcriptional gene silencing and genome defence. *Nature*, **457**, 413–420.
60. Morris,K.V., Chan,S.W., Jacobsen,S.E. and Looney,D.J. (2004) Small interfering RNA-induced transcriptional gene silencing in human cells. *Science*, **305**, 1289–1292.
61. Ting,A.H., Schuebel,K.E., Herman,J.G. and Baylin,S.B. (2005) Short double-stranded RNA induces transcriptional gene silencing in human cancer cells in the absence of DNA methylation. *Nat. Genet*., **37**, 906–910.
62. Janowski,B.A., Huffman,K.E., Schwartz,J.C., Ram,R., Hardy,D., Shames,D.S., Minna,J.D. and Corey,D.R. (2005) Inhibiting gene expression at transcription start sites in chromosomal DNA with antigene RNAs. *Nat. Chem. Biol*., **1**, 216–222.
63. Svoboda,P., Stein,P., Filipowicz,W. and Schultz,R.M. (2004) Lack of homologous sequence-specific DNA methylation in response to stable dsRNA expression in mouse oocytes. *Nucleic Acids Res*., **32**, 3601–3606.
64. Moses,J., Goodchild,A. and Rivory,L.P. (2010) Intended transcriptional silencing with siRNA results in gene repression through sequence-specific off-targeting. *RNA*, **16**, 430–441.
65. Haraguchi,T., Mizutani,T., Yamamichi,N., Ito,T., Minoguchi,S. and Iba,H. (2007) SiRNAs do not induce RNA-dependent transcriptional silencing of retrovirus in human cells. *FEBS Lett*., **581**, 4949–4954.
66. Wang,F., Koyama,N., Nishida,H., Haraguchi,T., Reith,W. and Tsukamoto,T. (2006) The assembly and maintenance of heterochromatin initiated by transgene repeats are independent of the RNA interference pathway in mammalian cells. *Mol. Cell Biol*., **26**, 4028–4040.
67. Svoboda,P. (2007) Off-targeting and other non-specific effects of RNAi experiments in mammalian cells. *Curr. Opin. Mol. Ther*., **9**, 248–257.
68. Nishikura,K. (2006) Editor meets silencer: crosstalk between RNA editing and RNA interference. *Nat. Rev. Mol. Cell Biol*., **7**, 919–931.
69. Hundley,H.A. and Bass,B.L. (2010) ADAR editing in double-stranded UTRs and other noncoding RNA sequences. *Trends Biochem. Sci*., **35**, 377–383.
70. Kumar,M. and Carmichael,G.G. (1997) Nuclear antisense RNA induces extensive adenosine modifications and nuclear retention of target transcripts. *Proc. Natl Acad. Sci. USA*, **94**, 3542–3547.
71. Polson,A.G. and Bass,B.L. (1994) Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase. *EMBO J*., **13**, 5701–5711.
72. Stefl,R., Oberstrass,F.C., Hood,J.L., Jourdan,M., Zimmermann,M., Skrisovska,L., Maris,C., Peng,L., Hofr,C., Emeson,R.B. *et al*. (2010) The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. *Cell*, **143**, 225–237.
73. MacRae,I.J., Ma,E., Zhou,M., Robinson,C.V. and Doudna,J.A. (2008) In vitro reconstitution of the human RISC-loading complex. *Proc. Natl Acad. Sci. USA*, **105**, 512–517.
74. Hock,J., Weinmann,L., Ender,C., Rudel,S., Kremmer,E., Raabe,M., Urlaub,H. and Meister,G. (2007) Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. *EMBO Rep*., **8**, 1052–1060.

SUPPLEMENTAL MATERIAL

dsRNA expression in the mouse elicits RNAi in oocytes and low adenosine deamination in somatic cells

Running title: dsRNA in mammals

Jana Nejepinska[1], Radek Malik[1], Jody Filkowski[2,3], Matyas Flemr[1], Witold Filipowicz [2], and Petr Svoboda[1+]

Keywords: dsRNA, RNA hairpin, RNAi, interferon, IFN, adenosine deamination, ADAR.

[1]Institute of Molecular Genetics, Videnska 1083, 14220 Praha 4, Czech Republic

[2] Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland

[2] current address: Department of Biological Sciences, University of Lethbridge, 4401 University Drive, Lethbridge, Alberta, T1K 3M4, Canada

[+]corresponding author:

Petr Svoboda (svobodap@img.cas.cz)

SUPPLEMENTARY FIGURES

Figure S1 Expression constructs

(A) A schematic overview of *Mos* transcribed sequences used for vector cloning, dsRNA expression and experimental analysis. *Mos* gene is intronless. The blue arrow depicts the coding sequence. The length of the dsRNA region is about 0.5 kb. Drawn approximately to the scale. (B) A schematic organization of MosIR, MosP, Mos3 trangenes. *Mos* sequence present in MosP and Mos3 vectors correspond to the sequence fragment shown in the panel A. The last two schemes depict expression cassettes of RL-Mos3 and FF reporter plasmids used in cell culture experiments. (C) EGFP fluorescence in MosP, Mos3 and MosIR transgenic lines used in the project. The fresh tail biopsies were photographed together.

Figure S2 Analysis of MosIR expression

(A) Relative MosIR levels in transgenic organs. The graph shows the MosIR expression displayed as $\log_2$ relative to *Hprt*, a routinely used internal standard in qPCR. MosIR expression was monitored with CagI primers (Table SIII). (B) Comparison of expression of MosIR and pCAGEGFP (parental plasmid into which was inserted the Mos inverted repeat) transcripts in cultured cells suggests similar mRNA expression from both plasmids. 293 cells grown in a 24-well plate were transfected with indicated amounts of MosIR and pCAGEGFP and the level of CAG-driven mRNA was determined 48 hours after transfection by qPCR using CagI qPCR primers. (C) MosIR produces significantly reduced levels of EGFP. 293 cells grown in a 24-well plate were transfected with indicated amounts of MosIR and pCAGEGFP and the amount of EGFP was determined 48 hours after transfection by flow cytometry. The X-axis shows green fluorescence intensity (FL1-H), the Y axis indicates the frequency of cells with a given intensity ($5x10^4$ cells were counted).

Figure S3 SOLiD analysis of transgenic organs and transfected 293 cells.

(A) Column graphs show absolute numbers of reads in three SOLiD small RNA libraries and counts of reads of different sequence lengths. Pie graphs show representation of several different classes of small RNAs in analyzed samples. (B) Distribution of 21-23 nt long reads along the MosIR transgene. Graph depicts cumulative counts of 5' ends of 21-23 nt long small RNAs. Sense and antisense RNAs are shown above and below the line, respectively. Note small RNAs generated from the plasmid backbone and antisense small RNAs from intronic and EGFP coding sequences, which indicate that the transfected MosIR plasmid is transcribed in a rather complex manner. (C) Length distribution of small RNAs from EGFP

and Mos inverted repeat (MosIR) regions of the MosIR transgene in the transgenic brain and in 293 cells transiently transfected with the MosIR transgene.

Figure S4 Methylation status of different *Mos* sequences in different genetic backgrounds

Mos DNA bisulfite sequencing was performed as described previously (1). Each row of circles represents one clone from bisulfite sequencing. White circles represent unmethylated cytosines, black circles represent methylated cytosines. Shaded area denotes a region of the *Mos* sequence that is complementary to the MosIR hairpin. Unframed data are derived from the endogenous *Mos* gene sequence while clones in red frames are derived from transgenic *Mos* sequences, which carry a mutation of one KpnI site (highlighted in the Supplementary Figure S1) allowing for distinguishing between endogenous and transgenic clones. The upper two datasets show methylation of endogenous *Mos* gene sequences in tail DNA from two independent MosIR transgenic lines. Tail EGFP fluorescence shows that the MosIR transgene is active in these tissues. In contrast to the endogenous *Mos* sequence, the *Mos* sequence inserted in the Mos3 transgene is hypermethylated in the genome. DNA methylation is not unusual for exonic sequences. *Mos* sequence in an active MosP transgene is hypomethylated. The last panel shows analysis of tail DNA from an animal showing signs of spontaneous silencing of the MosP transgene, which identified a hypermethylated clone derived from the MosP sequence.

Figure S5 Adenosine deamination in nuclear, cytoplasmic and polysomal fractions

(A) Analysis of adenosine deamination of MosIR RNA in nuclear and cytoplasmic fractions from MEFs derived from transgenic embryos. Adenosine deamination was analyzed by cloning RT-PCR products from the MosIR region indicated on the top. Lower panels show results of editing analysis of MosIR transcripts from nuclear and cytoplasmic fractions, which were prepared from MosIR MEFs as described previously (2). Each row of squares represents one sequenced clone. Each square represents an individual adenosine in the analyzed sequence. A part of the PCR product is derived from the stem (dsRNA) sequence (white squares) and a part from single stranded sequence (grey squares). A/G conversions are indicated by black squares. (B) Polysome profiling of MosIR-positive MEFs. Profiling was performed as described previously (3). Material sampling for qPCR is indicated by vertical lines. (C) qPCR analysis of relative abundance of *Hprt* and MosIR RNAs in the three fractions indicated in the panel A. Relative RNA content in the monosomal fraction was arbitrarily set to one. (D) Edited MosIR RNA is present in the polysomal fraction in transgenic MEFs. qPCR product obtained from the polysomal fraction was cloned and 23

individual clones were sequenced. Since the qPCR product localizes into the inverted repeat, two regions from both arms were apparently amplified (grey areas), which can be distinguished by a single nucleotide mutation in one of the arms. Sequenced clones were first sorted out according to their origin from sense or antisense arm and their editing (apparent as A/G conversion) was determined. The bottom panel shows five clones, in which we identified RNA editing (highlighted as black squares among white squares; each row of squares represents all adenosines in an individual sequenced clone). Origin of edited clones from the sense or antisense arms of the MosIR is indicated.

SUPPLEMENTARY TABLES

Table SI -.Probe sets differentially expressed upon dsRNA expression in 293 and HeLa cells

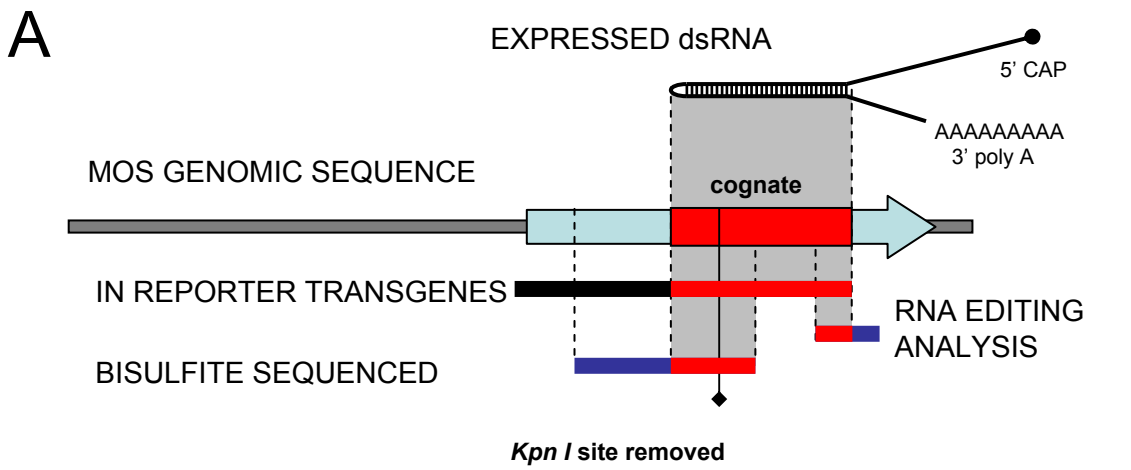
Table SII - Differentially expressed probe sets (>2-fold) in 293 and HeLa cells transfected with the MosIR plasmid.

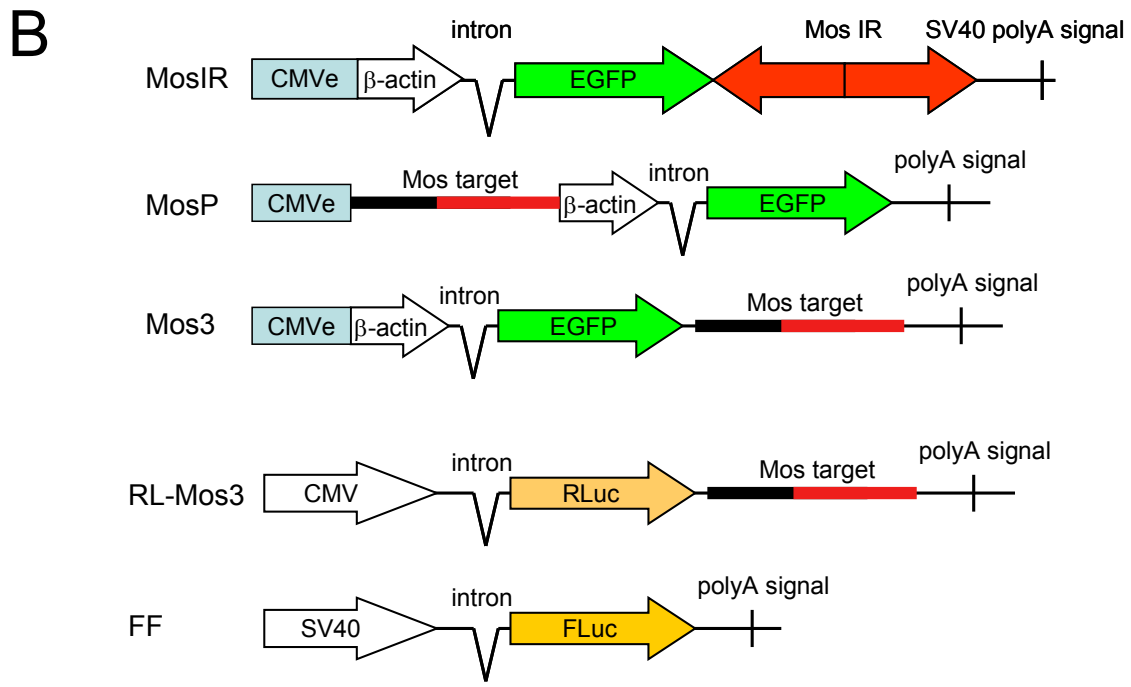
Table SIII - Sequences of primes used in this study

REFERENCES

1. Svoboda, P., Stein, P., Filipowicz, W. and Schultz, R.M. (2004) Lack of homologous sequence-specific DNA methylation in response to stable dsRNA expression in mouse oocytes. *Nucleic Acids Res*, **32**, 3601-3606.
2. Djouder, N., Metzler, S.C., Schmidt, A., Wirbelauer, C., Gstaiger, M., Aebersold, R., Hess, D. and Krek, W. (2007) S6K1-mediated disassembly of mitochondrial URI/PP1gamma complexes activates a negative feedback program that counters S6K1 survival signaling. *Mol Cell*, **28**, 28-40.
3. Chiu, W.L., Wagner, S., Herrmannova, A., Burela, L., Zhang, F., Saini, A.K., Valasek, L. and Hinnebusch, A.G. (2010) The C-terminal region of eukaryotic translation initiation factor 3a (eIF3a) promotes mRNA recruitment, scanning, and, together with eIF3j and the eIF3b RNA recognition motif, selection of AUG start codons. *Mol Cell Biol*, **30**, 4415-4434.

Figure S1


A
EXPRESSED dsRNA
5' CAP
AAAAAAAAA
3' poly A
MOS GENOMIC SEQUENCE
cognate
IN REPORTER TRANSGENES
RNA EDITING
ANALYSIS
BISULFITE SEQUENCED
Kpn I site removed


B
MosIR
CMVe
β-actin
intron
EGFP
Mos IR
SV40 polyA signal
MosP
Mos target
intron
β-actin
EGFP
polyA signal
Mos3
CMVe
β-actin
intron
EGFP
Mos target
polyA signal
RL-Mos3
CMV
intron
RLuc
Mos target
polyA signal
FF
SV40
intron
FLuc
polyA signal

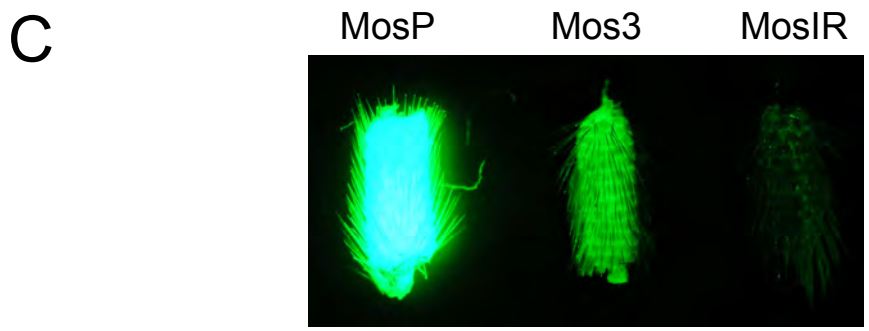

C
MosP
Mos3
MosIR

Figure S2

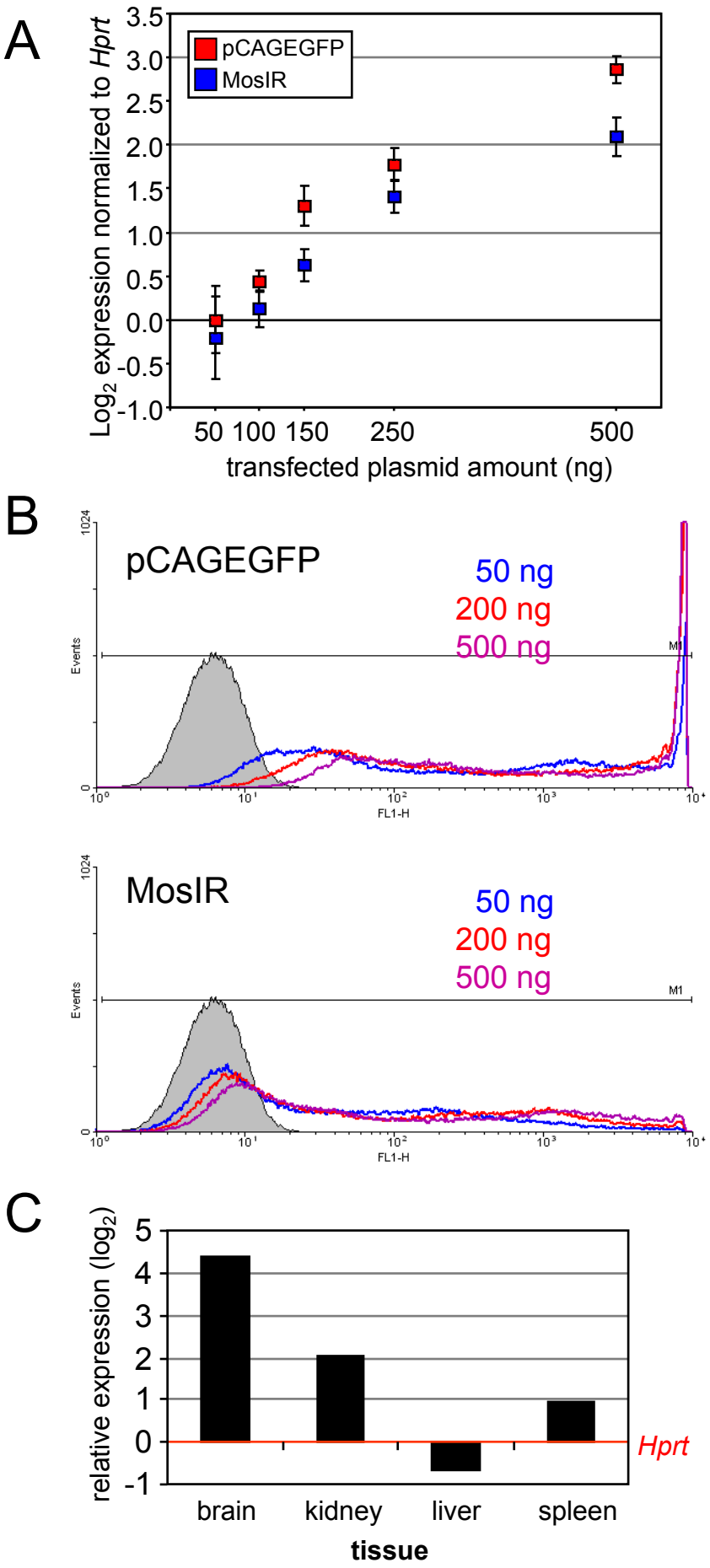

A
pCAGEGFP
MosIR
Log2 expression normalized to Hprt
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
50 100 150 250 500
transfected plasmid amount (ng)
B
pCAGEGFP
50 ng
200 ng
500 ng
M1
Events
1024
0
FL1-H
MosIR
50 ng
200 ng
500 ng
M1
Events
1024
0
FL1-H
C
relative expression (log2)
5
4
3
2
1
0
-1
Hprt
brain
kidney
liver
spleen
tissue

Figure S3

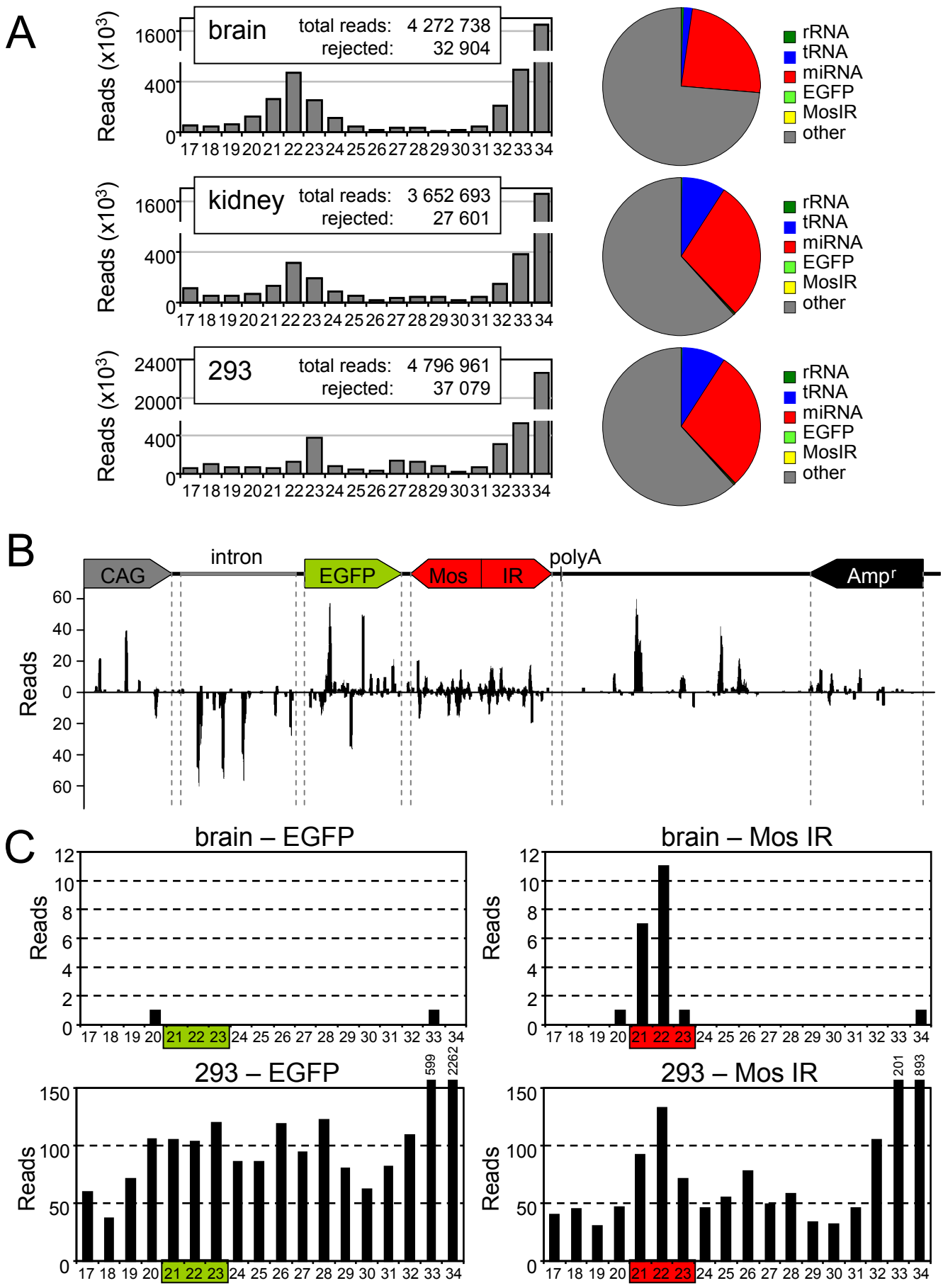

A
brain
total reads: 4 272 738
rejected: 32 904
kidney
total reads: 3 652 693
rejected: 27 601
293
total reads: 4 796 961
rejected: 37 079
Reads (x10³)
rRNA
tRNA
miRNA
EGFP
MosIR
other
B
CAG
intron
EGFP
Mos
IR
polyA
Ampʳ
Reads
C
brain – EGFP
brain – Mos IR
293 – EGFP
293 – Mos IR
599
2262
201
893

Figure S4

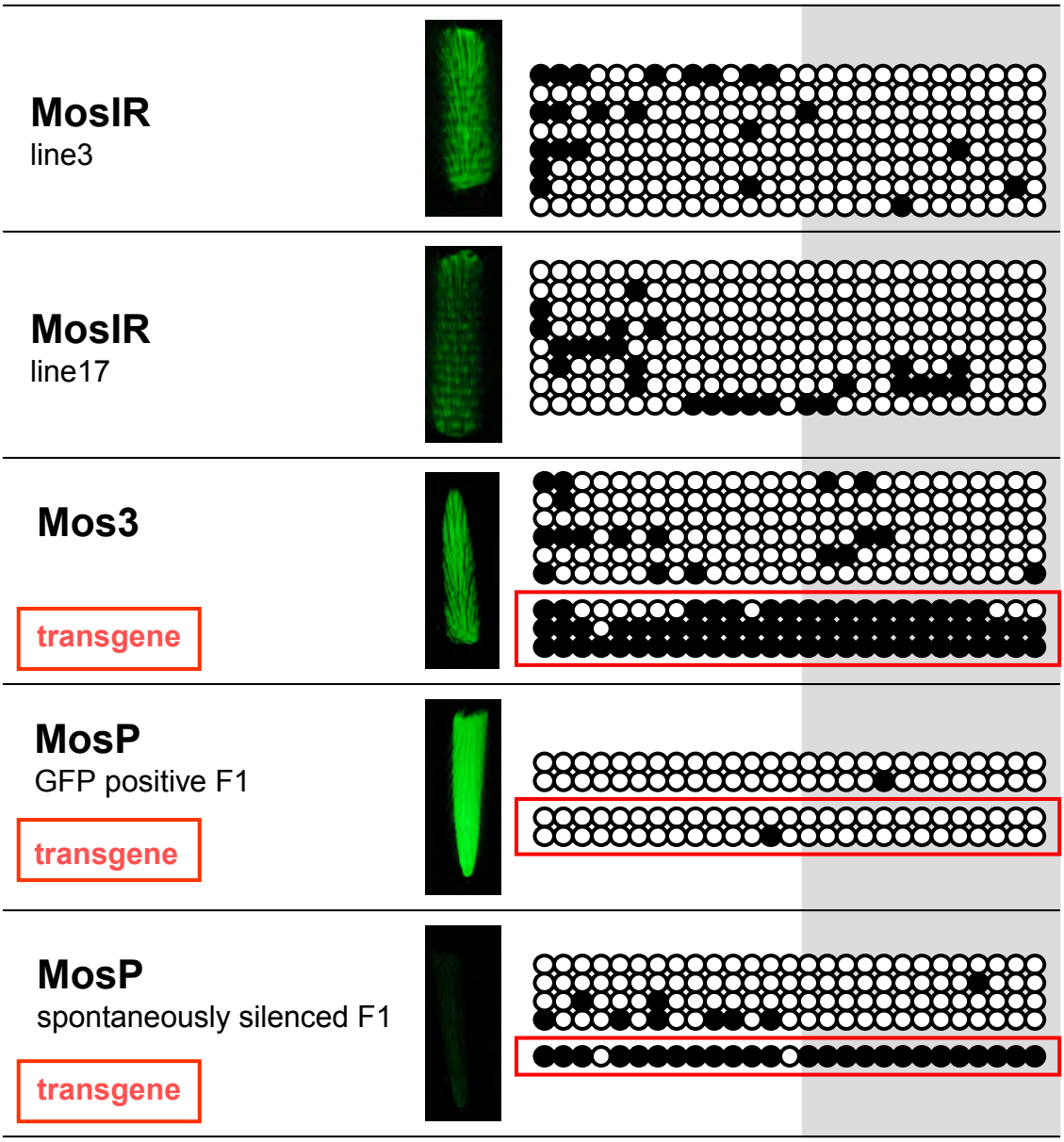

MosIR
line3
MosIR
line17
Mos3
transgene
MosP
GFP positive F1
transgene
MosP
spontaneously silenced F1
transgene

Figure S5

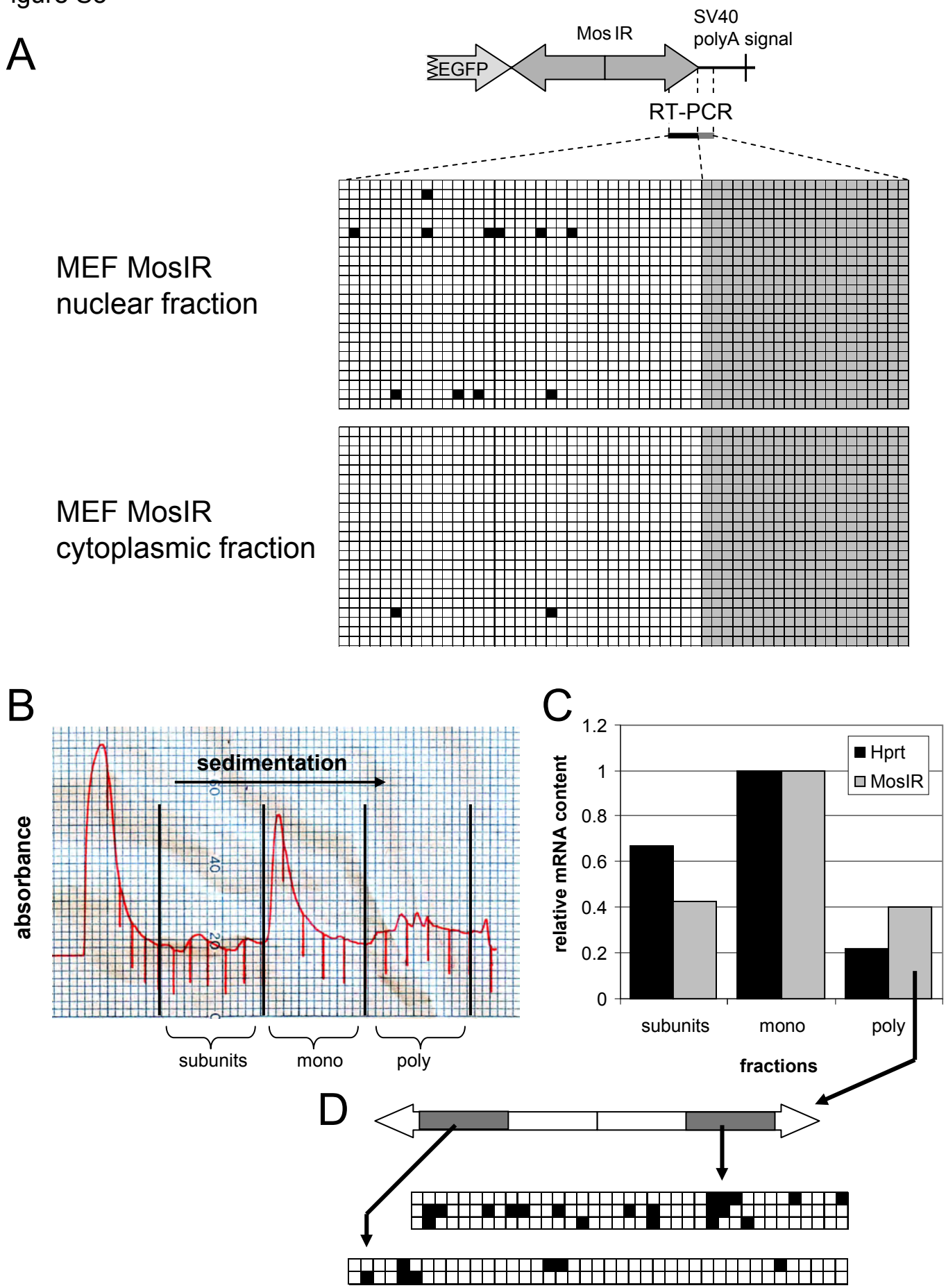

A
Mos IR
SV40
polyA signal
EGFP
RT-PCR
MEF MosIR
nuclear fraction
MEF MosIR
cytoplasmic fraction
B
sedimentation
absorbance
subunits
mono
poly
C
relative mRNA content
Hprt
MosIR
1.2
1
0.8
0.6
0.4
0.2
0
subunits
mono
poly
fractions
D

Table SI

Gene	Description	RefSeq	fold change HeLa	fold change 293
C1orf79	chromosome 1 open reading frame 79	XM_378848	5.51	4.61
CIP29	cytokine induced protein 29 kDa	NM_033082	2.53	3.35
DKFZp686L1418	hypothetical gene supported by BX538329	XM_499121	10.47	3.86
FAM24B	family with sequence similarity 24, member B	NM_152644	2.96	2.72
FAM80B	family with sequence similarity 80, member B	NM_020734	3.44	2.74
FLJ37953	hypothetical protein FLJ37953	NM_152382	2.28	2.21
HIST1H4H	histone 1, H4h	NM_003543	3.02	2.23
HIST2H4	histone 2, H4	NM_003548	2.06	5.73
INHBE	inhibin, beta E	NM_031479	2.83	3.32
KHDRBS1	KH domain containing, RNA binding, signal transduction associated	NM_006559	3.57	5.69
LOC441241	chaperonin containing TCP1, subunit 6A (zeta 1)-like	XM_496886	2.33	3.33
LOC441378	Pvt1 oncogene homolog, MYC activator (mouse), LOC441378	XM_372058	2.21	2.78
P8	p8 protein (candidate of metastasis 1)	NM_012385	2	2.61
SFRS7	splicing factor, arginine/serine-rich 7, 35kDa	NM_0010316	3.51	2.38
ST7L	suppression of tumorigenicity 7 like	NM_017744	2.02	2.37
TARS	Threonyl-tRNA synthetase	NM_152295	10.24	2.85
WDFY3	WD repeat and FYVE domain containing 3	NM_014991	2.4	2.53

Table SII

Name	Gene Symbol (Genedata An	RefSeq Transcrip	UniGene [A]	HeLa 500/0	Hek 500/0
210118_s_at	IL1A	NM_000575	Hs.1722	34.11	0.82
205476_at	CCL20	NM_004591	Hs.75498	20.72	0.97
207850_at	CXCL3	NM_002090	Hs.89690	11.91	1.11
204760_s_at	NR1D1, THRA	NM_003250	Hs.724	11.56	1.75
231412_at	DKFZp686L14188	XM_499121	Hs.150298	10.43	3.86
240206_at	TARS	NM_152295	Hs.481860	10.22	2.94
211506_s_at	IL8	NM_000584	Hs.624	9.43	1.07
205205_at	RELB	NM_006509	Hs.307905	9.38	1.76
204470_at	CXCL1	NM_001511	Hs.789	8.20	0.81
205599_at	TRAF1	NM_005658	Hs.531251	8.16	0.95
244070_at	SYNE1	NM_015293	Hs.12967	7.93	1.06
206157_at	PTX3	NM_002852	Hs.567326	7.74	1.18
209305_s_at	GADD45B	NM_015675	Hs.110571	7.63	0.81
31637_s_at	NR1D1, THRA	NM_003250	Hs.724	7.46	1.55
1569020_at	NEDD9	NM_006403	Hs.37982	7.27	0.91
239669_at	HIST1H3D	NM_003530	Hs.532144	7.09	1.56
1554980_a_at	ATF3	NM_001030287	Hs.460	7.04	1.23
227099_s_at	LOC387763	XM_373497	Hs.530443	6.87	1.04
206814_at	NGFB	NM_002506	Hs.2561	6.84	0.94
242329_at	LOC401317	XM_379479	Hs.585228	6.77	1.70
218066_at	SLC12A7	NM_006598	Hs.172613	6.68	0.96
209774_x_at	CXCL2	NM_002089	Hs.75765	6.49	1.00
1562153_a_at	PVT1	XM_372058	Hs.584794	6.48	1.25
209304_x_at	GADD45B	NM_015675	Hs.110571	6.36	1.10
223484_at	NMES1	NM_032413	Hs.112242	6.23	0.99
229518_at	FAM46B	NM_052943	Hs.59771	6.08	0.82
219270_at	MGC4504	NM_024111	Hs.155569	5.97	2.07
205409_at	FOSL2	NM_005253	Hs.568265	5.88	1.21
202859_x_at	IL8	NM_000584	Hs.624	5.83	0.74
206085_s_at	CTH	NM_001902	Hs.19904	5.76	1.87
202531_at	IRF1	NM_002198	Hs.436061	5.68	1.26
205193_at	MAFF	NM_012323	Hs.517617	5.67	1.03
224588_at	XIST	NR_001564	Hs.529901	5.66	1.00
207574_s_at	GADD45B	NM_015675	Hs.110571	5.66	1.06
223774_at	C1orf79	XM_378848	Hs.528699	5.48	4.61
202672_s_at	ATF3	NM_001030287	Hs.460	5.48	1.43
220987_s_at	C11orf17, NUAK2	NM_020642	Hs.131180	5.47	1.06
215495_s_at	SAMD4	NM_015589	Hs.98259	5.42	0.96
211414_at	GLS	NM_014905	Hs.116448	5.17	1.02
238725_at	IRF1	NM_002198	Hs.436061	5.10	1.18
202014_at	PPP1R15A	NM_014330	Hs.76556	5.09	1.31
202643_s_at	TNFAIP3	NM_006290	Hs.211600	5.06	1.07
230542_at	ZNF597	NM_152457	Hs.88630	4.85	0.86
236898_at	transcribed locus		Hs.550924	4.79	0.95
37028_at	PPP1R15A	NM_014330	Hs.76556	4.56	1.55
209636_at	NFKB2	NM_002502	Hs.73090	4.56	1.15
202638_s_at	ICAM1	NM_000201	Hs.515126	4.39	0.85
207113_s_at	TNF	NM_000594	Hs.241570	4.33	0.89
202023_at	EFNA1	NM_004428	Hs.516664	4.32	1.31
209457_at	DUSP5	NM_004419	Hs.2128	4.26	0.89
236987_at	IMAGE:2365727			4.14	1.13
221530_s_at	BHLHB3	NM_030762	Hs.177841	4.06	1.32
234799_at	ADARB1	NM_001033049	Hs.474018	3.98	0.88
214472_at	HIST1H3D	NM_003530	Hs.532144	3.92	1.76
221728_x_at	XIST	NR_001564	Hs.529901	3.87	1.06
36711_at	MAFF	NM_012323	Hs.517617	3.87	1.43
222802_at	EDN1			3.81	1.06
218880_at	FOSL2	NM_005253	Hs.220971	3.79	1.04
203725_at	GADD45A	NM_001924	Hs.80409	3.77	1.30
202644_s_at	TNFAIP3	NM_006290	Hs.211600	3.70	1.12

Name	Gene Symbol (Genedata An	RefSeq Transcrip	UniGene [A]	HeLa 500/0	Hek 500/0
223196_s_at	SESN2	NM_031459	Hs.469543	3.66	1.25
225142_at	KIAA1718	XM_376680	Hs.308710	3.57	1.79
219826_at	ZNF419	NM_024691	Hs.98593	3.57	1.42
36829_at	PER1	NM_002616	Hs.445534	3.53	0.71
213146_at	KIAA0346			3.52	1.13
213649_at	SFRS7	NM_001031684	Hs.309090	3.47	2.38
227458_at	CD274	NM_014143	Hs.521989	3.45	1.06
229520_s_at	C14orf118	NM_017926	Hs.410231	3.43	1.76
229344_x_at	FAM80B	NM_020734	Hs.504670	3.43	2.74
218995_s_at	EDN1	NM_001955	Hs.511899	3.40	0.87
214185_at	KHDRBS1	NM_006559	Hs.445893	3.38	5.69
1558305_at	TNRC15	NM_015575	Hs.435841	3.38	0.96
235857_at	KCTD11	NM_001002914	Hs.585037	3.35	0.77
209383_at	DDIT3	NM_004083	Hs.505777	3.33	1.34
210538_s_at	BIRC3	NM_001165	Hs.127799	3.33	0.94
225557_at	AXUD1	NM_033027	Hs.370950	3.33	1.64
204907_s_at	BCL3	NM_005178	Hs.31210	3.31	0.99
221805_at	NEFL	NM_006158	Hs.521461	3.26	0.94
203927_at	NFKBIE	NM_004556	Hs.458276	3.23	1.28
214447_at	ETS1	NM_005238	Hs.369438	3.20	1.10
1554283_at	CCRN4L	NM_012118	Hs.548091	3.19	1.21
212471_at	KIAA0241	NM_015060	Hs.128056	3.13	1.25
215243_s_at	GJB3	NM_001005752	Hs.522561	3.13	1.09
221602_s_at	FAIM3	NM_005449	Hs.58831	3.09	1.37
202637_s_at	ICAM1	NM_000201	Hs.515126	3.09	1.01
1559064_at	NUP153	NM_005124	Hs.121088	3.08	1.33
218280_x_at	HIST2H2AA	NM_003516	Hs.530461	3.07	1.59
233245_at	C6orf107	NM_017754	Hs.555954	3.04	1.18
235128_at	SYNPO	NM_007286	Hs.435228	3.04	0.91
238422_at	LOC151534		Hs.516124	3.03	1.13
1559060_a_at	KIAA1961	NM_001008738	Hs.483329	3.03	0.95
210069_at	CHKB, CPT1B	NM_004377	Hs.439777	3.02	1.51
202861_at	PER1	NM_002616	Hs.445534	3.00	0.87
208180_s_at	HIST1H4H	NM_003543	Hs.421737	3.00	2.23
202150_s_at	NEDD9	NM_006403	Hs.37982	2.99	1.05
214502_at	HIST1H2BJ	NM_021058	Hs.132550	2.99	1.22
AFFX-M27830_5	28S rRNA			2.98	1.25
201694_s_at	EGR1	NM_001964	Hs.326035	2.98	1.12
240982_at	CARS	NM_001014437	Hs.274873	2.98	1.12
223780_s_at	THRAP1	NM_005121	Hs.282678	2.98	1.72
218881_s_at	FOSL2	NM_005253	Hs.568265	2.96	1.00
223217_s_at	NFKBIZ	NM_001005474	Hs.319171	2.95	0.87
230283_at	NEURL2	NM_080749	Hs.517094	2.95	1.28
238476_at	LOC153222	NM_153607	Hs.484195	2.95	1.32
231146_at	FAM24B	NM_152644		2.94	2.76
206144_at	MAGI1	NM_001033057	Hs.567389	2.94	1.29
214522_x_at	HIST1H3D	NM_003530	Hs.532144	2.92	1.11
207535_s_at	NFKB2	NM_002502	Hs.73090	2.91	1.05
228153_at	IBRDC2	NM_182757	Hs.148741	2.90	1.10
204035_at	SCG2	NM_003469	Hs.516726	2.90	1.28
202821_s_at	LPP	NM_005578	Hs.444362	2.90	0.96
214455_at	HIST1H2BC	NM_003526	Hs.546314	2.89	0.92
209010_s_at	TRIO	NM_007118	Hs.130031	2.88	0.99
236947_at	SEMA3C	NM_006379	Hs.269109	2.88	1.15
1554420_at	ATF3	NM_001030287	Hs.460	2.87	1.15
58367_s_at	ZNF419	NM_024691	Hs.98593	2.87	1.30
223773_s_at	C1orf79	XM_378848	Hs.528699	2.85	3.01
205113_at	NEF3	NM_005382	Hs.458657	2.84	1.05
231468_at	IMAGE:3134456			2.84	1.15
210587_at	INHBE	NM_031479	Hs.279497	2.83	3.33

Name	Gene Symbol (Genedata An	RefSeq Transcrip	UniGene [A]	HeLa 500/0	Hek 500/0
225015_s_at	STK40	NM_032017	Hs.471768	2.81	0.93
230973_at	SH2D5	XM_375698	Hs.166270	2.80	0.96
206374_at	DUSP8	NM_004420	Hs.41688	2.78	1.37
235737_at	TSLP	NM_033035	Hs.389874	2.76	1.30
211286_x_at	CSF2RA	NM_006140	Hs.520937	2.74	1.15
227180_at	ELOVL7	NM_024930	Hs.274256	2.72	0.91
210070_s_at	CHKB, CPT1B	NM_004377	Hs.439777	2.70	1.39
230104_s_at	TPPP	NM_007030	Hs.584860	2.69	0.81
1554997_a_at	PTGS2	NM_000963	Hs.196384	2.68	0.95
209912_s_at	KIAA0415	XM_166527	Hs.558440	2.66	1.36
203372_s_at	SOCS2	NM_003877	Hs.485572	2.66	1.43
209908_s_at	TGFB2	NM_003238	Hs.133379	2.63	1.06
206411_s_at	ABL2	NM_005158	Hs.159472	2.63	0.67
214290_s_at	HIST2H2AA	NM_003516	Hs.530461	2.63	1.62
223195_s_at	SESN2	NM_031459	Hs.469543	2.62	1.40
228946_at	PDZK6	NM_015693	Hs.391481	2.60	1.77
1555614_at	IMAGE:3538202			2.59	1.11
217127_at	CTH	NM_001902	Hs.19904	2.59	1.90
220324_at	C6orf155	NM_024882	Hs.368337	2.59	1.11
206359_at	SOCS3	NM_003955	Hs.527973	2.58	0.77
215501_s_at	DUSP10	NM_007207	Hs.497822	2.58	1.50
213900_at	C9orf61	NM_004816	Hs.118003	2.56	0.74
1555355_a_at	ETS1	NM_005238	Hs.369438	2.55	0.55
226612_at	FLJ25076	XM_059689	Hs.126856	2.54	1.70
243894_at	SLC41A2	NM_032148	Hs.238996	2.54	1.28
205749_at	CYP1A1	NM_000499	Hs.72912	2.54	1.32
225840_at	TEF	NM_003216	Hs.181159	2.54	0.95
223878_at	INPP4B	NM_003866	Hs.480837	2.54	1.07
240454_at	TTC7A	NM_020458	Hs.370603	2.54	0.87
209401_s_at	SLC12A4	NM_005072	Hs.10094	2.53	1.25
236402_at	IMAGE:2664543			2.53	1.34
229069_at	CIP29	NM_033082	Hs.505676	2.53	3.30
215078_at	SOD2	NM_000636	Hs.487046	2.52	1.05
1555441_at	FLJ10808	NM_018227	Hs.212774	2.51	0.95
214138_at	ZNF79	NM_007135	Hs.522399	2.50	1.89
220994_s_at	STXBP6	NM_014178	Hs.508958	2.49	0.99
205816_at	ITGB8	NM_002214	Hs.285724	2.48	0.91
219256_s_at	SH3TC1	NM_018986	Hs.479116	2.48	0.99
201009_s_at	TXNIP	NM_006472	Hs.533977	2.47	1.07
244852_at	C18orf4	NM_032160	Hs.124673	2.47	1.19
220046_s_at	CCNL1	NM_020307	Hs.4859	2.46	1.47
202708_s_at	HIST2H2BE	NM_003528	Hs.2178	2.45	1.50
203373_at	SOCS2	NM_003877	Hs.485572	2.45	1.35
1559257_a_at	MAGI1	NM_001033057	Hs.567389	2.44	1.09
206864_s_at	HRK	NM_003806	Hs.87247	2.44	1.37
1552314_a_at	EYA3	NM_001990	Hs.185774	2.44	1.48
234902_s_at	ZNF416	NM_017879	Hs.247711	2.44	1.53
227599_at	LOC151963	NM_178496	Hs.151443	2.43	0.90
243861_at	KIAA1961	NM_001008738	Hs.483329	2.43	1.33
214482_at	ZBTB25	NM_006977	Hs.164347	2.43	1.17
216549_s_at	TBC1D22B	NM_017772	Hs.485270	2.42	0.81
1554229_at	LOC153222	NM_153607	Hs.484195	2.42	0.94
242316_at	transcribed locus		Hs.511399	2.42	1.12
226847_at	FST	NM_006350	Hs.9914	2.41	1.24
220671_at	CCRN4L	NM_012118	Hs.548091	2.40	0.98
201531_at	ZFP36	NM_003407	Hs.534052	2.40	1.05
238660_at	WDFY3	NM_014991	Hs.480116	2.39	2.57
204200_s_at	PDGFB	NM_002608	Hs.1976	2.39	0.94
203010_at	STAT5A	NM_003152	Hs.437058	2.39	1.15
207510_at	BDKRB1	NM_000710	Hs.553486	2.39	1.12

Name	Gene Symbol (Genedata An	RefSeq Transcrip	UniGene [A]	HeLa 500/0	Hek 500/0
241529_at	FLJ20397	NM_017802	Hs.521328	2.38	1.10
207085_x_at	CSF2RA	NM_006140	Hs.520937	2.38	1.17
227404_s_at	EGR1	NM_001964	Hs.326035	2.38	1.11
1559826_a_at	LOC401074	XM_376247	Hs.528540	2.36	1.23
227392_at	NISCH	NM_007184	Hs.435290	2.36	1.37
241396_at	IMAGE:4693718			2.36	0.82
204472_at	GEM	NM_005261	Hs.345139	2.36	1.01
222142_at	CYLD	NM_015247	Hs.432993	2.35	1.28
38037_at	HBEGF	NM_001945	Hs.799	2.35	0.73
204549_at	IKBKE	NM_014002	Hs.321045	2.34	1.11
1559051_s_at	C6orf150	NM_138441	Hs.344080	2.33	1.09
216200_at	PLEKHM1	NM_014798	Hs.514242	2.33	1.16
227301_at	LOC441241, LOC441244	XM_496886	Hs.488399	2.33	3.33
203249_at	EZH1	NM_001991	Hs.194669	2.33	0.87
204958_at	PLK3	NM_004073	Hs.153640	2.32	1.16
1555411_a_at	CCNL1	NM_020307	Hs.4859	2.32	1.43
203691_at	PI3	NM_002638	Hs.112341	2.31	1.13
233112_at	C9orf150	NM_203403	Hs.445356	2.30	1.00
208094_s_at	MGC10471	NM_030818	Hs.24998	2.30	1.69
206566_at	SLC7A1	NM_003045	Hs.14846	2.29	0.94
218000_s_at	PHLDA1	NM_007350	Hs.484885	2.29	1.01
213353_at	ABCA5	NM_018672	Hs.421474	2.28	1.16
221603_at	PEX16	NM_004813	Hs.100915	2.27	0.95
206926_s_at	IL11	NM_000641	Hs.467304	2.27	1.05
218847_at	IMP-2	NM_001007225	Hs.35354	2.27	1.02
206175_x_at	ZNF222	NM_013360	Hs.279840	2.27	1.82
224978_s_at	USP36	NM_025090	Hs.464243	2.27	1.04
1554986_a_at	SNX19	NM_014758	Hs.444024	2.27	1.12
210004_at	OLR1	NM_002543	Hs.412484	2.26	0.87
220606_s_at	NBLA03831	NM_020233	Hs.47668	2.25	1.29
237439_at	USP43	XM_371015	Hs.23935	2.25	0.99
232035_at	HIST1H4H	NM_003543	Hs.421737	2.25	2.10
235686_at	FLJ37953	NM_152382	Hs.204619	2.25	2.20
211038_s_at	MGC12760	XM_496351	Hs.522876	2.24	1.71
1568695_s_at	DDX26	NM_012141	Hs.439440	2.23	0.87
236117_at	transcribed locus		Hs.42747	2.23	1.12
209832_s_at	CDT1	NM_030928	Hs.122908	2.22	1.25
207219_at	ZNF643	NM_023070	Hs.133034	2.22	1.64
1558290_a_at	LOC441378, PVT1	XM_372058	Hs.133107	2.21	2.78
202500_at	DNAJB2	NM_006736	Hs.77768	2.21	1.42
203439_s_at	STC2	NM_003714	Hs.233160	2.21	1.59
209850_s_at	CDC42EP2	NM_006779	Hs.343380	2.21	0.81
227486_at	NT5E	NM_002526	Hs.153952	2.21	0.89
232832_at	DKFZp434J0226	XM_375629	Hs.466975	2.20	1.06
241433_at	RCOR3	NM_018254	Hs.356399	2.20	0.65
229264_at	PARP8	NM_024615	Hs.369581	2.20	1.65
235914_at	SYNPO	NM_007286	Hs.435228	2.20	0.97
209959_at	NR4A3	NM_006981	Hs.279522	2.20	1.09
205181_at	ZNF193	NM_006299	Hs.100921	2.20	1.52
214520_at	FOXC2	NM_005251	Hs.558329	2.19	1.04
1565579_at	TATDN2	NM_014760	Hs.475401	2.19	0.99
238910_at	MGC3020	NM_024048	Hs.460642	2.18	1.16
206673_at	GPR	NM_007223	Hs.37196	2.18	0.77
213139_at	SNAI2	NM_003068	Hs.360174	2.18	1.04
244441_at	USP31	NM_020718	Hs.183817	2.18	0.97
203602_s_at	ZBTB17	NM_003443	Hs.433764	2.18	1.13
231963_at	IMAGE:3869276		Hs.26039	2.18	0.66
201010_s_at	TXNIP	NM_006472	Hs.533977	2.17	1.08
202847_at	PCK2	NM_001018073	Hs.75812	2.17	1.31
225979_at	PLEKHG2	NM_022835	Hs.111217	2.17	0.80

Table SII

Name	Gene Symbol (Genedata An	RefSeq Transcrip	UniGene [A]	HeLa 500/0	Hek 500/0
205228_at	RBMS2	NM_002898	Hs.505729	2.17	1.04
1555007_s_at	WDR66	NM_144668	Hs.507125	2.16	1.08
203828_s_at	IL32	NM_001012631	Hs.943	2.16	1.23
205207_at	IL6	NM_000600	Hs.512234	2.16	0.75
213323_s_at	ZC3H7B	NM_017590	Hs.474970	2.16	0.91
226206_at	MAFK	NM_002360	Hs.584824	2.16	0.85
216271_x_at	SYDE1	NM_033025	Hs.528701	2.15	0.84
206744_s_at	ZMYM5	NM_014242	Hs.530988	2.15	1.81
208436_s_at	IRF7	NM_001572	Hs.166120	2.15	1.74
232076_at	ZNF707	NM_173831	Hs.521922	2.15	0.91
214443_at	PVR	NM_006505	Hs.171844	2.15	0.78
206523_at	PSCD3	NM_004227	Hs.487479	2.15	0.86
220121_at	LINS1	NM_018148	Hs.105633	2.15	1.39
243560_at	transcribed locus		Hs.553158	2.15	0.97
227140_at	INHBA	NM_002192	Hs.28792	2.14	1.04
237367_x_at	CFLAR	NM_003879	Hs.390736	2.14	0.88
221563_at	DUSP10	NM_007207	Hs.497822	2.14	1.41
210387_at	HIST1H2BG	NM_003518	Hs.240135	2.14	1.20
205490_x_at	GJB3	NM_001005752	Hs.522561	2.14	1.18
1554291_at	KIAA0701	NM_001006947	Hs.387336	2.14	0.88
219492_at	CHIC2	NM_012110	Hs.335393	2.13	1.35
209193_at	PIM1	NM_002648	Hs.81170	2.13	1.03
1560485_at	HIVEP1	NM_002114	Hs.567284	2.13	1.00
230013_s_at	ARHGAP23	XM_290799	Hs.374446	2.12	1.05
230450_at	IMAGE:2357664			2.12	1.10
218724_s_at	TGIF2	NM_021809	Hs.292281	2.12	1.13
202902_s_at	CTSS	NM_004079	Hs.181301	2.12	1.00
207110_at	KCNJ12	NM_021012	Hs.200629	2.12	0.70
229468_at	CDK3	NM_001258	Hs.584745	2.11	0.95
208284_x_at	GGT1	NM_001032364	Hs.444164	2.11	1.06
236657_at	cDNA YI37C01		Hs.432924	2.11	0.85
225220_at	ACA24 snoRNA gene		Hs.535762	2.11	1.75
1554014_at	CHD2	NM_001271	Hs.220864	2.11	1.80
210190_at	STX11	NM_003764	Hs.118958	2.10	0.91
235414_at	ZNF383	NM_152604	Hs.567750	2.10	1.31
211468_s_at	RECQL5	NM_001003715	Hs.514480	2.10	1.28
1558212_at	FLJ35024	XM_379622	Hs.416043	2.10	1.96
1557257_at	IMAGE:1628131			2.10	0.89
211605_s_at	RARA	NM_000964	Hs.535499	2.10	0.94
205807_s_at	TUFT1	NM_020127	Hs.489922	2.09	1.31
235062_at	LOC120379	NM_138789	Hs.420662	2.09	1.16
227347_x_at	HES4	NM_021170	Hs.154029	2.09	1.01
238624_at	NLK	NM_016231	Hs.208759	2.09	1.34
201008_s_at	TXNIP	NM_006472	Hs.533977	2.09	0.88
206044_s_at	BRAF	NM_004333	Hs.324250	2.08	1.39
235116_at	TRAF1	NM_005658	Hs.531251	2.08	1.01
202684_s_at	RNMT	NM_003799	Hs.8086	2.08	1.10
242827_x_at	transcribed locus		Hs.272159	2.08	0.88
204015_s_at	DUSP4	NM_001394	Hs.417962	2.08	1.13
220382_s_at	ARHGAP28	NM_001010000	Hs.183114	2.07	0.78
240983_s_at	CARS	NM_001014437	Hs.274873	2.07	1.44
228250_at	KIAA1961	NM_001008738	Hs.483329	2.07	0.99
227452_at	LOC440462	XM_498681	Hs.157726, H	2.07	1.17
205931_s_at	CREB5	NM_001011666	Hs.437075	2.07	1.23
226462_at	STXBP6	NM_014178	Hs.508958	2.07	0.93
209324_s_at	RGS16	NM_002928	Hs.413297	2.07	0.71
1554274_a_at	SSH1	NM_018984	Hs.199763	2.07	1.01
220370_s_at	USP36	NM_025090	Hs.464243	2.07	1.08
212665_at	TIPARP	NM_015508	Hs.12813	2.07	1.70
216262_s_at	TGIF2	NM_021809	Hs.292281	2.06	0.89

Name	Gene Symbol (Genedata An	RefSeq Transcrip	UniGene [A]	HeLa 500/0	Hek 500/0
242727_at	ARL8	NM_178815	Hs.25362	2.06	1.11
226991_at	NFATC2	NM_012340	Hs.356321	2.06	0.94
224739_at	PIM3	NM_001001852	Hs.530381	2.06	1.03
204420_at	FOSL1	NM_005438	Hs.283565	2.06	0.83
221308_at	FRS2	NM_006654	Hs.334831	2.06	0.88
207046_at	HIST2H4	NM_003548	Hs.534370	2.05	5.75
207082_at	CSF1	NM_000757	Hs.173894	2.05	0.96
235558_at	RBMS2	NM_002898	Hs.505729	2.05	0.84
228442_at	NFATC2	NM_012340	Hs.356321	2.04	0.92
1553986_at	RASEF	NM_152573	Hs.129136	2.04	1.08
211527_x_at	VEGF	NM_001025366	Hs.73793	2.04	1.01
31846_at	RHOD	NM_014578	Hs.15114	2.04	0.86
211416_x_at	GGTLA4	NM_080920	Hs.355394	2.04	1.20
228856_at	MGC2474	NM_023931	Hs.460604	2.03	1.60
206388_at	PDE3A	NM_000921	Hs.386791	2.03	1.11
236631_at	C21orf125	NM_194309	Hs.146127	2.03	1.00
1559929_at	IMAGE:4042121		Hs.350952	2.03	1.02
228230_at	PRIC285	NM_033405	Hs.517180	2.03	0.83
213556_at	LOC390940	XM_372732	Hs.22049	2.03	1.05
205590_at	RASGRP1	NM_005739	Hs.511010	2.03	1.05
223937_at	FOXP1	NM_001012505	Hs.431498	2.03	1.04
235456_at	HIST1H2BD	NM_021063	Hs.130853	2.02	1.49
202912_at	ADM	NM_001124	Hs.441047	2.01	0.99
228298_at	MGC16044	NM_138371	Hs.258002	2.01	1.02
201044_x_at	DUSP1	NM_004417	Hs.171695	2.01	0.96
239678_at	AP1GBP1	NM_007247	Hs.101480	2.01	1.27
209239_at	NFKB1	NM_003998	Hs.431926	2.01	1.03
236302_at	PPM1E	NM_014906	Hs.245044	2.01	1.25
1553962_s_at	RHOB	NM_004040	Hs.502876	2.01	0.81
209230_s_at	P8	NM_012385	Hs.513463	2.00	2.63
211674_x_at	CTAG1A, CTAG1B	NM_001327	Hs.534310	2.00	1.03
1554140_at	WDR78	NM_024763	Hs.49421	1.25	4.64
230763_at	LOC128153	NM_138796	Hs.171130	0.91	4.61
215071_s_at	HIST1H2AC	NM_003512	Hs.484950	1.52	4.05
230142_s_at	CIRBP	NM_001280	Hs.25489	1.62	3.86
1553494_at	TDH	NR_001578	Hs.130610	0.96	3.81
230815_at	LOC389765	XM_372122	Hs.160561	1.54	3.57
205278_at	GAD1	NM_000817	Hs.420036	1.27	3.54
223714_at	ZNF256	NM_005773	Hs.288736	1.01	3.41
1562309_s_at	PHF21B	NM_138415	Hs.254097	0.76	3.21
230861_at	CHRNA7			0.90	3.17
222612_at	PSPC1	NM_018282	Hs.213198	1.59	3.05
232611_at	LOC92497		Hs.524660	1.46	3.03
231945_at	FILIP1	NM_015687	Hs.526972	0.94	3.00
1555151_s_at	TDH	NR_001578	Hs.130610	0.84	2.97
236706_at	LOC129530	NM_174898	Hs.164589	1.28	2.97
207166_at	GNGT1	NM_021955	Hs.555871	1.23	2.97
218371_s_at	PSPC1	NM_018282	Hs.213198	1.45	2.96
230795_at	HIST2H4	NM_003548	Hs.55468	1.42	2.95
219455_at	FLJ21062	NM_024788	Hs.521012	1.49	2.89
1561759_at	LOC441038	XM_376342	Hs.535008	1.08	2.84
1561761_x_at	LOC441038	XM_376342	Hs.535008, H	1.61	2.81
239282_at	CCDC41	NM_016122	Hs.279209	0.91	2.73
241348_at	ZNF654	NM_018293	Hs.27595	1.00	2.72
209602_s_at	GATA3	NM_001002295	Hs.524134	1.00	2.65
242586_at	FSD1CL	NM_207647	Hs.136901	1.03	2.64
228209_at	clone NT2NE2013189		Hs.537031	1.24	2.60
242807_at	FSD1L	NM_207647	Hs.136901	1.38	2.59
202086_at	MX1	NM_002462	Hs.517307	1.71	2.57
220366_at	ELSPBP1	NM_022142	Hs.104894	1.12	2.55

Name	Gene Symbol (Genedata An	RefSeq Transcrip	UniGene [A]	HeLa 500/0	Hek 500/0
224165_s_at	FLJ12476	NM_001031715	Hs.444535	1.58	2.55
213938_at	CAST1	NM_015576		0.77	2.54
208886_at	H1F0	NM_005318	Hs.226117	1.04	2.49
234040_at	HELLS	NM_018063	Hs.463677	1.77	2.46
1562904_s_at	Clone PP1195		Hs.571462	1.04	2.46
204919_at	PRR4	NM_007244	Hs.533634	1.02	2.43
1564282_a_at	LOC285708			0.89	2.39
1558651_at	CCDC28A	NM_015439	Hs.412019	1.15	2.39
205185_at	SPINK5	NM_006846	Hs.331555	1.11	2.39
233141_s_at	ST7L	NM_017744	Hs.201921	1.98	2.37
226164_x_at	FAM80B	NM_020734	Hs.504670	1.88	2.35
214079_at	DHRS2	NM_005794	Hs.272499	1.16	2.35
1564474_at	LOC285711	XM_211988	Hs.161338	1.02	2.35
205896_at	SLC22A4	NM_003059	Hs.310591	1.15	2.33
242639_at	NARG2	NM_001018089	Hs.200943	1.19	2.31
222728_s_at	MGC5306	NM_024116	Hs.355750	1.95	2.31
215559_at	ABCC6	NM_001171	Hs.442182	1.13	2.31
232158_x_at	NPAL1	NM_207330	Hs.134190	0.63	2.29
230536_at	PBX4	NM_025245	Hs.466257	1.08	2.28
1554614_a_at	PTBP2	NM_021190	Hs.269895	1.72	2.26
219635_at	ZNF606	NM_025027	Hs.287629	1.01	2.25
223589_at	ZNF416	NM_017879	Hs.247711	1.85	2.24
1568838_at	clone CS0DF033YE17		Hs.498519	1.10	2.24
206463_s_at	DHRS2	NM_005794	Hs.272499	1.43	2.24
222161_at	NAALAD2	NM_005467	Hs.503560	0.90	2.23
214016_s_at	SFPQ	NM_005066	Hs.355934	1.68	2.23
231953_at	FALZ	NM_004459	Hs.444200	1.04	2.21
239812_s_at	FLJ12476	NM_001031715	Hs.444535	1.00	2.21
244190_at	THAP5	NM_182529	Hs.290259	0.87	2.20
1552738_a_at	ST7L	NM_017744	Hs.201921	1.30	2.19
205432_at	OVGP1	NM_002557	Hs.1154	1.14	2.19
238658_at	transcribed locus		Hs.444083	1.01	2.19
202508_s_at	SNAP25	NM_003081	Hs.167317	1.53	2.18
230600_at	LRRC46	NM_033413	Hs.130767	0.94	2.17
203861_s_at	ACTN2	NM_001103	Hs.498178	0.84	2.17
1569366_a_at	ZNF569	NM_152484	Hs.511848	1.02	2.16
234907_x_at	POLB	NM_002690	Hs.521563	1.04	2.16
1556006_s_at	IMAGE:4619723		Hs.529862	1.40	2.16
228497_at	SLC22A15	NM_018420	Hs.125482	1.56	2.15
1557261_at	LOC339005, LOC440253	XM_370838	Hs.558967	1.12	2.15
238458_at	EFHA2	NM_181723	Hs.403594	1.09	2.15
231500_s_at	LOC388221	XM_370939	Hs.444600	1.22	2.15
220361_at	FLJ12476	NM_001031715	Hs.444535	0.85	2.14
226614_s_at	C8orf13	NM_053279	Hs.124299	0.72	2.14
238695_s_at	RAB39B	NM_171998	Hs.24970	1.02	2.13
241745_at	clone BRAWH3003343		Hs.571199	0.95	2.12
213908_at	LOC339005, LOC440253	XM_370838	Hs.558967	1.10	2.12
1561396_at	EPHA6	XM_114973	Hs.292059	1.10	2.10
228953_at	KIAA1971	XM_058720	Hs.377360	1.84	2.09
203895_at	PLCB4	NM_000933	Hs.472101	1.78	2.09
203896_s_at	PLCB4	NM_000933	Hs.472101	1.45	2.08
204726_at	CDH13	NM_001257	Hs.436040	1.00	2.08
242522_at	transcribed locus		Hs.120040	1.02	2.08
220369_at	KIAA2010	NM_017936	Hs.533887	1.60	2.07
223791_at	FAM27B, FAM27C	XM_499139	Hs.545934	1.60	2.07
225699_at	LOC285958		Hs.25892	1.86	2.06
216303_s_at	MTMR1	NM_003828	Hs.347187	0.96	2.06
207156_at	HIST1H2AG	NM_021064	Hs.51011	0.95	2.05
1554743_x_at	PMS1	NM_000534	Hs.111749	1.20	2.04
1569974_x_at	LOC441220	XM_496867	Hs.520804	1.38	2.03

Name	Gene Symbol (Genedata An	RefSeq Transcrip	UniGene [A]	HeLa 500/0	Hek 500/0
205437_at	ZNF211	NM_006385	Hs.469694	1.40	2.03
209824_s_at	ARNTL	NM_001030272	Hs.65734	1.44	2.02
230837_at	IMAGE:286061			0.87	2.01
203862_s_at	ACTN2	NM_001103	Hs.498178	0.86	2.01
218683_at	PTBP2	NM_021190	Hs.269895	1.10	2.00
227359_at	C1orf102	NM_145047	Hs.202207	1.35	2.00
1554264_at	CKAP2	NM_018204	Hs.444028	0.81	2.00
203444_s_at	MTA2	NM_004739	Hs.173043	1.26	0.50
239351_at	IMAGE:1301821			0.95	0.49
228870_at	FAM84B	NM_174911	Hs.124951	1.02	0.49
237622_at	IMAGE:1690036			0.76	0.48
203505_at	ABCA1	NM_005502	Hs.429294	1.16	0.48
207390_s_at	SMTN	NM_006932	Hs.149098	1.22	0.48
1557062_at	IMAGE:4838183		Hs.434894	1.31	0.48
232282_at	WNK3	NM_001002838	Hs.92423	0.77	0.47
231259_s_at	CCND2	NM_001759	Hs.376071	0.89	0.46
1552658_a_at	NAV3	NM_014903	Hs.306322	1.09	0.46
216953_s_at	WT1	NM_000378	Hs.408453	0.66	0.43
236330_at	clone BRAMY3002120		Hs.444593	1.02	0.42
200952_s_at	CCND2	NM_001759	Hs.376071	0.95	0.40
235102_x_at	EPN2	NM_014964	Hs.514728	1.05	0.39
213854_at	SYNGR1	NM_004711	Hs.216226	0.90	0.37
238320_at	TncRNA		Hs.433995	0.28	0.69
232412_at	clone CAE10055		Hs.587911	0.33	0.95
230930_at	LOC338620		Hs.558072	0.34	1.12
117_at	HSPA6	NM_002155	Hs.3268	0.34	0.83
212177_at	C6orf111	NM_032870	Hs.520287	0.34	0.88
244503_at	BDNF	NM_001709	Hs.502182	0.35	0.88
234989_at	TncRNA		Hs.433995	0.36	0.89
213418_at	HSPA6	NM_002155	Hs.3268	0.38	0.88
234032_at	ZCCHC7	NM_032226	Hs.571599	0.39	0.97
214078_at	PAK3	NM_002578	Hs.390616	0.40	0.89
214587_at	COL8A1	NM_001850	Hs.134830	0.41	0.93
214807_at	clone DKFZp564O0862		Hs.99472	0.41	0.97
242191_at	LOC200030, LOC400781	NM_183372	Hs.125298	0.42	1.14
228620_at	PRKRA	NM_003690	Hs.405537	0.42	1.26
210602_s_at	CDH6	NM_004932	Hs.171054	0.43	1.04
239270_at	PLCXD3	NM_001005473	Hs.145404	0.44	1.00
235030_at	FAM55C	NM_145037	Hs.130195	0.44	0.93
235060_at	DKFZp547E087		Hs.531664	0.44	0.86
213517_at	PCBP2	NM_005016	Hs.546271	0.44	0.69
230629_s_at	EP400	NM_015409	Hs.507307	0.44	0.71
244043_at	IMAGE:1700929			0.45	0.85
215143_at	FLJ36166	NM_182634	Hs.148768	0.46	0.82
206100_at	CPM	NM_001005502	Hs.567251	0.46	0.77
240450_at	PAPPA	NM_002581	Hs.494928	0.47	1.08
236140_at	GCLM	NM_002061	Hs.315562	0.47	1.12
214605_x_at	GPR1	NM_005279	Hs.184907	0.47	1.11
1570482_at	SLC22A3	NM_021977	Hs.567337	0.48	0.94
224875_at	FLJ37562	NM_152409	Hs.406549	0.48	0.93
215470_at	GTF2H2	NM_001515	Hs.191356	0.48	0.98
1556331_a_at	IMAGE:5259142			0.48	1.21
203913_s_at	HPGD	NM_000860	Hs.77348	0.48	0.95
238635_at	FLJ21657	NM_022483	Hs.558531	0.49	1.40
239960_x_at	clone SKMUS2001014		Hs.508823	0.49	1.35
243308_at	MORN1	NM_024848	Hs.567585	0.49	1.06
204932_at	TNFRSF11B	NM_002546	Hs.81791	0.49	0.82
228353_x_at	STS-1	NM_032873	Hs.444075	0.49	1.17
1554513_s_at	FLJ14640	NM_032816	Hs.317590	0.49	0.95
1555778_a_at	POSTN	NM_006475	Hs.136348	0.49	1.02

Table SIII

name	sequence	product length (bp)	note
genotyping, T1 treatment and editing analysis			
MosE.fwd	GAGCAAGACGTTTGTAAGATCA	268	editing analysis
MosE.rev	TGCTCAAGGGGCTTCATGATGT		
Mos 3.fwd	CTCTTTTGGAATCACCCTAAGGG	317	transgene genotyping
MosIR+3.rev	GCTGTTTCATATACTGATGACCTCTTTAT		
MosIR.fwd	ACTCTTTTGGAATCACCCTAAGCC	316	
MosIR+3.rev	GCTGTTTCATATACTGATGACCTCTTTAT		
MosP.fwd	GAGCAAGACGTTTGTAAGATCA	241	
MosP.rev	AGATGGGGAGAGTGAAGACAAGA		
Egfp.fwd	GCACCATCTTCTTCAAGGACGAC	343	
Egfp.rev	TCTTTGCTCAGGGCGGACTG		
MosT1.Fwd	CAACGTGACTCTACACCAAGTC	204	amplification after RNase T1 treatment
MosT1.Rev	ATCTTACAAACGTCTTGCTCACTG		
qPCR mouse			
Mos.fwd	GGGAACAGGTATGTCTGATGCA	75	qPCR
Mos.rev	CACCGTGGTAAGTGGCTTTATACA		
Mos3.fwd	CTCTTTTGGAATCACCCTAAGGG	102	
MosIR+3q.rev	GCCAGAAGTCAGATGCTCAAGG		
CagI.fwd	GGCTCTGACTGACCGCGTTAC	176	
CagI.rev	GCTGAACTTGTGGCCGTTTACG		
Ifit.fwd	AGAGAGTCAAGGCAGGTTTCTGAG	190	
Ifit.fwd	TCTCACTTCCAAATCAGGTATGTCA		
Pkr.fwd	TCCTCAGAGAACGTGTTTACGAAC	169	
Pkr.rev	TGAAAACTTGACCAAATCCACCTA		
Oas1.fwd	AGACGTTGTGGAGTGAAGTTTGAG	335	
Oas1.rev	CCCAGCTTCTCCTTACACAGTTG		
Rig1.fwd	AGCTTACTCGGAGGTTTGAAGAAA	234	
Rig1.fwd	CAGTCAGTATGCCAGGCTTTAGAA		
Hprt.fwd	GCTACTGTAATGATCAGTCAACGG	213	qPCR - standard
Hprt.rev	CTGTATCCAACACTTCGAGAGGTC		
globin.fwd	GCAGCCACGGTGGCGAGTAT	257	
globin.rev	GTGGGAVAGGAGVTTGAAAT		
qPCR human			
IRF1.fwd	CTCTGAAGCTACAACAGATGAGGA	169	qPCR
IRF1.rev	CCTCCTTACAGCTAAAGTCTCCAT		
NMI.fwd	CCTTTGAAAAAGAAGAAGTTGCTC	130	
NMI.rev	ATAAACCTGGAATCTGACTCCTGAA		
GCH1.fwd	ATCTCAGATGTCCTAAACGATGCT	185	
GCH1.fev	GATTTCTACAATCCTCGCAAGTTT		
C1S.fwd	TACTATGTTGCCACAGACATAAATG	148	
C1S.rev	CAATTAACTCCGCAATTCTTCAT		
IFIT1.fwd	CTAAGCAAAACCCTGCAGAACG	179	
IFIT1.rev	GGAATTCAATCTGATCCAAGACTC		
PKR.fwd	GTGACCAGCACACTCGCTTCTG	217	
PKR.rev	CCATGCCAAACCTCTTGTCCACAG		
RIG-I.fwd	CATGTCCACCTTCAGAAGTGTCTG	137	
RIG-I.rev	GGTTTTTCCACAACCTGTAGGAGC		
HPRT.fwd	TGACCTTGATTTATTTTGCATACC	102	qPCR - standard
HPRT.rev	CGAGCAAGACGTTCAGTCCT		

Supplement 3

Nejepinska J, Malik R, Moravec M, Svoboda P.

Deep sequencing reveals complex spurious transcription from transiently transfected plasmids.

PLoS One

2012

7(8):e43283

OPEN ACCESS Freely available online

PLOS ONE

Deep Sequencing Reveals Complex Spurious Transcription from Transiently Transfected Plasmids

Jana Nejepinska[◔], Radek Malik[◔], Martin Moravec, Petr Svoboda*

Institute of Molecular Genetics AS CR, Prague, Czech Republic

Abstract

Transient plasmid transfection is a common approach in studies in cultured mammalian cells. To examine behavior of transfected plasmids, we analyzed their transcriptional landscape by deep sequencing. We have found that the entire plasmid sequence is transcribed at different levels. Spurious transcription may have undesirable effects as some plasmids, when co-transfected, inhibited expression of luciferase reporters in a dose-dependent manner. In one case, we attributed this effect to a Kan/Neo resistance cassette, which generated a unique population of edited sense and antisense small RNAs. The unexpected complexity of expression from transiently transfected plasmids underscores the importance of appropriate experimental controls.

Citation: Nejepinska J, Malik R, Moravec M, Svoboda P (2012) Deep Sequencing Reveals Complex Spurious Transcription from Transiently Transfected Plasmids. PLoS ONE 7(8): e43283. doi:10.1371/journal.pone.0043283

Editor: Thomas Preiss, The John Curtin School of Medical Research, Australia

Received March 5, 2012; **Accepted** July 23, 2012; **Published** August 16, 2012

Funding: This research was supported by the Czech Science Foundation (grant GACR 204/09/0085), institutional support AV0Z50520514 and RVO 68378050, and the European Molecular Biology Organization Strategic Development Installation Grant (program 0001488, partial). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: petr.svoboda@img.cas.cz

◔ These authors contributed equally to this work.

Introduction

Transient plasmid transfection is a routine approach to study gene expression in mammalian cells. However, transient plasmid transfections are often used without appropriate attention to potential artifacts. Several factors contribute to this situation. Ancestors of currently used reporter plasmids were developed one or two decades ago (firefly luciferase –1987 [1], *Renilla* luciferase –1996 [2], green fluorescent protein –1994 [3]) when available technologies limited detailed analysis of plasmid expression and its behavior in transfected cells. At the same time, it is difficult to find information about unexpected effects in transfection experiments in the literature. Seminal articles published more than a decade ago [4–9] are cited rather infrequently and they are difficult to find among hundreds of irrelevant hits in literature database searches. In addition, the majority of common plasmids used in laboratories are commercial plasmids or their derivatives. Information about these plasmids is often restricted into company's technical notes (e.g. [10,11]) while peer-reviewed information about potential problems with these plasmids is limited or even absent.

Reports demonstrating false reporter response in transient transfections suggest three main sources of artifacts. First, reporter plasmids carry artificial cis-regulatory sequences, whose recognition by cellular proteins affects reporter expression at the transcription level [12–17]. In addition, recognition of artificial cis-regulatory sequences can be affected by experimental conditions, creating a false treatment response [18]. The second source of problems can be lesions in transfected DNA [7]. The third source of possible artifacts are post-transcriptional regulations. They could involve, for example, expression of complementary RNAs [19] or activation of the protein kinase R (PKR) and the type I interferon (IFN) response [4,5,9,20]. Upon activation by double-stranded RNA (dsRNA), PKR blocks protein synthesis by phosphorylating the eukaryotic initiation factor eIF2 (reviewed in detail in [21]). Taken together, an outcome of a transient transfection experiment can be biased in numerous ways. Yet, published results often include only a minimal set of controls and further details concerning the specificity, reproducibility, dynamic range of the assay *etc*. are not disclosed.

The presented work arose from an internal discussion in the lab regarding which plasmid DNA should be used to maintain a constant amount of transfected DNA per sample. In most cases, one can use a parental plasmid with the same backbone sequence. However, sometimes one needs to select a "neutral" DNA. We previously noticed that different co-transfected plasmids may influence each other. Therefore, we used deep sequencing to characterize transcriptomes of four common plasmids and we further tested sensitivity of luciferase reporters to co-transfection of a larger set of plasmids. We show here that 1) transfected plasmids are a rich source of spurious RNA transcription, 2) pEGFP and its derivatives carrying the same Kan/Neo backbone have strong negative effects on luciferase activity when compared to pBluescript (pBS), and 3) the most likely cause of luciferase reporter inhibition is dsRNA originating from co-transfected plasmids.

Materials and Methods

Plasmids

The following plasmids were used in this study: pBluescript II KS (+) (Stratagene), pEGFP-C1 (Clontech), phRL-SV40 (Promega), pCR2.1 (Invitrogen), pJET (Fermentas), pLMP (Open Biosystems), pTMP (Open Biosystems), pSUPER (OligoEngine), pTER [22],

and pCAG-EGFP [23]. pGL4-SV40 plasmid was produced by moving SV40 promoter from phRL-SV40 into pGL4.10 (Promega) backbone using BglII-HindIII restriction sites. pRFP-T expressing red fluorescent protein (RFP) was constructed by replacing the EGFP coding sequence in pEGFP-C1 by PCR-amplified monomeric TagRFP sequence (Evrogen) using NheI-BglII restriction sites. pCI-RFPT was constructed by a similar strategy by inserting the TagRFP sequence into the multiple cloning site of the pCI-Neo. To obtain pEGFP_Amp and pRFP_Amp, the kanamycin cassettes (flanked by BspHI restriction sites) in pEGFP-C1 and pRFP-T were replaced by BspHI-released β-lactamase sequence from pCAG-EGFP. Modified plasmids were verified by restriction digest and sequencing.

Cell Culture and Transfection

Human HEK-293 cells (ATCC no. CRL-1573) obtained from our collaborators [21] were maintained in DMEM (Sigma) supplemented with 10% fetal calf serum (Sigma), penicillin (100 U/mL, Invitrogen), and streptomycin (100 µg/mL, Invitrogen) at 37°C and 5% CO_2 atmosphere. For transfection, cells were plated on a 24-well plate, grown to 50% density and transfected using Turbofect *in vitro* Transfection Reagent (Fermentas) according to the manufacturer's protocol. Cells were co-transfected with 100 ng/well of each pGL4-SV40 and phRL-SV40 reporter plasmids and various amounts of the tested plasmid (50–250 ng per well). The total amount of transfected DNA was kept constant by adding pBS. After 48 hours, cells were washed with phosphate-buffered saline (PBS) and lysed with the Passive Lysis Buffer (Promega). Luciferase reporter activity was assessed using the Dual-Luciferase Reporter Assay (Promega) and luminiscence intensity was measured by Modulus Microplate Multimode Reader (Turner Biosystems). Luminometric data were adjusted to the total protein amount in lysates measured by Bradford Protein assay (Bio-Rad) according to manufacturer's instructions.

Next Generation Sequencing (SOLiD)

HEK-293 cells were plated on 6-well plates and grown to 50% density. Cells were transfected with pGL4-SV40 (3 µg/well), phRL-SV40 (3 µg/well), or pBS and pEGFP-C1 (each 1.5 µg/well in one well), cultured for 48 hours, washed with PBS, and total RNA was isolated using RNAzol (MRC, Inc.) according to the manufacturer's protocol. RNA quality was verified by Agilent 2100 Bioanalyzer.

The library construction from total RNA and deep sequencing of RNA transcriptome (RNA-seq) were performed by Seqomics (Szeged, Hungary) using SOLiD (version 4.0) sequencing platform. Bioinformatic analysis was performed as described previously [24]. We used sequence tags (reads) in datasets converted from color-space into the fasta format (fastq output files by Seqomics). The 3′ adaptor sequences were removed from raw reads using an algorithm requiring at least 3-nt exact matches between 50-nt reads and the adaptor sequence.

Sequence reads of 18–50 nt in length (after subtracting the adaptor sequence) were obtained and mapped onto specified

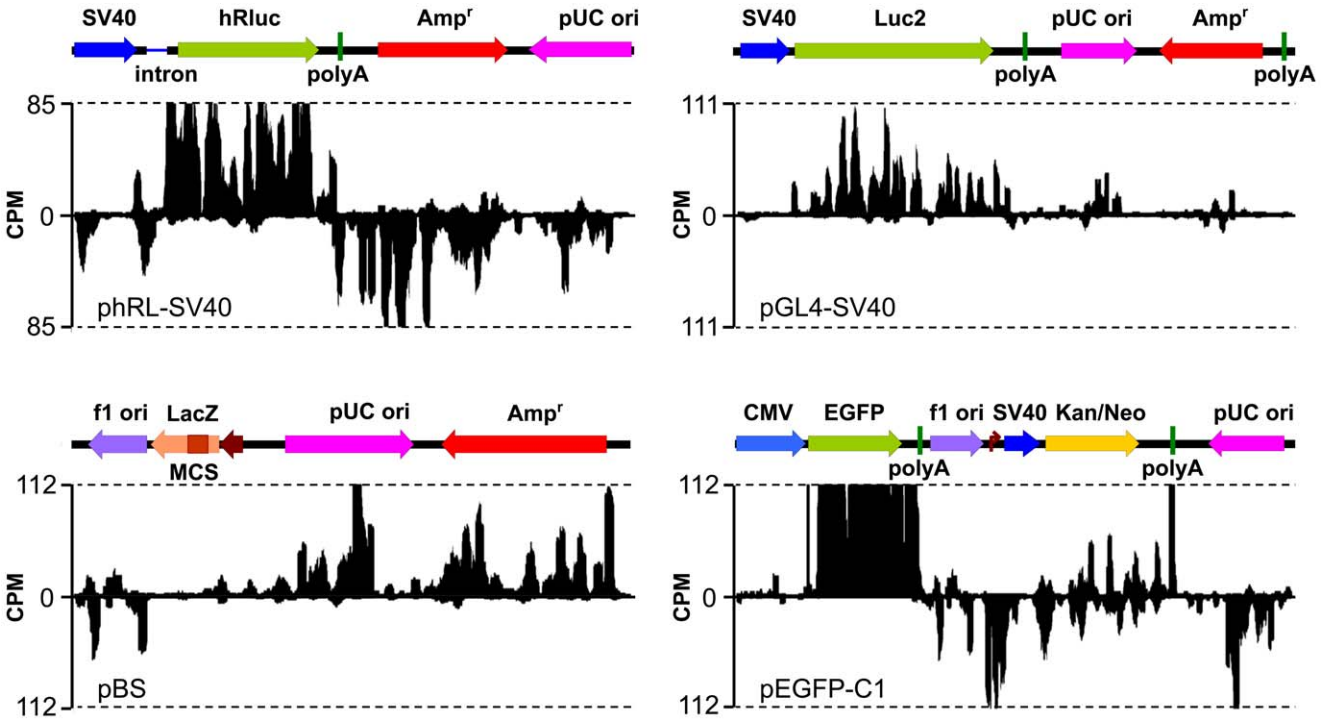


Figure 1. SOLiD sequencing of total RNAs derived from transfected plasmids. Each graph shows the read density along each of the studied plasmids in sense (above X-axis) and antisense (below X-axis) directions. Y-scale represents normalized read density (aligned 18–50 nt reads on a counts per million (CPM) scale). The Y-scale maximum corresponds to 1000 counts of mapped reads and was set to reveal read density in plasmid backbones. Cut-offs for CPM display are indicated by dashed lines. Plasmid annotation is based on maps provided by the manufacturers. Abbreviations: SV40, SV40 early promoter; hRLuc, "humanized" *Renilla* luciferase coding sequence; Luc2, modified firefly luciferase coding sequence; polyA, polyadenylation signal; pUC ori, pUC plasmid replication origin; Amp^r, β-lactamase open reading frame encoding for ampicillin resistance; f1 ori, f1 single-strand DNA replication origin; LacZ, 5′-terminal part of the lacZ gene encoding the N-terminal fragment of β-galactosidase; MCS, multiple cloning site; CMV, cytomegalovirus promoter; EGFP, enhanced green fluorescence protein; Kan/Neo, Kanamycin/neomycin resistance gene coding sequence. Note that these data offer only an indirect comparison of RNA expression originating from individual plasmids because the amount of an actively transcribed transfected plasmid in a sample cannot be precisely quantified. Source data for all four graphs are provided in the Table S4.
doi:10.1371/journal.pone.0043283.g001

Table 1. Deep sequencing results (18–50 nt reads) for different classes of RNAs.

Selected endogenous transcripts							
	length	**phRL-SV40**		**pGL4-SV40**		**pBS/pEGFP-C1**	
	(nt)	**RPKM**	**total**	**RPKM**	**total**	**RPKM**	**total**
Hprt1	1,407	**9.78**	776	**9.23**	595	**12.72**	820
Tbp	1,844	**3.71**	386	**3.86**	326	**4.44**	375
Alas1	2,430	**2.30**	315	**2.44**	272	**2.74**	305
β_2-microglobulin	1,715	**9.04**	874	**7.35**	578	**10.37**	815
β-actin	1,917	**21.75**	2,350	**23.09**	2,029	**17.55**	1,542
γ-tubulin	1,949	**2.68**	294	**3.27**	292	**2.98**	266
Gapdh	1,875	**199.19**	21,055	**201.25**	17,295	**186.88**	16,060
Selected plasmid-derived transcripts							
	length	**phRL-SV40**		**pGL4-SV40**		**pBS/pEGFP-C1**	
	(nt)	**RPKM**	**total**	**RPKM**	**total**	**RPKM**	**total**
EGFP	798	**–**	–	**–**	–	**5517.95**	201,821
***Renilla* luciferase**	936	**351.44**	18,544	**–**	–	**–**	–
Firefly luciferase	1653	**–**	–	**137.60**	10,425	**–**	–
Ampicillin CDS	861	**203.04**	9,855	**30.92**	1,220	**140.26**	5,535
Kan/Neo CDS	795	**–**	–	**–**	–	**140.57**	5,122
Selected microRNAs							
		phRL-SV40		**pGL4-SV40**		**pBS/pEGFP-C1**	
		RPM	**total**	**RPM**	**total**	**RPM**	**total**
miR-19b		**14.40**	812	**26.49**	1,214	**12.46**	571
miR-29a		**17.85**	1,006	**16.25**	745	**17.26**	791
miR-29b		**25.86**	1,458	**26.25**	1,203	**21.99**	1,008
miR-31		**20.43**	1,152	**17.02**	780	**12.48**	572
miR-125b		**50.13**	2,826	**49.92**	2,288	**45.80**	2,099

doi:10.1371/journal.pone.0043283.t001

plasmid allowing only perfectly matched reads. Each read was allowed to be mapped only once (it was removed from the dataset when it was successfully annotated). Putative polyadenylation sites controlling termination of plasmid-derived transcripts were analyzed as follows: All reads with removed adaptor sequence not aligning perfectly to the plasmid in the first round of mapping were tested for the presence of "A" residue at the 3′ end of the read. When present, all homopolymeric A residues at the 3′ end were trimmed and all resulting read sequences >17 nt in length were mapped on the plasmid. For RNA editing (A>I) analysis, unannotated reads were re-mapped to plasmid sequences allowing for up to five mismatches. Mapping scripts for the analysis were programmed using the Visual Basic 2010 platform (Microsoft).

Sequencing quality and depth of all samples was comparable (Table S1). To confirm that a similar fraction of all reads could be mapped to human genome sequences in all samples, we blasted all reads (original dataset in fastq format) against the human genome (UCSC, built hg18) using memory-efficient short read aligner Bowtie (version 0.12.7) [25] allowing to map reads with up to three mismatches over the entire alignment (using the "V" mode of the Bowtie algorithm). Minor differences in sequencing depths of samples were normalized using the number of reads per kilobase per million mapped reads (RPKM). The RPKM calculation was based on reads perfectly mapping to the genome or the transfected plasmid. We did not find evidence that genome-derived reads would be mistaken for plasmid-derived reads. A control mapping in color-space format was performed to validate results of direct mapping using the fastq format. High throughput sequencing data in color-space format were deposited in the GEO database (GSE36062).

Flow Cytometry

pCI-RFPT plasmid was labeled using Label IT® Tracker™ Intracellular Nucleic Acid Localization Kit, Cy5 (Mirus) according to the manufacturer's instructions. HEK-293 cells plated in 24-well plates were co-transfected with 150 ng/well of pCI-RFPT plasmid and 350 ng/well of pBS or pEGFP-C1 plasmid. Cells were collected 36 hours post-transfection and analyzed using LSRII cytometer (BD Bioscience). Data analysis was performed by FlowJo software (Treestar, Inc.). The bleedthrough signal in collected uncompensated FACS data was eliminated by appro-

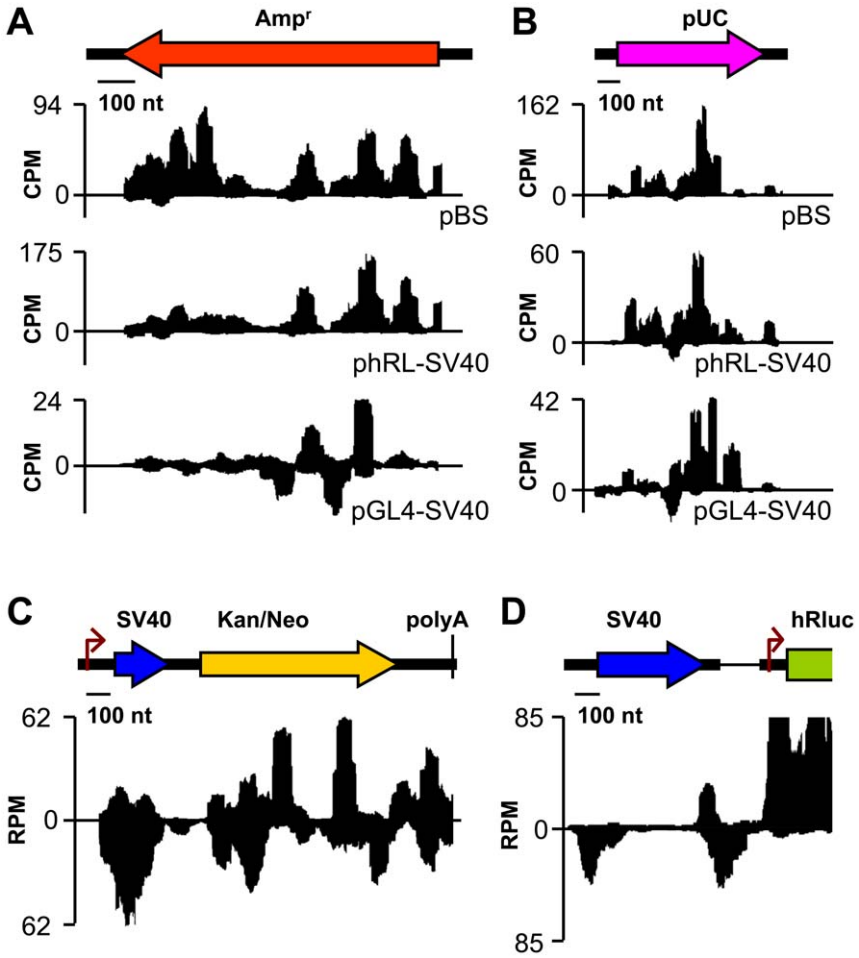


Figure 2. Read densities in specific regions of transfected plasmids. (A) Amp^r, (B) pUC ori. Note the similar pattern of read density for identical sequences in different plasmids and lower read density and altered patterns in pGL4-SV40. (C) Kan/Neo is the main locus where overlapping sense and antisense expression is detected. (D) SV40 promoter and intron-derived RNAs. Y-scale represents normalized read density – aligned reads on a counts per million (CPM) scale.
doi:10.1371/journal.pone.0043283.g002

priate compensation using controls transfected with only one fluorescent reporter.

Results

Deep Sequencing Reveals Complex Expression of Transfected Plasmids

phRL-SV40, pGL4-SV40, pEGFP-C1, and pBS plasmids were selected for deep sequencing. phRL-SV40 and pGL4-SV40 represent common *Renilla* and firefly luciferase reporter plasmids. pGL4-SV40 represents a newer generation of firefly luciferase reporters where putative mammalian transcription factor-binding sites in the plasmid backbone have been extensively mutated to minimize spurious expression [10]. pEGFP-C1 belongs to a family of plasmids for expressing protein fusions with the enhanced green fluorescent protein (EGFP). pBS is a common small cloning plasmid without any annotated eukaryotic transcription unit. All four plasmids utilize pUC prokaryotic origin of replication. phRL-SV40, pGL4-SV40, and pBS carry β-lactamase gene providing ampicillin resistance (Amp^r) while pEGFP-C1 encodes kanamycin/neomycin (Kan/Neo) resistance for selection in bacteria as well as in mammalian cells.

Deep sequencing revealed a surprisingly complex picture of transcription originating from transfected phRL-SV40, pGL4-SV40, pEGFP-C1, and pBS plasmids (Fig. 1 and Table 1). Luciferase and EGFP mRNAs were clearly dominating transcriptional landscapes of the reporter plasmids. The most prominent transcript was EGFP mRNA from pEGFP-C1, whose 18–50 nt read frequency was 28,354 RPKM. For comparison, *GAPDH* and β-actin reads were identified in the same sample in 978 RPKM and 91 RPKM, respectively. *Renilla* and firefly luciferase mRNA levels were comparable to endogenous genes (1679 and 699 RPKM, respectively). Strikingly, specific regions in all plasmid backbones yielded read densities within the same order of magnitude as were read densities of Kan/Neo resistance or Firefly and *Renilla* luciferase mRNAs (Fig. 1). These data document that transfected plasmids generate significant amounts of RNA from regions that are not expected to be transcribed in mammalian cells.

Plasmid backbone-derived RNAs usually accumulated in distinct clusters in sense or antisense directions, but clusters in opposing directions rarely overlapped (Fig. 1 and 2). The origin and fate of the spurious transcripts remain unknown. While the pattern of read densities would suggest multiple transcription start sites, distribution of reads along reporter mRNAs indicates that the read distribution is also biased by the deep sequencing procedure. Thus, one cannot unequivocally determine transcription start sites. Likewise, termination of the anomalous transcription could not be

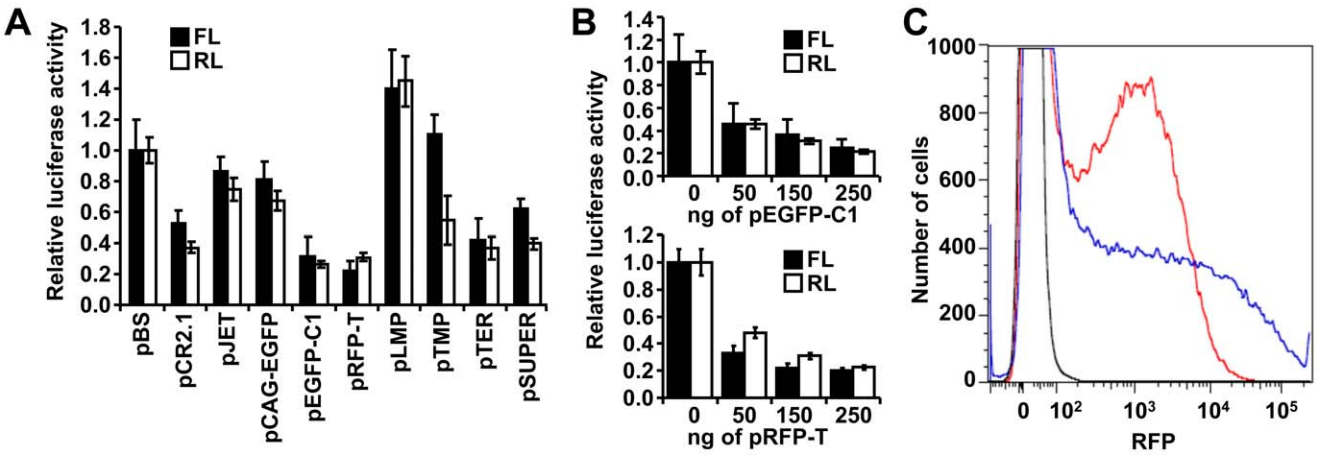


Figure 3. Effects of co-transfected plasmids on expression of luciferase reporters. (A) Different plasmids have different effects on luciferase reporters. HEK-293 cells were co-transfected with 100 ng/well of each luciferase reporter and 150 ng of a tested plasmid. *Renilla* luciferase (RL) and firefly luciferase (FL) activities in pBS co-transfection were set to one. Data represent results of four transfection experiments performed in triplicates. Error bars = SEM. (B) Dose-dependent suppression of luciferase activities by co-transfected pEGFP-C1 (upper panel) and pRFP-T (lower panel). HEK-293 cells were co-transfected with 100 ng/well of each luciferase reporter and 0–250 ng/well of pEGFP-C1 or pRFP-T. The amount of transfected DNA was kept constant by adding pBS. Error bars = SEM. Data represent results of four transfection experiments performed in triplicates. (C) pEGFP-C1 negatively affects RFP reporter expression. HEK-293 cells were co-transfected with 150 ng/well of pCI-RFPT plasmid and 350 ng/well of pBS or pEGFP-C1 plasmid. RFP expression was analyzed 36 hours post-transfection by flow cytometry. X axis = RFP fluorescence intensity. Y axis = cell count. Colored curves show distribution of RFP signal as follows: black curve = untransfected cells; blue curve = pCI-RFPT + pBS co-transfection, and red curve = pCI-RFPT + pEGFP-C1 co-transfection. Total counts of transfected (RFP-positive) cells were identical in both samples (Fig. S1C). The shape of the red curve suggests that pEGFP-C1 reduces RFP fluorescence in transfected cells. The experiment has been performed three times, results from a representative experiment are shown.
doi:10.1371/journal.pone.0043283.g003

precisely determined. Analyses of 3′ oligoadenylated sequence tags did not provide conclusive evidence that putative polyadenylation signal sequences are functional. On the other hand, polyadenylation following canonical AAUAAA polyadenylation sites of reporter transcripts was readily detected (data not shown). Expression patterns were typically reproducible for the same sequences. For example, Amp^r region yielded RNAs mostly in the antisense orientation relative to the β-lactamase coding sequence and read density along the Amp^r or pUC sequences in phRL-SV40 and pBS yielded similar patterns (Fig. 2A, 2B). Even pGL4-SV40, which has a modified backbone sequence, yielded partially similar patterns along the related sequences (Fig. 2A, 2B). The lowest level of spurious transcription was found in the pGL4 backbone, which is modified to have a minimal number of putative transcription factor binding sites (Fig. 1, 2A, 2B). However, we cannot rule out that transfection efficiency contributed to this outcome. The highest level of overlapping sense and antisense transcripts was found in the Kan/Neo resistance gene in pEGFP-C1 (Fig. 2C). A robust antisense transcription originating from the SV40 promoter (Fig. 2C, 2D) and from an intron upstream of *Renilla* luciferase in phRL-SV40 (Fig. 2D) was also remarkable.

Taken together, deep sequencing analysis of transfected plasmids demonstrated an unexpectedly complex and robust spurious plasmid expression during common transfection experiments. The bulk of the spurious expression occurs reproducibly in specific regions in specific directions. Finally, strong overlaps between abundant sense and antisense expression within individual plasmids are rather uncommon.

Co-transfected Plasmids Influence Expression

To complement the analysis of spurious transcription from transfected plasmids, we tested how different plasmids influence luciferase reporter expression in co-transfection experiments. We co-transfected ten different plasmids from our plasmid collection with firefly (pGL4-SV40) and *Renilla* (phRL-SV40) luciferases reporters. Tested plasmids represented a diverse set of plasmids used for cloning (pBS, pCR2.1, pJET), as reporters (pCAG-EGFP, pEGFP-C1, pRFP-T), or for RNAi (pLMP, pTMP, pTER and pSUPER). This experiment revealed that majority of plasmids (except for pLMP) inhibited luciferases when compared to pBS (Fig. 3A). Notably, both luciferases were similarly affected suggesting a common mechanism affecting plasmid-derived reporter expression. The strongest repression of luciferase reporters was observed in co-transfection with pEGFP-C1 and pRFP-T plasmids carrying the Kan/Neo resistance gene. The inhibitory effect of pEGFP-C1 and pRFP-T was dose-dependent (Fig. 3B) and was not restricted to luciferase reporters because pCI-RFPT-dependent red fluorescence was also inhibited by co-transfected pEGFP-C1 (Fig. 3C and S1).

Reduction of luciferase activities was apparently not due to cellular toxicity (no reduced growth or increased cell mortality were observed (Fig. S2)). Since pRFP-T exerted the same inhibitory effect as pEGFP-C1, we hypothesized that the inhibition is mediated by the plasmid backbone, which is common for both plasmids, and not by EGFP sequence or its expression. To get further insights into possible causes of inhibitory effects of pEGFP-C1, we re-examined deep sequencing data searching for any transcriptome features unique to pEGFP-C1. The analysis included size distribution of RNA fragments and frequency of A/G conversion (which is a hallmark of edited dsRNA [24]). Consistently with our previous analysis of small RNAs derived from long dsRNA, we have identified a population of RNAs carrying A/G conversions residing within the Kan/Neo cassette of pEGFP-C1 (Fig. 4A). While RNAs carrying A/G conversion were occasionally found elsewhere, none of the other plasmids produced a similar cluster of edited RNAs (Fig. 4A, 4B). Since the Kan/Neo cassette is also a region with a prominent presence of overlapping sense and antisense transcripts (Fig. 1), we conclude that the Kan/Neo cassette generates dsRNA.

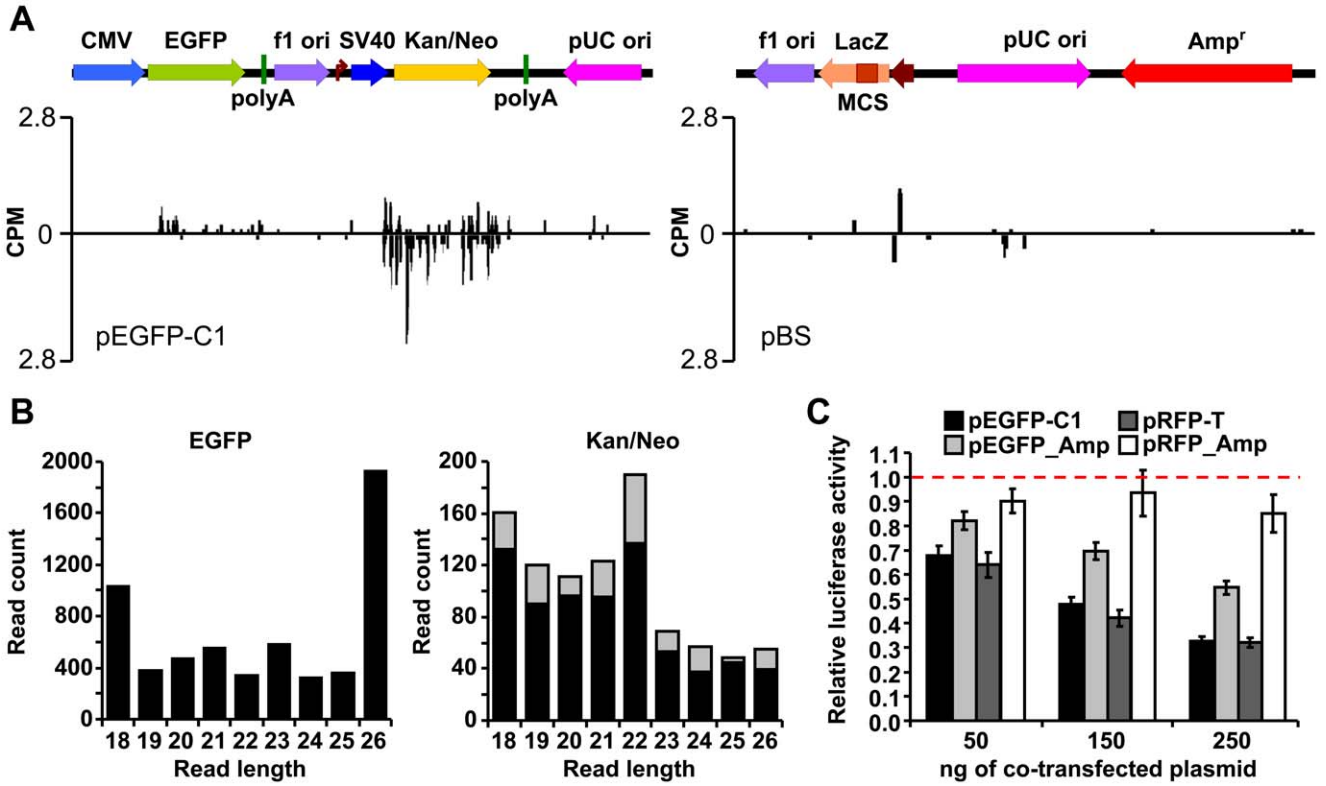


Figure 4. Kan/Neo cassette has a unique small RNA signature and contributes to downregulated expression of luciferase reporters. (A) Analysis of putative adenosine-deaminated small RNAs derived from Kan/Neo cassette (left panel) and pBS (right panel). The distribution of 20–24 nt reads with A/G conversions along pEGFP-C1 and pBS sequences is shown. (B) Size distribution of RNAs originating from EGFP CDS and Kan/Neo CDS sequences in HEK-293 cells. Small RNAs are sorted along the X-axis according to their length (18–26 nt long reads are shown). The Y-axis in both graphs shows the absolute number of reads carrying EGFP- (left) or Kan/Neo-derived sequences (right). The gray portion of each column indicates the fraction of reads carrying up to five A/G sequence changes. Note the absence of edited reads from EGFP CDS region. (C) Replacement of the Kan/Neo cassette by Amp^r (denoted by _Amp) relieves repression of luciferase reporters. HEK-293 cells were co-transfected with 100 ng/well of each luciferase reporter and 0–250 ng of one of the four plasmids shown above the graph. The total amount of transfected DNA was kept constant by adding pBS. *Renilla* luciferase activity relative to the sample co-transfected with pBS (dashed line) is shown. Error bars = SEM. Data represent two independent experiments done in quadruplicates.
doi:10.1371/journal.pone.0043283.g004

To test if the presence of the Kan/Neo cassette is necessary for the inhibitory effect of pEGFP-C1 and pRFP-T on co-transfected plasmids, we replaced the Kan/Neo resistance in pEGFP-C1 and pRFP-T with ampicillin resistance (placed in the same orientation as Kan/Neo). Co-transfection with pGL4-SV40 and phRL-SV40 luciferase reporters showed that the presence of the kanamycin resistance cassette in pEGFP-C1 and pRFP-T plasmids is associated with the reporter inhibition (Fig. 4C). In the case of pRFP-T, exchange of resistance completely abolished the inhibitory effect while there was a residual inhibition observed in the case of pEGFP-C1. This suggests that other factors also contribute to the inhibition by pEGFP-C1 in co-transfection experiments. The relief of repression upon the replacement of the kanamycin cassette was not due to a lower transfection efficiency of ampicillin-bearing plasmids since the number of RFP or EGFP-positive cells remained the same upon resistance exchange (data not shown). Because expression of fluorescent proteins could contribute to inhibitory effects, we analyzed EGFP and RFP fluorescence produced by pEGFP-C1, pEGFP_Amp, pRFP-T, and pRFPT_Amp. Remarkably, cells transfected with pEGFP_Amp exhibited lower EGFP fluorescence while RFP fluorescence of pRFP_Amp was the same as of pRFP-T (Fig. S3). Thus, neither different transfection efficiency nor different expression of a fluorescent protein explains the relief of inhibition when the Kan/Neo resistance is replaced with the ampicillin resistance in pRFP. We propose a model where RNA transcribed from the Kan/Neo sequence in pEGFP-C1 (and likely in other plasmids utilizing the same backbone) inhibits expression of co-transfected plasmids in a process likely involving the formation of dsRNA.

Discussion

We provide the first deep sequencing-based mapping of transcriptional landscapes of transiently transfected plasmids. Although the level of spurious transcription can be estimated only indirectly, it is clear that it reaches significant levels since read density of spurious transcripts is within the same order of magnitude as that of transcripts produced by annotated eukaryotic genes. Co-transfection experiments suggest that spurious transcription effects depend more on the origin of spurious transcription rather than its level. While pBS and pEGFP-C1 yield comparable levels of spurious transcription, their effects on co-transfected plasmids are in a stark contrast. We cannot rule out that overexpressed CDS in a co-transfected plasmid may also exert inhibitory effects, perhaps related to translational factors consumption. However, the analysis of RFP expression in cells transfected with pRFP-T and pRFP_Amp suggests that a presence of an overexpressed CDS is not the main cause of inhibitory effects on co-transfected luciferase reporters (Figs. S1C and 4C).

The mechanism by which spurious transcription exerts its effect on co-transfected plasmids is unclear at the moment. We ruled out toxic effects on cells or global repression of translation. Some evidence suggests that dsRNA expression triggers the inhibition. First, the Kan/Neo cassette, which is implicated in the inhibition, is the main locus where we observe significant amount of overlapping sense and antisense transcripts (Fig. 2). Second, small RNAs carrying A/G conversions cluster in the Kan/Neo cassette (Fig. 4B). Three lines of evidence suggest that observed A/G conversion is due to the activity of adenosine deamination of dsRNA. First, 21–26 nt-long reads mapping to pEGFP-C1 have the highest A/G nucleotide change rate among all plasmids and among all possible nucleotide changes (Table S2). Second, the editing should result in more frequent appearance of repeated A/G conversions in a single read when compared to random nucleotide changes, which would presumably occur independently of each other in one read. Indeed, A/G conversion was the most frequent repeated nucleotide change in single reads (>80% of all of them) and was mostly found in 21–26 nt-long reads mapping to pEGFP-C1 (Table S3). Third, similar small RNAs were found to be derived from a long RNA hairpin but not single-stranded RNA [24].

It should be noted that our deep sequencing of total RNA does not allow for a good quantitative estimate of the abundance of plasmid-derived small RNAs relative to endogenous mRNAs. While a library preparation causes variable biases [26], we also cannot distinguish, which short tags (<30 nt) existed as small RNAs in cells and which were produced during the library preparation. Also, knowing frequencies of reads matching endogenous miRNAs (Table 1C) does not alleviate the problem since they have a unique origin and their quantitative analysis by deep sequencing is also problematic [27]. Thus, miRNA data may provide an insight into the composition of our deep sequencing libraries but cannot serve as a precise reference point to estimate the abundance of other classes of small RNAs.

dsRNA can induce sequence-specific RNA interference (RNAi) effect. In fact, a study of integrated plasmid repeats in *Drosophila* cells revealed short interfering RNAs (siRNAs) produced from long dsRNA precursors originating from plasmid sequences [28]. We cannot formally exclude RNA silencing because the effect was not directly tested in cells devoid of Dicer. However, it is unlikely that the inhibition of co-transfected reporters involves RNA silencing. First, our previous study of long dsRNA fate in cultured mammalian cells suggests that long dsRNA does not efficiently enter RNAi there [24]. Second, RNAi targets homologous sequences, thus Kan/Neo derived siRNAs could not explain down-regulation of luciferase reporters. Third, a miRNA-like inhibition of luciferase reporters by Kan/Neo-derived siRNAs (i.e. with a less extensive homology), is also unlikely because a large portion of the endogenous gene expression should be strongly affected as well.

It has been reported previously that expression from non-viral and viral vectors is suppressed in mammalian cells in a PKR-dependent manner [9]. Furthermore, while plasmid expression was negatively influenced by the presence of PKR, translation of endogenous proteins was not strongly affected [9]. PKR and dsRNA were also implicated in the selective inhibition of transfected genes in the past [4,5,8]. Notably, negative effects of pEGFP-C1 co-transfection are similar to effects of long dsRNA expression in co-transfection experiments (our unpublished observation). At the same time, we have reported that long dsRNA expression has a negligible impact on expression of the genome and does not provoke robust PKR phosphorylation nor the typical type I IFN response to dsRNA [24]. Therefore, we speculate that pEGFP-C1-induced repression of co-transfected luciferase reporters involves PKR-mediated selective inhibition of luciferase expression, which is induced by dsRNA produced by spurious transcription from the Kan/Neo cassette.

Our work has important implications for transient transfection experiments. First, control experiments must be performed to test how different ratios of co-transfected plasmids influence the outcome. Second, experiments should be designed such that they would not require adding of any extra DNA to equalize the amount of transfected DNA. If extra DNA needs to be added, a parental plasmid lacking the "experimental factor" would be the most reasonable choice as it would neutralize the effect caused by the plasmid backbone. However, if one needs a to add a different "neutral DNA", we recommend using pBS, which has a very common backbone carrying pUC ori and resistance to ampicillin. Third, when combining luciferase and EGFP reporters in one experiment, we recommend avoiding plasmids with the Kan/Neo resistance backbone related to that of pEGFP-C1. Fourth, since several experiments suggest that integrated reporters are insensitive to effects observed during transient transfection [8,24], a production of stable cell lines should be considered as an alternative or a complementary approach to transient transfections whenever possible.

Supporting Information

Figure S1 Co-transfection of pEGFP-C1 inhibits expression of a reporter plasmid without affecting transfection efficiency. HEK-293 cells were co-transfected with 150 ng of pCI-RFPT reporter plasmid labeled with Cy5 (using Label IT® Tracker™ Intracellular Nucleic Acid Localization Kit, Cy5, Mirus) and 350 ng of either pBluescript (pBS) or pEGFP-C1 plasmid. (**A**) Percentage and (**B**) fluorescence intensity of Cy5-positive cells are similar in pBS and pEGFP-C1-transfected cells demonstrating a similar transfection efficiency. Cy5 fluorescence was estimated as a geometric mean of Cy5 fluorescence intensity shown relative to that of the pBS-transfected sample. Cy5 fluorescence in transfected cells was examined 12 hours post-transfection by flow cytometry. (**C**) Percentage of RFP-positive cells is similar in cells transfected either with pEGFP-C1 or pBS. (**D**) A decrease in RFP-fluorescence in pEGFP-C1 transfected cells indicates an inhibition of a pCI-RFPT reporter expression. RFP expression was analyzed 36 hours post-transfection by flow cytometry. RFP fluorescence was estimated as a geometric mean of RFP fluorescence intensity shown relative to that of the pBS-transfected sample. The experiment was performed three times; the graph shows results of a representative experiment (data from the same experiment are shown in Figure 3C).
(PDF)

Figure S2 Transfection with pEGFP-C1 or pRFP-T plasmid does not have toxic effects on transfected cells. (**A**) Percentage of dead (Hoechst 33258-positive) cells after transfection with different plasmids. HEK-293 cells in a 24-well plate were transfected either with pEGFP-C1 or pRFP-T plasmid (150 ng per well, pBluescript was added to 500 ng per well) or pBluescript (500 ng per well). Cells were analyzed by flow cytometry for the incorporation of Hoechst 33258 dye to visualize dead cells 48 hours post-transfection. Cells treated with Puromycin served as a positive control for Hoechst 33258 staining. There was no increase in a percentage of dead cells in cells transfected either with pEGFP-C1 or pRFP-T plasmids (tested plasmids) compared to Bluescript (pBS)-transfected or untransfected cells. (**B**) Relative protein amount in lysates of transfected cells is not significantly affected by different amounts of EGFP and RFP-expressing plasmids. HEK-293 cells were co-transfected with

100 ng of each phRL-SV40 and pGL4-SV40 reporter plasmids and an increasing amount of indicated plasmid (nanograms of plasmid per well are indicated in parentheses). The total amount of transfected DNA was kept constant by adding pBS. After 48 hours, cells were washed with phosphate-buffered saline (PBS) and lysed in the Passive Lysis Buffer (Promega). Total protein amount in lysates was estimated by Bradford Protein assay (Bio-Rad). Data show a result of a representative experiment performed in quadruplicates. Error bars = SEM.
(PDF)

Figure S3 Effects of replacement of a Kan/Neo resistance cassette by an Amp resistance cassette on reporter expression. HEK-293 cells were co-transfected with increasing amounts of pEGFP-C1 (GFP-Kanamycin) or pRFP-T plasmid (RFP-Kanamycin) or their derivatives where Kan/Neo^R cassette was replaced by Amp^R cassette (GFP- and RFP-Ampicillin). The total amount of transfected DNA was maintained constant by adding pBS. The EGFP and RFP fluorescence were analyzed by flow cytometry. Geometric mean fluorescence intensity of **(A)** EGFP-positive and **(B)** RFP-positive cells in a representative experiment is shown. Note that replacement of the resistance cassette in pEGFP-C1 results in a mild reduction of EGFP fluorescence level in EGFP-positive cells while RFP-expressing plasmids yield the same levels of RFP fluorescence regardless of the resistance cassette.
(PDF)

Table S1 Library metrics. Reads were mapped using fastq output files from Seqomics as described in Material and Methods.
(DOCX)

Table S2 Nucleotide change frequencies in transcriptome of cells transiently transfected with selected plasmids. Reads of 18–50 nt in length were mapped to plasmid sequences allowing for up to 5 mismatches. Frequencies of all possible nucleotide changes were evaluated for short (21–26 nt) or long (50 nt) reads separately. Putative A-to-I RNA editing (represented as A-to-G change) is highlighted in yellow. Frequency of A-to-I RNA editing is at least two-times higher compared to frequencies of other nucleotide changes in short (21–26 nt) reads derived from pEGFP-C1-transfected cells.
(DOCX)

Table S3 Multiple A-to-G and other conversions within individual reads. Reads were mapped as indicated in the Table S2. Number of short (21–26 nt) reads mapped with multiple (2–4) identical nucleotide conversions to indicated plasmids is shown in grey. Note the increased number of reads containing multiple A-to-G conversions in pEGFP-C1 sample.
(DOCX)

Table S4 Figure 1 CPM source data.
(XLS)

Acknowledgments

Authors thank to Zdenek Cimburek for help with FACS analysis.

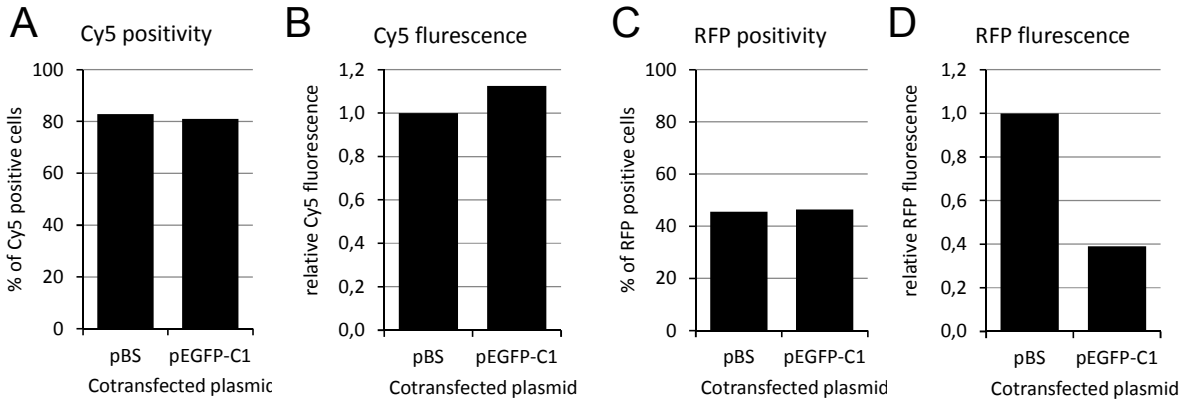
Author Contributions

Conceived and designed the experiments: JN RM PS. Performed the experiments: JN RM. Analyzed the data: JN RM MM. Contributed reagents/materials/analysis tools: JN RM. Wrote the paper: JN RM PS.

References

1. de Wet JR, Wood KV, DeLuca M, Helinski DR, Subramani S (1987) Firefly luciferase gene: structure and expression in mammalian cells. Mol Cell Biol 7: 725–737.
2. Lorenz WW, Cormier MJ, O'Kane DJ, Hua D, Escher AA, et al. (1996) Expression of the Renilla reniformis luciferase gene in mammalian cells. J Biolumin Chemilumin 11: 31–37.
3. Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC (1994) Green fluorescent protein as a marker for gene expression. Science 263: 802–805.
4. Kaufman RJ, Davies MV, Pathak VK, Hershey JW (1989) The phosphorylation state of eucaryotic initiation factor 2 alters translational efficiency of specific mRNAs. Mol Cell Biol 9: 946–958.
5. Kalvakolanu DV, Bandyopadhyay SK, Tiwari RK, Sen GC (1991) Enhancement of expression of exogenous genes by 2-aminopurine. Regulation at the post-transcriptional level. J Biol Chem 266: 873–879.
6. Farr A, Roman A (1992) A pitfall of using a second plasmid to determine transfection efficiency. Nucleic Acids Res 20: 920.
7. Sorscher DH, Cordeiro-Stone M (1994) Inhibition of reporter gene expression in mammalian cells. Effects of distinct carcinogen lesions in DNA. Carcinogenesis 15: 1093–1096.
8. Kaufman RJ (1997) DNA transfection to study translational control in mammalian cells. Methods 11: 361–370.
9. Terenzi F, deVeer MJ, Ying H, Restifo NP, Williams BR, et al. (1999) The antiviral enzymes PKR and RNase L suppress gene expression from viral and non-viral based vectors. Nucleic Acids Res 27: 4369–4375.
10. Paguio A, Almond B, Fan F, Stecha P, Garwin D, et al. (2005) pGL4 vectors: A new generation of luciferase reporter vectors. Promega Notes 89: 7–10.
11. Schagat T, Paguio A, Kopish K (2007) Normalizing genetic reporter assays: approaches and considerations for increasing consistency and statistical significance. Cell Notes: 9–12.
12. Hong SJ, Chae H, Kim KS (2002) Promoterless luciferase reporter gene is transactivated by basic helix-loop-helix transcription factors. Biotechniques 33: 1236–1238, 1240.
13. Osborne SA, Tonissen KF (2002) pRL-TK induction can cause misinterpretation of gene promoter activity. Biotechniques 33: 1240–1242.
14. Schule R, Muller M, Kaltschmidt C, Renkawitz R (1988) Many transcription factors interact synergistically with steroid receptors. Science 242: 1418–1420.
15. Thavathiru E, Das GM (2001) Activation of pRL-TK by 12S E1A oncoprotein: drawbacks of using an internal reference reporter in transcription assays. Biotechniques 31: 528–530, 532.
16. Thirunavukkarasu K, Miles RR, Halladay DL, Onyia JE (2000) Cryptic enhancer elements in luciferase reporter vectors respond to the osteoblast-specific transcription factor Osf2/Cbfa1. Biotechniques 28: 506–510.
17. Vopalensky V, Masek T, Horvath O, Vicenova B, Mokrejs M, et al. (2008) Firefly luciferase gene contains a cryptic promoter. RNA 14: 1720–1729.
18. Shifera AS, Hardin JA (2010) Factors modulating expression of Renilla luciferase from control plasmids used in luciferase reporter gene assays. Anal Biochem 396: 167–172.
19. Chang LJ, Stoltzfus CM (1985) Gene expression from both intronless and intron-containing Rous sarcoma virus clones is specifically inhibited by antisense RNA. Mol Cell Biol 5: 2341–2348.
20. Ghazawi I, Cutler SJ, Low P, Mellick AS, Ralph SJ (2005) Inhibitory effects associated with use of modified Photinus pyralis and Renilla reniformis luciferase vectors in dual reporter assays and implications for analysis of ISGs. J Interferon Cytokine Res 25: 92–102.
21. Sinkkonen L, Hugenschmidt T, Filipowicz W, Svoboda P (2010) Dicer is associated with ribosomal DNA chromatin in mammalian cells. PLoS One 5: e12175.
22. van de Wetering M, Oving I, Muncan V, Pon Fong MT, Brantjes H, et al. (2003) Specific inhibition of gene expression using a stably integrated, inducible small-interfering-RNA vector. EMBO Rep 4: 609–615.
23. Kaname T, Huxley C (2001) Simple and efficient vectors for retrofitting BACs and PACs with mammalian neoR and EGFP marker genes. Gene 266: 147–153.
24. Nejepinska J, Malik R, Filkowski J, Flemr M, Filipowicz W, et al. (2012) dsRNA expression in the mouse elicits RNAi in oocytes and low adenosine deamination in somatic cells. Nucleic Acids Res 40: 399–413.
25. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10: R25.
26. Fang Z, Cui X (2011) Design and validation issues in RNA-seq experiments. Brief Bioinform 12: 280–287.
27. Linsen SE, de Wit E, Janssens G, Heater S, Chapman L, et al. (2009) Limitations and possibilities of small RNA digital gene expression profiling. Nat Methods 6: 474–476.
28. Hartig JV, Esslinger S, Bottcher R, Saito K, Forstemann K (2009) Endo-siRNAs depend on a new isoform of loquacious and target artificially introduced, high-copy sequences. EMBO J 28: 2932–2944.

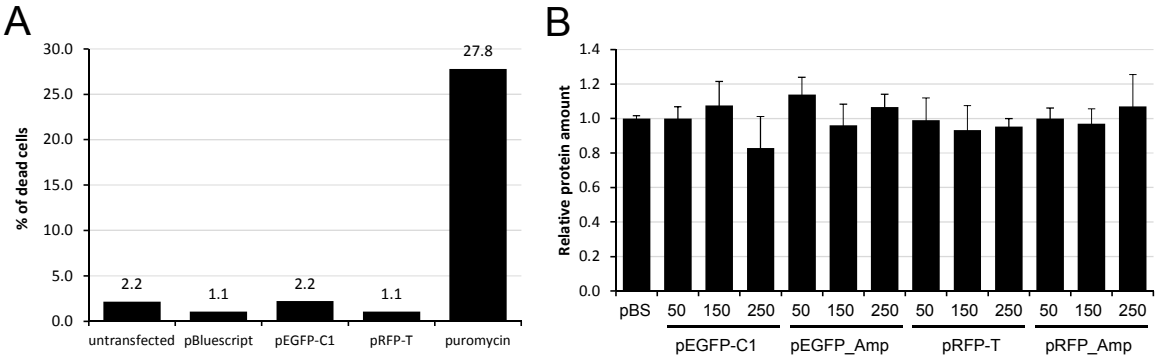
Figure S1



Supplementary Figure S1: Co-transfection of pEGFP-C1 inhibits expression of a reporter plasmid without affecting transfection efficiency.

HEK-293 cells were co-transfected with 150 ng of pCI-RFPT reporter plasmid labeled with Cy5 (using Label IT® Tracker™ Intracellular Nucleic Acid Localization Kit, Cy5, Mirus) and 350 ng of either pBluescript (pBS) or pEGFP-C1 plasmid.. **(A)** Percentage and **(B)** fluorescence intensity of Cy5-positive cells are similar in pBS and pEGFP-C1-transfected cells demonstrating a similar transfection efficiency. Cy5 fluorescence was estimated as a geometric mean of Cy5 fluorescence intensity shown relative to that of the pBS-transfected sample. Cy5 fluorescence in transfected cells was examined 12 hours post-transfection by flow cytometry. **(C)** Percentage of RFP-positive cells is similar in cells transfected either with pEGFP-C1 or pBS. **(D)** A decrease in RFP-fluorescence in pEGFP-C1 transfected cells indicates an inhibition of a pCI-RFPT reporter expression. RFP expression was analyzed 36 hours post-transfection by flow cytometry. RFP fluorescence was estimated as a geometric mean of RFP fluorescence intensity shown relative to that of the pBS-transfected sample. The experiment was performed three times; the graph shows results of a representative experiment (data from the same experiment are shown in Figure 3C).

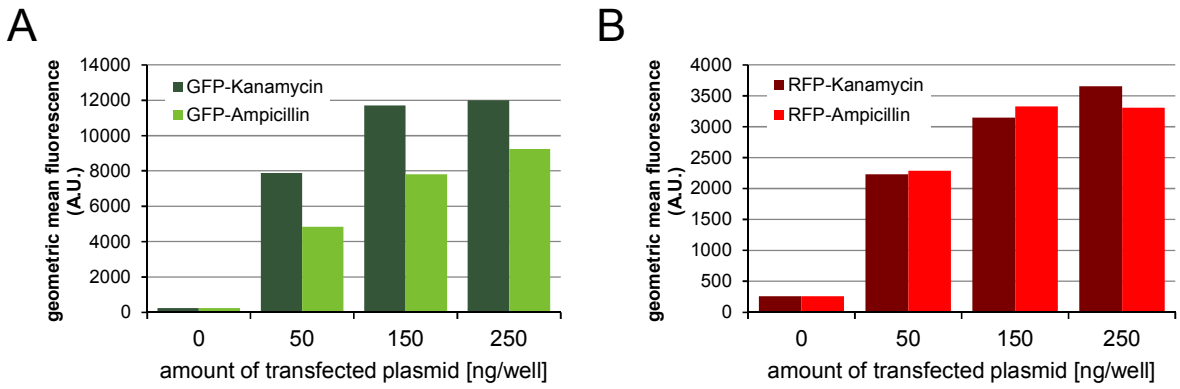
Figure S2



Supplementary Figure S2: Transfection with pEGFP-C1 or pRFP-T plasmid does not have toxic effects on transfected cells.

(A) Percentage of dead (Hoechst 33258-positive) cells after transfection with different plasmids. HEK-293 cells in a 24-well plate were transfected either with pEGFP-C1 or pRFP-T plasmid (150 ng per well, pBluescript was added to 500 ng per well) or pBluescript (500 ng per well). Cells were analyzed by flow cytometry for the incorporation of Hoechst 33258 dye to visualize dead cells 48 hours post-transfection. Cells treated with Puromycin served as a positive control for Hoechst 33258 staining. There was no increase in a percentage of dead cells in cells transfected either with pEGFP-C1 or pRFP-T plasmids (tested plasmids) compared to Bluescript (pBS)-transfected or untransfected cells. **(B)** Relative protein amount in lysates of transfected cells is not significantly affected by different amounts of EGFP and RFP-expressing plasmids. HEK-293 cells were co-transfected with 100 ng of each phRL-SV40 and pGL4-SV40 reporter plasmids and an increasing amount of indicated plasmid (nanograms of plasmid per well are indicated in parentheses). The total amount of transfected DNA was kept constant by adding pBS. After 48 hours, cells were washed with phosphate-buffered saline (PBS) and lysed in the Passive Lysis Buffer (Promega). Total protein amount in lysates was estimated by Bradford Protein assay (Bio-Rad). Data show a result of a representative experiment performed in quadruplicates. Error bars = SEM.

Figure S3



Supplementary Figure S3: Effects of replacement of a Kan/Neo resistance cassette by an Amp resistance cassette on reporter expression.

HEK-293 cells were co-transfected with increasing amounts of pEGFP-C1 (GFP-Kanamycin) or pRFP-T plasmid (RFP-Kanamycin) or their derivatives where Kan/Neo^R cassette was replaced by Amp^R cassette (GFP- and RFP-Ampicillin). The total amount of transfected DNA was maintained constant by adding pBS. The EGFP and RFP fluorescence were analyzed by flow cytometry. Geometric mean fluorescence intensity of **(A)** EGFP-positive and **(B)** RFP-positive cells in a representative experiment is shown. Note that replacement of the resistance cassette in pEGFP-C1 results in a mild reduction of EGFP fluorescence level in EGFP-positive cells while RFP-expressing plasmids yield the same levels of RFP fluorescence regardless of the resistance cassette.

Table S1

sample	phRL-SV40	pGL4-SV40	pBS/pEGFP-C1
number of reads in library	56374066	45833776	55581560
plasmid mapped (perfect)	41555	15417	251458
genome mapped reads (perfect)	11757994	9003682	8668128
genome mapped reads (1x mismatch)	3351205	2248047	2420240
genome mapped reads (2x mismatch)	1355861	1013057	1030870
genome mapped reads (3x mismatch)	1670962	1347401	1240010
genome mapped reads (all)	18136022	13612187	13359248
genome and plasmid mapped (perfect)	11799549	9019099	8919586

Supplementary Table S1: Library metrics.

Reads were mapped using fastq output files from Seqomics as described in Material and methods.

Table S2

pGL4-SV40	21-26 nt	A	T	C	G	50 nt	A	T	C	G
	A	/	6.30%	9.56%	6.97%	A	/	8.19%	9.93%	10.47%
	T	8.62%	/	7.79%	5.75%	T	6.05%	/	6.05%	5.39%
	C	9.62%	7.41%	/	8.07%	C	7.94%	7.23%	/	7.90%
	G	9.12%	9.23%	11.55%	/	G	10.85%	10.00%	9.98%	/
phRL-SV40	21-26 nt	A	T	C	G	50 nt	A	T	C	G
	A	y	6.93%	7.55%	10.44%	A	/	8.60%	9.36%	12.77%
	T	9.10%	/	11.27%	6.62%	T	6.87%	/	7.68%	7.21%
	C	6.41%	8.69%	/	8.17%	C	6.22%	6.89%	/	7.08%
	G	8.17%	6.72%	9.93%	/	G	9.55%	8.99%	8.78%	/
pEGFP-C1	21-26 nt	A	T	C	G	50 nt	A	T	C	G
	A	/	7.67%	7.01%	**18.90%**	A	/	8.66%	9.46%	10.58%
	T	6.32%	/	6.84%	6.18%	T	5.31%	/	5.65%	5.46%
	C	9.45%	7.49%	/	7.32%	C	9.13%	9.22%	/	9.54%
	G	8.49%	6.32%	8.01%	/	G	9.10%	8.69%	9.22%	/
pBS	21-26 nt	A	T	C	G	50 nt	A	T	C	G
	A	/	6.47%	6.71%	10.31%	A	/	9.77%	9.87%	13.73%
	T	8.27%	/	8.15%	6.83%	T	7.89%	/	8.26%	6.63%
	C	8.63%	7.67%	/	4.56%	C	5.89%	6.27%	/	6.76%
	G	11.03%	8.75%	12.59%	/	G	8.78%	7.75%	8.38%	/

Supplementary Table S2: Nucleotide change frequencies in transcriptome of cells transiently transfected with selected plasmids.

Reads of 18-50 nt in length were mapped to plasmid sequences allowing for up to 5 mismatches. Frequencies of all possible nucleotide changes were evaluated for short (21-26 nt) or long (50 nt) reads separately. Putative A-to-I RNA editing (represented as A-to-G change) is highlighted in yellow. Frequency of A-to-I RNA editing is at least two-times higher compared to frequencies of other nucleotide changes in short (21-26 nt) reads derived from pEGFP-C1-transfected cells.

Table S3

identical conversions		phRL-SV40		pGL4-SV40		pBS		pEGFP-C1	
		A/G	other	A/G	other	A/G	other	A/G	other
2x	number of reads	**10**	**2**	**0**	**2**	**6**	**2**	**90**	**11**
	%	83.3	16.7	0	100	75	25	89.1	10.9
3x	number of reads	**9**	**0**	**3**	**0**	**0**	**0**	**48**	**9**
	%	100	0	100	0	0	0	84.2	15.8
4x	number of reads	**0**	**0**	**4**	**0**	**4**	**0**	**44**	**1**
	%	0	0	100	0	100	0	97.8	2.2

Supplementary Table S3: Multiple A-to-G and other conversions within individual reads.

Reads were mapped as indicated in the Table S2. Number of short (21-26 nt) reads mapped with multiple (2-4) identical nucleotide conversions to indicated plasmids is shown in grey. Note the increased number of reads containing multiple A-to-G conversions in pEGFP-C1 sample.

Supplement 4

Nejepinska J, Malik R, Svoboda P.

Expressed dsRNA provokes selective sequence-independent PKR response in transient co-transfection experiments

Manuscript attached

Expressed dsRNA provokes selective sequence-independent PKR response in transient co-transfection experiments

Authors:

Jana Nejepinska, Radek Malik, and Petr Svoboda[+]

Affiliation:

Institute of Molecular Genetics, Videnska 1083, 14220 Prague 4, Czech Republic

[+]corresponding author:

Petr Svoboda (svobodap@img.cas.cz)

Running title:

Inhibitory effects of dsRNA-expression

Keywords:

dsRNA, RNA hairpin, PKR, plasmid

Abstract

In mammals, double-stranded RNA (dsRNA) plays roles in sequence-specific RNA interference, sequence-independent interferon response, and RNA editing by adenosine deaminases. We have previously shown that long hairpin dsRNA expression in cultured cells does not activate the interferon response, it is poorly processed into siRNAs, and it is partially edited. Here, we demonstrate that dsRNA expressed from transiently transfected plasmids strongly inhibits expression of co-transfected reporter plasmids but not expression of endogenous genes or reporters stably integrated in the genome. The inhibition is concentration-dependent and independent of a cell type, transfection method, or dsRNA sequence. The inhibition occurs at the level of translation and is mediated by protein kinase R (PKR). PKR binds the expressed dsRNA, becomes phosphorylated and changes its distribution along polysome fractions. In conclusion, we demonstrate that expression from plasmids is selectively repressed if one of co-transfected plasmids produces dsRNA. Our results highlight the importance of proper controls and careful interpretation of co-transfection experiments.

Introduction

Double-stranded RNA (dsRNA) can be either endogenous or exogenous. Endogenous dsRNA is formed from cellular transcripts upon basepairing of complementary RNA sequences (either by intramolecular pairing, forming a hairpin, or by pairing between two molecules). A common source of exogenous dsRNA is a viral infection. Most viruses form dsRNA in a certain phase of their life cycle; therefore, dsRNA is recognized by a vertebrate cell as a hallmark of viral presence (reviewed in [1]). In mammalian cells, dsRNA can enter three pathways: RNA interference (RNAi), the interferon (IFN) response, and RNA editing. RNAi causes a sequence-specific RNA degradation mediated by ~22 nt small interfering RNAs (siRNAs) produced from long dsRNA by RNase III Dicer (reviewed in [2]). The IFN response is a complex of vertebrate pathways involved in the innate immune response against viruses (reviewed in [3]). Protein kinase R (PKR), a key enzyme in the IFN pathway, is activated upon binding of dsRNA to its dsRNA-binding domain. Activated PKR phosphorylates α-subunit of eukaryotic initiation factor 2 (eIF2α), which stabilizes the GEF-eIF2-GDP complex and, consequently, causes the inhibition of translation initiation [4,5]. In addition to PKR, the IFN response involves coordinated action of other molecules, such as oligoadenylate synthetase, RNase L, RIG-I, or NF-κB [1]. The inhibition of proteosynthesis via PKR is commonly considered as sequence-independent and global (non-selective). Nevertheless, several groups observed local PKR effects and the inhibition of specific mRNAs [6-8]. RNA editing is mediated by adenosine deaminase acting on RNA (ADAR) family of enzymes. ADARs are nuclear and cytoplasmic enzymes activated by dsRNA and converting adenosines to inosines (which are recognized as guanosines during translation). Editing of dsRNA can cause target RNA degradation or modify its coding potential (reviewed in [9]).

Transient co-transfection is a common approach to deliver an experimental plasmid together with appropriate reporters into mammalian cells. Among the most common reporters are luciferases where a dual luciferase reporter system allows for using one luciferase as a targeted experimental reporter and the other one as a non-targeted

control for normalization. We have previously shown that long dsRNA expression neither activates the IFN pathway nor induces RNAi [10]. However, we noticed that expressed dsRNA had a common inhibitory effect on co-transfected luciferase reporters; hence, we decided to explore this inhibitory phenomenon in more detail. Here, we show that transient co-transfection of a hairpin-expressing plasmid inhibits expression of co-transfected reporter plasmids in a sequence-independent manner. The inhibition occurs posttranscriptionally, affects expression from co-transfected plasmids but neither endogenous genes nor stably integrated reporters. The repression occurs at the level of translation and is PKR dependent. Our data suggest that PKR can elicit a selective repression of translation of mRNAs from transfected plasmids upon appearance of dsRNA in a transient transfection. However, this PKR activity does not induce the IFN response and does not suppress translation of transcripts from the genome. We propose that PKR response has two forms: the first one, which results in the global repression of translation and is associated with the IFN response and the second one, which blocks translation more selectively and does not induce the IFN response. Furthermore, our results provide an important framework for correct interpretation of experiments based on transient transfections.

Results and Discussion

Plasmids expressing dsRNA inhibit co-transfected reporters

To express long double-stranded RNA in mammalian cells, we used a plasmid (pCAGEGFP-MosIR), in which the inverted repeat of the *Mos* gene sequence (MosIR) was inserted into the 3'UTR of an EGFP reporter controlled by a strong chimeric (CMV/β-actin) promoter (Fig. 1A) [10]. The plasmid-derived transcript forms an intramolecular duplex (~500 bp) downstream of the EGFP coding sequence; formation of dsRNA structure has been demonstrated *in vitro* [11] and *in vivo* [10]. *Mos* expression and function are restricted to oocytes [12,13]; hence, effects observed in other cell types are sequence-independent

and are not mediated by changes in endogenous *Mos* mRNA levels. Furthermore, dsRNA derived from the *Mos* sequence is one of the most studied long dsRNAs and its processing and effects have been described to considerable detail in mammalian cells [10,11,14-16]. pCAGEGFP-MosIR plasmid did not induce efficient RNAi in cultured cells, presumably because of inefficient processing into siRNAs [10]. However, we noticed a unique, sequence-independent inhibitory effect of pCAGEGFP-MosIR on luciferase activities produced from co-transfected luciferase reporters and we decided to investigate this phenomenon in more detail.

For initial characterization of inhibitory effects of pCAGEGFP-MosIR, we co-transfected HEK293 cells with a constant amount of *Renilla* luciferase (RL) and firefly luciferase (FL) reporters and with increasing amounts of pCAGEGFP-MosIR plasmid. As a control, we used the parental pCAGEGFP lacking the inverted repeat or pCAGEGFP-MosMos, in which the Mos sequences were inserted as a tandem repeat in a head-to-tail orientation to produce a plasmid with the same size and sequence composition but not producing hairpin dsRNA (Fig. 1A). pCAGEGFP-MosIR caused a strong (up to 90 %) concentration-dependent decrease in both luciferase activities while the inhibitory effect of the two control plasmids was small (Fig. 1B). These data suggest that the inverted repeat inserted into pCAGEGFP was responsible for the strong suppression of co-transfected luciferase reporters. The inhibition was independent of transfection procedure used, because it was also observed using Nanofectin transfection reagent or calcium phosphate transfection (Fig. S1).

Observed suppression was presumably sequence-independent because *Mos* sequence lacks similarity to luciferase reporters. A minor reduction of luciferase activities by co-transfected pCAGEGFP or pCAGEGFP-MosMos was not surprising because we have previously reported that plasmids may affect expression of co-transfected reporters to various degrees, presumably as a result of their complex transcription [17]. In any case, pCAGEGFP and pCAGEGFP-MosMos effects can be used as a baseline, to which pCAGEGFP-MosIR results can be compared.

To test whether the inhibitory effect of pCAGEGFP-MosIR is dependent on the activity of the CMV/β-actin promoter, we linearized the pCAGEGFP-MosIR plasmid upstream of the CMV/β-actin promoter (without affecting the coding sequence) or downstream of it (disrupting MosIR transcription) (Fig. 1C). Interestingly, pCAGEGFP-MosIR linearization downstream of the CMV/β-actin promoter relieved the suppression of luciferase activities to levels comparable to pCAGEGFP and pCAGEGFP-MosMos plasmids (Fig. 1B). At the same time, pCAGEGFP-MosIR linearized upstream of the CMV/β-actin promoter caused clearly stronger suppression of luciferase activities than linearization downstream of the promoter (Fig. 1C) although not as strong as the intact pCAGEGFP-MosIR. These data indicate that the active transcription of the inverted repeat is necessary for the suppression of the co-transfected luciferase reporters. A weaker effect of the pCAGEGFP-MosIR linearized upstream of the CMV/β-actin promoter when compared to the intact plasmid could be explained by reduced expression from the linearized plasmid (data not shown).

Next, we wanted to investigate whether the observed inhibition is restricted to luciferase reporters and luminometric measurement. Therefore, we employed flow cytometry analysis, in which RFP-expressing plasmid was used as a reporter and co-transfected with either pCAGEGFP or pCAGEGFP-MosIR plasmid. As the RFP expression was inhibited in cells co-transfected with pCAGEGFP-MosIR plasmid (Fig. S2) compared to pCAGEGFP-co-transfected cells, we assume that the inhibition of reporter activity caused by co-transfection with hairpin-expressing plasmid is independent on both RL and FL-based reporters and luminometric assay.

To analyze dsRNA expression from pCAGEGFP-MosIR in depth, we performed high-throughput sequencing of HEK293 cells transfected either with pCAGEGFP or pCAGEGFP-MosIR using SOLiD technology. The plasmid backbone yielded similar coverage pattern in both plasmids (Fig. 2A), while the absolute number of reads derived from the EGFP coding sequence of pCAGEGFP-MosIR was approximately three times lower compared to pCAGEGFP. To confirm that pCAGEGFP-MosIR produces dsRNA, we performed analysis of adenosine deamination of short SOLiD sequence tags as previously described [10,17]. As

predicted, RNA editing of small RNAs, manifested as adenosine/guanosine (A/G) conversion, clustered to the MosIR region (Fig. 2B). Quantitative analysis of A/G conversion revealed massive editing of small RNAs within the MosIR region (Fig. 2C). On the other hand, A/G conversion was negligible in small RNA reads mapping to EGFP region of the same plasmid (Fig. 2D). These data serve as an indirect evidence of dsRNA formation from MosIR region and go along with our previous data [10,17].

The observed suppression of co-transfected luciferase reporters appears to be a general phenomenon because we have observed the inhibitory effect of pCAGEGFP-MosIR in HeLa cells (Fig. 3A), P19 cells (Fig. S3), mouse embryonic stem cells, primary mouse fibroblasts, and HepG2 cells (data not shown). We also tested two other hairpin sequences. Inverted repeats made of mouse *Lin28a/b* and *Elavl2* sequences were inserted into the same position in pCAGEGFP as MosIR hairpin and the resulting plasmids pCAGEGFP-Lin28IR and pCAGEGFP-Elavl2IR were co-transfected with luciferase reporters in HEK293 cells. Plasmids in which the CMV/β-actin promoter was replaced with ZP3 promoter (pZP3EGFP-Lin28IR and pZP3EGFP-Elavl2IR) served as controls. As ZP3 promoter is oocyte-specific, dsRNA should not be produced when pZP3EGFP-Lin28IR or pZP3EGFP-Elavl2IR are transfected into HEK293 cells. Similarly to pCAGEGFP-MosIR, pCAGEGFP-Lin28IR and pCAGEGFP-Elavl2IR caused concentration-dependent decrease in both luciferase activities (Fig. 3B, C). At the same time, co-transfection of ZP3 derivatives did not have any detectable inhibitory effect, independently confirming that the inhibition is mediated by actively transcribed dsRNA. Taken together, dsRNA produced from a transiently transfected plasmid has a general, concentration-dependent, and sequence-independent inhibitory effect on the expression from co-transfected reporters.

dsRNA expression does not suppress stably integrated reporters

The inhibitory effect of Mos dsRNA on co-transfected reporters was surprising because our previous analysis did not reveal any marked changes in endogenous gene expression [10]. Transient transfection of pCAGEGFP-MosIR showed only minor effect on the cellular transcriptome and no activation of the IFN pathway as transfected cells had normal

proliferation rate and morphology. In contrast to results from transient transfections, we did not observe inhibition of RL and FL luciferase reporters integrated in the genome (Fig. 4A). This led us to the hypothesis that the silencing phenomenon predominantly affects expression from transiently transfected plasmids. To confirm it, we created HEK293 cell line with a stable integration of RL reporter alone. This cell line was co-transfected with FL reporter and an increasing amount of pCAGEGFP-MosIR plasmid. Consistently with previous results, FL activity was strongly reduced in a concentration-dependent manner while RL activity remained unaffected (Fig. 4B).

As both transfected plasmids and cellular DNA are transcribed in the nucleus, it is striking that cells are able to recognize the expression source of the same reporter sequence and suppress expression of transiently co-transfected plasmids while ignoring the expression from plasmids integrated in the genome.

dsRNA-mediated suppression occurs at the level of translation

To get a better insight into this phenomenon, we examined at which level is the plasmid expression inhibited. The suppression might take place at the level of plasmid DNA (entry into/exclusion from the nucleus, plasmid stability), transcription, or it can occur post-transcriptionally. Thus, we examined steady-state levels of transcripts originating from co-transfected plasmids using real-time PCR. These results showed that steady state levels of transcripts from constant amounts of co-transfected luciferase reporters remained constant while the level of transcripts from pCAGEGFP-MosIR was rising proportionally to the amount of co-transfected pCAGEGFP-MosIR (Fig. 5A). These data indicate that the mechanism suppressing co-transfected reporters is post-transcriptional. In addition, we did not observe any effect at the level of co-transfected plasmid DNA (data not shown).

Next, we used polysome profiling to analyze behavior of luciferase reporter transcripts and endogenous mRNA in HEK293 cells co-transfected with luciferase reporters and either pCAGEGFP-MosIR or pCAGEGFP plasmid. We isolated RNA from fractions corresponding to monosomes (80S) and polysomes and analyzed the abundance of mRNA in each fraction using real-time PCR (Fig. 5B). Consistently with our previous

results, we observed lower relative abundance of luciferase mRNAs in polysome fractions from cells co-transfected with dsRNA-expressing plasmid compared to control pCAGEGFP plasmid. Remarkably, while there was a high amount of pCAGEGFP mRNA in polysome fractions, the amount of pCAGEGFP-MosIR mRNA on polysomes was relatively low, suggesting that pCAGEGFP-MosIR inhibits translation of its own transcript. In contrast, endogenous *Hprt1* and *$\beta_2$microglobulin* (B2M) mRNAs remained abundant in polysome fractions regardless of co-transfected plasmid. These data support the notion that suppression selectively inhibits all co-transfected reporters, while endogenous mRNAs (represented by *Hprt1 and B2M*) remain efficiently translated.

dsRNA-mediated suppression is PKR-dependent

In the search for the mechanism inhibiting co-transfected luciferase reporters, we revisited the possible role of PKR. Our previous data excluded the activation of the IFN response because we did not observe activation of IFN-stimulated genes and the PKR phosphorylation in transfected cells was low [10]. However, our results were consistent with data of Terenzi *et al*., describing that expression from non-viral and viral vectors is suppressed in mammalian cells in a PKR-dependent manner while translation of endogenous proteins is not strongly affected [18]. PKR and dsRNA were also implicated in the selective inhibition of transfected genes by other groups [7,19]. Consistently, our previous study of spurious RNA expression from transfected plasmid showed that some plasmids unexpectedly produce dsRNA and inhibit expression of co-transfected reporter plasmids [17].

To test the role of PKR, we generated a stable PKR knockdown cell line expressing shRNA targeting *Pkr* mRNA (Fig. S4). Transient transfection of selected cell lines with pCAGEGFP-MosIR plasmid resulted in the relief of repression of luciferase reporters (Fig. 6A) indicating that repression is PKR-dependent. As PKR activation involves direct dsRNA binding to dsRNA-binding domain of PKR, we tested whether PKR directly binds dsRNA expressed from pCAGEGFP-MosIR by RNA immunoprecipitation followed by real-time PCR analysis. The plasmid-derived RNA was significantly enriched in pCAGEGFP-MosIR-

transfected sample but not in the control (pCAGEGFP-transfected) sample (Fig. 6B). This suggests that MosIR hairpin is directly bound by PKR. Accordingly, we detected increased PKR phosphorylation exclusively in pCAGEGFP-MosIR-transfected cells, with a concentration-dependent pattern (Fig. 6C). On the other hand, phosphorylated PKR was completely absent in pCAGEGFP-treated cells and control cells. The phosphorylation of eIF2α, a main PKR substrate, was increased only marginally. Therefore, our results demonstrate that PKR directly binds dsRNA and mediates the suppression of co-transfected reporters.

In addition, we analyzed PKR distribution along polysome profiles of HEK293 cells transfected either with pCAGEGFP or pCAGEGFP-MosIR. First, we noticed that expression of long dsRNA results in the increased amount of monosomes/free ribosomes and reduced amount of polysomes (Fig. 6D). This indicates that long dsRNA expression significantly inhibits a considerable part of proteosynthesis. It is possible that transcripts from co-transfected plasmids occupy a large portion of translation machinery. Alternatively, cellular translation could be affected in general but the effect does not affect cell viability, has a minimal impact on translation of cellular mRNAs (Fig. 5B) and on the transcriptome [10].

Interestingly, distribution of PKR and eIF2α and their phosphorylated forms along the polysome profile differed between fractions from cells transfected with pCAGEGFP-MosIR or pCAGEGFP (Fig. 6E). This is apparent when compared with distribution of the ribosomal protein S14 (RPS14), which should reflect amount of ribosomes in individual fractions. Consistently with the higher monosome peak in polysome profiles from cells transfected with pCAGEGFP-MosIR, we observed higher signal from RPS14 in the same fractions suggesting that dsRNA expression leads to accumulation of monosomes. Remarkably, PKR was mostly found in soluble and 40S fractions but not in monosome (80S) or polysome fractions in samples expressing dsRNA. Phosphorylated PKR was almost negligible in all pCAGEGFP fractions, consistently with its absence in the unfractionated lysate (Fig. 6C). In cells transfected with pCAGEGFP-MosIR, both PKR and its phosphorylated form were strongly redistributed from soluble to monosome and polysome fractions. On the other

hand, distribution of eIF2α and its phosphorylated form was affected only marginally. These data might indicate that upon dsRNA binding, PKR is activated and selectively redistributed from the soluble pool towards monosome and polysome fractions. Whether this reflects translational inhibition or PKR binding to the translated MosIR transcript is under investigation.

The inhibition of translation of plasmid-borne transcripts raises several questions. First, which feature of dsRNA triggers the response. While we use dsRNA, which forms a hairpin in the 3'UTR of a translatable mRNA, a similar inhibitory effect was observed upon transfection of pEGFP-C1 plasmid, where dsRNA presumably forms upon basepairing of sense and antisense transcripts from the kanamycin/neomycin region [17]. Thus, it needs to be tested which features of dsRNA are essential for the effect. Second, it is unclear how cells distinguish between transcripts originating from a transiently-transfected plasmid and a plasmid stably integrated in the genome (Fig. 4B). One possibility is that plasmid-born transcripts are somehow differentiated either by RNA properties (such as poly(A) length or a covalent modification [20,21]) or by different protein factors bound to these transcripts. Another possibility is that the sensitivity is circumstantial. Several reports demonstrated a negative correlation between plasmid-derived translation and the amount of PKR [7,18,22] or RNase L [18] and, vice versa, enhanced plasmid expression after treatment with PKR inhibitors [18,23,24]. It is possible that dsRNA appearance in a co-transfection experiment makes newly synthesized mRNAs more prone to translational repression. Indeed, PKR inhibition was demonstrated to accelerate synthesis of newly synthesized mRNA [23]. In addition, depletion of eIF4G (a translation initiation factor necessary for de novo translation, but dispensable for subsequent rounds of translation) suppresses translation of plasmid-derived RNAs much more than ongoing cellular translation [25]. Our preliminary experiments would favor the latter option since plasmid-borne transcripts appear not to be dsRNA sensitive when they are transfected 24 hours before the dsRNA-expressing plasmid (data not shown). However, this hypothesis needs further testing using different types of dsRNA and inducible expression of dsRNA allowing for precise timing of dsRNA appearance during expression of luciferase reporters.

Our results suggest that dsRNA can induce two types of PKR responses: (1) the all-out response, marked by global repression of translation and activation of IFN-stimulated genes and (2) a partial response, marked by selective translation repression of specific “unfavorable” transcripts. We provide a clear example of selective repression of transcripts from plasmids when one of co-transfected plasmids produces dsRNA. It becomes apparent that spurious transcription of plasmids can inadvertently produce dsRNA. Thus, it is important to use proper controls and carefully interpret inhibitory effects in co-transfection experiments.

Material and Methods

Plasmids

Schematic structures of the relevant parts of plasmid constructs used in the project are shown in Fig. 1A and described in the text. Plasmids were purchased from the manufacturers specified in the parentheses: pBluescript II KS (+) (Stratagene), pGL4-SV40 (Promega; for simplicity referred to as FL) and phRL-SV40 (Promega; for simplicity referred to as RL). The construction of plasmids pCAGEGFP-MosIR [10] and pCAGEGFP [26] was described previously. ZP3EGFP-Lin28IR (Flemr *et al.*, unpublished) and ZP3EGFP-Elavl2IR (Chalupnikova *et al.*, unpublished) plasmids containing mouse *Lin28a/b* and *Elavl2* sequences, respectively, were constructed similarly as pCAGEGFP-MosIR plasmid and will be described in detail elsewhere. For pCAGEGFP-Lin28IR and pCAGEGFP-Elavl2IR plasmids construction, parental pCAGEGFP plasmid was modified by inserting NotI site downstream of EGFP coding sequence and Lin28IR and Elavl2IR fragments were cloned into NotI site. pCAGEGFP-MosMos plasmid was produced by inserting the *Mos* coding sequence into BglII site of pCAGEGFP plasmid. Screening was performed to select the plasmid containing a head-to-tail double Mos insertion. All plasmids were verified by sequencing.

Cell culture and transfection

Human HEK293 and HeLa cells were maintained in DMEM (Sigma) supplemented with 10 % fetal calf serum (Sigma), penicillin (100 U/mL, Invitrogen), and streptomycin (100 µg/mL, Invitrogen) at 37 °C and 5 % CO_2 atmosphere. For transfection, cells were plated on a 24-well plate, grown to 50 % density and transfected using Turbofect *in vitro* Transfection Reagent (Thermo Scientific).

Cells were co-transfected with 100 ng of each FL and RL reporter plasmids and various amount of tested plasmid (50-250 ng per well). The total amount of transfected DNA was kept constant by adding promoterless pBluescript or parental pCAGEGFP plasmid. After 48 hours, cells were washed with phosphate-buffered saline (PBS) and lysed with the Passive Lysis Buffer (Promega). Luciferase reporter activity was assessed using the Dual-Luciferase Reporter Assay (Promega) and luminiscence intensity was measured by Modulus

Microplate Multimode Reader (Turner Biosystems). Variability in cell densities was minimized by adjustment to total protein concentration measured by Protein assay kit (Bio-Rad) according to the manufacturer's protocol.

For transfection of linearized plasmids, pCAGEGFP-MosIR was digested either by SalI or EagI (New England Biolabs), dephosphorylated by Shrimp alkaline phosphatase (Fermentas), and purified using QIAquick PCR Purification Kit (Qiagen); all steps were performed according to the manufacturer's instructions.

HEK293 cells stably expressing both RL and FL reporters were described elsewhere [10]. Similar procedure was used for establishing RL-only-expressing stable HEK293 cells. HeLa PKR knock-down cells were generated by a similar procedure except that Zeocin (250 µg/ml) was used for selection of positive clones. shRNA (5'-GATCCCCGGCAGTTAGTCCTTTATTATTCAAGAGATAATAAAGGACTAACTGCCTTTTTA-3') targeting PKR was cloned into pTER plasmid [27]. A stable cell line expressing empty pTER plasmid was used as a control. The positively selected individual clones were tested for PKR knockdown by western blotting.

Next generation sequencing (Solid)

HEK293 cells were plated on 6-well plates and grown to 50 % density. Cells were transfected with 2.7 µg/well of either pCAGEGFP or pCAGEGFP-MosIR plasmid, cultured for 48 hours, washed with PBS, and total RNA was isolated using RNAzol (MRC, Inc.) according to the manufacturer's protocol. RNA quality was verified by Agilent 2100 Bioanalyzer. The library construction from total RNA and deep sequencing of RNA transcriptome were performed by Seqomics (Szeged, Hungary) using SOLiD (version 4.0) sequencing platform. Bioinformatic analysis was performed as described previously [10]. High throughput sequencing data in color-space format were deposited in the GEO database.

Western blotting

Parental or transfected HEK293 cells were grown in 6-well plates. Before collection, cells were washed with PBS and lysed in Whole-cell lysis buffer (10% glycerol, 0.5 mM EDTA, 1 mM DTT, 2 mM natrium fluoride, 0.2 % Triton X-100 in PBS pH 7.4) supplemented with protease and phosphatase inhibitors (Calciochem). Proteins were separated on 10 % polyacrylamide gel and transferred to PVFD membrane (Millipore). Following antibodies were used for detection: total PKR (Abcam #ab32052, 1:5000 dilution), PKR [pT^{446}] (Abcam #ab32036, 1:1000), total EIF2α (Santa-Cruz #sc-11386, 1:1000), EIF2α [pS^{52}] (Life Techonologies #44728G, 1:1000), RPS14 (Santa-Cruz #sc-68873, 1:1000), and tubulin (Sigma #T6074, 1:5000). SuperSignal West Femto Chemiluminescent Substrate (Pierce) was used for detection.

Polysome profiling and real-time PCR

Cells were plated on 15-cm dishes, grown to 50 % density and co-transfected with FL (8 μg per dish), RL (8 μg per dish), and either pCAGEGFP or pCAGEGFP-MosIR plasmid (12 μg per dish) using polyethyleneimine. After 24 hours, cells were harvested with PBS in the presence of cycloheximide (100 μg/ml). Whole-cell extracts were prepared in breaking buffer (20 mM Tris-HCl, pH 7.5, 50 mM KCl, 10 mM $MgCl_2$, 10 % glycerol, 1 % Triton X-100, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, Protease Inhibitor Cocktail Set III (Calbiochem), RiboLock RNase inhibitor (Thermo Scientific) (0.3 U/μl)). Ten or fifteen A_{260} units of whole-cell extracts were separated by velocity sedimentation on a 5 % to 45 % sucrose gradient by centrifugation at 39,000 rpm for 2.5 h in SW41Ti rotor (Beckman). Gradient fractions were collected and scanned at 254 nm to visualize ribosomal species.

For RNA isolation, each fraction was mixed with 1 ml of ice cold 96 % ethanol and 3 M sodium acetate (20:1) and 1 μl of glycogen (RNA grade, Fermentas) and incubated overnight at -20 °C. Samples were centrifuged (16,000 g, 30 min, 4 °C) and RNA from pellets was isolated using RNAzol (MRC) according to the manufacturer's protocol. Residual DNA was degraded by Turbo DNase (Ambion) and the same portion of each fraction was reverse-transcribed using RevertAid Premium Reverse Transcriptase (Thermo

Scientific) according to the manufacturer's instructions. Reverse transcriptase was omitted in control (-RT) samples. Real-time PCR was performed on LC480 machine (Roche) using Maxima SYBR green qPCR Master Mix (Fermentas). Values of crossing points were evaluated and corrected according to PCR efficiency for each reaction. For protein isolation, each fraction was ethanol precipitated according to [8]. The equivalent portion of each fraction was used for SDS-PAGE and western blotting.

RNA immunoprecipitation

HEK293 cells were transfected by pCAGEGFP or pCAGEGFP-MosIR plasmids using polyethyleneimine. After 48 hours, cells were washed with PBS and scraped into IP buffer (20 mM HEPES, 1 mM EDTA, 0.1 M NaCl, 0.5 % NP-40 (Nonidet), 1 mM DTT, 0.2 mM Na_3VO_4, 50 mM NaF, 1X Protease inhibitor cocktail set (Calbiochem), 1X Phosphatase inhibitor cocktail set (Calbiochem), and RiboLock RNase Inhibitor (Thermo Scientific)) and incubated on ice for 15 minutes with occasional vortexing. Lysates were passed several times through 21G and 27G needles, respectively and incubated on ice for 15 minutes with occasional vortexing. Samples were centrifuged for 20 minutes at 12,000 g and 4 °C and the supernatant was diluted 5X with IP buffer without NP-40. Samples were mixed with Protein-A-Sepharose beads 4B Fast Flow (Sigma) and 5 µg of either PKR antibody (Abcam #ab32052) or control IgG antibody (Abcam #37415) and incubated for 2 hours at 4 °C on a rotator. Samples were washed 5X with IP buffer with 0.1 % NP-40, centrifuged for 4 minutes at 4000 g and 4 °C. RNA was isolated from the pellet using RNAzol according to manufacturer's protocol and RNA amount was analyzed by real-time PCR.

Acknowledgements

We thank Zdenek Cimburek, Ondrej Horvath, Susan Wagner, and Ilona Urbarova for technical assistance. This research was supported by the Czech Science Foundation (grant GACR 204/09/0085), institutional support AV0Z50520514 and RVO 68378050, and the European Molecular Biology Organization Strategic Development Installation Grant (program 0001488, partial).

Figure legends

Figure 1 Expression of hairpin RNA affects the luciferase activity of transiently transfected reporter plasmids. (A) Schematic composition of pCAGEGFP-MosIR, pCAGEGFP, and pCAGEGFP-MosMos plasmids. (B) Reporter activity is inhibited by hairpin RNA in a concentration-dependent manner. HEK293 cells were transiently transfected with a constant amount of firefly luciferase (square), *Renilla* luciferase (triangle) reporter plasmids, and increasing amount of a tested plasmid. Luciferase activities were measured 48 hours post-transfection. (C) The inhibition of luciferase activity is released when hairpin RNA expression is diminished. HEK293 cells were transiently transfected with intact or SalI- (upstream of the promoter) or EagI- (downstream of the promoter) digested pCAGEGFP-MosIR plasmid. Approximate sites of restriction digest are shown on the scheme. pBluescript plasmid was added to maintain the amount of transfected DNA constant. Both luciferase activities are shown relative to cells transfected with 0 ng of the pCAGEGFP-MosIR plasmid. Data are shown as an average of at least 3 experiments made in triplicates. Error bars = SEM.

Figure 2 Deep sequencing analysis of total RNAs derived from pCAGEGFP and pCAGEGFP-MosIR plasmids. (A, B) Histograms show coverage of pCAGEGFP (upper panel) and pCAGEGFP-MosIR (lower panel) plasmids by (A) total reads 18-50 nt in length or (B) reads 20-24 nt in length with up to 5 adenosine-to-guanosine mismatches. Y-scale represents normalized read density (counts per million, CPM); y-scale maximum corresponds to (A) 1500 or (B) 200 CPM. Schematic representation of plasmids' features is shown below the histogram. (C, D) Length distribution of edited (gray) and non-edited (black) reads derived from MosIR (C) or EGFP (D) region of pCAGEGFP-MosIR plasmid. Note the different y-scale for each graph.

Figure 3 The inhibition of luciferase reporter activity is independent of a cell line and dsRNA sequence and dependent on dsRNA expression. (A) Reporter activity is inhibited in

a concentration-dependent manner by pCAGEGFP-MosIR plasmid compared to parental pCAGEGFP plasmid in HeLa cells. (B, C) The inhibition of reporter activity is independent of the hairpin RNA sequence and it is absent when the inverted repeat is placed downstream of non-active ZP3 promoter. (B) HEK293 ells were transiently transfected with 100 ng/well of each RL (triangle) and FL (square) reporter plasmids and increasing amount (0-250 ng/well) of either pCAGEGFP-Lin28IR (containing an active promoter) or pZP3EGFP-Lin28IR plasmid (containing an inactive promoter). Luciferase activity was analyzed 48 hours after transfection. pBluescript plasmid was added to maintain the amount of transfected DNA constant. Both luciferase activities are shown relative to cells transfected with 0 ng of the hairpin-expressing plasmid. (C) Similar to (B) except using Elavl2IR-expressing (pCAGEGFP-MosIR or pZP3EGFP-Elavl2IR) plasmids. Data are shown as an average of at least 3 experiments made in triplicates. Error bars = SEM.

Figure 4 A hairpin-expressing plasmid affects the luciferase activity of transiently transfected but not stably integrated reporter plasmids. HEK293 cells with stably integrated RL and FL reporters (A), or HEK293 cells with stably integrated RL reporter only (B) were transiently transfected with increasing amount of pCAGEGFP-MosIR plasmid and FL plasmid (if not stably integrated). pBluescript plasmid was added to maintain the amount of transfected DNA constant. Both FL (squares) and RL (triangles) luciferase activities were analyzed 48 hours post-transfection. Luciferase activities in cells transfected with 0 ng of the pCAGEGFP-MosIR plasmid were set to one. Data show an average of at least 3 experiments done in triplicates. Error bars = SEM.

Figure 5 dsRNA inhibits translation of transcripts from transiently-transfected plasmids. (A) Transiently transfected RL and FL reporters are not inhibited at RNA level. HEK293 cells were transiently transfected with FL and RL reporters (100 ng/well) and increasing doses of pCAGEGFP-MosIR plasmid (0-250 ng/well). Amount of mRNA was analyzed by real-time PCR. Expression was normalized to *Hprt1* housekeeping gene and expression levels in cells transfected with 50 ng of MosIR plasmid were set to 1. Error bars

= SEM. (B) dsRNA-dependent inhibition of translation affects more transiently transfected plasmids than endogenous genes. HEK293 cells were transfected with RL, FL, and either pCAGEGFP or pCAGEGFP-MosIR plasmid. Distribution of mRNA in fractions collected during polysome profiling was analyzed by real-time PCR. For each sample, a fraction representing monosomes (80S) and early (poly1) and late (poly2) polysomes (depicted on the scheme) was included in the quantification. Expression levels in polysome fractions of pCAGEGFP- (black bars) and pCAGEGFP-MosIR- (white bars) transfected cells are normalized to 80S fraction. Panels show expression profiles for endogenous genes (HPRT1 and B2M), plasmid-expressed transcripts (FL, RL) and either pCAGEGFP or pCAGEGFP-MosIR plasmids (pCAG). B2M, β_2 microglobulin; HPRT1, hypoxantine phosphoryltransferase; FL, firefly luciferase; RL, *Renilla* luciferase.

Figure 6 The inhibition of luciferase reporter activity is partly dependent on PKR protein level. (A) Parental HeLa cells (ctrl) or HeLa cells with stably down-regulated PKR expression (KD) were co-transfected with 100 ng/well of each RL (light colors) and FL (dark colors) reporter plasmids and 0 or 50 ng/well of pCAGEGFP-MosIR plasmid. Parental pCAGEGFP plasmid was used to maintain the amount of transfected DNA constant. Data are shown as an average of 2 experiments performed in quadruplicates; data from 2 independent KD clones (B4 and E2, Fig. S4) were pooled together. Error bars = SEM. (B) pCAGEGFP-MosIR transcript associates with PKR. HEK293 cells were transfected with pCAGEGFP or pCAGEGFP-MosIR plasmid. Cell lysates (48 hours post-transfection) were used for immunoprecipitation with anti-PKR antibody or control IgG antibody. Immunoprecipitated RNA was isolated, reverse-transcribed and used for real-time PCR analysis. Data are expressed as a percentage of input. (C) pCAGEGFP-MosIR expression activates PKR. Western blotting analysis of HEK293 cells transfected with increasing amount of pCAGEGFP and pCAGEGFP-MosIR plasmids (50-150-250 ng per well). pBluescript was added to maintain the amount of transfected DNA constant. pBS, pBluescript only. UN, untransfected cells; MosIR, pCAGEGFP-MosIR. (D) pCAGEGFP-MosIR expression affects overall translation rates. Polysome profile overlap from HEK293 cells

transfected with pCAGEGFP (blue) or pCAGEGFP-MosIR (red) plasmid. (E) pCAGEGFP-MosIR expression affects the distribution of PKR and its phosphorylated form in polyribosome analysis. Polysome profile (upper part) and western blotting analysis (lower part) of respective fractions. The alignment of fractions on polysome profile and lanes on western blotting is exemplified by dashed lines for the fraction 11. HEK293 cells transfected with either pCAGEGFP or pCAGEGFP-MosIR plasmids were subjected to polysome profiling. Proteins were isolated from each fraction by ethanol precipitation and the same volume aliquots from each fraction were analyzed by polyacrylamide gel electrophoresis and western blotting. The last lane in each sample represents 1 % of input. γ, phosphorylated form of protein; RPS14, ribosomal protein S14.

References

1. Gantier MP, Williams BR (2007) The response of mammalian cells to double-stranded RNA. Cytokine Growth Factor Rev 18: 363-371.
2. Carthew RW, Sontheimer EJ (2009) Origins and Mechanisms of miRNAs and siRNAs. Cell 136: 642-655.
3. Sadler AJ, Williams BR (2008) Interferon-inducible antiviral effectors. Nat Rev Immunol 8: 559-568.
4. Sudhakar A, Ramachandran A, Ghosh S, Hasnain SE, Kaufman RJ, et al. (2000) Phosphorylation of serine 51 in initiation factor 2 alpha (eIF2 alpha) promotes complex formation between eIF2 alpha(P) and eIF2B and causes inhibition in the guanine nucleotide exchange activity of eIF2B. Biochemistry 39: 12929-12938.
5. Kaufman RJ (1997) DNA transfection to study translational control in mammalian cells. Methods 11: 361-370.
6. Ben-Asouli Y, Banai Y, Pel-Or Y, Shir A, Kaempfer R (2002) Human interferon-gamma mRNA autoregulates its translation through a pseudoknot that activates the interferon-inducible protein kinase PKR. Cell 108: 221-232.
7. Kaufman RJ, Davies MV, Pathak VK, Hershey JW (1989) The phosphorylation state of eucaryotic initiation factor 2 alters translational efficiency of specific mRNAs. Mol Cell Biol 9: 946-958.
8. Valasek L, Szamecz B, Hinnebusch AG, Nielsen KH (2007) In vivo stabilization of preinitiation complexes by formaldehyde cross-linking. Methods Enzymol 429: 163-183.
9. Nishikura K (2010) Functions and regulation of RNA editing by ADAR deaminases. Annu Rev Biochem 79: 321-349.
10. Nejepinska J, Malik R, Filkowski J, Flemr M, Filipowicz W, et al. (2012) dsRNA expression in the mouse elicits RNAi in oocytes and low adenosine deamination in somatic cells. Nucleic Acids Res 40: 399-413.
11. Svoboda P, Stein P, Schultz RM (2001) RNAi in mouse oocytes and preimplantation embryos: effectiveness of hairpin dsRNA. Biochem Biophys Res Commun 287: 1099-1104.
12. Colledge WH, Carlton MB, Udy GB, Evans MJ (1994) Disruption of c-mos causes parthenogenetic development of unfertilized mouse eggs. Nature 370: 65-68.
13. Hashimoto N, Watanabe N, Furuta Y, Tamemoto H, Sagata N, et al. (1994) Parthenogenetic activation of oocytes in c-mos-deficient mice. Nature 370: 68-71.
14. Stein P, Svoboda P, Schultz RM (2003) Transgenic RNAi in mouse oocytes: a simple and fast approach to study gene function. Dev Biol 256: 187-193.
15. Stein P, Zeng F, Pan H, Schultz RM (2005) Absence of non-specific effects of RNA interference triggered by long double-stranded RNA in mouse oocytes. Dev Biol 286: 464-471.
16. Svoboda P, Stein P, Filipowicz W, Schultz RM (2004) Lack of homologous sequence-specific DNA methylation in response to stable dsRNA expression in mouse oocytes. Nucleic Acids Res 32: 3601-3606.
17. Nejepinska J, Malik R, Moravec M, Svoboda P (2012) Deep sequencing reveals complex spurious transcription from transiently transfected plasmids. PLoS One 7: e43283.

18. Terenzi F, deVeer MJ, Ying H, Restifo NP, Williams BR, et al. (1999) The antiviral enzymes PKR and RNase L suppress gene expression from viral and non-viral based vectors. Nucleic Acids Res 27: 4369-4375.
19. Kalvakolanu DV, Bandyopadhyay SK, Tiwari RK, Sen GC (1991) Enhancement of expression of exogenous genes by 2-aminopurine. Regulation at the post-transcriptional level. J Biol Chem 266: 873-879.
20. Motorin Y, Lyko F, Helm M (2010) 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. Nucleic Acids Res 38: 1415-1430.
21. Squires JE, Patel HR, Nousch M, Sibbritt T, Humphreys DT, et al. (2012) Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res 40: 5023-5033.
22. Wang Y, Samuel CE (2009) Adenosine deaminase ADAR1 increases gene expression at the translational level by decreasing protein kinase PKR-dependent eIF-2alpha phosphorylation. J Mol Biol 393: 777-787.
23. Rubin CM, Kimura RH, Schmid CW (2002) Selective stimulation of translational expression by Alu RNA. Nucleic Acids Res 30: 3253-3261.
24. Kaufman RJ, Murtha P (1987) Translational control mediated by eucaryotic initiation factor-2 is restricted to specific mRNAs in transfected cells. Mol Cell Biol 7: 1568-1571.
25. Novoa I, Carrasco L (1999) Cleavage of eukaryotic translation initiation factor 4G by exogenously added hybrid proteins containing poliovirus 2Apro in HeLa cells: effects on gene expression. Mol Cell Biol 19: 2445-2454.
26. Kaname T, Huxley C (2001) Simple and efficient vectors for retrofitting BACs and PACs with mammalian neoR and EGFP marker genes. Gene 266: 147-153.
27. van de Wetering M, Oving I, Muncan V, Pon Fong MT, Brantjes H, et al. (2003) Specific inhibition of gene expression using a stably integrated, inducible small-interfering-RNA vector. EMBO Rep 4: 609-615.

Figure 1

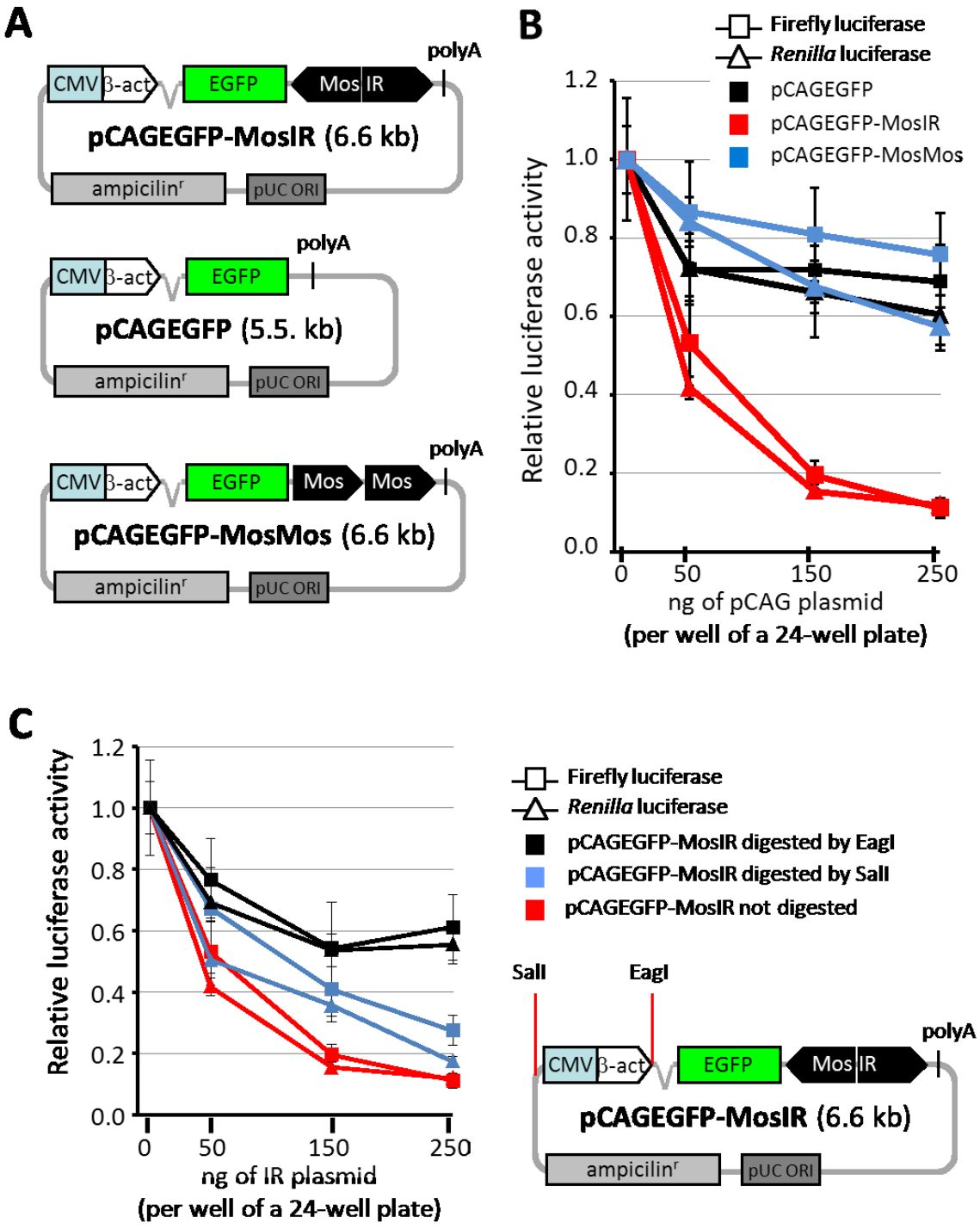

A
polyA
CMV
β-act
EGFP
Mos
IR
pCAGEGFP-MosIR (6.6 kb)
ampicilin^r
pUC ORI
polyA
CMV
β-act
EGFP
pCAGEGFP (5.5. kb)
ampicilin^r
pUC ORI
polyA
CMV
β-act
EGFP
Mos
Mos
pCAGEGFP-MosMos (6.6 kb)
ampicilin^r
pUC ORI
B
Firefly luciferase
Renilla luciferase
pCAGEGFP
pCAGEGFP-MosIR
pCAGEGFP-MosMos
Relative luciferase activity
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
50
150
250
ng of pCAG plasmid
(per well of a 24-well plate)
C
Relative luciferase activity
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
50
150
250
ng of IR plasmid
(per well of a 24-well plate)
Firefly luciferase
Renilla luciferase
pCAGEGFP-MosIR digested by EagI
pCAGEGFP-MosIR digested by SalI
pCAGEGFP-MosIR not digested
SalI
EagI
polyA
CMV
β-act
EGFP
Mos
IR
pCAGEGFP-MosIR (6.6 kb)
ampicilin^r
pUC ORI

Figure 2

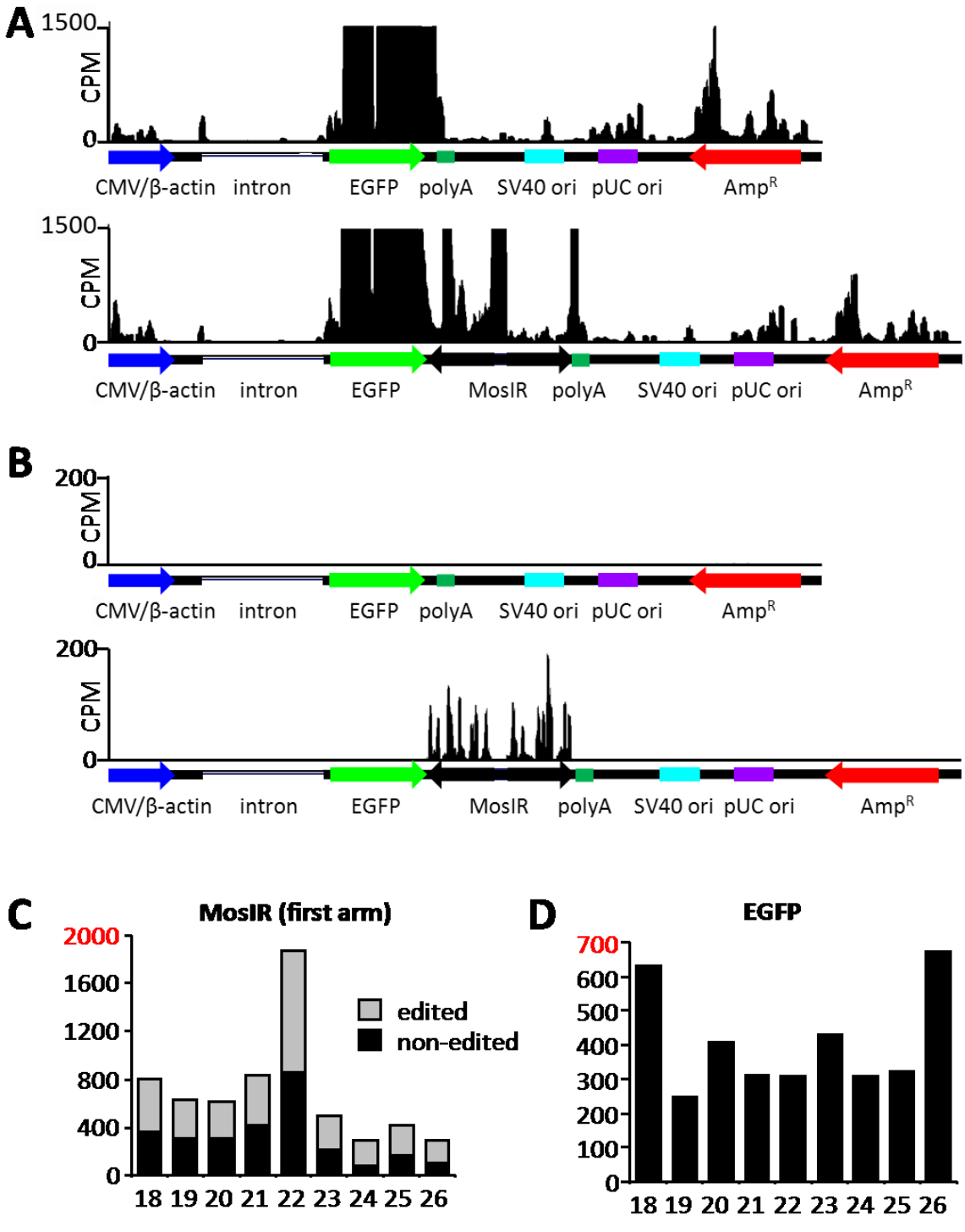

A
1500
CPM
0
CMV/β-actin
intron
EGFP
polyA
SV40 ori
pUC ori
Amp^R
1500
CPM
0
CMV/β-actin
intron
EGFP
MosIR
polyA
SV40 ori
pUC ori
Amp^R
B
200
CPM
0
CMV/β-actin
intron
EGFP
polyA
SV40 ori
pUC ori
Amp^R
200
CPM
0
CMV/β-actin
intron
EGFP
MosIR
polyA
SV40 ori
pUC ori
Amp^R
C
MosIR (first arm)
2000
1600
1200
800
400
0
edited
non-edited
18 19 20 21 22 23 24 25 26
D
EGFP
700
600
500
400
300
200
100
0
18 19 20 21 22 23 24 25 26

Figure 3

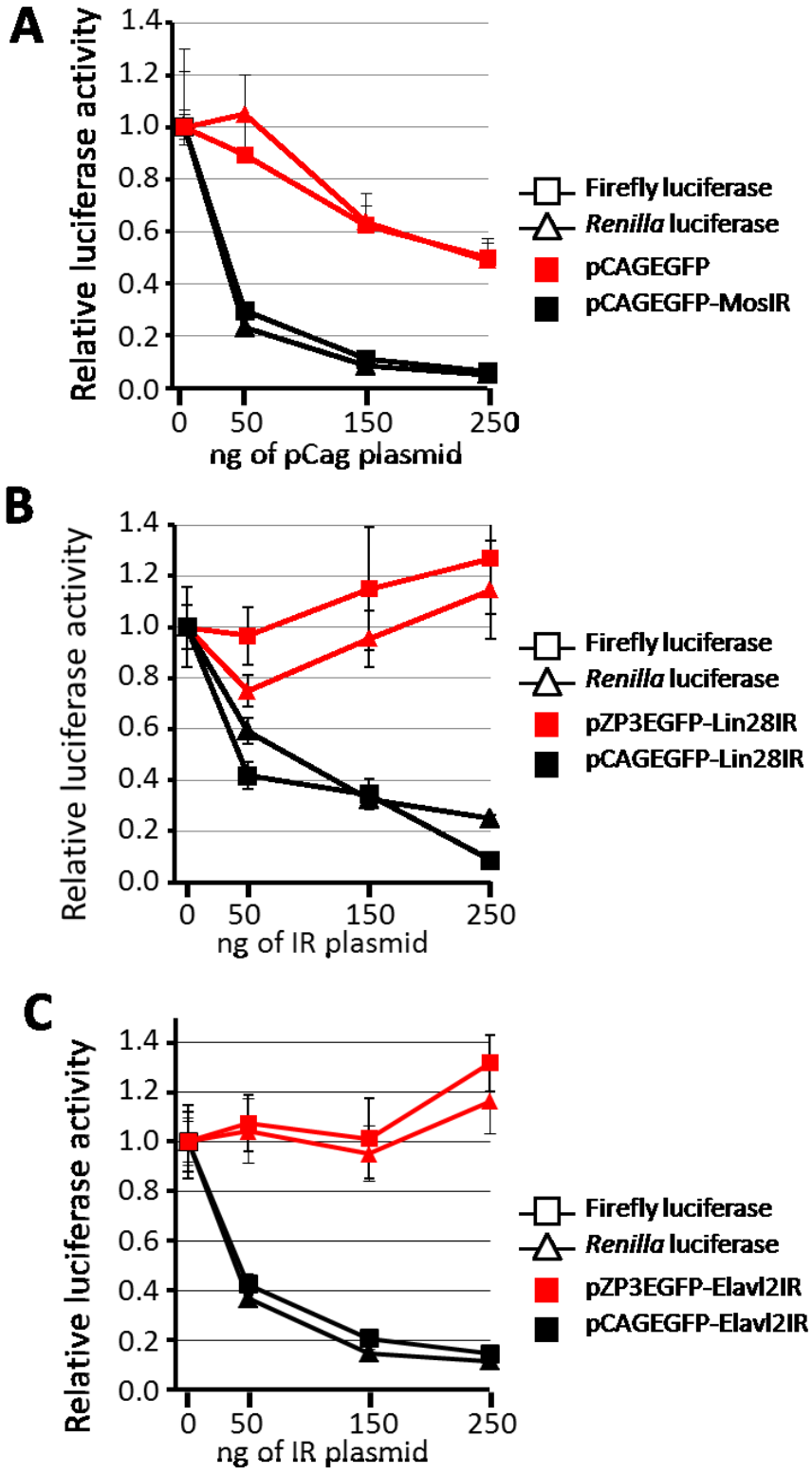

A
Relative luciferase activity
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
50
150
250
ng of pCag plasmid
Firefly luciferase
Renilla luciferase
pCAGEGFP
pCAGEGFP-MosIR
B
Relative luciferase activity
ng of IR plasmid
Firefly luciferase
Renilla luciferase
pZP3EGFP-Lin28IR
pCAGEGFP-Lin28IR
C
Relative luciferase activity
ng of IR plasmid
Firefly luciferase
Renilla luciferase
pZP3EGFP-Elavl2IR
pCAGEGFP-Elavl2IR

Figure 4

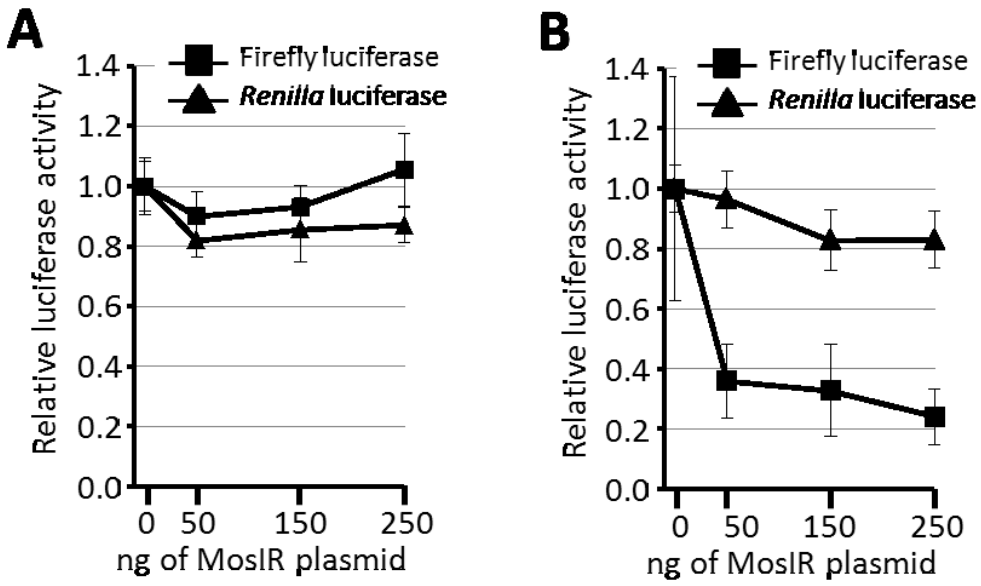

A
B
Firefly luciferase
Renilla luciferase
Relative luciferase activity
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
50
150
250
ng of MosIR plasmid

Figure 5

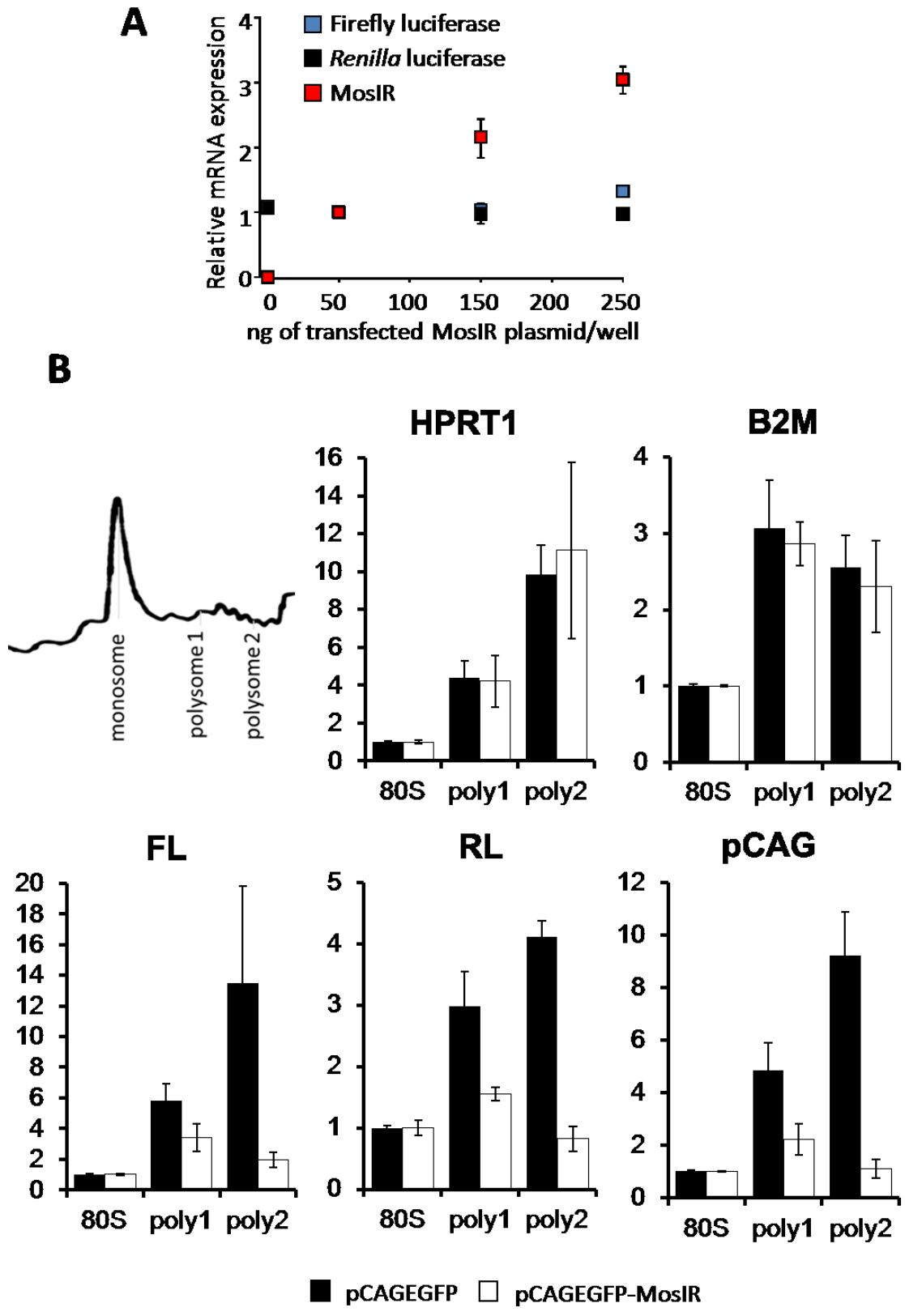
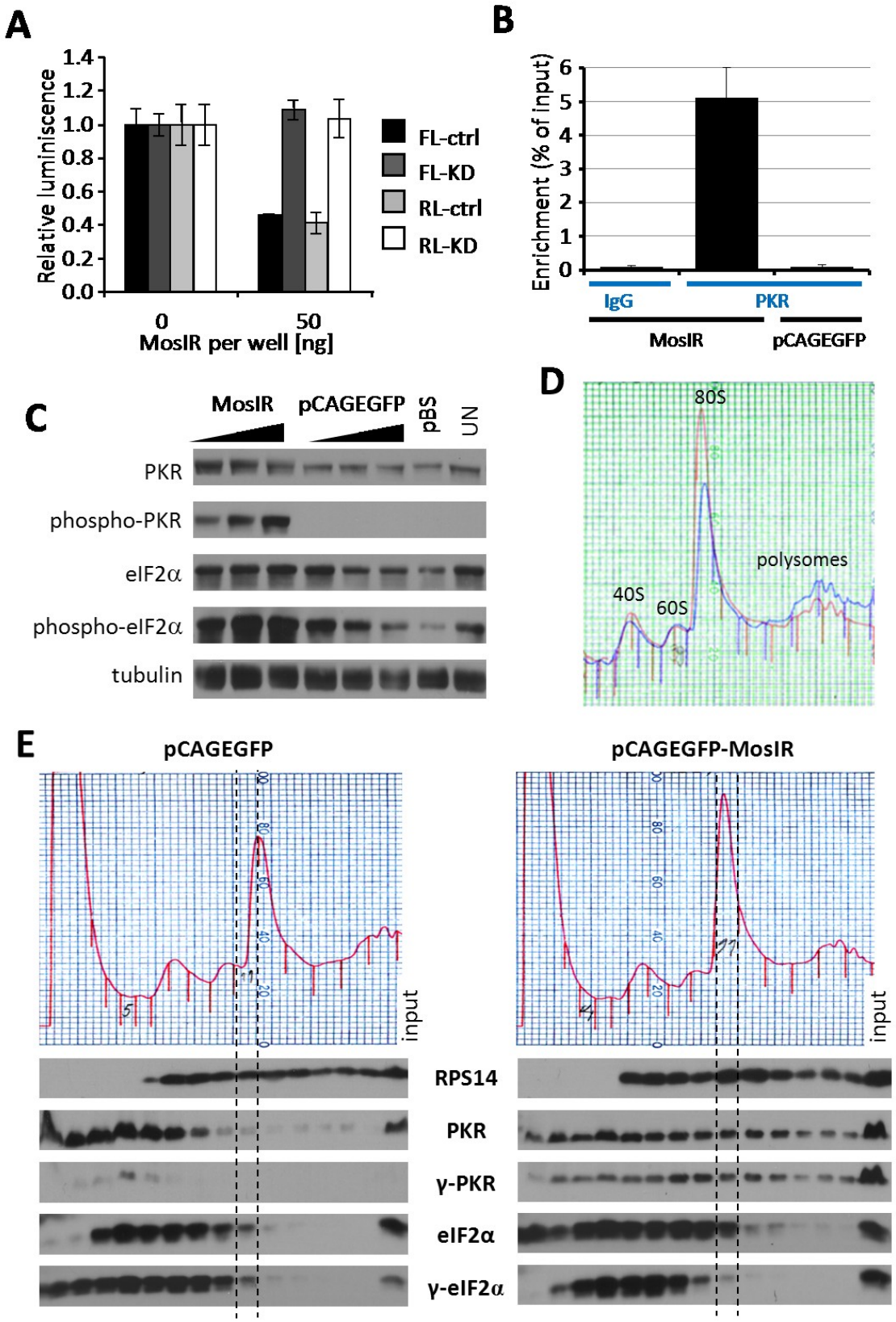

A
Firefly luciferase
Renilla luciferase
MosIR
Relative mRNA expression
ng of transfected MosIR plasmid/well
B
monosome
polysome 1
polysome 2
HPRT1
B2M
FL
RL
pCAG
80S
poly1
poly2
pCAGEGFP
pCAGEGFP-MosIR

Figure 6


A
Relative luminiscence
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
50
MosIR per well [ng]
FL-ctrl
FL-KD
RL-ctrl
RL-KD
B
Enrichment (% of input)
6
5
4
3
2
1
0
IgG
PKR
MosIR
pCAGEGFP
C
MosIR
pCAGEGFP
pBS
UN
PKR
phospho-PKR
eIF2α
phospho-eIF2α
tubulin
D
80S
polysomes
40S
60S
E
pCAGEGFP
pCAGEGFP-MosIR
input
input
RPS14
PKR
γ-PKR
eIF2α
γ-eIF2α

Expressed dsRNA provokes selective sequence-independent PKR response in transient co-transfection experiments

Authors:

Jana Nejepinska, Radek Malik, and Petr Svoboda[+]

Supplementary methods:

Nanofectin and calcium phosphate transfection

For transfection, cells were plated on a 24-well plate, grown to 70 % density and transfected using Nanofectin (PAA) reagent according to the manufacturer's protocol. Calcium phosphate transfection was performed according to a standard protocol.

Flow cytometry

HEK293 cells plated in 24-well plates were co-transfected with 150 ng/well of RFP reporter plasmid (pCI-RFP) and 350 ng/well of either pCAGEGFP-MosIR or pCAGEGFP plasmid. Cells were collected 24 and 36 hours post-transfection and analyzed by flow cytometry using LSRII cytometer (BD Bioscience). Data analysis was performed by FlowJo software (Treestar, Inc.).

Suplementary figures S1-S4

Figure S1

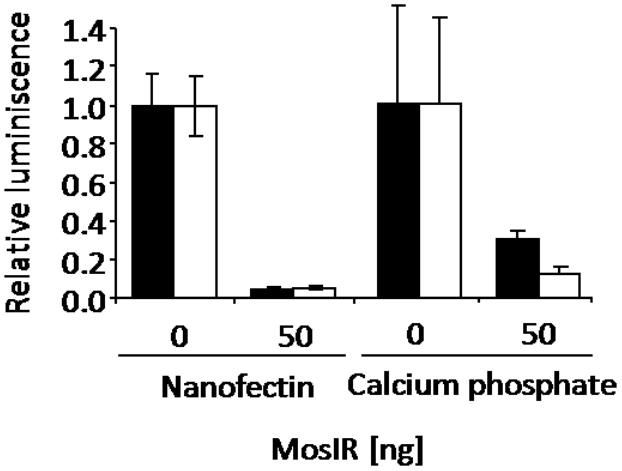


Figure S1

Inhibition of luciferase reporter activity is independent of a transfection procedure. HEK293 cells were transiently transfected with a constant amount of firefly luciferase (black bars), *Renilla* luciferase (white bars) reporter plasmids, and 50 ng/well of pCAGEGFP-MosIR plasmid using Nanofectin transfection reagent or calcium phosphate method. Luciferase activities were measured 48 hours post-transfection and normalized to cells transfected with 0 ng pCAGEGFP-MosIR (pBluescript plasmid was used to maintain the constant amount of transfected DNA).

Figure S2

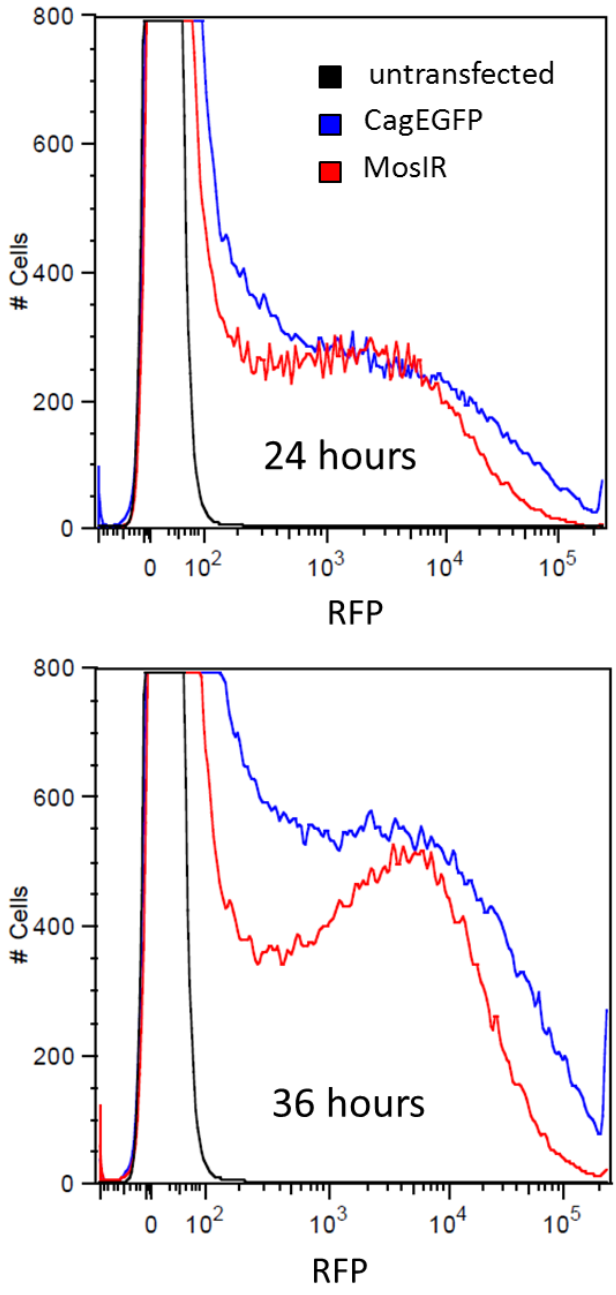


Figure S2

pCAGEGFP-MosIR plasmid inhibits expression of the RFP reporter. Untransfected cells (black) or cells cotransfected with pCI-RFP reporter plasmid and either pCAGEGFP (blue) or pCAGEGFP-MosIR (red) plasmid were analyzed by flow cytometry. Histogram of RFP fluorescence 24 and 36 hours post-transfection is shown. Data show results of a typical experiment.

Figure S3

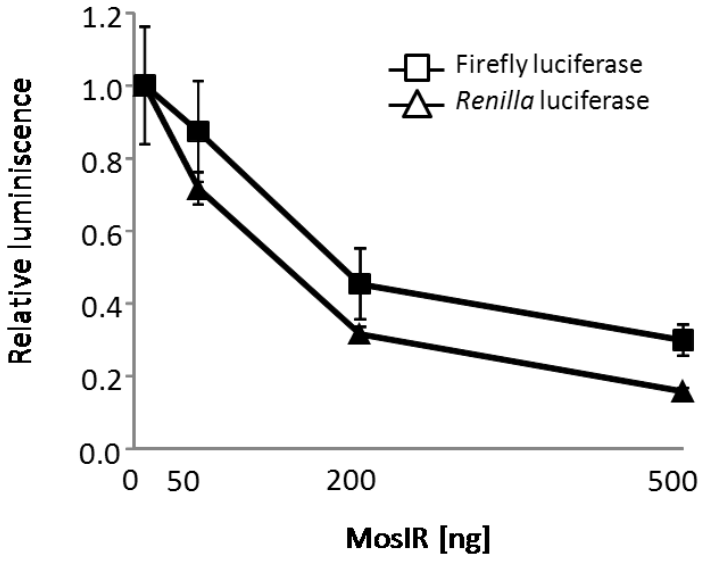


Figure S3

Reporter activity is inhibited in a concentration-dependent manner by pCAGEGFP-MosIR plasmid compared to parental pCAGEGFP plasmid in P19 cell line. P19 cells were transiently transfected with a constant amount of firefly luciferase (square), *Renilla* luciferase (triangle) reporter plasmids, and increasing amount of a pCAGEGFP-MosIR plasmid. Luciferase activities were measured 48 hours post-transfection. pBluescript plasmid was used to maintain the constant amount of transfected DNA. Data are normalized to cells transfected with 0 ng pCAGEGFP-MosIR. The average of two independent experiments done in triplicates.

Figure S4

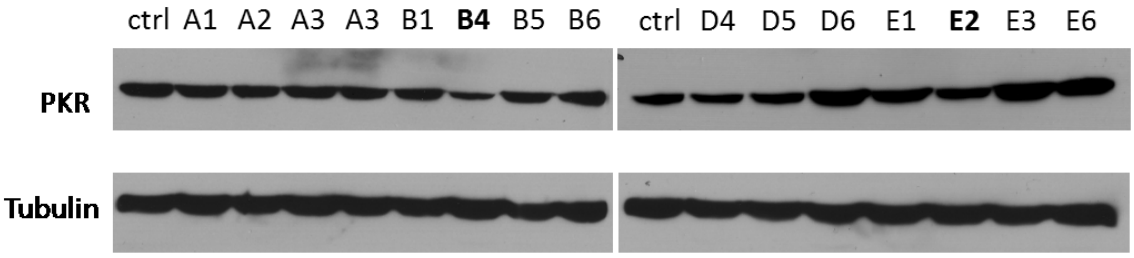


Figure S4
Western blotting showing PKR expression in selected HeLa clones stably expressing shRNA against PKR or control empty pTer plasmid (ctrl). Tubulin was used as a loading control. Clones #B4 and #E2 were used for experiments shown in Fig. 6A.

Supplement 5

Nejepinska J, Flemr, Svoboda P.

The canonical RNA interference pathway in animals

Regulatory RNAs

2012

111-150

Chapter 5
The Canonical RNA Interference Pathway in Animals

Jana Nejepinska, Matyas Flemr, and Petr Svoboda

Abstract The canonical RNA interference (RNAi) pathway is defined as a sequence-specific mRNA degradation mediated by short RNA molecules which are generated from long double-stranded RNA. Since its discovery in 1998, RNAi has become a popular tool for experimental silencing of gene expression. On the other hand, its natural role received less attention. Recent studies in animal systems, particularly the use of the next generation sequencing and analysis of animals defective in some aspect of small RNA biogenesis, revealed novel functions of RNAi and cross talks between RNAi and other pathways employing small RNAs. This chapter provides a comprehensive view of the natural canonical RNAi pathway in animals including its molecular mechanism and different biological roles.

Keywords Argonaute • Dicer • dsRNA • RNAi • siRNA

5.1 Introduction

RNA silencing is a common term for repression guided by small RNA molecules (20–30 nucleotides (nt) long). Suppressive effects of RNA silencing include mRNA degradation, translational repression, formation of repressive chromatin, and DNA deletion (reviewed in Czech and Hannon 2011; Ketting 2011). Some forms of RNA silencing exist in almost all eukaryotes. RNA interference (RNAi) is one of the best characterized RNA silencing pathways. The term RNAi has been coined for sequence-specific mRNA degradation mediated by small RNAs produced from a long double-stranded RNA (dsRNA). Although the term RNAi is sometimes used

P. Svoboda (✉)
Institute of Molecular Genetics AS CR, Prague, Czech Republic
e-mail: svobodap@img.cas.cz

B. Mallick (eds.), *Regulatory RNAs*,
DOI 10.1007/978-3-642-22517-8_5,

for a broad range of RNA silencing pathways (Ketting 2011), we will use it in its original connotation here. Since our aim is to provide a detailed overview of RNAi in animals, we will not discuss data from plants, fungi, and protists. However, it must be acknowledged that research in these models provided a fundamental contribution to understanding RNAi and related pathways in animals. Likewise, other animal RNA silencing pathways, such as the microRNA (miRNA, reviewed in Kim 2005) pathway and piwi-associated RNA (piRNA, reviewed in Aravin et al. 2007) pathway, will be discussed only in the context of the canonical RNAi pathway.

RNAi was first described in the nematode *Caenorhabditis elegans* (*C. elegans*) by Andrew Fire and Craig Mello, who observed that injected long dsRNA induced sequence-specific mRNA degradation in the whole animal (Fire et al. 1998). Initial studies of effects of dsRNA in animals showed RNAi effects in a wide range of animal taxa, including mammals (Svoboda et al. 2000; Wianny and Zernicka-Goetz 2000; Lohmann et al. 1999; Sanchez Alvarado and Newmark 1999; Kennerdell and Carthew 1998). The molecular mechanism of RNAi was deciphered using a combination of genetic and biochemical approaches. An important step was the establishment of the biochemical model in *Drosophila* embryo lysates, which revealed that long dsRNA is processed by an RNase III Dicer into short interfering RNAs (siRNA) that guide specific cleavage of cognate mRNAs in positions corresponding to the center of the siRNA:miRNA duplex (Bernstein et al. 2001; Tuschl et al. 1999; Zamore et al. 2000). Genetic studies in *C. elegans* discovered numerous components of RNAi, including the Argonaute protein family and RNA-dependent RNA polymerases (RdRPs) (Ketting et al. 1999; Smardon et al. 2000; Vastenhouw et al. 2003). The final gap in understanding the RNAi mechanism was bridged by structural analyses of Argonaute 2, which revealed that this protein carries the endonucleolytic activity, which cleaves cognate mRNAs (Liu et al. 2004; Meister et al. 2004; Song et al. 2004). The following text summarizes the current knowledge of the RNAi mechanism and its function in the three most studied animal model systems: *C. elegans*, *Drosophila*, and mammals.

5.2 The Mechanism of RNAi

In this section, we summarize current understanding of ribonucleoprotein complexes involved in RNAi. The RNAi pathway (Fig. 5.1) can be divided into three steps (1) cleavage of long dsRNA by Dicer into siRNAs, (2) loading of small RNAs on the effector complex known as the RNA-induced silencing complex (RISC), and (3) recognition and cleavage of cognate RNAs by the RISC. In addition to the core pathway, two extensions of the pathway, which are restricted to some animal species, should be mentioned (1) an amplification step, in which RdRPs generate secondary siRNAs (transitive RNAi) and (2) the systemic RNAi, where an RNAi response can spread across cellular boundaries.

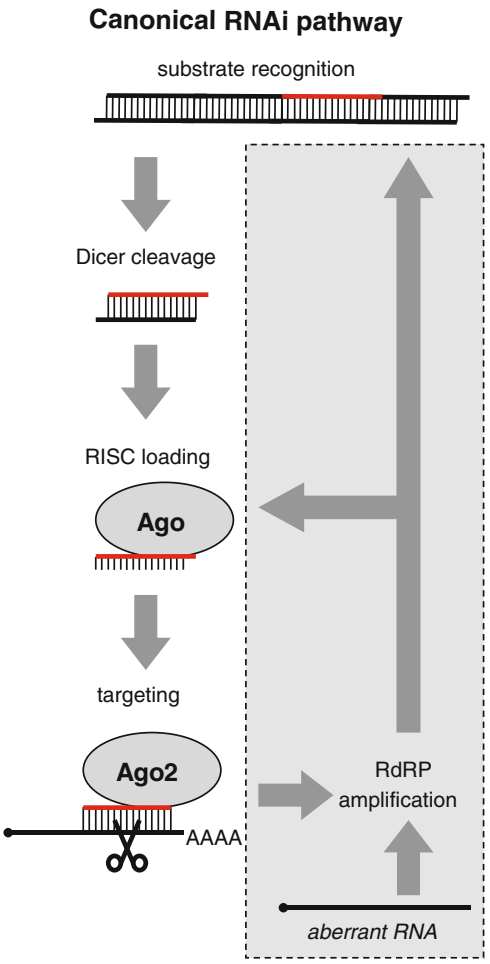


Fig. 5.1 A schematic view of the RNAi pathway. The "core" RNAi pathway involves a processing of long dsRNA into siRNAs by Dicer and loading of siRNAs on the RISC complex, which cleaves cognate transcripts. The *gray* area represents an amplification loop where RdRPs generate secondary siRNAs or dsRNA templates for Dicer processing

5.2.1 dsRNA Recognition and Cleavage

RNAi is triggered by dsRNA, a unique helical structure formed by two antiparallel RNA strands. dsRNA is frequently formed by repetitive sequences or during viral infections. dsRNA can form by base pairing two single-stranded RNAs, as an intramolecular duplex (a hairpin, reviewed in Svoboda and Cara 2006), or the second strand can be synthesized on a single-stranded RNA template by an RdRP (reviewed in Ng et al. 2008).

Recognition of dsRNA is mediated by the dsRNA-binding domain (dsRBD), which is found in a diverse group of proteins (dsRNA-binding proteins, dsRBPs) involved in various responses to dsRNA. Some proteins contain a single dsRBD while others carry multiple copies of the domain. Specific structural features of dsRNA (such as dsRNA termini or mismatches) and individual dsRBDs present in specific

proteins contribute to the routing of dsRNA molecules into specific pathways (reviewed in Doyle and Jantsch 2002; Gantier and Williams 2007; Tian et al. 2004).

The key protein for dsRNA processing into siRNAs is Dicer, which typically carries a dsRBD at the C-terminus and can directly recognize and cleave dsRNA in vitro (Ketting et al. 2001; Zhang et al. 2002; Provost et al. 2002). However, the production of small RNAs in the RNAi pathway is also assisted by proteins carrying tandemly arrayed dsRBDs (for more details, see Sect. 2.1.2).

5.2.1.1 Dicer Structure and Function

Dicer generates small RNAs in RNAi and many other RNA silencing pathways (reviewed for example in Jaskiewicz and Filipowicz (2008)). Dicer is a large (~200 kDa), multi-domain RNase III endonuclease, which cleaves both strands of a duplex dsRNA (Fig. 5.2). This cleavage produces small (21–27 nt long) RNA duplexes with two nucleotide 3′ overhangs and 5′-monophosphate and 3′-hydroxyl groups at RNA termini. Dicer is a member of a class of RNase III enzymes, which carry two RNase III domains. Several other domains are typically found in Dicer-like proteins in multicellular eukaryotes. These include N-terminal DEAD-like (DExD) and helicase superfamily C domains, piwi/argonaute/zwille (PAZ) domain, domain of unknown function DUF283, and C-terminal dsRBD.

The ribonuclease activity of Dicer requires magnesium ions. Dicer can efficiently cleave dsRNA longer than 30 base pairs (bp), yielding siRNA of approximately 20 bp (Provost et al. 2002; Zhang et al. 2002). Dicer preferentially cleaves dsRNA at the termini but it can also cleave internally with low efficiency (Zhang et al. 2002). Dicer functions as a molecular ruler, measuring the length of the substrate from the PAZ domain to RNase III domains where it is cleaved. The PAZ

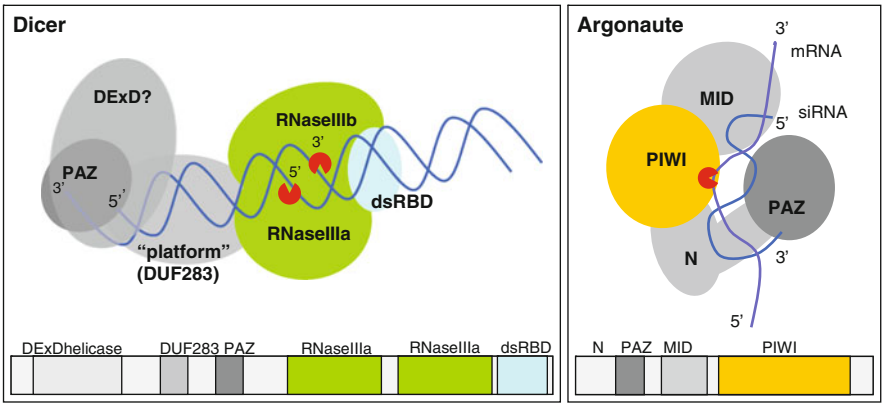


Fig. 5.2 Key components of RNA silencing. A schematic domain organization of Dicer and Argonaute. The schematic structural organization based on crystalography and single-particle electron microscopy structural analysis of Dicer (*left*) and Argonaute (*right*) is shown (Du et al. 2008; MacRae et al. 2007; Macrae et al. 2006; Wang et al. 2009a)

domain binds the end of dsRNA with high affinity to 3′ protruding overhangs (Ma et al. 2004; Lingel et al. 2003; Song et al. 2003; Yan et al. 2003). RNase IIIa and IIIb domains form a single processing center containing two catalytic "half-sites," each cleaving one strand of the duplex and producing short dsRNA with two nt 3′ overhang. The RNase IIIa domain processes the protruding 3′-OH-bearing strand, and the RNase IIIb cuts the opposite 5′-phosphate-containing strand (Zhang et al. 2004). This model was validated by the crystal structure of the full length Dicer from *Giardia intestinalis* (Macrae et al. 2006; MacRae et al. 2007), which showed that the RNase III domains form a catalytic center connected with the PAZ domain by a long α-helix ("connector" helix), which is implicated in determining the product length. The connector helix is supported by a platform-like structure containing the DUF283 domain, which has a dsRBD-like fold (Dlakic 2006) and perhaps mediates protein–protein interaction (Qin et al. 2010).

Most Dicer enzymes contain a DExD helicase domain suggesting that it might be involved in ATP-dependent binding and remodeling of nucleic acids. Although *Drosophila* DCR-2 seems to require ATP for processive cleavage of dsRNA, mammalian Dicer proteins do not require ATP (Provost et al. 2002; Zhang et al. 2002; Nykanen et al. 2001). A kinetic analysis of Dicer mutants showed that the DExD domain could have an autoinhibitory role because a deletion or a mutation of this domain increased catalytic efficiency of Dicer in both single- and multiple-turnover assays (Ma et al. 2008). A modest stimulation of Dicer catalysis was observed in the presence of TRBP, which interacts with the DExD domain (Ma et al. 2008). These and other data (Soifer et al. 2008) suggest that the DExD domain may be a part of the system selecting dsRNA substrates that efficiently enter the RNAi pathway. Further studies will be necessary to establish if and how the DExD domain participates in dsRNA processing and loading of small RNAs on the RISC.

Some organisms, like mammals, *C.elegans* or *Trypanosoma*, utilize a single Dicer protein to produce both siRNAs and miRNAs. In contrast, *Drosophila* utilizes two Dicer paralogs, DCR-1 to produce miRNAs and DCR-2 to produce siRNAs (Lee et al. 2004). DCR-2 contains an N-terminal helicase motif and hydrolyzes ATP. ATP hydrolysis is needed for the processing of dsRNA but not miRNA precursors, and a model has been proposed that the helicase domain is important for DCR-2 processivity (Cenik et al. 2011). Some species utilize even more Dicer paralogs with distinct functions and different cleavage product lengths (e.g., four Dicer paralogs in *Arabidopsis thaliana* (reviewed in Meins et al. 2005)).

Animal Dicer usually localizes to the cytoplasm but it can be also found in the nucleus (Billy et al. 2001; Sinkkonen et al. 2010). However, the significance of the nuclear localization of mammalian Dicer remains unknown at the moment, and it is commonly accepted that siRNA production in animals takes place in the cytoplasm.

5.2.1.2 Proteins with Multiple dsRBDs

RDE-4, the first RNAi protein with multiple dsRBDs (RNAi dsRBP), was identified by a systematic screen for *C. elegans R*NAi-*de*ficient mutants (Tabara et al. 1999).

The *rde-4* mutant was completely deficient in RNAi but failed to show any discernible phenotype, including the absence of transposon activation, which was observed in some other *rde* mutants (Tabara et al. 1999). *Rde-4* encodes a 385-amino acid protein carrying two N-terminal dsRBDs and a third degenerate dsRBD at the C-terminus. A similar organization is found in other proteins implicated in RNA silencing, including *Drosophila* R2D2 and Loquacious and mammalian TRBP and PACT. However, while the phylogenetic analysis suggests that these proteins probably evolved from a common ancestral protein, they significantly diverged (Murphy et al. 2008) and play distinct roles in dsRNA processing and RISC loading.

Biochemical characterization of recombinant RDE-4 showed that it preferentially binds long dsRNA, and its dimerization is necessary for the cleavage of dsRNA into siRNAs (Parker et al. 2006). According to the model supported by mutants and biochemical analyses, RDE-4 dimers bind cooperatively to dsRNA and, together with Dicer, RDE-1 (an Argonaute protein), and DRH-1/2 (*D*icer-*r*elated *h*elicase from the DExH helicase superfamily), forms a complex initiating the RNAi (Tabara et al. 2002; Parker et al. 2006, 2008). The presence of an Argonaute protein in the complex suggests that dsRNA recognition, processing into siRNA, and loading of the Argonaute-containing effector complex could be integrated in one complex. RDE-4 is involved in siRNA production from dsRNA but is not essential for later steps of RNAi because RDE-4 immunoprecipitates with trigger dsRNA but not siRNA (Tabara et al. 2002) and *rde-4* loss can be rescued with injection of synthetic siRNA (Parrish and Fire 2001).

It should be mentioned that this model of RDE-4 function is based on the analysis of the RNAi responding to exogenous dsRNA (exo-RNAi), which is one of the several RNAi-related pathways in *C. elegans*. RNA silencing in *C. elegans* evolved into an extremely complex system, which utilizes 27 Argonaute proteins and numerous classes of small RNAs (Ketting 2011). The endogenous RNAi pathway (endo-RNAi), which targets endogenous genes (Ambros et al. 2003), employs a distinct mechanism of siRNA production involving DCR-1 and RDE-4 but not RDE-1 and DRH-1/2 (Gent et al. 2010; Lee et al. 2006).

Drosophila employs two RNAi-RBPs: R2D2 and Loquacious. R2D2 was co-purified with Dicer-2 during purifying siRNA-generating activity from *Drosophila* S2 cell lysates (Liu et al. 2003). R2D2 bears 33% identity to RDE-4 but its role is different. DCR-2/R2D2 association does not affect DCR-2 processing. Instead, DCR-2/R2D2 bind siRNAs to promote AGO2-containing RISC loading (Liu et al. 2006) (for more details, see the Sect. 2.2.2). The second RNAi dsRBP in *Drosophila* is Loquacious, which was found to associate with DCR-1, suggesting that the miRNA pathway in *Drosophila* employs a distinct dsRBP in substrate routing (Saito et al. 2005; Forstemann et al. 2005). Surprisingly, deep sequencing of small RNAs in *Drosophila* revealed that siRNAs in the endogenous RNAi pathway are produced by DCR-2 but depend preferentially on Loquacious and not on R2D2, the canonical DCR-2 partner (Czech et al. 2008). It was further shown that endo-siRNAs in *Drosophila* predominantly bind AGO2 and can arise from perfect duplexes formed from overlapping sense and antisense transcripts as well as from

long hairpins containing bulges and mismatches (Czech et al. 2008). Detailed analysis of Loquacious gene expression revealed specific protein isoforms, which associate with miRNA (LOQS-PB) and RNAi (LOQS-PD) pathways (Hartig et al. 2009; Zhou et al. 2009). Two studies proposed a model where Loquacious and R2D2 function sequentially and non-redundantly in the endogenous RNAi pathway (Marques et al. 2010; Hartig and Forstemann 2011). LOQS-PD functions upstream, stimulating DCR-2-mediated processing of dsRNA whereas R2D2 acts during RISC loading.

TRBP, one of the two known mammalian RNAi dsRBPs, was identified as a Dicer-interacting partner involved in miRNA processing and RISC loading (Chendrimada et al. 2005; Haase et al. 2005). Notably, the role of TRBP in RNA silencing has been studied in cells where the physiological substrate for Dicer processing and RISC loading are miRNA precursors and where long dsRNA readily activates the protein kinase R (PKR) and interferons (IFN). Thus, while the RISC loading role of TRBP may be common for miRNA and RNAi pathways, it is not clear if an isoform of TRBP plays any specific role in recognition and processing of long dsRNA in the canonical mammalian RNAi pathway. Since TRBP also interacts with and inhibits PKR (Cosentino et al. 1995; Park et al. 1994), it was speculated that TRBP could be a component of a network of protein–protein interactions underlying a reciprocal regulation of RNAi/miRNA and IFN-PKR pathway (Haase et al. 2005). This notion is further supported by PACT, a paralog of TRBP, which exerts a positive effect on PKR. PACT was shown to interact with TRBP and Dicer and to facilitate siRNA production (Kok et al. 2007).

5.2.1.3 Helicases and Other Auxiliary Factors

There are several other protein factors which interact with Dicer and facilitate dsRNA recognition and cleavage. In *C. elegans*, numerous DCR-1-interacting proteins were identified by proteomic analyses (Tabara et al. 2002; Duchaine et al. 2006). Among these proteins are the aforementioned Dicer-related helicases (DRH-1, DRH-2, and DRH-3), which interact with DCR-1. DRH-1/2 interact with RDE-4 and DCR-1, and they are essential for exo-RNAi (Tabara et al. 2002). DRH-3 is essential for viability and it is involved in the endo-RNAi pathway (Duchaine et al. 2006). Thus, distinct DRH/DCR-1 complexes recognize different dsRNA triggers and mediate processing and loading on distinct primary Argonaute proteins.

The closest mammalian homologues of *Drh* genes are helicases *Ddx58*, *Dhx58*, and *Ifih1*, which are expressed in the immune system. DDX58, which is most similar to DRH1, is a helicase also known as RIG-I, which recognizes blunt-ended dsRNA and induces the interferon response (Marques et al. 2006). In fact, the presence of 3′ overhangs in Dicer products impairs RIG-I ability to unwind the dsRNA substrate and activates downstream signaling to the transcription factor IRF-3. The porcine ortholog of *Ddx58*, *Rhiv-1*, was initially identified as a locus responding to porcine reproductive and respiratory syndrome virus (PPRSV)

infection (Zhang et al. 2000). Thus, while DRH proteins facilitate Dicer processing in *C. elegans*, their closest mammalian homologues also respond to dsRNA but acquired novel roles in the immune system, which include an ability to tolerate Dicer products but respond to blunt-ended nonself dsRNAs, such as by-products of viral replication.

While other Dicer-associated proteins were found in *C. elegans*, *Drosophila*, and mammals, they seem to function downstream of Dicer cleavage and will be discussed further below.

5.2.2 RISC Complex Formation

The next step after siRNA production is selection and loading of one of its strand onto the RISC. The key component of RISC is an Argonaute family protein (AGO), which binds the selected siRNA strand and uses it as a sequence-specific guide for recognizing mRNAs that will be degraded. Upon formation of a perfect duplex, AGO protein endonucleolytically cleaves the cognate RNA in the middle of the duplex.

5.2.2.1 Argonaute Proteins

Argonaute proteins have a molecular weight of ~100 kDa and carry two distinct domains: the central PAZ domain and the PIWI (P-element induced wimpy testis) domain at the carboxy-terminus. Two additional domains are recognized, the N-terminal domain and the MID domain between PAZ and PIWI domains (Fig. 5.2). The PAZ domain binds the 3′ end of a short RNA in a sequence-independent manner (Lingel et al. 2003, 2004; Ma et al. 2004; Song et al. 2003). Structural studies of archeal Argonaute homologues showed that the PIWI domain has an RNase H-like fold (Song et al. 2004; Ma et al. 2005; Yuan et al. 2005; Parker et al. 2004). siRNA is anchored with its 3′ end in the PAZ domain. The 5′-phosphate of the siRNA is buried in a pocket at the interface between the MID domain and the PIWI domain (reviewed in Jinek and Doudna 2009). The 5′ end of the base pairing cognate mRNA enters between the N-terminal and PAZ domains, and its 3′ end exits between the PAZ and MID domains. Argonaute was identified to be a "slicer" (Liu et al. 2004; Meister et al. 2004; Song et al. 2004), i.e., the enzyme catalyzing the cleavage of the cognate mRNA in the canonical RNAi pathway. The active site in the PIWI domain is positioned to cleave the mRNA opposite the middle of the siRNA guide (Song et al. 2004).

Argonaute proteins can be divided into three distinct groups (reviewed in Faehnle and Joshua-Tor 2007) (1) AGO proteins, found in all kingdoms, (2) PIWI proteins, found in animals, and (3) WAGO proteins, found only in worms. The WAGO subfamily was described only recently (Yigit et al. 2006), so it is not recognized in the older literature, which typically divides Argonaute proteins into

AGO and PIWI subgroups (Carmell et al. 2002). Specific Argonaute proteins functioning in the RNAi pathway include RDE-1 (exo-RNAi) and ERGO-1 (endo-RNAi) in *C. elegans*, AGO-2 in *Drosophila*, and AGO2 in mammals. Other Argonautes act in the miRNA and other pathways employing small RNAs.

5.2.2.2 RISC Assembly

The key step in RISC formation is the loading of a short RNA produced by Dicer into the complex. In vitro experiments with mammalian proteins suggest that Dicer, TRBP, and AGO2 are critical for RISC loading, and the minimal RISC is composed of the AGO2 protein loaded with an siRNA (Martinez et al. 2002; Gregory et al. 2005; MacRae et al. 2008). A complex of Dicer, an RNAi dsRBP, and an Argonaute protein participates in RISC formation in all studied animal models (Tabara et al. 2002; Tomari et al. 2004b; Gregory et al. 2005).

RISC assembly is best understood in *Drosophila* and human models (Tomari et al. 2004a, b; Pham et al. 2004; Gregory et al. 2005; MacRae et al. 2008). The model of RISC loading in *Drosophila* suggests that RISC assembly occurs in several steps which involve a number of described complexes (Tomari and Zamore 2005). The first complex, formed by siRNA, R2D2, and DCR-2, is also known as R1 or R2/D2/DCR-2 initiator (RDI) complex (Pham et al. 2004; Kim et al. 2007) and develops into a mature form of the RISC loading complex RLC (Tomari and Zamore 2005). The RLC determines strand selection and recruits AGO2 (and other proteins) to form a pre-RISC (Kim et al. 2007), which contains duplex siRNA. Finally, the release of the passenger strand from the duplex produces holo-RISC, which can base pair with complementary mRNA substrates.

The coupling of dsRNA cleavage and RISC assembly is a matter of debate. It was suggested that, after cleavage, small-RNA duplexes need to dissociate from Dicer and then rebind to a sensor of the thermodynamic asymmetry of the duplex, because the siRNA guide strand will be at random orientation (Tomari et al. 2004a). Indeed, small RNA sorting in *Drosophila* suggests that dicing and RISC assembly are uncoupled (Tomari et al. 2007). In contrast, the immunopurified or reconstituted human AGO2 complex can use pre-miRNAs but not siRNA duplexes for a target cleavage suggesting that Dicer cleavage and RISC assembly are functionally coupled in humans (Gregory et al. 2005). However, newer data indicate that, just as in flies, human RISC assembly is uncoupled from dicing and ATP facilitates RISC loading of small-RNA duplexes (Yoda et al. 2010). Interestingly, all four human AGO proteins showed similar structural preferences for small-RNA duplexes, which were highly reminiscent of *Drosophila* AGO1 but not of AGO2 (Yoda et al. 2010).

Fly AGO1 and AGO2 require ATP for the RISC loading (Nykanen et al. 2001; Pham et al. 2004; Tomari et al. 2004a; Kawamata et al. 2009). ATP is presumably used to trigger the dynamic conformational opening of AGO proteins so that they can accept small RNA duplexes (Kawamata et al. 2009). Earlier studies suggested a difference between fly and human systems because human RISC assembly using

immunopurified or reconstituted human RLC containing AGO2, Dicer, and TRBP did not require ATP hydrolysis (Gregory et al. 2005; MacRae et al. 2008; Maniataki and Mourelatos 2005). Recent data suggest that ATP facilitates also human RISC loading while it is dispensable for unwinding (Yoda et al. 2010).

Strand selection in the fly RLC is controlled by R2D2. Analysis of the interaction of DCR-2/R2D2 complex with siRNA duplexes showed that R2D2 orients the complex according to thermodynamic stabilities of siRNA strands and binds the 5′-phosphate of the passenger strand at the thermodynamically more stable end (Tomari et al. 2004b). Thus, R2D2 functions as a licensing factor for routing siRNAs into the RNAi pathway. Interestingly, a thorough analysis of AGO2 complexes revealed that, unlike mature miRNAs which are loaded on AGO1, complementary strands of mature miRNAs (miRNA*) are efficiently loaded on AGO2 in DCR2/R2D2-dependent manner (Ghildiyal et al. 2010; Okamura et al. 2011). Thus, the role of R2D2 in sorting small RNAs is wider and extends into the miRNA pathway.

Mammals differ from *Drosophila* because they do not separate Dicer and Argonaute proteins dedicated to RNAi and miRNA pathways, and it is assumed that both pathways use a similar if not the same RLC. However, it should be kept in mind that our knowledge of the mammalian RLC comes either from cells where RLC normally loads miRNAs or from in vitro reconstitution of the RLC with purified proteins. The mammalian RLC is functionally similar to that of *Drosophila*. It is composed of Dicer, TRBP, and AGO2 (Gregory et al. 2005; MacRae et al. 2008). In vitro reconstituted mammalian RLC contains one copy of each protein and has dicing, guide-strand selection, AGO2-loading, and slicing activities.

Biochemical and structural analysis suggests that TRBP is flexibly bound to the Dicer DExH/D domain (Wang et al. 2009a; Daniels et al. 2009). TRBP seems to bridge the release of the siRNA by Dicer and the loading of the duplex onto AGO2. Binding by TRBP may allow the siRNA intermediate to stay associated with RLC after being released from Dicer and may also help in orientation of the siRNA for AGO2 loading. It was also predicted that TRBP acts as a sensor of the thermodynamic stability of 5′ siRNA in strand selection during RISC loading, similarly to DCR-2 and R2D2 (a TRBP homologue) in *Drosophila* (Wang et al. 2009a). However, the supporting evidence is inconclusive (Haase et al. 2005) although some argue that TRBP can indeed act as a sensor (Gredell et al. 2010). Furthermore, while TRBP function is similar to that of R2D2, TRBP sequence is more closely related to Loquacious than R2D2 (Murphy et al. 2008). It is tempting to speculate that the closer evolutionary distance between TRBP and Loquacious and the lack of an R2D2 ortholog reflect the fact that the mammalian endogenous RNAi is restricted to a few cell types and employs proteins normally functioning in the miRNA pathway.

The final step in the assembly of an active RISC is the release of the passenger strand from the siRNA duplex. A helicase activity was proposed to separate the two siRNA strands while the guide remains bound to AGO2 (Sontheimer 2005; Tomari and Zamore 2005; Meister and Tuschl 2004). A candidate for such a helicase in *Drosophila* is Armitage helicase (Tomari et al. 2004a). However, experimental data

support a simple solution where passenger strand cleavage by AGO2 slicer activity liberates the single-stranded guide siRNA strand from the pre-RISC complex (Matranga et al. 2005; Miyoshi et al. 2005; Kim et al. 2007). Removal of siRNA passenger strand cleavage products is assisted by C3PO endoribonuclease, which was identified as a RISC-enhancing factor that promotes RISC activation (Liu et al. 2009). The cleavage-assisted mechanism is typical for AGO2-loaded fly and human siRNAs in the RNAi pathway while passenger strand cleavage is not important for loading miRNAs (Matranga et al. 2005).

5.2.2.3 RISC Composition: AGO2-Interacting Proteins

While the minimal active RISC contains only the "slicing" Argonaute protein and the guide siRNA strand (Martinez et al. 2002; MacRae et al. 2008, Rand, 2004 #424), RISC activity was found in different models and cell types to reside in ~200 kDa, ~500 kDa, or 80S complexes (Martinez et al. 2002; Nykanen et al. 2001; Mourelatos et al. 2002; Pham et al. 2004). Various protein components of RISC complexes either contribute to RISC formation or might regulate RISC activity, stability, target selection, mode of repression, or otherwise contribute to RISC function.

In *C. elegans*, the RISC complex was not biochemically purified, so protein complexes containing Argonaute proteins RDE-1 (exo-RNAi) and ERGO-1 (endo-RNAi) remain uncharacterized.

Analysis of RISC in *Drosophila* embryo lysate identified the several additional components of AGO2-containing complexes:

dFMR1: *Drosophila* ortholog of human fragile X mental retardation protein (FMRP) (Caudy et al. 2002; Ishizuka et al. 2002; Pham et al. 2004). dFMR1 is associated with ribosomes through interaction with ribosomal proteins L5 and L1 and with complexes containing miRNAs (Ishizuka et al. 2002). dFMR1 is not a conserved RISC component involved in RNAi. While depletion of dFMR1 reduces RNAi efficiency in *Drosophila* S2 cells (Caudy et al. 2002), the loss of mammalian FMRP has no apparent direct impact on RISC function (Didiot et al. 2009).
VIG: Vasa Intronic Gene (Caudy et al. 2002; Pham et al. 2004). VIG is a conserved protein, which encodes for a putative RNA-binding protein, whose depletion reduces RNAi efficiency (Caudy et al. 2002). *Vig* mutants are more susceptible to viral infections in *Drosophila* (Zambon et al. 2006). Whether this role of VIG is coupled with its presence in the RISC complex is not known. There is no evidence that SERBP1, the closest mammalian VIG homologue, would be associated with RISC.
Armitage: RNA helicase, which was identified as a maternal effect gene required for RNAi (Tomari et al. 2004a). Armitage is probably not required for the RISC activity. Instead, it was proposed to facilitate the removal of the passenger strand during RISC formation (Tomari et al. 2004a). The mammalian homologue of Armitage is an Argonaute-associated protein MOV10 (Meister et al. 2005).

TSN: Tudor Staphylococcal Nuclease is a protein containing five staphylococcal/micrococcal nuclease domains and a tudor domain. It is a component of the RISC in *C. elegans*, *Drosophila*, and mammals (Caudy et al. 2003; Pham et al. 2004). The role of TSN in RISC RNAi remains enigmatic. TSN is not the "slicer" (Schwarz et al. 2004), and it has been implicated in promoting cleavage of hyper-edited dsRNA (Scadden 2005). This observation is surprising considering that RNA editing and RNAi pathways appear mutually antagonistic (Tonkin and Bass 2003; Yang et al. 2005). TSN is also connected with the miRNA pathway, where it mediates degradation of edited miRNA precursors (Yang et al. 2006; Kawahara et al. 2007a).
DMP68 (RM62): This conserved helicase was co-purified with AGO1 and dFMR1 (Ishizuka et al. 2002). It seems to be required for RNAi in S2 cells where depletion of DMP68 results in inhibition of RNAi (Ishizuka et al. 2002). Whether DMP68 is needed for RISC formation or for RISC activity/stability is not known.

Many AGO2-associated proteins were identified in mammalian cells (reviewed in detail in Peters and Meister 2007). These include MOV10, DDX6 (Rck/p54), DDX20 (Gemin3), TNRC6A (GW182), and many others. However, experiments concerning mammalian AGO2-associated proteins were performed in somatic cells where these proteins are loaded with miRNAs. Accordingly, some of the AGO-associated proteins clearly associate with miRNA-mediated repression, while it is not known if any AGO2-associated protein is required for RNAi.

5.2.3 Target Recognition and Cleavage

The RISC complex uses the loaded siRNA as a guide for recognizing its target for cleavage. As mentioned above, the first cleavage actually targets the passenger strand of a loaded siRNA duplex to free the guiding strand, so it can base pair to cognate mRNAs (Matranga et al. 2005; Kim et al. 2007). Whether the first cleavage has any further effects on the RISC structure is not known.

Target recognition requires that the loaded RISC finds and hybridizes to the complementary targets. Target recognition by siRNAs exhibits a distinct 5′ bias. Analysis of miRNA-targeted mRNAs revealed that miRNA bases 2–8 form a distinct "seed," which base pairs perfectly to the target transcript (Lewis et al. 2003; Enright et al. 2003). It was found that the 5′ half of a small RNA provides most of the binding energy that tethers RISC to a target RNA (Doench et al. 2003; Haley and Zamore 2004). Biochemical analysis of target recognition by mammalian RISC showed that the RISC is apparently not systematically scanning transcripts and it is unable to unfold structured RNA. Thus, RISC randomly transiently contacts single-stranded RNA and promotes siRNA-target base pairing (Ameres et al. 2007). The 5′ end of the loaded siRNA creates a thermodynamic threshold for a stable association of RISC with its target (Ameres et al. 2007).

The fact that 5′ and 3′ ends of an siRNA are bound by distinct binding pockets and that both ends contribute differently to binding to the target lead to a "two-state

model of Argonaute function" (Tomari and Zamore 2005). In this model, the 3′ end is bound in the PAZ domain and the 5′ end in a pocket at the interface between the MID and the PIWI domains. The 5′ end is pre-organized to interact with the cognate mRNA and, upon binding, the 3′ end is dislodged from the binding pocket to allow for base pairing of the 3′ end. Base pairing in the middle of siRNA results in a correct orientation and cleavage of the cognate strand in the active site. This model is supported by recent structural data (Wang et al. 2009b).

Notably, siRNAs can mediate other silencing effects than the cleavage. An imperfect complementarity in the middle to the base pairing may result in translational repression, which is a typical effect of miRNAs (Doench et al. 2003). In addition, out of four mammalian AGO proteins, which can associate with siRNAs, only AGO2 has the "slicer" activity (Liu et al. 2004; Meister et al. 2004; Song et al. 2004).

5.2.4 *RdRP Amplifier and Its Loss During Animal Evolution*

RdRP is an ancestral component of RNAi because RdRP orthologs were identified in RNA silencing pathways in plants, fungi, and some animals: QDE-1 in *Neurospora crassa* (Cogoni and Macino 1999), EGO-1 and RRF-1 in *C. elegans* (Grishok et al. 2001; Smardon et al. 2000), SDE1/SGS2 in *Arabidopsis* (Dalmay et al. 2001; Mourrain et al. 2000), and Rdp1 in *Schizosaccharomyces pombe* (Hall et al. 2002; Volpe et al. 2002).

C. elegans is the main animal model for studying RdRPs. While an earlier study of RdRP in *C. elegans* suggested that dsRNA synthesis can be primed by primary siRNAs (a model of "degradative PCR") (Sijen et al. 2001), later studies demonstrated that RdRPs do not require the priming by primary siRNAs and produce short RNAs using RISC-targeted mRNAs as templates (Sijen et al. 2007; Pak and Fire 2007). Surprisingly, a sequencing of small RNAs associated with ongoing RNAi in *C. elegans* showed that Dicer-independent secondary siRNAs constitute the majority of cloned siRNAs (Pak and Fire 2007). These secondary siRNAs are only antisense, carry 5′-di- or triphosphates, and are not bound by RDE-1 but by other Argonaute proteins (Sijen et al. 2007; Pak and Fire 2007). *C. elegans* genome encodes for 27 Argonaute proteins (Yigit et al. 2006) and four putative RdRPs, three of which were implicated in RNA silencing (Duchaine et al. 2006; Lee et al. 2006; Sijen et al. 2001; Smardon et al. 2000). A systematic analysis of small RNAs combining different cloning strategies with the next generation sequencing and analysis of Argonaute proteins revealed an amazing complexity of RNA silencing pathways in *C. elegans*. Different RNA substrates are processed in Dicer-dependent and Dicer-independent manner to produce numerous classes of small RNAs, which are loaded on different AGO proteins (Gent et al. 2010; Yigit et al. 2006; Correa et al. 2010; Vasale et al. 2010). Thus, the core RNAi pathway (Sects. 2.1–2.3) in *C. elegans* can be seen as a starter followed by the main course made by RdRPs.

Homologues of these RdRPs exist in numerous eumetazoan phyla, including *Nematoda* (e.g., *C. elegans*), *Cnidaria* (*Hydra*), *Chelicerata* (tick), *Hemichordata* (acorn worm), *Urochordata* (sea squirt), and *Cephalochordata* (lancelet), but appear absent in others, including *Platyhelminthes* (*Planaria*), *Mandibulata* (*Drosophila*), and *Craniata* (vertebrates). Phylogenetic data suggest that RdRPs in RNA silencing pathways have a monophyletic origin, i.e., evolved from a single ancestral RdRP (Murphy et al. 2008; Cerutti and Casas-Mollano 2006). The fact that RdRP orthologs are found in other protostomes and deuterostomes but not in *Drosophila* or mammals suggests a repeated loss of the ancestral RdRP component of RNA silencing. Whether RdRP activity completely disappeared from RNAi in *Drosophila* and mammals is unclear. The missing RdRP orthologs in RNA silencing in *Drosophila* or vertebrates could be replaced by another RdRP, for example by horizontal transfer of some viral RdRP. In fact, there is evidence for RdRP analogs in *Drosophila* and vertebrates (Lipardi and Paterson 2009; Sam et al. 1998; Maida et al. 2009; Pelczar et al. 2010). Whether and how these activities participate in RNAi remains unresolved. An earlier report of RdRP activity in *Drosophila* (Lipardi et al. 2001) was contradicted by experiments demonstrating the absence of transitive RNAi generating secondary sequences upstream of the region targeted by siRNAs (Roignant et al. 2003; Schwarz et al. 2002). A similar lack of transitive RNAi was observed in mammals (Stein et al. 2003).

Two recent reports propose that two different RdRP activities in *Drosophila* and human cells can generate dsRNA that can be processed by Dicer. One of the RdRPs is ELP1, a noncanonical RdRP conserved in all eukaryotes, which associates with DCR-2, and its loss results in reduction of endo-siRNAs and upregulation of transposon transcripts (Lipardi and Paterson 2009). The second RdRP is a ribonucleoprotein complex of the human telomerase reverse transcriptase (TERT) and the RNA component of mitochondrial RNA processing endoribonuclease (RMRP). RMRP shows a strong preference for substrates that have 3′ fold-back structures and produces dsRNA that can be processed by Dicer (Maida et al. 2009). In both cases, additional experiments are needed to confirm that ELP1 or TERT-RMRP participates in RNAi and what their exact role is. These include analysis of small RNAs in mutants lacking *Elp1* or TERT and further characterization of complexes containing ELP1 or TERT and RNAi components. Since the transitive RNAi was not detected in *Drosophila* and mammals, it is possible that analogous RdRP activities are not involved in the production of secondary siRNAs. Instead, they could play a role in siRNA-independent production of dsRNA substrates for RNAi. That could explain why ELP1 was not found in any of the previous biochemical studies of RNAi in *Drosophila* and mammals.

5.2.5 Systemic and Environmental RNAi

RNAi can either act in a cell autonomous manner, i.e., affecting only cells directly exposed to dsRNA, or can propagate across cell boundaries (Fig. 5.3). The non-cell autonomous RNAi was observed already during the first RNAi experiments in

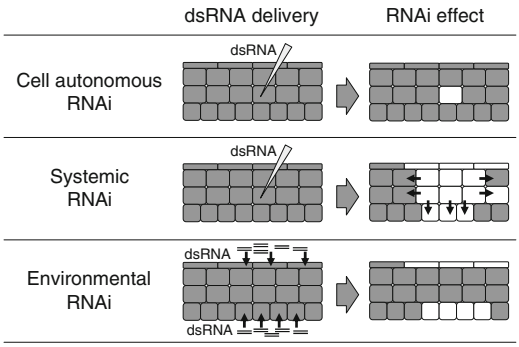


Fig. 5.3 A schematic overview of the different types of RNAi explained on a model of silencing a "*green*" gene in cells. The *first row* shows the cell-autonomous RNAi where the RNAi agent targeting the "*green*" gene is delivered into a cell (e.g., by injection or viral transduction). The silencing effect remains only in the cells where the RNAi was induced initially. Systemic RNAi (the *middle row*) includes processes where the silencing is transported from the cell where RNAi was induced to other cells or even to different tissues where the silencing also takes place. In case of environmental RNAi (the *bottom row*), the dsRNA is taken up from the environment of the cell (external or internal environment in relation to the animal). The silencing effect is observed in all cells which can take up the dsRNA

C. elegans (Fire et al. 1998). It worked like a magic trick: when animals were microinjected with dsRNA into head, tail, intestine, or gonad arm, or if they were even just soaked in dsRNA solution or fed by bacteria expressing dsRNA, a specific null phenotype was induced in the whole animal and even in its progeny, demonstrating a surprising ability of dsRNA to cross cellular boundaries (Fire et al. 1998; Tabara et al. 1998; Timmons and Fire 1998). Two modes of non-cell autonomous RNAi are recognized (1) environmental RNAi involves processes in which dsRNA is taken up by a cell from the environment; (2) systemic RNAi includes processes where a silencing signal spreads from a cell across cellular boundaries into other cells. The studies of RNAi in *C. elegans* show that both modes can be combined and environmental RNAi can be followed by systemic RNAi.

At least two pathways for dsRNA uptake were described (1) a specific transmembrane channel-mediated uptake and (2) an alternative endocytosis-mediated uptake (reviewed in Huvenne and Smagghe 2010; Whangbo and Hunter 2008).

The best understood systemic RNA mechanism is that of *C. elegans* where the transport of the silencing signal to neighboring cells is controlled by dsRNA-transporting channels. *Sid-1* and *sid-2* genes were identified in forward genetic screen to be responsible for systemic RNAi in *C. elegans* (Winston et al. 2002).

SID-1 (systemic RNAi deficient-1) is a conserved transmembrane protein that has homologues in a wide range of animals, including mammals. *Sid-1* mutants have intact cell autonomous RNAi, but are unable to perform either systemic RNAi or environmental RNAi in response to feeding, soaking, or injection of dsRNA (Winston et al. 2002). SID-1 sensitizes *Drosophila* cells to RNAi induced by soaking, enabling concentration-dependent cellular uptake of dsRNA suggesting

that SID-1 forms a dsRNA channel (Feinberg and Hunter 2003; Shih et al. 2009). SID-2 is a transmembrane protein localized to an apical membrane of intestinal cells. It is necessary for the initial import of dsRNA from gut lumen, but not for the systemic spread of silencing signals among cells. *Sid-2* homologues have been identified only in two other *Caenorhabditis* species (Winston et al. 2007).

To date, non-cell autonomous RNAi has also been discovered in parasitic nematodes (Geldhof et al. 2007), *Hydra* (Chera et al. 2006), *Planaria* (Newmark et al. 2003; Orii et al. 2003), or insects (Tomoyasu et al. 2008; Xu and Han 2008). However, non-cell autonomous RNAi is not present uniformly. For example, only one of eight tested *Caenorhabditis* species showed efficient environmental RNAi (Winston et al. 2007). Diverse non-cell autonomous RNAi also exists in insects where different taxa have up to three expressed or silent *sid-1* orthologs (Huvenne and Smagghe 2010; Whangbo and Hunter 2008). Some insects, such as red flour beetle *Tribolium*, have efficient systemic RNAi where injection of adults causes RNAi effects in the progeny (Bucher et al. 2002). In *Drosophila*, the natural role of non-autonomous RNAi seems to be coupled with antiviral role of RNAi in adult flies (Saleh et al. 2009). Experimentally induced non-cell autonomous RNAi in *Drosophila* was achieved in some (Dzitoyeva et al. 2003; Eaton et al. 2002) but not all cases (Roignant et al. 2003).

Non-cell autonomous RNA with an extent similar to that of *C. elegans* or of some insects is highly unlikely in vertebrates. However, a limited environmental or systemic RNAi may exist therein as the homologues of *sid-1* have been found in all sequenced vertebrate genomes (Jose and Hunter 2007). Two *sid-1* homologues (*SidT1* and *SidT2*) are present in mice and humans with a documented role for *SidT1* in dsRNA uptake in humans (Duxbury et al. 2005; Wolfrum et al. 2007). Furthermore, experimental overexpression of human *SidT1* significantly facilitated cellular uptake of siRNAs and resulted in increased RNAi efficacy (Duxbury et al. 2005). At the same time, it should be kept in mind that the mammalian immune system employs a number of proteins responding to dsRNA independently of RNAi (Gantier and Williams 2007), the canonical RNAi pathway efficiently operates in a limited number of cell types, and mammalian RNAi does not seem to participate in the innate immunity (Cullen 2006). Thus, the primary role of a dsRNA uptake mechanism in mammals is likely not involving RNAi even though it could have served such a role in an ancestral organism.

5.3 Roles of RNAi in Animals

Because dsRNA often originates from harmful sources, such as viruses and mobile elements, the role of RNAi is often viewed as a form of innate immunity. While this role is experimentally supported, analysis of RNAi mutants suggests additional roles. This part summarizes the current knowledge of the role of RNAi in combating viruses, maintaining genome integrity, and control of gene expression (see also the chapter Diversity, overlap, and relationships in the small RNA landscape by Michelle Scott in this volume).

5.3.1 Antiviral Role of RNAi

RNA viruses generate dsRNA during their replication cycle in host cells; DNA viruses often produce complementary sense and antisense transcripts, which can form dsRNA upon annealing. Thus, dsRNA is a common marker of viral infection and it is recognized by various mechanisms mediating an innate immune response. A role of RNA silencing in the innate immunity is supported by several lines of evidence, which were first found in plants and later also in invertebrates (reviewed in Xie and Guo 2006; Marques and Carthew 2007) (1) siRNAs derived from viral sequences were found in infected organisms (Hamilton and Baulcombe 1999), (2) inhibition of RNA silencing results in increased viral replication (Mourrain et al. 2000), and (3) some viruses produce suppressors of RNA silencing (SRS) (Voinnet et al. 1999).

Several studies addressed the role of RNAi in viral suppression in *C. elegans*. As endogenous viral pathogens of *C. elegans* were unknown, this problem was bypassed by using an "artificial" infection with viruses, which had a broad host range and could infect *C. elegans* under laboratory conditions. Model viral infections were based on the (+)ssRNA flock house virus (FHV) (Lu et al. 2005) or the (−)ssRNA vesicular stomatitis virus (VSV) (Schott et al. 2005; Wilkins et al. 2005). Infection with the recombinant VSV was augmented in strong RNAi mutant animals (*rde-1* and *rde-4*), and mutants produced higher viral titers. Furthermore, VSV infection was attenuated in *rrf-3* and *eri-1* mutants that are hypersensitive to RNAi (Wilkins et al. 2005). Similar results were obtained from infected cultured cells (Schott et al. 2005) and from an FHV infection of *rde-1* mutants (Lu et al. 2005). The antiviral role of exo-RNAi in nematodes was recently demonstrated also for a newly discovered natural viral infection of *C. elegans* and *C. briggsae* (Felix et al. 2011).

An antiviral role of RNAi has also been demonstrated in insects. It was shown that FHV is an initiator and a target of RNA silencing in *Drosophila* host cells (Li et al. 2002). Infection of 14 different *Drosophila* RNA silencing mutants with a dsRNA X virus (DXV) showed that all but three lines were significantly more susceptible to viral infection (reduced survival and elevated viral titers) than normal flies. Moreover, replication of DXV was sequence-specifically inhibited (but not absolutely blocked) by "immunizing" *Drosophila* S2 cells with dsRNA from the coding region of DXV before infection (Zambon et al. 2006). Interestingly, increased susceptibility was observed not only for mutants of the RNAi pathway, such as *r2d2*, *armi*, or *ago2*, but also for mutants of the piRNA pathway (*aubergine* and *piwi*), suggesting that RNAi is not the only RNA silencing pathway in *Drosophila* dedicated to the antiviral response. A number of studies provides ample evidence that RNAi plays an essential role in antiviral response in insects (Galiana-Arnoux et al. 2006; Keene et al. 2004; Nayak et al. 2010; Sanchez-Vargas et al. 2009; Wang et al. 2006).

In contrast to nematodes and insects, data supporting involvement of mammalian RNAi in antiviral defense is weak (reviewed in detail in Cullen 2006). It is

unlikely that RNAi substantially acts as an antiviral mechanism in mammals where long dsRNA induces a complex sequence-independent antiviral response, commonly known as the interferon response (reviewed in Gantier and Williams 2007). Consistent with this, no siRNAs of viral origin have been found in human cells infected with a wide range of viruses (Pfeffer et al. 2005). Occasional observations, such as detection of a single siRNA in HIV-1 infected cells (Bennasser et al. 2005) does not provide any conclusive evidence that RNAi is processing viral dsRNA and suppresses viruses under physiological conditions in vivo.

It must be stressed that any circumstantial evidence suggesting the role of RNAi in viral suppression must be carefully examined and interpreted. Since viruses coevolve with different hosts and explore all possible strategies to maintain and increase their fitness, it is not surprising that viral reproductive strategies come into contact with mammalian RNA silencing pathways, particularly the miRNA pathway, which shares components with the RNAi pathway. For example, Epstein-Barr virus (EBV) and several other viruses encode their own miRNAs (Pfeffer et al. 2004, 2005; Sullivan et al. 2005) or take advantage of host cell miRNAs to enhance their replication (Jopling et al. 2005; Pfeffer et al. 2005).

Another evidence for an interaction between viruses and RNA silencing is the presence of putative SRS in various viruses. As viral genomes rapidly evolve, SRS should be functionally relevant. For example, B2 protein in Nodaviruses (e.g., FHV) is essential for replication and inhibits Dicer function, and B2-deficient FHV can be rescued by artificial inhibition of RNAi response (Li et al. 2002). B2 protein also enhances the accumulation of Nodaviral RNA in infected mammalian cells (Fenner et al. 2006; Johnson et al. 2004). Other potential SRS molecules have been identified in viruses infecting vertebrates, such as Adenovirus VA1 noncoding RNA (Lu and Cullen 2004), Influenza NS1 protein (Li et al. 2004), Vaccinia virus E3L protein (Li et al. 2004), Ebola virus VP35 protein (Haasnoot et al. 2007), Tas protein in primate foamy virus (Lecellier et al. 2005), or HIV-1 Tat protein (Bennasser et al. 2005).

The existence of SRS in viruses infecting mammals does not prove that these viruses are targeted by RNAi in mammalian cells. First, viruses may have a broader range of hosts (or vectors), e.g., blood sucking insects. Thus, a virus can be targeted by RNAi in one host and by a different defense mechanism in another one. For example, the Dengue virus, whose life cycle takes place in humans and mosquitoes, is targeted by RNAi in mosquitoes and it likely evolved some adaptation to circumvent the response (Sanchez-Vargas et al. 2009). Second, viral SRS in mammalian cells may have other purpose than counteracting viral suppression by RNAi. Since biogenesis and mechanism of action of mammalian miRNAs overlaps with RNAi, it is possible that the role of such SRS is to modify cellular gene expression by suppressing the activity of miRNAs. Third, the primary effect of SRS may be aimed at other defense mechanisms recognizing and responding to dsRNA and, as a consequence, SRS effects on RNA silencing are observed.

5.3.2 RNAi and Suppression of Mobile Elements

In *C. elegans*, mutant screens for components of the RNAi pathway and for the repressors of Tc1 transposon revealed that RNAi controls the activity of transposable elements (TEs) (Ketting et al. 1999; Tabara et al. 1999). Consistent with these results, primary endo-siRNAs derived from Tc1 transposon were observed in the germline (Sijen and Plasterk 2003). High-throughput sequencing identified additional small RNA species involved in TE silencing. These include Dicer-independent piRNAs expressed in all animal phyla (reviewed in Malone and Hannon 2009a) and secondary siRNAs generated by RdRPs in *C. elegans* (Pak and Fire 2007; Sijen et al. 2007). An unexpected connection linking piRNAs to endo-siRNA biogenesis has been observed in *C. elegans* germline (Das et al. 2008), where endo-siRNAs targeting Tc3 element were dependent on piRNAs while Tc1-derived endo-siRNAs were still present in piwi mutants. These findings suggest a cross talk between small RNA pathways in worms to ensure efficient TE silencing.

In *Drosophila*, the small RNA pathways are believed to play distinct roles in TE silencing in the germline and in somatic tissues. While piRNAs are responsible for genome surveillance predominantly in the germline, TE-derived endo-siRNAs have been identified in somatic tissues and cultured cell lines (Czech et al. 2008; Okamura et al. 2008; Kawamura et al. 2008; Ghildiyal et al. 2008; Chung et al. 2008). However, piRNAs from ovarian somatic cells have also been described, arising specifically from the flamenco locus (Malone et al. 2009; Li et al. 2009). As endo-siRNAs from the flamenco locus and several other piRNA loci were also detected, it is likely that the *Drosophila* piRNA and endo-siRNA pathways might be interdependent, similarly to *C. elegans*, in their task to efficiently silence TEs (Ghildiyal et al. 2008).

Abundant piRNAs and potential endo-siRNAs derived from transposons and repetitive elements were also found in *Xenopus tropicalis* (Armisen et al. 2009). Here, the piRNAs are restricted solely to the germline while endo-siRNAs originating from similar genomic loci were found in both oocytes and somatic tissues, with most endo-siRNAs mapping to the palindromic sequences of Polinton DNA transposons.

RNAi-mediated TE silencing has also been documented in the mouse germline (Watanabe et al. 2006, 2008; Tam et al. 2008). Mutations in the piRNA pathway components are detrimental to sperm development, suggesting that piRNAs are a dominant class of small RNAs controlling TE activity in the male germline (Malone and Hannon 2009b). In contrast, female mice lacking functional piRNA pathway are fertile with no obvious defects in oocytes (Carmell et al. 2007). Endo-siRNAs suppress TE silencing in mammalian oocytes as documented by derepression of some retrotransposons in oocytes depleted of Dicer or AGO2 (Murchison et al. 2007; Watanabe et al. 2008). As already proposed for invertebrates, the piRNA and endo-siRNA pathways likely cooperate in creating a complex silencing network against TEs in the mammalian germline. Long terminal repeat MT elements and SINE elements are strongly upregulated in *Dicer-/-* oocytes, while

the levels of IAP retrotransposon are elevated in the absence of Mili protein but not in *Dicer*-/- oocytes (Watanabe et al. 2008; Murchison et al. 2007). Still many loci composed of other types of TEs, e.g., LINE retrotransposons, give rise to both piRNAs and endo-siRNAs, again suggesting that the biogenesis of these small RNAs is interdependent. The role of endogenous RNAi in TE silencing extends from germ cells to preimplantation embryo stages. Apart from maternally derived piRNAs and endo-siRNAs, which persist in the embryos for a large part of preimplantation development, zygotic endo-siRNAs are generated de novo, mainly to control the activity of zygotically activated MuERV-L retrotransposon (Ohnishi et al. 2010; Svoboda et al. 2004). SINE-derived endo-siRNAs also increase in abundance in early embryo stages, which is consistent with the observation that B1/Alu SINE endo-siRNAs account for a vast majority of endo-siRNAs sequenced from mouse embryonic stem cells (mESCs) (Babiarz et al. 2008). Whether these SINE endo-siRNAs play an active role in TE silencing in mESCs similarly to other TE-derived endo-siRNAs in oocytes remains to be determined. RNAi-dependent silencing of LINE transposons has also been described in cultured HeLa cells, where endo-siRNAs derived from bidirectional transcripts of sense and antisense L1 promoter were proposed to control L1 activity (Yang and Kazazian 2006). Although some evidence for retrotransposon-derived endo-siRNAs from mammalian somatic cells was obtained from deep sequencing data (Kawaji et al. 2008), a convincing support for the function of endo-siRNAs in TE silencing in mammalian somatic tissues has yet to be provided. Further reading on small RNAs and TEs is provided in the chapter Diversity, overlap, and relationship in the small RNA landscape by Michelle Scott (this volume).

5.3.3 *RNAi and Regulation of Endogenous Genes and Development*

A role of endogenous RNAi in shaping the protein-coding transcriptomes during development has been challenged by mutant worms and flies lacking essential components of the RNAi pathway, which were viable and produced healthy offspring (Tabara et al. 1999; Lee et al. 2004; Okamura et al. 2004). In view of those findings, RNAi had been viewed solely as a defense mechanism against invasive nucleic acids. However, deep sequencing analyses revealed that endo-siRNAs with sequence complementarity to hundreds of protein-coding mRNAs are present in *C. elegans* (Ambros et al. 2003; Ruby et al. 2006). Gene regulating endo-siRNAs in *C. elegans* differ in biogenesis and in requirements for functional components of the RNAi exerting machinery (Pak and Fire 2007; Sijen et al. 2007). This is consistent with microarray analysis of mutant worms lacking various RNAi-related factors, which identified non-overlapping sets of differentially expressed genes, supporting the idea of functionally distinct RNAi pathways in nematodes (Lee et al. 2006). In search for the cellular and developmental processes,

which might be controlled by endo-siRNAs in *C. elegans*, spermatogenesis-associated genes were found enriched in the group of transcripts matching endo-siRNAs (Ruby et al. 2006). Further studies revealed that mutations in RNAi-related genes result in defects in meiotic chromosome disjunction, spindle formation, or microtubule organization during sperm development and ultimately lead to male sterility or embryonic lethality of the offspring (Han et al. 2008, 2009; Pavelec et al. 2009; Gent et al. 2009).

A fraction of *Drosophila* endo-siRNAs maps to protein-coding regions (Ghildiyal et al. 2008; Kawamura et al. 2008; Okamura et al. 2008; Czech et al. 2008). However, only endo-siRNAs derived from a small number of loci are produced in sufficient amount to reduce target mRNA levels, as exemplified by the *esi-2* locus-derived endo-siRNAs targeting DNA damage-response gene *Mus-308* (Czech et al. 2008; Okamura et al. 2008). The second type of *Drosophila* endo-siRNAs arise from overlapping antisense transcripts observed in hundreds of protein-coding loci. Abundance of these endo-siRNAs is generally low, likely reflecting the fact that RdRP does not significantly contribute to endo-siRNA biogenesis in flies. In addition, their potential mRNA targets are not upregulated in *Ago2*-deficient flies, suggesting that these endo-siRNAs are not involved in posttranscriptional control of mRNA levels under physiological conditions (Czech et al. 2008). Interestingly, a dsRNA/endo-siRNA-binding protein Blanks, which associates with DCR-2 and forms an alternative Argonaute-independent functional RISC complex, has a role in spermatogenesis (Gerbasi et al. 2011). As Blanks deletion does not affect transposon activity, this finding would imply a role for endo-siRNAs in regulation of protein-coding mRNAs in *Drosophila* sperm development.

Although *Drosophila* mutants lacking *Dcr-2* or *Ago-2* develop to normal adults with no specific phenotype under standard laboratory conditions, severe defects in embryonic development have been noted in these mutants upon exposure to temperature perturbations (Lucchetta et al. 2009). When two halves of a living embryo were maintained at different temperatures, the mutant embryos were not able to compensate for faster development in the anterior part exposed to elevated temperature, which lead to segmentation abnormalities. This indicates that endo-siRNA pathway is needed to stabilize embryonic development under environmental stress (Lucchetta et al. 2009). Strikingly, the endo-siRNA-linked defects in *C. elegans* sperm development were also observed upon elevated temperature (Duchaine et al. 2006; Kennedy et al. 2004; Han et al. 2008; Pavelec et al. 2009; Gent et al. 2009). It is currently unknown whether the temperature-sensitive endo-siRNA pathway mobilization might be linked to increased endo-siRNA production or to enhanced activity of the RNAi machinery. However, it is tempting to speculate that support of development during unfavorable conditions is another role for endo-siRNA pathway in animals.

In mice, perturbation of the endo-siRNA pathway in oocytes is responsible for severe meiotic defects and resulting female infertility. Oocyte-specific knockout of either *Dicer* or *Ago2* leads to similar phenotypes, including chromosome misalignment and defective spindle (Murchison et al. 2007; Kaneda et al. 2009; Tang et al.

2007). These effects were originally attributed to the loss of maternal miRNAs. However, miRNA pathway is suppressed in mouse oocytes and nonessential as oocytes lacking canonical miRNAs are fertile (Ma et al. 2010; Suh et al. 2010). Interestingly, transcriptomes of oocytes lacking either *Dicer* or *Ago2* are similarly affected and genes matching pseudogene-derived endo-siRNAs (Tam et al. 2008; Watanabe et al. 2008) are enriched in the group of upregulated genes in both knockouts (Kaneda et al. 2009; Ma et al. 2010; Tang et al. 2009). In addition, putative endo-siRNA targets are enriched in cell cycle regulators and genes involved in microtubule organization and dynamics (Tam et al. 2008). These findings suggest that regulation of protein-coding genes by endo-siRNAs controls the equilibrium of protein factors required for proper spindle formation, chromosome segregation, and meiosis progression in mouse oocytes. As pseudogenes are a rapidly evolving source of dsRNA for endo-siRNA production, it will be interesting to investigate whether the role of RNAi in spindle formation during meiotic maturation of oocytes is conserved in mammals.

Endo-siRNAs have also been proposed to contribute to the self-renewal and proliferation of mESCs, since the proliferation and differentiation defects observed in *Dicer*-/- mESCs are stronger than in *Dgcr8*-/- mESCs (Kanellopoulou et al. 2005; Murchison et al. 2005; Wang et al. 2007). A population of endo-siRNAs derived mostly from hairpin-forming B1/Alu subclass of SINE elements was identified in mESCs (Babiarz et al. 2008). Fragments of SINE elements are commonly present in untranslated regions of protein-coding transcripts; therefore, it is possible that SINE-derived endo-siRNAs participate in posttranscriptional gene silencing in mESCs. However, this hypothesis has not been tested experimentally.

Scarce evidence is available for potential role of endo-siRNAs in the regulation of protein-coding mRNAs in mammalian somatic tissues. Endo-siRNAs derived from natural antisense transcripts of *Slc34a* gene were identified in mouse kidney, where this sodium/phosphate cotransporter exerts its physiological function (Carlile et al. 2009). However, changes in expression levels of *Slc34a* upon suppression of the endo-siRNA pathway have not been addressed. Deep sequencing revealed a set of potential endo-siRNAs generated from overlapping sense/antisense transcripts and from hairpin structures within introns of protein-coding genes in the mouse hippocampus (Smalheiser et al. 2011). The most abundant endo-siRNAs from *SynGAP1* gene locus were also found in complexes with AGO proteins and FMRP in vivo. Interestingly, a large part of potential hippocampal endo-siRNA targets encode for proteins involved in the control of synaptic plasticity, and the number of endo-siRNAs derived from these gene loci increased significantly during olfactory discrimination training (Smalheiser et al. 2011). Given the fact that the vast majority of identified endo-siRNA sequences mapped to intronic regions, the endo-siRNAs could act co-transcriptionally on nuclear pre-mRNAs, perhaps similarly to the mechanism of RNAi-mediated inhibition of RNA polymerase II elongation recently described in *C. elegans* (Guang et al. 2010). Alternatively, endo-siRNAs could control a correct distribution of target mRNAs as unspliced pre-mRNA can be exported from the neuronal nucleus and transported to dendrites for processing (Glanzer et al. 2005). In any case, these findings open an

attractive hypothesis that endo-siRNAs participate in synaptic plasticity during learning process, and the neuronal endo-siRNA pathway might be also linked to various neurodegenerative disorders (Smalheiser et al. 2011).

5.3.4 Interaction of RNAi with Other dsRNA-Induced Pathways

Diversity of dsRBPs shows that dsRNA plays other roles apart from serving as a trigger in RNAi. In this section, we will discuss two dsRNA responding pathways: A-to-I editing and interferon response, which coexist and interact with RNAi.

5.3.4.1 A-to-I Editing

A-to-I editing is mediated by adenosine deaminases acting on RNA (ADARs), enzymes that carry dsRBD and recognize both inter- and intramolecular dsRNAs longer than 20–30 bp (Nishikura et al. 1991). ADARs convert adenosines to inosines, which are interpreted as guanosines during translation. It was predicted that more than 85% of pre-mRNAs may be edited, predominantly in the noncoding regions (Athanasiadis et al. 2004). Many long perfect dsRNAs (>100 bp) undergo extensive editing with a conversion of approximately 50% of adenosines to inosines (Nishikura et al. 1991; Polson and Bass 1994). On the other hand, short RNAs (~20–30 bp) or imperfect long dsRNAs are edited selectively; usually only a few adenines at specific sites are deaminated (Lehmann and Bass 1999).

RNA editing can negatively influence RNAi in several ways. First, ADARs can compete with RNAi for dsRNA substrates including siRNAs. The ADAR1 isoform (ADAR1p150) strongly binds siRNA and reduces the availability of dsRNA for RNAi, resulting in less efficient RNAi in normal cells compared to *Adar1*-/- cells (Yang et al. 2005). Interestingly, injection of high doses of siRNAs enhances ADAR1 expression, suggesting a role of ADAR1 in a cellular feedback mechanism in response to siRNA (Hong et al. 2005).

A change of a single base in a sequence may result either in destabilization of dsRNA structure (inosine–uridine pair) or in its stabilization (inosine–cytidine pair) (Nishikura 2010). This transition in the local and global stability of dsRNA structure can influence further processing of dsRNA, such as the selection of the effective siRNA strand (Bartel 2004; Du and Zamore 2005; Meister and Tuschl 2004). While moderate deamination (one I–U pair per siRNA) does not prevent Dicer processing to siRNAs (Zamore et al. 2000), hyperditing (~50% of deaminated adenosines) can make dsRNA resistant to Dicer processing (Scadden and Smith 2001). Hyperedited dsRNA is also degraded by TSN (Scadden 2005).

Moreover, editing affects target recognition as a single nucleotide mismatch between siRNA and target mRNA can reduce RNAi efficacy (Scadden and Smith 2001) or modify target specificity (Kawahara et al. 2007b). This has been well documented for miRNAs. Several pri-miRNAs (e.g., miR-142) are known to undergo

editing, which inhibits Drosha cleavage or even causes degradation of pri-miRNA by TSN (Nishikura 2010; Scadden 2005; Yang et al. 2006). In other cases, pri-miRNA editing does not influence Drosha activity but inhibits processing of pre-miRNA by Dicer (e.g., miR-151) (Kawahara et al. 2007a). Last but not least, RNA editing might also inhibit export of miRNAs from the nucleus (Nishikura 2010).

An important connection between A-to-I editing and RNAi has been documented in *C. elegans*. In contrast to mice, where *Adar1* or *Adar2* deletion is lethal, in *Drosophila* and *C. elegans Adr* null phenotype causes only weak phenotypic alterations (Palladino et al. 2000; Tonkin et al. 2002). *Adr-1* or *adr-2* mutant worms exhibit a defective chemotaxis, but the phenotype is reverted when worms lacking *Adar* are crossed with RNAi-defective strains (Tonkin and Bass 2003). This indicates a necessity of a balance between RNAi and ADAR-mediated editing.

5.3.4.2 Interferon Response

Mammalian somatic cells can respond to dsRNA in a sequence-independent manner. A pioneering work by Hunter et al. showed that different types of dsRNA can block translation in reticulocyte lysates (Hunter et al. 1975). Analysis of the phenomenon identified PKR that is activated upon binding to dsRNA and blocks translation by phosphorylating the alpha subunit of eukaryotic initiation factor 2 (eIF2α) (Meurs et al. 1990). Activation of PKR represents a part of a complex response to foreign molecules known as the interferon response (reviewed in Sadler and Williams 2007), which includes an activation of the NFκB transcription factor and a large number of interferon-stimulated genes (ISGs) (Geiss et al. 2001). In addition to PKR, several other proteins recognizing dsRNA are integrated to the interferon response, including helicases RIG-I and MDA5, which sense cytoplasmic dsRNA and activate interferon expression, and the 2′,5′-oligoadenylate synthetase (OAS), which produces 2′,5′-linked oligoadenylates that induce general degradation of RNAs by activating latent RNase L (reviewed in Gantier and Williams 2007; Sadler and Williams 2007).

Interactions between RNAi and interferon response are poorly understood. There are two clear mechanistical connections between these two pathways. First, TRBP and PACT, two dsRNA-binding proteins, which were mentioned earlier as Dicer-interacting proteins, interact also with PKR. Notably, while TRBP inhibits PKR (Cosentino et al. 1995; Park et al. 1994), PACT activates it (Patel and Sen 1998). While cytoplasmic long dsRNA in somatic cells apparently triggers the interferon response, it is not clear if the same dsRNA is also routed into the RNAi pathways. Experiments in oocytes and undifferentiated mESCs (Stein et al. 2005; Yang et al. 2001) suggest that RNAi dominates in response to cytoplasmic long dsRNA in the absence of a strong interferon response and that the interferon pathway dominates when its relevant components are present. Nevertheless, this view may be too simplistic because there are several reports showing induction of RNAi with experimental intracellular expression of long dsRNA in transformed and primary somatic cells (Elbashir et al. 2001; Diallo et al. 2003; Gan et al. 2002; Shinagawa and Ishii 2003; Tran et al. 2004; Yi et al. 2003). A recent study of effects of ubiquitous long

dsRNA expression in transgenic mice shows that an expressed long dsRNA is well tolerated in mammalian somatic cells and a robust RNAi response is observed only in oocytes (Nejepinska et al. in press). Further research is needed to understand mechanisms routing long dsRNA into RNAi and interferon pathways.

The second connection between RNAi and interferon response is evolutionary. As mentioned earlier, mammalian RNA helicases *Ddx58*, *Dhx58*, and *Ifih1*, which are involved in immune response, are the closest homologues of helicases involved in the processing of long dsRNA during RNAi in *C. elegans*. Notably, DDX58, also known as RIG-I, is an established component of the interferon response to long dsRNA (Yoneyama et al. 2004). This suggests that the interferon response, which has a common trigger and evolved after the RNAi pathway, adopted several components from the latter pathway. It remains to be determined whether these and other components of RNAi lost their function in RNAi entirely or mediate a certain form of a cross talk between RNAi and the interferon response.

5.4 Closing Remarks

Most biologists likely consider RNAi as an excellent tool to study gene function. At the same time, RNAi is a complex natural phenomenon with numerous physiological functions. We have summarized the current knowledge of the molecular mechanism of RNAi and its biological roles. RNAi in animals likely originated from an ancient innate immune response allowing dealing with viral infections and parasitic sequences in the genome, which have a capacity to generate dsRNA. The original role of RNAi still persists in invertebrates. In mammals, the defensive role of RNAi was largely replaced by a more recent form of immune system and RNAi retained its traditional function in suppression of mobile elements in the female germline. While mammals do not combat viruses by RNAi anymore, they adopted RNAi for a regulation of endogenous genes in the female germline. The story of RNAi provides another example of ingenuity of nature in evolving novel functions for old mechanisms.

Acknowledgments We thank Radek Malik for help with manuscript preparation. Our research is supported by the following grants: EMBO SDIG project 1483, GACR 204/09/0085, GACR P305/10/2215, and Kontakt ME09039. P.S. is a holder of the J.E. Purkyne Fellowship.

References

Ambros V, Lee RC, Lavanway A, Williams PT, Jewell D (2003) MicroRNAs and other tiny endogenous RNAs in *C. elegans*. Curr Biol 13(10):807–818. doi:S0960982203002872 [pii]

Ameres SL, Martinez J, Schroeder R (2007) Molecular basis for target RNA recognition and cleavage by human RISC. Cell 130(1):101–112. doi:S0092–8674(07)00583–1 [pii], 10.1016/j.cell.2007.04.037

Aravin AA, Hannon GJ, Brennecke J (2007) The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science 318(5851):761–764. doi:318/5851/761 [pii], 10.1126/science.1146484

Armisen J, Gilchrist MJ, Wilczynska A, Standart N, Miska EA (2009) Abundant and dynamically expressed miRNAs, piRNAs, and other small RNAs in the vertebrate Xenopus tropicalis. Genome Res 19(10):1766–1775. doi:gr.093054.109 [pii]. 10.1101/gr.093054.109

Athanasiadis A, Rich A, Maas S (2004) Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. PLoS Biol 2(12):e391

Babiarz JE, Ruby JG, Wang Y, Bartel DP, Blelloch R (2008) Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. Genes Dev 22(20):2773–2785. doi:22/20/2773 [pii]. 10.1101/gad.1705308

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116 (2):281–297. doi:S0092867404000455 [pii]

Bennasser Y, Le SY, Benkirane M, Jeang KT (2005) Evidence that HIV-1 encodes an siRNA and a suppressor of RNA silencing. Immunity 22(5):607–619. doi:S1074–7613(05)00105–6 [pii]. 10.1016/j.immuni.2005.03.010

Bernstein E, Caudy AA, Hammond SM, Hannon GJ (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409(6818):363–366

Billy E, Brondani V, Zhang H, Muller U, Filipowicz W (2001) Specific interference with gene expression induced by long, double-stranded RNA in mouse embryonal teratocarcinoma cell lines. Proc Natl Acad Sci USA 98(25):14428–14433

Bucher G, Scholten J, Klingler M (2002) Parental RNAi in Tribolium (Coleoptera). Curr Biol 12 (3):R85–86. doi:S0960982202006668 [pii]

Carlile M, Swan D, Jackson K, Preston-Fayers K, Ballester B, Flicek P, Werner A (2009) Strand selective generation of endo-siRNAs from the Na/phosphate transporter gene Slc34a1 in murine tissues. Nucleic Acids Res 37(7):2274–2282. doi:gkp088 [pii]. 10.1093/nar/gkp088

Carmell MA, Xuan Z, Zhang MQ, Hannon GJ (2002) The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. Genes Dev 16 (21):2733–2742

Carmell MA, Girard A, van de Kant HJ, Bourc'his D, Bestor TH, de Rooij DG, Hannon GJ (2007) MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev Cell 12(4):503–514. doi:S1534–5807(07)00100–1 [pii]. 10.1016/j.devcel.2007.03.001

Caudy AA, Myers M, Hannon GJ, Hammond SM (2002) Fragile X-related protein and VIG associate with the RNA interference machinery. Genes Dev 16(19):2491–2496

Caudy AA, Ketting RF, Hammond SM, Denli AM, Bathoorn AM, Tops BB, Silva JM, Myers MM, Hannon GJ, Plasterk RH (2003) A micrococcal nuclease homologue in RNAi effector complexes. Nature 425(6956):411–414

Cenik ES, Fukunaga R, Lu G, Dutcher R, Wang Y, Hall TM, Zamore PD (2011) Phosphate and R2D2 Restrict the Substrate Specificity of Dicer-2, an ATP-Driven Ribonuclease. Mol Cell. doi:S1097–2765(11)00178-X [pii]. 10.1016/j.molcel.2011.03.002

Cerutti H, Casas-Mollano JA (2006) On the origin and functions of RNA-mediated silencing: from protists to man. Curr Genet 50(2):81–99. doi:10.1007/s00294–006–0078-x

Chendrimada TP, Gregory RI, Kumaraswamy E, Norman J, Cooch N, Nishikura K, Shiekhattar R (2005) TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature 436(7051):740–744

Chera S, de Rosa R, Miljkovic-Licina M, Dobretz K, Ghila L, Kaloulis K, Galliot B (2006) Silencing of the hydra serine protease inhibitor Kazal1 gene mimics the human SPINK1 pancreatic phenotype. J Cell Sci 119(Pt 5):846–857. doi:jcs.02807 [pii]. 10.1242/jcs.02807

Chung WJ, Okamura K, Martin R, Lai EC (2008) Endogenous RNA interference provides a somatic defense against Drosophila transposons. Curr Biol 18(11):795–802. doi:S0960–9822 (08)00603–9 [pii]. 10.1016/j.cub.2008.05.006

Cogoni C, Macino G (1999) Gene silencing in Neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase. Nature 399(6732):166–169. doi:10.1038/20215

Correa RL, Steiner FA, Berezikov E, Ketting RF (2010) MicroRNA-directed siRNA biogenesis in *Caenorhabditis elegans*. PLoS Genet 6(4):e1000903. doi:10.1371/journal.pgen.1000903

Cosentino GP, Venkatesan S, Serluca FC, Green SR, Mathews MB, Sonenberg N (1995) Double-stranded-RNA-dependent protein kinase and TAR RNA-binding protein form homo- and heterodimers in vivo. Proc Natl Acad Sci USA 92(21):9445–9449

Cullen BR (2006) Is RNA interference involved in intrinsic antiviral immunity in mammals? Nat Immunol 7(6):563–567. doi:ni1352 [pii]. 10.1038/ni1352

Czech B, Hannon GJ (2011) Small RNA sorting: matchmaking for Argonautes. Nat Rev Genet 12 (1):19–31. doi:nrg2916 [pii]. 10.1038/nrg2916

Czech B, Malone CD, Zhou R, Stark A, Schlingeheyde C, Dus M, Perrimon N, Kellis M, Wohlschlegel JA, Sachidanandam R, Hannon GJ, Brennecke J (2008) An endogenous small interfering RNA pathway in Drosophila. Nature 453(7196):798–802. doi:nature07007 [pii]. 10.1038/nature07007

Dalmay T, Horsefield R, Braunstein TH, Baulcombe DC (2001) SDE3 encodes an RNA helicase required for post-transcriptional gene silencing in Arabidopsis. EMBO J 20(8):2069–2078. doi:10.1093/emboj/20.8.2069

Daniels SM, Melendez-Pena CE, Scarborough RJ, Daher A, Christensen HS, El Far M, Purcell DF, Laine S, Gatignol A (2009) Characterization of the TRBP domain required for dicer interaction and function in RNA interference. BMC Mol Biol 10:38. doi:1471–2199–10–38 [pii]. 10.1186/1471–2199–10–38

Das PP, Bagijn MP, Goldstein LD, Woolford JR, Lehrbach NJ, Sapetschnig A, Buhecha HR, Gilchrist MJ, Howe KL, Stark R, Matthews N, Berezikov E, Ketting RF, Tavare S, Miska EA (2008) Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the Caenorhabditis elegans germline. Mol Cell 31(1):79–90. doi:S1097–2765(08)00392–4 [pii]. 10.1016/j.molcel.2008.06.003

Diallo M, Arenz C, Schmitz K, Sandhoff K, Schepers U (2003) Long endogenous dsRNAs can induce complete gene silencing in mammalian cells and primary cultures. Oligonucleotides 13 (5):381–392. doi:10.1089/154545703322617069

Didiot MC, Subramanian M, Flatter E, Mandel JL, Moine H (2009) Cells lacking the fragile X mental retardation protein (FMRP) have normal RISC activity but exhibit altered stress granule assembly. Mol Biol Cell 20(1):428–437. doi:E08–07–0737 [pii]. 10.1091/mbc.E08–07–0737

Dlakic M (2006) DUF283 domain of Dicer proteins has a double-stranded RNA-binding fold. Bioinformatics 22(22):2711–2714. doi:btl468 [pii]. 10.1093/bioinformatics/btl468

Doench JG, Petersen CP, Sharp PA (2003) siRNAs can function as miRNAs. Genes Dev 17 (4):438–442. doi:10.1101/gad.1064703

Doyle M, Jantsch MF (2002) New and old roles of the double-stranded RNA-binding domain. J Struct Biol 140(1–3):147–153. doi:S1047847702005440 [pii]

Du T, Zamore PD (2005) microPrimer: the biogenesis and function of microRNA. Development 132(21):4645–4652. doi:132/21/4645 [pii]. 10.1242/dev.02070

Du Z, Lee JK, Tjhen R, Stroud RM, James TL (2008) Structural and biochemical insights into the dicing mechanism of mouse Dicer: a conserved lysine is critical for dsRNA cleavage. Proc Natl Acad Sci USA 105(7):2391–2396. doi:0711506105 [pii]. 10.1073/pnas.0711506105

Duchaine TF, Wohlschlegel JA, Kennedy S, Bei Y, Conte D Jr, Pang K, Brownell DR, Harding S, Mitani S, Ruvkun G, 3rd Yates JR, Mello CC (2006) Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. Cell 124(2):343–354, S0092–8674(05)01394–2 [pii]. 10.1016/j.cell.2005.11.036

Duxbury MS, Ashley SW, Whang EE (2005) RNA interference: a mammalian SID-1 homologue enhances siRNA uptake and gene silencing efficacy in human cells. Biochem Biophys Res Commun 331(2):459–463. doi:S0006–291X(05)00672–8 [pii]. 10.1016/j.bbrc.2005.03.199

Dzitoyeva S, Dimitrijevic N, Manev H (2003) Gamma-aminobutyric acid B receptor 1 mediates behavior-impairing actions of alcohol in Drosophila: adult RNA interference and pharmacological evidence. Proc Natl Acad Sci USA 100(9):5485–5490. doi:10.1073/pnas.0830111100.0830111100 [pii]

Eaton BA, Fetter RD, Davis GW (2002) Dynactin is necessary for synapse stabilization. Neuron 34(5):729–741. doi:S0896627302007213 [pii]

Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature 411 (6836):494–498

Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS (2003) MicroRNA targets in Drosophila. Genome Biol 5(1):R1. doi:10.1186/gb-2003–5–1-r1. gb-2003–5–1-r1 [pii]

Faehnle CR, Joshua-Tor L (2007) Argonautes confront new small RNAs. Curr Opin Chem Biol 11 (5):569–577. doi:S1367–5931(07)00116–0 [pii]. 10.1016/j.cbpa.2007.08.032

Feinberg EH, Hunter CP (2003) Transport of dsRNA into cells by the transmembrane protein SID-1. Science 301(5639):1545–1547. doi:10.1126/science.1087117. 301/5639/1545 [pii]

Felix MA, Ashe A, Piffaretti J, Wu G, Nuez I, Belicard T, Jiang Y, Zhao G, Franz CJ, Goldstein LD, Sanroman M, Miska EA, Wang D (2011) Natural and experimental infection of Caenorhabditis nematodes by novel viruses related to nodaviruses. PLoS Biol 9(1): e1000586. doi:10.1371/journal.pbio.1000586

Fenner BJ, Thiagarajan R, Chua HK, Kwang J (2006) Betanodavirus B2 is an RNA interference antagonist that facilitates intracellular viral RNA accumulation. J Virol 80(1):85–94. doi:80/1/85 [pii]. 10.1128/JVI.80.1.85–94.2006

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 391 (6669):806–811

Forstemann K, Tomari Y, Du T, Vagin VV, Denli AM, Bratu DP, Klattenhoff C, Theurkauf WE, Zamore PD (2005) Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. PLoS Biol 3(7):e236. doi:05-PLBI-RA-0205R1 [pii]. 10.1371/journal.pbio.0030236

Galiana-Arnoux D, Dostert C, Schneemann A, Hoffmann JA, Imler JL (2006) Essential function in vivo for Dicer-2 in host defense against RNA viruses in Drosophila. Nat Immunol 7 (6):590–597. doi:ni1335 [pii]. 10.1038/ni1335

Gan L, Anton KE, Masterson BA, Vincent VA, Ye S, Gonzalez-Zulueta M (2002) Specific interference with gene expression and gene function mediated by long dsRNA in neural cells. J Neurosci Methods 121(2):151–157. doi:S0165027002002303 [pii]

Gantier MP, Williams BR (2007) The response of mammalian cells to double-stranded RNA. Cytokine Growth Factor Rev 18(5–6):363–371. doi:S1359–6101(07)00085–8 [pii]. 10.1016/j.cytogfr.2007.06.016

Geiss G, Jin G, Guo J, Bumgarner R, Katze MG, Sen GC (2001) A comprehensive view of regulation of gene expression by double-stranded RNA-mediated cell signaling. J Biol Chem 276(32):30178–30182. doi:276/32/30178 [pii]

Geldhof P, Visser A, Clark D, Saunders G, Britton C, Gilleard J, Berriman M, Knox D (2007) RNA interference in parasitic helminths: current situation, potential pitfalls and future prospects. Parasitology 134(Pt 5):609–619. doi:S0031182006002071 [pii]. 10.1017/S0031182006002071

Gent JI, Schvarzstein M, Villeneuve AM, Gu SG, Jantsch V, Fire AZ, Baudrimont A (2009) A Caenorhabditis elegans RNA-directed RNA polymerase in sperm development and endogenous RNA interference. Genetics 183(4):1297–1314. doi:genetics.109.109686 [pii]. 10.1534/genetics.109.109686

Gent JI, Lamm AT, Pavelec DM, Maniar JM, Parameswaran P, Tao L, Kennedy S, Fire AZ (2010) Distinct phases of siRNA synthesis in an endogenous RNAi pathway in C. elegans soma. Mol Cell 37(5):679–689. doi:S1097-–2765(10)00041-–9 [pii]. 10.1016/j.molcel.2010.01.012

Gerbasi VR, Preall JB, Golden DE, Powell DW, Cummins TD, Sontheimer EJ (2011) Blanks, a nuclear siRNA/dsRNA-binding complex component, is required for Drosophila spermiogenesis. Proc Natl Acad Sci USA 108(8):3204–3209. doi:1009781108 [pii]. 10.1073/pnas.1009781108

Ghildiyal M, Seitz H, Horwich MD, Li C, Du T, Lee S, Xu J, Kittler EL, Zapp ML, Weng Z, Zamore PD (2008) Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. Science 320(5879):1077–1081. doi:1157396 [pii]. 10.1126/science.1157396

Ghildiyal M, Xu J, Seitz H, Weng Z, Zamore PD (2010) Sorting of Drosophila small silencing RNAs partitions microRNA* strands into the RNA interference pathway. RNA 16(1):43–56. doi:rna.1972910 [pii]. 10.1261/rna.1972910

Glanzer J, Miyashiro KY, Sul JY, Barrett L, Belt B, Haydon P, Eberwine J (2005) RNA splicing capability of live neuronal dendrites. Proc Natl Acad Sci USA 102(46):16859–16864. doi:0503783102 [pii]. 10.1073/pnas.0503783102

Gredell JA, Dittmer MJ, Wu M, Chan C, Walton SP (2010) Recognition of siRNA asymmetry by TAR RNA binding protein. Biochemistry 49(14):3148–3155. doi:10.1021/bi902189s

Gregory RI, Chendrimada TP, Cooch N, Shiekhattar R (2005) Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. Cell 123(4):631–640. doi:S0092–8674(05)01109–8 [pii]. 10.1016/j.cell.2005.10.022

Grishok A, Pasquinelli AE, Conte D, Li N, Parrish S, Ha I, Baillie DL, Fire A, Ruvkun G, Mello CC (2001) Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. Cell 106(1):23–34

Guang S, Bochner AF, Burkhart KB, Burton N, Pavelec DM, Kennedy S (2010) Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. Nature 465 (7301):1097–1101. doi:nature09095 [pii]. 10.1038/nature09095

Haase AD, Jaskiewicz L, Zhang H, Laine S, Sack R, Gatignol A, Filipowicz W (2005) TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. EMBO Rep 6(10):961–967

Haasnoot J, de Vries W, Geutjes EJ, Prins M, de Haan P, Berkhout B (2007) The Ebola virus VP35 protein is a suppressor of RNA silencing. PLoS Pathog 3(6):e86. doi:06-PLPA-RA-0347 [pii]. 10.1371/journal.ppat.0030086

Haley B, Zamore PD (2004) Kinetic analysis of the RNAi enzyme complex. Nat Struct Mol Biol 11(7):599–606. doi:10.1038/nsmb780. nsmb780 [pii]

Hall IM, Shankaranarayana GD, Noma K, Ayoub N, Cohen A, Grewal SI (2002) Establishment and maintenance of a heterochromatin domain. Science 297(5590):2232–2237

Hamilton AJ, Baulcombe DC (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. Science 286(5441):950–952. doi:7953 [pii]

Han W, Sundaram P, Kenjale H, Grantham J, Timmons L (2008) The Caenorhabditis elegans rsd-2 and rsd-6 genes are required for chromosome functions during exposure to unfavorable environments. Genetics 178(4):1875–1893. doi:178/4/1875 [pii]. 10.1534/genetics.107.085472

Han T, Manoharan AP, Harkins TT, Bouffard P, Fitzpatrick C, Chu DS, Thierry-Mieg D, Thierry-Mieg J, Kim JK (2009) 26G endo-siRNAs regulate spermatogenic and zygotic gene expression in Caenorhabditis elegans. Proc Natl Acad Sci USA 106(44):18674–18679. doi:0906378106 [pii]. 10.1073/pnas.0906378106

Hartig JV, Forstemann K (2011) Loqs-PD and R2D2 define independent pathways for RISC generation in Drosophila. Nucleic Acids Res. doi:gkq1324 [pii]. 10.1093/nar/gkq1324

Hartig JV, Esslinger S, Bottcher R, Saito K, Forstemann K (2009) Endo-siRNAs depend on a new isoform of loquacious and target artificially introduced, high-copy sequences. EMBO J 28 (19):2932–2944. doi:emboj2009220 [pii]. 10.1038/emboj.2009.220

Hong J, Qian Z, Shen S, Min T, Tan C, Xu J, Zhao Y, Huang W (2005) High doses of siRNAs induce eri-1 and adar-1 gene expression and reduce the efficiency of RNA interference in the mouse. Biochem J 390(Pt 3):675–679. doi:BJ20050647 [pii]. 10.1042/BJ20050647

Hunter T, Hunt T, Jackson RJ, Robertson HD (1975) The characteristics of inhibition of protein synthesis by double-stranded ribonucleic acid in reticulocyte lysates. J Biol Chem 250 (2):409–417

Huvenne H, Smagghe G (2010) Mechanisms of dsRNA uptake in insects and potential of RNAi for pest control: a review. J Insect Physiol 56(3):227–235. doi:S0022–1910(09)00342–4 [pii]. 10.1016/j.jinsphys.2009.10.004

Ishizuka A, Siomi MC, Siomi H (2002) A Drosophila fragile X protein interacts with components of RNAi and ribosomal proteins. Genes Dev 16(19):2497–2508

Jaskiewicz L, Filipowicz W (2008) Role of Dicer in posttranscriptional RNA silencing. Curr Top Microbiol Immunol 320:77–97

Jinek M, Doudna JA (2009) A three-dimensional view of the molecular machinery of RNA interference. Nature 457(7228):405–412. doi:nature07755 [pii]. 10.1038/nature07755

Johnson KL, Price BD, Eckerle LD, Ball LA (2004) Nodamura virus nonstructural protein B2 can enhance viral RNA accumulation in both mammalian and insect cells. J Virol 78 (12):6698–6704. doi:10.1128/JVI.78.12.6698–6704.2004. 78/12/6698 [pii]

Jopling CL, Yi M, Lancaster AM, Lemon SM, Sarnow P (2005) Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. Science 309(5740):1577–1581. doi:309/5740/1577 [pii]. 10.1126/science.1113329

Jose AM, Hunter CP (2007) Transport of sequence-specific RNA interference information between cells. Annu Rev Genet 41:305–330. doi:10.1146/annurev.genet.41.110306.130216

Kaneda M, Tang F, O'Carroll D, Lao K, Surani MA (2009) Essential role for Argonaute2 protein in mouse oogenesis. Epigenetics Chromatin 2(1):9. doi:1756–8935–2–9 [pii]. 10.1186/1756–8935–2–9

Kanellopoulou C, Muljo SA, Kung AL, Ganesan S, Drapkin R, Jenuwein T, Livingston DM, Rajewsky K (2005) Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev 19(4):489–501

Kawahara Y, Zinshteyn B, Chendrimada TP, Shiekhattar R, Nishikura K (2007a) RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer–TRBP complex. EMBO Rep 8 (8):763–769. doi:7401011 [pii]. 10.1038/sj.embor.7401011

Kawahara Y, Zinshteyn B, Sethupathy P, Iizasa H, Hatzigeorgiou AG, Nishikura K (2007b) Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. Science 315 (5815):1137–1140. doi:315/5815/1137 [pii]. 10.1126/science.1138050

Kawaji H, Nakamura M, Takahashi Y, Sandelin A, Katayama S, Fukuda S, Daub CO, Kai C, Kawai J, Yasuda J, Carninci P, Hayashizaki Y (2008) Hidden layers of human small RNAs. BMC Genomics 9:157. doi:1471–2164–9–157 [pii]. 10.1186/1471–2164–9–157

Kawamata T, Seitz H, Tomari Y (2009) Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. Nat Struct Mol Biol 16(9):953–960. doi:nsmb.1630 [pii]. 10.1038/nsmb.1630

Kawamura Y, Saito K, Kin T, Ono Y, Asai K, Sunohara T, Okada TN, Siomi MC, Siomi H (2008) Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. Nature 453 (7196):793–797. doi:nature06938 [pii]. 10.1038/nature06938

Keene KM, Foy BD, Sanchez-Vargas I, Beaty BJ, Blair CD, Olson KE (2004) RNA interference acts as a natural antiviral response to O'nyong-nyong virus (Alphavirus; Togaviridae) infection of Anopheles gambiae. Proc Natl Acad Sci USA 101(49):17240–17245. doi:0406983101 [pii]. 10.1073/pnas.0406983101

Kennedy S, Wang D, Ruvkun G (2004) A conserved siRNA-degrading RNase negatively regulates RNA interference in *C. elegans*. Nature 427(6975):645–649. doi:10.1038/nature02302. nature02302 [pii]

Kennerdell JR, Carthew RW (1998) Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in the wingless pathway. Cell 95(7):1017–1026. doi:S0092–8674(00)81725–0 [pii]

Ketting RF (2011) The many faces of RNAi. Dev Cell 20(2):148–161. doi:S1534–5807(11) 00040–2 [pii]. 10.1016/j.devcel.2011.01.012

Ketting RF, Haverkamp TH, van Luenen HG, Plasterk RH (1999) Mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. Cell 99(2):133–141

Ketting RF, Fischer SE, Bernstein E, Sijen T, Hannon GJ, Plasterk RH (2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. Genes Dev 15(20):2654–2659. doi:10.1101/gad.927801

Kim VN (2005) MicroRNA biogenesis: coordinated cropping and dicing. Nat Rev Mol Cell Biol 6 (5):376–385

Kim K, Lee YS, Carthew RW (2007) Conversion of pre-RISC to holo-RISC by Ago2 during assembly of RNAi complexes. RNA 13(1):22–29. doi:rna.283207 [pii]. 10.1261/rna.283207

Kok KH, Ng MH, Ching YP, Jin DY (2007) Human TRBP and PACT directly interact with each other and associate with dicer to facilitate the production of small interfering RNA. J Biol Chem 282(24):17649–17657. doi:M611768200 [pii]. 10.1074/jbc.M611768200

Lecellier CH, Dunoyer P, Arar K, Lehmann-Che J, Eyquem S, Himber C, Saib A, Voinnet O (2005) A cellular microRNA mediates antiviral defense in human cells. Science 308 (5721):557–560. doi:308/5721/557 [pii]. 10.1126/science.1108784

Lee YS, Nakahara K, Pham JW, Kim K, He Z, Sontheimer EJ, Carthew RW (2004) Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell 117 (1):69–81

Lee RC, Hammell CM, Ambros V (2006) Interacting endogenous and exogenous RNAi pathways in Caenorhabditis elegans. RNA 12(4):589–597. doi:rna.2231506 [pii]. 10.1261/rna.2231506

Lehmann KA, Bass BL (1999) The importance of internal loops within RNA substrates of ADAR1. J Mol Biol 291(1):1–13. doi:10.1006/jmbi.1999.2914. S0022–2836(99)92914–5 [pii]

Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB (2003) Prediction of mammalian microRNA targets. Cell 115(7):787–798

Li H, Li WX, Ding SW (2002) Induction and suppression of RNA silencing by an animal virus. Science 296(5571):1319–1321. doi:10.1126/science.1070948. 296/5571/1319 [pii]

Li WX, Li H, Lu R, Li F, Dus M, Atkinson P, Brydon EW, Johnson KL, Garcia-Sastre A, Ball LA, Palese P, Ding SW (2004) Interferon antagonist proteins of influenza and vaccinia viruses are suppressors of RNA silencing. Proc Natl Acad Sci USA 101(5):1350–1355. doi:10.1073/pnas.0308308100. 0308308100 [pii]

Li C, Vagin VV, Lee S, Xu J, Ma S, Xi H, Seitz H, Horwich MD, Syrzycka M, Honda BM, Kittler EL, Zapp ML, Klattenhoff C, Schulz N, Theurkauf WE, Weng Z, Zamore PD (2009) Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. Cell 137 (3):509–521. doi:S0092–8674(09)00452–8 [pii]. 10.1016/j.cell.2009.04.027

Lingel A, Simon B, Izaurralde E, Sattler M (2003) Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. Nature 426(6965):465–469. doi:10.1038/nature02123. nature02123 [pii]

Lingel A, Simon B, Izaurralde E, Sattler M (2004) Nucleic acid 3′-end recognition by the Argonaute2 PAZ domain. Nat Struct Mol Biol 11(6):576–577. doi:10.1038/nsmb777. nsmb777 [pii]

Lipardi C, Paterson BM (2009) Identification of an RNA-dependent RNA polymerase in Drosophila involved in RNAi and transposon suppression. Proc Natl Acad Sci USA 106 (37):15645–15650. doi:0904984106 [pii]. 10.1073/pnas.0904984106

Lipardi C, Wei Q, Paterson BM (2001) RNAi as random degradative PCR: siRNA primers convert mRNA into dsRNAs that are degraded to generate new siRNAs. Cell 107(3):297–307. doi: S0092–8674(01)00537–2 [pii]

Liu Q, Rand TA, Kalidas S, Du F, Kim HE, Smith DP, Wang X (2003) R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. Science 301(5641):1921–1925. doi:10.1126/science.1088710. 301/5641/1921 [pii]

Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song JJ, Hammond SM, Joshua-Tor L, Hannon GJ (2004) Argonaute2 is the catalytic engine of mammalian RNAi. Science 305 (5689):1437–1441

Liu X, Jiang F, Kalidas S, Smith D, Liu Q (2006) Dicer-2 and R2D2 coordinately bind siRNA to promote assembly of the siRISC complexes. RNA 12(8):1514–1520. doi:rna.101606 [pii]. 10.1261/rna.101606

Liu Y, Ye X, Jiang F, Liang C, Chen D, Peng J, Kinch LN, Grishin NV, Liu Q (2009) C3PO, an endoribonuclease that promotes RNAi by facilitating RISC activation. Science 325 (5941):750–753. doi:325/5941/750 [pii]. 10.1126/science.1176325

Lohmann JU, Endl I, Bosch TC (1999) Silencing of developmental genes in Hydra. Dev Biol 214 (1):211–214. doi:10.1006/dbio.1999.9407. S0012–1606(99)99407–1 [pii]

Lu S, Cullen BR (2004) Adenovirus VA1 noncoding RNA can inhibit small interfering RNA and MicroRNA biogenesis. J Virol 78(23):12868–12876. doi:78/23/12868 [pii]. 10.1128/JVI.78.23.12868–12876.2004

Lu R, Maduro M, Li F, Li HW, Broitman-Maduro G, Li WX, Ding SW (2005) Animal virus replication and RNAi-mediated antiviral silencing in *Caenorhabditis elegans*. Nature 436 (7053):1040–1043

Lucchetta EM, Carthew RW, Ismagilov RF (2009) The endo-siRNA pathway is essential for robust development of the Drosophila embryo. PLoS One 4(10):e7576. doi:10.1371/journal.pone.0007576

Ma JB, Ye K, Patel DJ (2004) Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. Nature 429(6989):318–322. doi:10.1038/nature02519. nature02519 [pii]

Ma JB, Yuan YR, Meister G, Pei Y, Tuschl T, Patel DJ (2005) Structural basis for 5′-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. Nature 434(7033):666–670. doi:nature03514 [pii]. 10.1038/nature03514

Ma E, MacRae IJ, Kirsch JF, Doudna JA (2008) Autoinhibition of human dicer by its internal helicase domain. J Mol Biol 380(1):237–243. doi:S0022–2836(08)00547–0 [pii]. 10.1016/j.jmb.2008.05.005

Ma J, Flemr M, Stein P, Berninger P, Malik R, Zavolan M, Svoboda P, Schultz RM (2010) MicroRNA activity is suppressed in mouse oocytes. Curr Biol 20(3):265–270. doi:S0960–9822(09)02205–2 [pii]. 10.1016/j.cub.2009.12.042

Macrae IJ, Zhou K, Li F, Repic A, Brooks AN, Cande WZ, Adams PD, Doudna JA (2006) Structural basis for double-stranded RNA processing by Dicer. Science 311(5758):195–198. doi:311/5758/195 [pii]. 10.1126/science.1121638

MacRae IJ, Zhou K, Doudna JA (2007) Structural determinants of RNA recognition and cleavage by Dicer. Nat Struct Mol Biol 14(10):934–940. doi:nsmb1293 [pii]. 10.1038/nsmb1293

MacRae IJ, Ma E, Zhou M, Robinson CV, Doudna JA (2008) In vitro reconstitution of the human RISC-loading complex. Proc Natl Acad Sci USA 105(2):512–517. doi:0710869105 [pii]. 10.1073/pnas.0710869105

Maida Y, Yasukawa M, Furuuchi M, Lassmann T, Possemato R, Okamoto N, Kasim V, Hayashizaki Y, Hahn WC, Masutomi K (2009) An RNA-dependent RNA polymerase formed by TERT and the RMRP RNA. Nature 461(7261):230–235. doi:nature08283 [pii]. 10.1038/nature08283

Malone CD, Hannon GJ (2009a) Molecular evolution of piRNA and transposon control pathways in Drosophila. Cold Spring Harb Symp Quant Biol 74:225–234. doi:sqb.2009.74.052 [pii]. 10.1101/sqb.2009.74.052

Malone CD, Hannon GJ (2009b) Small RNAs as guardians of the genome. Cell 136(4):656–668. doi:S0092–8674(09)00127–5 [pii]. 10.1016/j.cell.2009.01.045

Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, Hannon GJ (2009) Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell 137(3):522–535. doi:S0092–8674(09)00377–8 [pii]. 10.1016/j.cell.2009.03.040

Maniataki E, Mourelatos Z (2005) A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. Genes Dev 19(24):2979–2990. doi:19/24/2979 [pii]. 10.1101/gad.1384005

Marques JT, Carthew RW (2007) A call to arms: coevolution of animal viruses and host innate immune responses. Trends Genet 23(7):359–364. doi:S0168–9525(07)00152–7 [pii]. 10.1016/j.tig.2007.04.004

Marques JT, Devosse T, Wang D, Zamanian-Daryoush M, Serbinowski P, Hartmann R, Fujita T, Behlke MA, Williams BR (2006) A structural basis for discriminating between self and nonself double-stranded RNAs in mammalian cells. Nat Biotechnol 24(5):559–565. doi:nbt1205 [pii]. 10.1038/nbt1205

Marques JT, Kim K, Wu PH, Alleyne TM, Jafari N, Carthew RW (2010) Loqs and R2D2 act sequentially in the siRNA pathway in Drosophila. Nat Struct Mol Biol 17(1):24–30. doi:nsmb.1735 [pii]. 10.1038/nsmb.1735

Martinez J, Patkaniowska A, Urlaub H, Luhrmann R, Tuschl T (2002) Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell 110(5):563–574. doi:S009286740200908X [pii]

Matranga C, Tomari Y, Shin C, Bartel DP, Zamore PD (2005) Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. Cell 123 (4):607–620. doi:S0092–8674(05)00922–0 [pii]. 10.1016/j.cell.2005.08.044

Meins F Jr, Si-Ammour A, Blevins T (2005) RNA silencing systems and their relevance to plant development. Annu Rev Cell Dev Biol 21:297–318. doi:10.1146/annurev.cellbio.21.122303.114706

Meister G, Tuschl T (2004) Mechanisms of gene silencing by double-stranded RNA. Nature 431 (7006):343–349. doi:10.1038/nature02873. nature02873 [pii]

Meister G, Landthaler M, Patkaniowska A, Dorsett Y, Teng G, Tuschl T (2004) Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Mol Cell 15 (2):185–197

Meister G, Landthaler M, Peters L, Chen PY, Urlaub H, Luhrmann R, Tuschl T (2005) Identification of novel argonaute-associated proteins. Curr Biol 15(23):2149–2155. doi:S0960–9822(05)01301–1 [pii]. 10.1016/j.cub.2005.10.048

Meurs E, Chong K, Galabru J, Thomas NS, Kerr IM, Williams BR, Hovanessian AG (1990) Molecular cloning and characterization of the human double-stranded RNA-activated protein kinase induced by interferon. Cell 62(2):379–390. doi:0092–8674(90)90374-N [pii]

Miyoshi K, Tsukumo H, Nagami T, Siomi H, Siomi MC (2005) Slicer function of Drosophila Argonautes and its involvement in RISC formation. Genes Dev 19(23):2837–2848. doi:gad.1370605 [pii]. 10.1101/gad.1370605

Mourelatos Z, Dostie J, Paushkin S, Sharma A, Charroux B, Abel L, Rappsilber J, Mann M, Dreyfuss G (2002) miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. Genes Dev 16(6):720–728

Mourrain P, Beclin C, Elmayan T, Feuerbach F, Godon C, Morel JB, Jouette D, Lacombe AM, Nikic S, Picault N, Remoue K, Sanial M, Vo TA, Vaucheret H (2000) Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. Cell 101(5):533–542. doi:S0092–8674(00)80863–6 [pii]

Murchison EP, Partridge JF, Tam OH, Cheloufi S, Hannon GJ (2005) Characterization of Dicer-deficient murine embryonic stem cells. Proc Natl Acad Sci USA 102(34):12135–12140. doi:0505479102 [pii]. 10.1073/pnas.0505479102

Murchison EP, Stein P, Xuan Z, Pan H, Zhang MQ, Schultz RM, Hannon GJ (2007) Critical roles for Dicer in the female germline. Genes Dev 21(6):682–693. doi:21/6/682 [pii]. 10.1101/gad.1521307

Murphy D, Dancis B, Brown JR (2008) The evolution of core proteins involved in microRNA biogenesis. BMC Evol Biol 8:92. doi:1471–2148–8–92 [pii]. 10.1186/1471–2148–8–92

Nayak A, Berry B, Tassetto M, Kunitomi M, Acevedo A, Deng C, Krutchinsky A, Gross J, Antoniewski C, Andino R (2010) Cricket paralysis virus antagonizes Argonaute 2 to modulate antiviral defense in Drosophila. Nat Struct Mol Biol 17(5):547–554. doi:nsmb.1810 [pii]. 10.1038/nsmb.1810

Nejepinska J, Malik R, Filkowski J, Flemr M, Filipowicz W, Svoboda P. dsRNA expression in the mouse elicits RNAi in oocytes and low adenosine deamination in somatic cells. Nucleic Acids Res, published online September 8, 2011 doi:10.1093/nar/gkr702

Newmark PA, Reddien PW, Cebria F, Sanchez Alvarado A (2003) Ingestion of bacterially expressed double-stranded RNA inhibits gene expression in planarians. Proc Natl Acad Sci USA 100(Suppl 1):11861–11865. doi:10.1073/pnas.1834205100. 1834205100 [pii]

Ng KK, Arnold JJ, Cameron CE (2008) Structure–function relationships among RNA-dependent RNA polymerases. Curr Top Microbiol Immunol 320:137–156

Nishikura K (2010) Functions and regulation of RNA editing by ADAR deaminases. Annu Rev Biochem 79:321–349. doi:10.1146/annurev-biochem-060208–105251

Nishikura K, Yoo C, Kim U, Murray JM, Estes PA, Cash FE, Liebhaber SA (1991) Substrate specificity of the dsRNA unwinding/modifying activity. EMBO J 10(11):3523–3532

Nykanen A, Haley B, Zamore PD (2001) ATP requirements and small interfering RNA structure in the RNA interference pathway. Cell 107(3):309–321. doi:S0092–8674(01)00547–5 [pii]

Ohnishi Y, Totoki Y, Toyoda A, Watanabe T, Yamamoto Y, Tokunaga K, Sakaki Y, Sasaki H, Hohjoh H (2010) Small RNA class transition from siRNA/piRNA to miRNA during pre-implantation mouse development. Nucleic Acids Res 38(15):5141–5151. doi:gkq229 [pii]. 10.1093/nar/gkq229

Okamura K, Ishizuka A, Siomi H, Siomi MC (2004) Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. Genes Dev 18(14):1655–1666. doi:10.1101/gad.1210204. 1210204 [pii]

Okamura K, Chung WJ, Ruby JG, Guo H, Bartel DP, Lai EC (2008) The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. Nature 453(7196):803–806. doi:nature07015 [pii]. 10.1038/nature07015

Okamura K, Robine N, Liu Y, Liu Q, Lai EC (2011) R2D2 organizes small regulatory RNA pathways in Drosophila. Mol Cell Biol 31(4):884–896. doi:MCB.01141–10 [pii]. 10.1128/MCB.01141–10

Orii H, Mochii M, Watanabe K (2003) A simple "soaking method" for RNA interference in the planarian Dugesia japonica. Dev Genes Evol 213(3):138–141. doi:10.1007/s00427–003–0310–3

Pak J, Fire A (2007) Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. Science 315(5809):241–244. doi:1132839 [pii]. 10.1126/science.1132839

Palladino MJ, Keegan LP, O'Connell MA, Reenan RA (2000) A-to-I pre-mRNA editing in Drosophila is primarily involved in adult nervous system function and integrity. Cell 102 (4):437–449

Park H, Davies MV, Langland JO, Chang HW, Nam YS, Tartaglia J, Paoletti E, Jacobs BL, Kaufman RJ, Venkatesan S (1994) TAR RNA-binding protein is an inhibitor of the interferon-induced protein kinase PKR. Proc Natl Acad Sci USA 91(11):4713–4717

Parker JS, Roe SM, Barford D (2004) Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. EMBO J 23(24):4727–4737. doi:7600488 [pii]. 10.1038/sj.emboj.7600488

Parker GS, Eckert DM, Bass BL (2006) RDE-4 preferentially binds long dsRNA and its dimerization is necessary for cleavage of dsRNA to siRNA. RNA 12(5):807–818. doi:rna.2338706 [pii]. 10.1261/rna.2338706

Parker GS, Maity TS, Bass BL (2008) dsRNA binding properties of RDE-4 and TRBP reflect their distinct roles in RNAi. J Mol Biol 384(4):967–979. doi:S0022–2836(08)01261–8 [pii]. 10.1016/j.jmb.2008.10.002

Parrish S, Fire A (2001) Distinct roles for RDE-1 and RDE-4 during RNA interference in *Caenorhabditis elegans*. RNA 7(10):1397–1402

Patel RC, Sen GC (1998) PACT, a protein activator of the interferon-induced protein kinase, PKR. EMBO J 17(15):4379–4390. doi:10.1093/emboj/17.15.4379

Pavelec DM, Lachowiec J, Duchaine TF, Smith HE, Kennedy S (2009) Requirement for the ERI/DICER complex in endogenous RNA interference and sperm development in *Caenorhabditis elegans*. Genetics 183(4):1283–1295. doi:genetics.109.108134 [pii]. 10.1534/genetics.109.108134

Pelczar H, Woisard A, Lemaitre JM, Chachou M, Andeol Y (2010) Evidence for an RNA polymerization activity in axolotl and xenopus egg extracts. PLoS One 5(12):e14411. doi:10.1371/journal.pone.0014411

Peters L, Meister G (2007) Argonaute proteins: mediators of RNA silencing. Mol Cell 26 (5):611–623. doi:S1097–2765(07)00257–2 [pii]. 10.1016/j.molcel.2007.05.001

Pfeffer S, Zavolan M, Grasser FA, Chien M, Russo JJ, Ju J, John B, Enright AJ, Marks D, Sander C, Tuschl T (2004) Identification of virus-encoded microRNAs. Science 304(5671):734–736. doi:10.1126/science.1096781. 304/5671/734 [pii]

Pfeffer S, Sewer A, Lagos-Quintana M, Sheridan R, Sander C, Grasser FA, van Dyk LF, Ho CK, Shuman S, Chien M, Russo JJ, Ju J, Randall G, Lindenbach BD, Rice CM, Simon V, Ho DD, Zavolan M, Tuschl T (2005) Identification of microRNAs of the herpesvirus family. Nat Methods 2(4):269–276. doi:nmeth746 [pii]. 10.1038/nmeth746

Pham JW, Pellino JL, Lee YS, Carthew RW, Sontheimer EJ (2004) A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in Drosophila. Cell 117(1):83–94. doi:S0092867404002582 [pii]

Polson AG, Bass BL (1994) Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase. EMBO J 13(23):5701–5711

Provost P, Dishart D, Doucet J, Frendewey D, Samuelsson B, Radmark O (2002) Ribonuclease activity and RNA binding of recombinant human Dicer. EMBO J 21(21):5864–5874

Qin H, Chen F, Huan X, Machida S, Song J, Yuan YA (2010) Structure of the Arabidopsis thaliana DCL4 DUF283 domain reveals a noncanonical double-stranded RNA-binding fold for protein–protein interaction. RNA 16(3):474–481. doi:rna.1965310 [pii]. 10.1261/rna.1965310

Roignant JY, Carre C, Mugat B, Szymczak D, Lepesant JA, Antoniewski C (2003) Absence of transitive and systemic pathways allows cell-specific and isoform-specific RNAi in Drosophila. RNA 9(3):299–308

Ruby JG, Jan C, Player C, Axtell MJ, Lee W, Nusbaum C, Ge H, Bartel DP (2006) Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. Cell 127(6):1193–1207. doi:S0092—8674(06)01468—1 [pii]. 10.1016/j.cell.2006.10.040

Sadler AJ, Williams BR (2007) Structure and function of the protein kinase R. Curr Top Microbiol Immunol 316:253–292

Saito K, Ishizuka A, Siomi H, Siomi MC (2005) Processing of pre-microRNAs by the Dicer-1-Loquacious complex in Drosophila cells. PLoS Biol 3(7):e235. doi:05-PLBI-RA-0198R1 [pii]. 10.1371/journal.pbio.0030235

Saleh MC, Tassetto M, van Rij RP, Goic B, Gausson V, Berry B, Jacquier C, Antoniewski C, Andino R (2009) Antiviral immunity in Drosophila requires systemic RNA interference spread. Nature 458(7236):346–350. doi:nature07712 [pii]. 10.1038/nature07712

Sam M, Wurst W, Kluppel M, Jin O, Heng H, Bernstein A (1998) Aquarius, a novel gene isolated by gene trapping with an RNA-dependent RNA polymerase motif. Dev Dyn 212(2):304–317. doi:10.1002/(SICI)1097–0177(199806)212:2<304::AID-AJA15>3.0.CO;2–3 [pii]. 10.1002/(SICI)1097–0177(199806)212:2<304::AID-AJA15>3.0.CO;2–3

Sanchez Alvarado A, Newmark PA (1999) Double-stranded RNA specifically disrupts gene expression during planarian regeneration. Proc Natl Acad Sci USA 96(9):5049–5054

Sanchez-Vargas I, Scott JC, Poole-Smith BK, Franz AW, Barbosa-Solomieu V, Wilusz J, Olson KE, Blair CD (2009) Dengue virus type 2 infections of aedes aegypti are modulated by the mosquito's RNA interference pathway. PLoS Pathog 5(2):e1000299. doi:10.1371/journal.ppat.1000299

Scadden AD (2005) The RISC subunit Tudor-SN binds to hyper-edited double-stranded RNA and promotes its cleavage. Nat Struct Mol Biol 12(6):489–496

Scadden AD, Smith CW (2001) RNAi is antagonized by A→I hyper-editing. EMBO Rep 2 (12):1107–1111. doi:10.1093/embo-reports/kve244. kve244 [pii]

Schott DH, Cureton DK, Whelan SP, Hunter CP (2005) An antiviral role for the RNA interference machinery in *Caenorhabditis elegans*. Proc Natl Acad Sci USA 102(51):18420–18424. doi:0507123102 [pii]. 10.1073/pnas.0507123102

Schwarz DS, Hutvagner G, Haley B, Zamore PD (2002) Evidence that siRNAs function as guides, not primers, in the Drosophila and human RNAi pathways. Mol Cell 10(3):537–548

Schwarz DS, Tomari Y, Zamore PD (2004) The RNA-induced silencing complex is a Mg2+-dependent endonuclease. Curr Biol 14(9):787–791. doi:10.1016/j.cub.2004.03.008. S0960982204001769 [pii]

Shih JD, Fitzgerald MC, Sutherlin M, Hunter CP (2009) The SID-1 double-stranded RNA transporter is not selective for dsRNA length. RNA 15(3):384–390. doi:rna.1286409 [pii]. 10.1261/rna.1286409

Shinagawa T, Ishii S (2003) Generation of Ski-knockdown mice by expressing a long double-strand RNA from an RNA polymerase II promoter. Genes Dev 17(11):1340–1345. doi:10.1101/gad.1073003. 17/11/1340 [pii]

Sijen T, Plasterk RH (2003) Transposon silencing in the *Caenorhabditis elegans* germ line by natural RNAi. Nature 426(6964):310–314. doi:10.1038/nature02107. nature02107 [pii]

Sijen T, Fleenor J, Simmer F, Thijssen KL, Parrish S, Timmons L, Plasterk RH, Fire A (2001) On the role of RNA amplification in dsRNA-triggered gene silencing. Cell 107(4):465–476

Sijen T, Steiner FA, Thijssen KL, Plasterk RH (2007) Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. Science 315(5809):244–247. doi:1136699 [pii]. 10.1126/science.1136699

Sinkkonen L, Hugenschmidt T, Filipowicz W, Svoboda P (2010) Dicer is associated with ribosomal DNA chromatin in mammalian cells. PLoS One 5(8):e12175. doi:10.1371/journal.pone.0012175

Smalheiser NR, Lugli G, Thimmapuram J, Cook EH, Larson J (2011) Endogenous siRNAs and noncoding RNA-derived small RNAs are expressed in adult mouse hippocampus and are up-regulated in olfactory discrimination training. RNA 17(1):166–181. doi:rna.2123811 [pii]. 10.1261/rna.2123811

Smardon A, Spoerke JM, Stacey SC, Klein ME, Mackin N, Maine EM (2000) EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. Curr Biol 10(4):169–178, S0960–9822(00)00323–7 [pii]

Soifer HS, Sano M, Sakurai K, Chomchan P, Saetrom P, Sherman MA, Collingwood MA, Behlke MA, Rossi JJ (2008) A role for the Dicer helicase domain in the processing of thermodynamically unstable hairpin RNAs. Nucleic Acids Res 36(20):6511–6522. doi:gkn687 [pii]. 10.1093/nar/gkn687

Song JJ, Liu J, Tolia NH, Schneiderman J, Smith SK, Martienssen RA, Hannon GJ, Joshua-Tor L (2003) The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. Nat Struct Biol 10(12):1026–1032. doi:10.1038/nsb1016. nsb1016 [pii]

Song JJ, Smith SK, Hannon GJ, Joshua-Tor L (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. Science 305(5689):1434–1437

Sontheimer EJ (2005) Assembly and function of RNA silencing complexes. Nat Rev Mol Cell Biol 6(2):127–138

Stein P, Svoboda P, Anger M, Schultz RM (2003) RNAi: mammalian oocytes do it without RNA-dependent RNA polymerase. RNA 9(2):187–192

Stein P, Zeng F, Pan H, Schultz RM (2005) Absence of non-specific effects of RNA interference triggered by long double-stranded RNA in mouse oocytes. Dev Biol 286(2):464–471

Suh N, Baehner L, Moltzahn F, Melton C, Shenoy A, Chen J, Blelloch R (2010) MicroRNA function is globally suppressed in mouse oocytes and early embryos. Curr Biol 20(3):271–277. doi:S0960–9822(09)02207–6 [pii]. 10.1016/j.cub.2009.12.044

Sullivan CS, Grundhoff AT, Tevethia S, Pipas JM, Ganem D (2005) SV40-encoded microRNAs regulate viral gene expression and reduce susceptibility to cytotoxic T cells. Nature 435(7042):682–686. doi:nature03576 [pii]. 10.1038/nature03576

Svoboda P, Cara AD (2006) Hairpin RNA: a secondary structure of primary importance. Cell Mol Life Sci 63(7–8):901–908

Svoboda P, Stein P, Hayashi H, Schultz RM (2000) Selective reduction of dormant maternal mRNAs in mouse oocytes by RNA interference. Development 127(19):4147–4156

Svoboda P, Stein P, Anger M, Bernstein E, Hannon GJ, Schultz RM (2004) RNAi and expression of retrotransposons MuERV-L and IAP in preimplantation mouse embryos. Dev Biol 269(1):276–285

Tabara H, Grishok A, Mello CC (1998) RNAi in *C. elegans*: soaking in the genome sequence. Science 282(5388):430–431

Tabara H, Sarkissian M, Kelly WG, Fleenor J, Grishok A, Timmons L, Fire A, Mello CC (1999) The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. Cell 99 (2):123–132

Tabara H, Yigit E, Siomi H, Mello CC (2002) The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in *C. elegans*. Cell 109(7):861–871

Tam OH, Aravin AA, Stein P, Girard A, Murchison EP, Cheloufi S, Hodges E, Anger M, Sachidanandam R, Schultz RM, Hannon GJ (2008) Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature 453(7194):534–538. doi: nature06904 [pii]. 10.1038/nature06904

Tang F, Kaneda M, O'Carroll D, Hajkova P, Barton SC, Sun YA, Lee C, Tarakhovsky A, Lao K, Surani MA (2007) Maternal microRNAs are essential for mouse zygotic development. Genes Dev 21(6):644–648. doi:21/6/644 [pii]. 10.1101/gad.418707

Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X, Bodeau J, Tuch BB, Siddiqui A, Lao K, Surani MA (2009) mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods 6(5):377–382. doi:nmeth.1315 [pii]. 10.1038/nmeth.1315

Tian B, Bevilacqua PC, Diegelman-Parente A, Mathews MB (2004) The double-stranded-RNA-binding motif: interference and much more. Nat Rev Mol Cell Biol 5(12):1013–1023. doi: nrm1528 [pii]. 10.1038/nrm1528

Timmons L, Fire A (1998) Specific interference by ingested dsRNA. Nature 395(6705):854. doi:10.1038/27579

Tomari Y, Zamore PD (2005) Perspective: machines for RNAi. Genes Dev 19(5):517–529. doi:19/5/517 [pii]. 10.1101/gad.1284105

Tomari Y, Du T, Haley B, Schwarz DS, Bennett R, Cook HA, Koppetsch BS, Theurkauf WE, Zamore PD (2004a) RISC assembly defects in the Drosophila RNAi mutant armitage. Cell 116 (6):831–841. doi:S0092867404002181 [pii]

Tomari Y, Matranga C, Haley B, Martinez N, Zamore PD (2004b) A protein sensor for siRNA asymmetry. Science 306(5700):1377–1380. doi:306/5700/1377 [pii]. 10.1126/science.1102755

Tomari Y, Du T, Zamore PD (2007) Sorting of Drosophila small silencing RNAs. Cell 130 (2):299–308. doi:S0092–8674(07)00761–1 [pii]. 10.1016/j.cell.2007.05.057

Tomoyasu Y, Miller SC, Tomita S, Schoppmeier M, Grossmann D, Bucher G (2008) Exploring systemic RNA interference in insects: a genome-wide survey for RNAi genes in Tribolium. Genome Biol 9(1):R10. doi:gb-2008–9–1-r10 [pii]. 10.1186/gb-2008–9–1-r10

Tonkin LA, Bass BL (2003) Mutations in RNAi rescue aberrant chemotaxis of ADAR mutants. Science 302(5651):1725

Tonkin LA, Saccomanno L, Morse DP, Brodigan T, Krause M, Bass BL (2002) RNA editing by ADARs is important for normal behavior in *Caenorhabditis elegans*. Embo J 21 (22):6025–6035

Tran N, Raponi M, Dawes IW, Arndt GM (2004) Control of specific gene expression in mammalian cells by co-expression of long complementary RNAs. FEBS Lett 573(1–3):127–134. doi:10.1016/j.febslet.2004.07.075. S0014579304009573 [pii]

Tuschl T, Zamore PD, Lehmann R, Bartel DP, Sharp PA (1999) Targeted mRNA degradation by double-stranded RNA in vitro. Genes Dev 13(24):3191–3197

Vasale JJ, Gu W, Thivierge C, Batista PJ, Claycomb JM, Youngman EM, Duchaine TF, Mello CC, Conte D Jr (2010) Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. Proc Natl Acad Sci USA 107 (8):3582–3587. doi:0911908107 [pii]. 10.1073/pnas.0911908107

Vastenhouw NL, Fischer SE, Robert VJ, Thijssen KL, Fraser AG, Kamath RS, Ahringer J, Plasterk RH (2003) A genome-wide screen identifies 27 genes involved in transposon silencing in *C. elegans*. Curr Biol 13(15):1311–1316. doi:S0960982203005396 [pii]

Voinnet O, Pinto YM, Baulcombe DC (1999) Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. Proc Natl Acad Sci USA 96 (24):14147–14152

Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, Martienssen RA (2002) Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science 297 (5588):1833–1837

Wang XH, Aliyari R, Li WX, Li HW, Kim K, Carthew R, Atkinson P, Ding SW (2006) RNA interference directs innate immunity against viruses in adult Drosophila. Science 312 (5772):452–454. doi:1125694 [pii]. 10.1126/science.1125694

Wang Y, Medvid R, Melton C, Jaenisch R, Blelloch R (2007) DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat Genet 39(3):380–385. doi: ng1969 [pii]. 10.1038/ng1969

Wang HW, Noland C, Siridechadilok B, Taylor DW, Ma E, Felderer K, Doudna JA, Nogales E (2009a) Structural insights into RNA processing by the human RISC-loading complex. Nat Struct Mol Biol 16(11):1148–1153. doi:nsmb.1673 [pii]. 10.1038/nsmb.1673

Wang Y, Juranek S, Li H, Sheng G, Wardle GS, Tuschl T, Patel DJ (2009b) Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. Nature 461 (7265):754–761. doi:nature08434 [pii]. 10.1038/nature08434

Watanabe T, Takeda A, Tsukiyama T, Mise K, Okuno T, Sasaki H, Minami N, Imai H (2006) Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. Genes Dev 20 (13):1732–1743. doi:gad.1425706 [pii]. 10.1101/gad.1425706

Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, Obata Y, Chiba H, Kohara Y, Kono T, Nakano T, Surani MA, Sakaki Y, Sasaki H (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature 453(7194):539–543. doi: nature06908 [pii]. 10.1038/nature06908

Whangbo JS, Hunter CP (2008) Environmental RNA interference. Trends Genet 24(6):297–305. doi:S0168–9525(08)00126–1 [pii]. 10.1016/j.tig.2008.03.007

Wianny F, Zernicka-Goetz M (2000) Specific interference with gene function by double-stranded RNA in early mouse development. Nat Cell Biol 2(2):70–75

Wilkins C, Dishongh R, Moore SC, Whitt MA, Chow M, Machaca K (2005) RNA interference is an antiviral defence mechanism in *Caenorhabditis elegans*. Nature 436(7053):1044–1047

Winston WM, Molodowitch C, Hunter CP (2002) Systemic RNAi in *C. elegans* requires the putative transmembrane protein SID-1. Science 295(5564):2456–2459. doi:10.1126/science.1068836. 1068836 [pii]

Winston WM, Sutherlin M, Wright AJ, Feinberg EH, Hunter CP (2007) Caenorhabditis elegans SID-2 is required for environmental RNA interference. Proc Natl Acad Sci USA 104 (25):10565–10570. doi:0611282104 [pii]. 10.1073/pnas.0611282104

Wolfrum C, Shi S, Jayaprakash KN, Jayaraman M, Wang G, Pandey RK, Rajeev KG, Nakayama T, Charrise K, Ndungo EM, Zimmermann T, Koteliansky V, Manoharan M, Stoffel M (2007) Mechanisms and optimization of in vivo delivery of lipophilic siRNAs. Nat Biotechnol 25 (10):1149–1157. doi:nbt1339 [pii]. 10.1038/nbt1339

Xie Q, Guo HS (2006) Systemic antiviral silencing in plants. Virus Res 118(1–2):1–6. doi:S0168–1702(05)00351–5 [pii]. 10.1016/j.virusres.2005.11.012

Xu W, Han Z (2008) Cloning and phylogenetic analysis of sid-1-like genes from aphids. J Insect Sci 8:1–6. doi:10.1673/031.008.3001

Yan KS, Yan S, Farooq A, Han A, Zeng L, Zhou MM (2003) Structure and conserved RNA binding of the PAZ domain. Nature 426(6965):468–474. doi:10.1038/nature02129. nature02129 [pii]

Yang N, Kazazian HH Jr (2006) L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. Nat Struct Mol Biol 13(9):763–771. doi: nsmb1141 [pii]. 10.1038/nsmb1141

Yang S, Tutton S, Pierce E, Yoon K (2001) Specific double-stranded RNA interference in undifferentiated mouse embryonic stem cells. Mol Cell Biol 21(22):7807–7816

Yang W, Wang Q, Howell KL, Lee JT, Cho DS, Murray JM, Nishikura K (2005) ADAR1 RNA deaminase limits short interfering RNA efficacy in mammalian cells. J Biol Chem 280 (5):3946–3953

Yang W, Chendrimada TP, Wang Q, Higuchi M, Seeburg PH, Shiekhattar R, Nishikura K (2006) Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. Nat Struct Mol Biol 13(1):13–21

Yi CE, Bekker JM, Miller G, Hill KL, Crosbie RH (2003) Specific and potent RNA interference in terminally differentiated myotubes. J Biol Chem 278(2):934–939. doi:10.1074/jbc.M205946200. M205946200 [pii]

Yigit E, Batista PJ, Bei Y, Pang KM, Chen CC, Tolia NH, Joshua-Tor L, Mitani S, Simard MJ, Mello CC (2006) Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. Cell 127(4):747–757. doi:S0092-–8674(06)01293-–1 [pii]. 10.1016/j.cell.2006.09.033

Yoda M, Kawamata T, Paroo Z, Ye X, Iwasaki S, Liu Q, Tomari Y (2010) ATP-dependent human RISC assembly pathways. Nat Struct Mol Biol 17(1):17–23. doi:nsmb.1733 [pii]. 10.1038/nsmb.1733

Yoneyama M, Kikuchi M, Natsukawa T, Shinobu N, Imaizumi T, Miyagishi M, Taira K, Akira S, Fujita T (2004) The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. Nat Immunol 5(7):730–737. doi:10.1038/ni1087. ni1087 [pii]

Yuan YR, Pei Y, Ma JB, Kuryavyi V, Zhadina M, Meister G, Chen HY, Dauter Z, Tuschl T, Patel DJ (2005) Crystal structure of *A. aeolicus* argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. Mol Cell 19(3):405–419. doi:S1097-–2765(05)01475-–9 [pii]. 10.1016/j.molcel.2005.07.011

Zambon RA, Vakharia VN, Wu LP (2006) RNAi is an antiviral immune response against a dsRNA virus in Drosophila melanogaster. Cell Microbiol 8(5):880–889. doi:CMI688 [pii]. 10.1111/j.1462–5822.2006.00688.x

Zamore PD, Tuschl T, Sharp PA, Bartel DP (2000) RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell 101(1):25–33

Zhang X, Wang C, Schook LB, Hawken RJ, Rutherford MS (2000) An RNA helicase, RHIV-1, induced by porcine reproductive and respiratory syndrome virus (PRRSV) is mapped on porcine chromosome 10q13. Microb Pathog 28(5):267–278. doi:10.1006/mpat.1999.0349. S0882–4010(99)90349–2 [pii]

Zhang H, Kolb FA, Brondani V, Billy E, Filipowicz W (2002) Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. EMBO J 21(21):5875–5885

Zhang H, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W (2004) Single processing center models for human Dicer and bacterial RNase III. Cell 118(1):57–68. doi:10.1016/j.cell.2004.06.017. S009286740400618X [pii]

Zhou R, Czech B, Brennecke J, Sachidanandam R, Wohlschlegel JA, Perrimon N, Hannon GJ (2009) Processing of Drosophila endo-siRNAs depends on a specific Loquacious isoform. RNA 15(10):1886–1895. doi:rna.1611309 [pii]. 10.1261/rna.1611309

Supplement 6

Nejepinska J, Flemr M, Svoboda P.

Control of the interferon response in RNAi experiments

Methods in Molecular Biology

2012

820:133-61

Chapter 9

Control of the Interferon Response in RNAi Experiments

Jana Nejepinska, Matyas Flemr, and Petr Svoboda

Abstract

The RNA interference (RNAi) and interferons have been an uneasy marriage. Ever since the discovery of RNAi in mammals, the interferon response has been a feared problem. While RNAi became an efficient and widespread method for gene silencing in mammals, numerous studies recognized several obstacles, including undesirable activation of the interferon response, which need to be overcome to achieve a specific and robust RNAi effect. The aim of this text is to provide theoretical and practical information for scientists who want to control interferon response and other adverse effects in their RNAi experiments.

Key words: RNA interference, Small interfering RNA, Short hairpin RNA, Double-stranded RNA, Interferon

1. Introduction

RNAi is an excellent tool for selective inhibition of gene expression and studies of gene function(s) – when properly used. Otherwise, it is an excellent tool to generate confusing results. While RNAi became a standard tool, the lack of appropriate controls and/or ignorance of nonspecific effects undermined its efficient use in various cases. This text gives a brief overview of adverse effects found in RNAi experiments in mammalian cells and provides guidelines for designing RNAi experiments and identifying one of the frequently encountered undesirable effects – the interferon response.

1.1. RNA Silencing in Mammals and Its Experimental Use

RNA interference (RNAi) and the microRNA (miRNA) pathways regulate gene expression by inducing sequence-specific degradation and/or translational repression of target mRNAs (reviewed for example in refs. 1–4). A common feature of both pathways is 21–22 nucleotide-long RNA molecules serving as sequence-specific

Marc De Ley (ed.), *Cytokine Protocols*, Methods in Molecular Biology, vol. 820,
DOI 10.1007/978-1-61779-439-1_9,

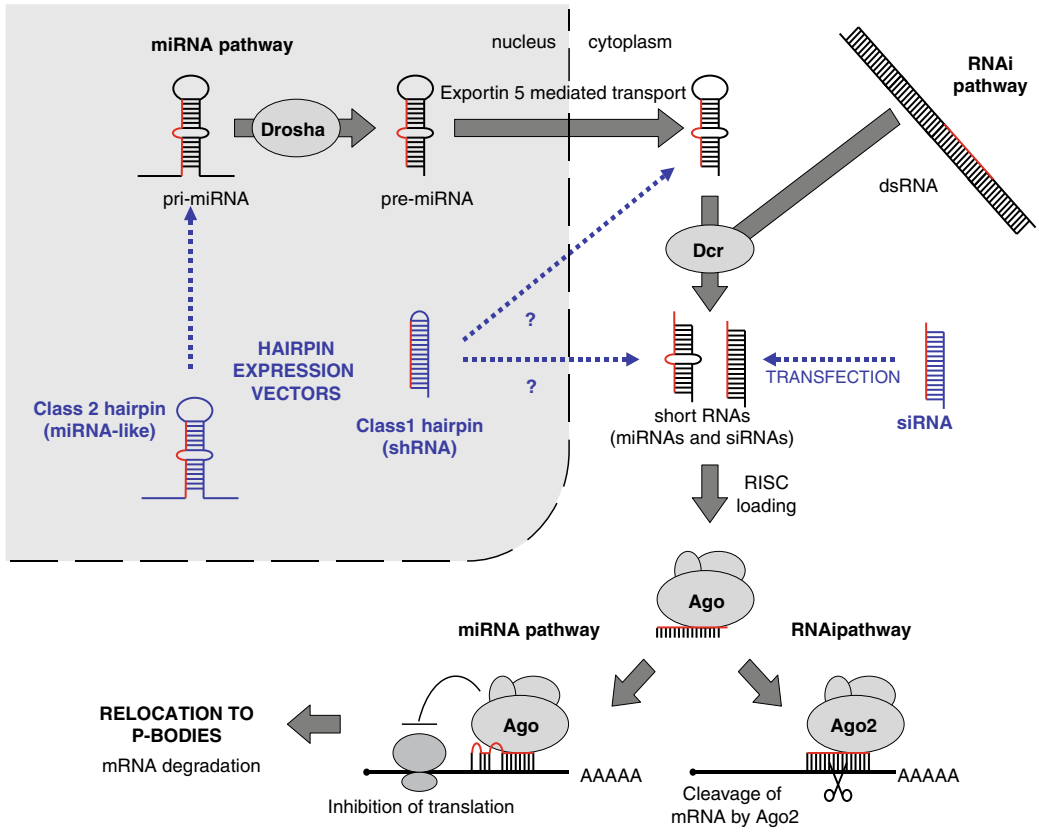


Fig. 1. RNA silencing in mammalian cells and its experimental use. The miRNA and RNAi pathways and two mechanisms of post-transcriptional silencing are depicted. Note that following Dicer cleavage, the miRNAs and siRNAs share a common pathway. Thus, the final silencing effect is dependent on the degree of homology with a cognate mRNA and the nature of an AGO protein rather than on the origin of the short RNA. The common entry points for experimental activation of RNAi are indicated in *blue.*

guides for silencing (Fig. 1). These short RNAs are released by Dicer, an RNase III family endonuclease, from various forms of double-stranded RNA (dsRNA). Mammals have only one Dicer protein, common for both RNAi and miRNA pathways.

The classical RNAi is initiated by long perfect dsRNA, which is processed by Dicer into double-stranded short interfering RNAs (siRNAs). siRNAs perfectly base-pair with a cognate RNA and guide its cleavage in the middle of the base-pairing sequence. Delivery of chemically synthesized siRNAs into mammalian cells also induces sequence-specific knockdown (5). However, long dsRNA is likely not a natural substrate of Dicer in mammalian somatic cells as dsRNA >30 bp is known to trigger sequence-independent pathways such as the protein kinase R (PKR) pathway (6). Long dsRNA induces RNAi only in oocytes, early embryos, embryonic stem cells, and possibly a few other mammalian cell types (7). Small RNA cloning experiments discovered that virtually all endogenous short RNAs linked to RNA silencing in somatic mammalian cells

are miRNAs (e.g., refs. 8, 9). miRNAs are transcribed as long primary transcripts (pre-miRNAs) with local hairpin structures that are processed by a nuclear RNase III Drosha-containing complex into short hairpin intermediates (pre-miRNAs). Pre-miRNAs are transported to the cytoplasm where Dicer releases a duplex containing a miRNA. A typical mammalian miRNA imperfectly base-pairs with a cognate 3′UTR and inhibits protein translation.

Despite certain distinctions, mammalian RNAi and miRNA pathways can be seen as one biochemical RNA silencing pathway because the effector complexes loaded by Dicer products appear functionally similar if not identical. Both siRNAs and miRNAs are loaded onto an Argonaute-containing effector ribonucleoprotein (RNP) complex, referred to as miRNP or RISC (RNA-induced Silencing Complex), which executes silencing. Four mammalian AGO proteins (AGO1 through AGO4) associate with miRNAs and are implicated in translational repression (10–12). In addition, AGO2 can mediate endonucleolytic cleavage of a target mRNA in the middle of the base-paired sequence (10, 11, 13). The AGO2-mediated cleavage requires formation of a perfect RNA duplex, while imperfect base-pairing, typical for most miRNAs, generally results in translational repression (14, 15). However, examples of miRNAs inducing RNAi-like cleavage also exist (16). Therefore, whether a short RNA will cause RNAi-like endonucleolytic cleavage or will induce the translational repression as a miRNA depends on the degree of complementarity and the AGO protein present, rather than on the origin of the short RNA. Despite this overlap between miRNA pathway and RNAi in mammals, we use the term RNAi for any experimental induction of sequence-specific cleavage, even if triggered by miRNA-like RNAs.

Structurally, there are three categories of short RNAs inducing RNAi, commonly referred here as RNAi triggers (Fig. 2): (1) *siRNA* – Duplexes of 21-22-mers with two nucleotide 3′ overhangs. (2) *Class I short hairpin* – Based on covalent linking of strands carrying functional siRNA sequences. The minimal class I hairpin contains a 19-bp dsRNA stem and 4–9 nucleotide loop and it is probably not processed like a classical miRNA (17–20). (3) *Class II hairpin* – Directly modeled after pre-miRNA (18, 21, 22).

RNA triggers can be prepared in vitro or expressed from DNA. The two most common experimental designs are (1) transient transfection of commercially obtained siRNAs and (2) transient or stable transfection of vectors expressing class I short hairpins from a pol III promoter. Class II hairpins were used less frequently but they have recently come into focus because they can be expressed together with a reporter from a single pol II promoter, thus providing more versatility than pol III-driven systems. It is important to mention these strategies because some of the nonspecific effects in RNAi experiments can be attributed to the carrying vector or delivery method.

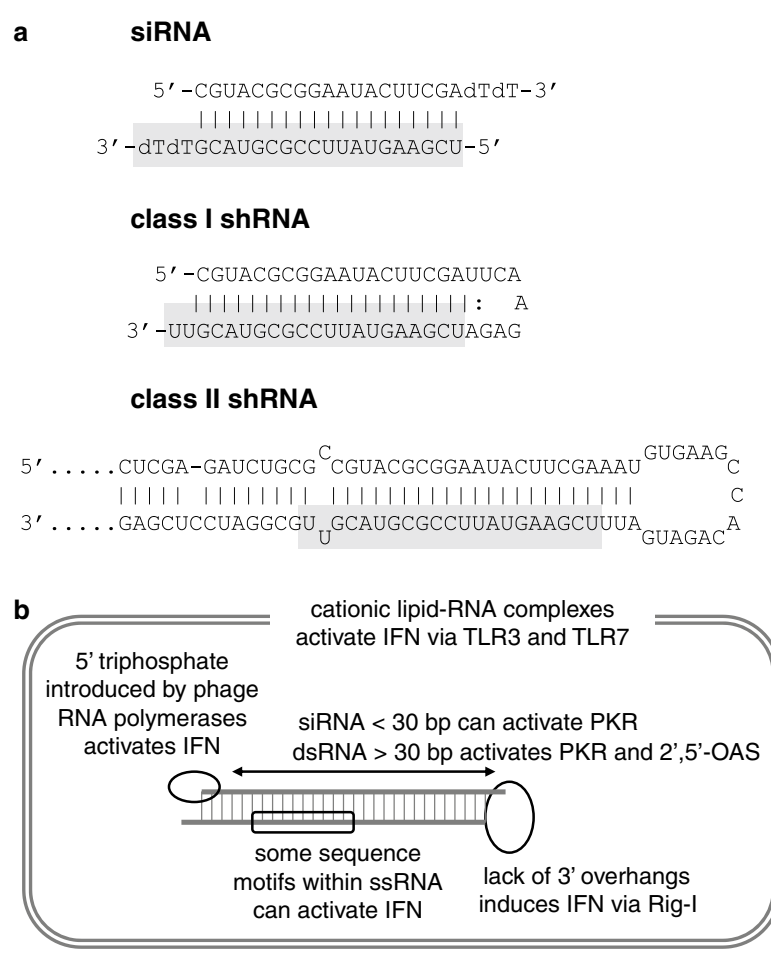


Fig. 2. Structural features of RNAi triggers and their relationship to interferon activation. (**a**) Three different types of RNAi triggers targeting firefly luciferase. siRNA and class II shRNAs were found in the literature (5, 56). Class I hairpin was modeled according to Brummelkamp et al. (17). Note that class I shRNA sequence represents the whole transcript produced by pol III from a vector while the class II shRNA corresponds to the part of the pol II transcript processed by Drosha and Dicer. Sequence within the *gray box* indicates the RNA fragment incorporated into the RISC complex. (**b**) Schematic display of structural features of an siRNA, which can trigger interferon response. See text for original references.

1.2. Nonspecific Effects in RNAi Experiments

Nonspecific effects can be divided into three categories. The first category includes the effects caused by activating interferon-related pathways. These effects are usually independent of siRNA sequence and they involve transcriptional activation of interferon-stimulated genes (ISGs). Although we generally refer to these effects as an interferon response, it should be kept in mind that an interferon response is broader and includes other pathways and effects not discussed here. The second category, referred to as off-targeting (see Subheading 1.4), includes sequence-dependent effects caused by targeting unintended transcripts. Finally, excessive use of RNAi

triggers may cause unintended effects by saturating the endogenous miRNA pathway resulting in relief of repression of genes repressed by miRNAs.

1.3. Interferon Response

Mammals have a complex system for responding to dsRNA in the cytoplasm. Various forms of cytosolic dsRNA can interact with several proteins that take part in the innate immune response. The interferon pathway is the most ubiquitous sequence-independent pathway induced by dsRNA in mammalian cells (reviewed in detail for example in ref. 23). One of the best-characterized effects of dsRNA is activation of PKR, which phosphorylates translation initiation factor eIF2α and causes general repression of translation. PKR is also involved in the regulation of NF-κB, which plays a key role in interferon induction. Interferon and dsRNA also activate 2′,5′-oligoadenylate synthetase (2′,5′-OAS) that produces 2′,5′-oligoadenylates with 5′-terminal triphosphate residues that subsequently induce activation of RNase L, a protein responsible for general RNA degradation (23).

Experimental induction of RNAi may result in activation of interferon response through some of the aforementioned proteins but there are also mechanisms activating interferons in dsRNA-independent manner. Different stimuli can lead to activation of overlapping but distinct sets of ISGs (24), and in specific cell types, particularly immune cells, the interferon response can be elicited by additional pathways (reviewed in ref. 25).

There are diverse features of RNAi triggers that can lead to the interferon activation through some of the proteins recognizing dsRNA. In 2003, two groups reported that transfection of siRNAs as well as pol III-driven shRNA expression can activate the interferon response (26, 27). Sledz et al. reported that transfection of siRNAs using Oligofectamine into different mammalian cells induced an RNAi effect as well as activation of PKR and expression of numerous ISGs (26). Microarray profiling of cells transfected with different concentrations (10, 25, 50 and 100 nM) of siRNAs revealed approximately 50 ISGs induced more than twofold at 48 h post-transfection. Some ISGs were induced at all siRNA concentrations while others only at higher ones (26). However, even mock transfection alone can sometimes have a stimulatory effect on ISGs (28). A detailed analysis of these effects revealed that siRNAs lacking 2-nt 3′ overhangs activate the interferon system via RNA helicase RIG-I (29). Thus, transfection of higher siRNA amounts may result in appearance of interferon-stimulating small RNAs lacking 2-nt 3′ overhangs, which seem to be a structural basis for discriminating between Dicer products and other short dsRNAs.

Another mechanism of activating interferons in RNAi experiments was described for short hairpins or siRNAs generated by in vitro transcription with phage polymerases (30). In this case, the interferon response is induced by the 5′ triphosphate GTP created by the

phage polymerase. Interferon induction by the triphosphate RNA ends is also mediated by RIG-I (31, 32). Therefore, whenever T7 or other phage polymerases are used to produce siRNAs or shRNAs, it is advisable to appropriately process their 5′ termini and include controls for measuring the interferon induction.

The interferon response was also induced by type I shRNAs expressed from H1 or U6 promoters (27, 33). Analysis of U6-promoter vectors, which induced strong ISG activation, identified a critical AA dinucleotide motif near the transcription start site suggesting a design flaw in U6-promoter based vectors (33). The exact mechanism of how the presence of the AA motif induced ISGs is unknown but these results provide a rationale for a better design of U6-driven shRNA vectors. Plasmids with H1 promoter too is not immune to ISG induction, but its exact cause remains unknown (33). Notably, interferon activation may not be caused only by shRNA expression only as lentiviral vectors seem to be less likely to induce OAS (one of the ISGs) than plasmid vectors (27).

Finally, interferon induction can be siRNA-specific as several motifs within single-stranded RNAs (ssRNA) from siRNAs, such as UGUGU and GUCCUUCAA, stimulate the interferon response in immune cells (34, 35). In addition, there are also immunostimulatory siRNAs without defined sequence motifs (36). Activation of the interferon response in these cases was likely mediated by TLR receptors and linked to endosomes (reviewed in detail in ref. 37).

Most researchers aim to avoid interferon activation while achieving an RNAi effect. Their interest in mechanisms of interferon activation ends when the interferon response is absent in their experiments. One of the common approaches to detect induction of the interferon pathways in cell lysates or even culture media is to employ antibodies recognizing ISGs (see Subheading 3.5.2). However, this approach is less sensitive than the analysis of transcripts of the interferon pathway and it may become costly. Therefore, RT-PCR analysis of some of the ISGs is the most accessible approach to detect if the interferon response was elicited in an experiment (see Subheading 3.5.1). Finally, if one is using microarrays to analyze results of an RNAi experiment, examination of probes detecting ISGs will provide a good indication whether or not the interferon response occurred (see Subheading 3.5.3).

1.4. Off-Targeting

Off-targeting occurs because short RNAs intended to induce Ago2-mediated cleavage of perfectly base-pairing targets imperfectly hybridize to other transcripts. One can view a short RNA introduced into a cell as a novel, abundant miRNA for which the cellular transcriptome is not adapted. This implies that off-targeting is common and can be found in most if not all RNAi experiments. Even worse, off-targeting cannot be effectively predicted as the reliability of computational miRNA target prediction is poor and

heavily dependent on conservation of miRNA binding sites. But, as we discuss later, off-targeting can be reduced by certain modifications of siRNAs and experimental design and its potentially destructive effect can be reduced by appropriate controls and careful interpretation of results.

The extent of off-targeting was not recognized in initial RNAi experiments. The early studies suggested that RNAi silencing requires perfect base-pairing. A sufficient control for RNAi specificity seemed to be an unrelated gene, typically a well-expressed housekeeping gene. Such control would indicate a global repression of gene expression caused for example by global repression of translation and/or nonspecific mRNA degradation: the hallmarks of PKR and 2′,5′-OAS activation. However, such controls are unlikely to detect off-targeting. Even if off-targeting would affect expression levels of hundreds of genes in a cell, the chance that one selected marker gene would be affected is slim.

The first strong evidence of off-targeting effects was demonstrated when mammalian cells transfected with siRNAs were systematically analyzed using microarrays (38). Jackson et al. found siRNA-specific expression patterns in transfected cells with only a few genes regulated in common by different siRNAs against the same gene. Although the effect was decreased when siRNA concentrations were lowered, the off-target regulation could not be eliminated completely and many of the off-targeted genes showed similar kinetics of targeting as the intended target. A similar effect has been observed with an siRNA targeting a luciferase sequence that has no homology in human genome. Moreover, off-target effects directed by the passive siRNA strand have also been detected. Although off-target regulation could not be completely explained, a portion of it appeared to be caused by partial complementarity between an siRNA and its target, reminiscent of the 5′ seeding regions of miRNAs (39). Off-targeting with various forms of RNAi triggers has been repeatedly demonstrated (38, 40) and some degree of off-targeting is likely widespread in RNAi experiments. However, improved understanding of siRNA and miRNA target recognition, better siRNA chemistry and pooling provide measures allowing for elimination of off-target effects in future experiments (41).

1.5. Controlling Nonspecific Effects in RNAi Experiments

The nonspecific effects in RNAi experiments are common and the best way to deal with them is to proper experimental design (42, 43). It should become a common policy to accept only those RNAi experiments that include controls truly decreasing the probability of misinterpretation due to nonspecific effects. An ideal RNAi experiment should (1) include one or more sensitive markers for the interferon response, (2) use two or more different siRNAs

targeting the same gene, (3) contain a rescue control by expressing an RNAi resistant version of the targeted gene, and (4) use phenotypic analysis designed to yield as uncommon phenotype(s) as possible.

2. Materials

2.1. Cell Lines and Culture Media

Selection of cell lines depends on individual needs or preferences. Common cell lines can be obtained from the American Type Culture Collection or other commercial sources. The following protocols are based on experience with the commonly used HeLa and HEK293 cell lines. These cells were maintained in Dulbecco's modified Eagle's medium (DMEM, Invitrogen) containing 10% fetal calf serum (FCS, Invitrogen), penicillin (100 U/mL, Invitrogen), and streptomycin (100 μg/mL, Invitrogen).

2.2. siRNAs

There are a number of commercial sources offering siRNA synthesis. They also provide predesigned, and in some cases even verified siRNAs. Among the most common providers are Thermo Fisher Scientific (former Dharmacon), Ambion, Qiagen, and Sigma. We recommend searching Web sites of these providers for information concerning predesigned and validated siRNAs targeting gene(s) of interest. One of the attractive options for RNAi knockdown is to use pools of siRNAs. These can be prepared in vitro from long dsRNA substrates (esiRNA) using own protocol or some of the commercial kits. Or, one can purchase a pool directly from a vendor. Particularly attractive option is to use ON-TARGET plus SMART pool siRNAs (Thermo Fisher Scientific), which reduce off-targeting by pooling siRNAs, which are in addition modified at their 5′end to reduce miRNA-like behavior (44).

2.3. DNA Oligonucleotides

These should be obtained from a local provider. As mutations in in vitro synthesized oligonucleotides are a common problem during cloning, we highly recommend using purified oligonucleotides from a provider with a good record of producing long DNA oligonucleotides.

2.4. shRNA Expressing Vectors

There are a number of different commercial and noncommercial plasmid vectors for RNAi that are accessible to individual researchers, and it is certainly a worthwhile investment to test several different vectors before committing resources to a specific one. The choice of the vector depends on whether one plans to make transient or stable transfections or to have an inducible or tissue-specific knockdown. Thus, refer to the recent literature and information at manufacturer's Web sites to choose a suitable experimental setup. The following protocols are designed for inserting oligonucleotides to produce Type I small hairpin from pSuper (OligoEngine) and

its derivates, such as pTer (20) (*Bgl* II/*Hin*d III cloning sites) or Type II miRNA-like small hairpins from pTMP or pLMP plasmids (Open Biosystems) (*Eco*RI/*Xho*I cloning sites). However, we want to point out that, while we and our collaborators routinely use these vectors, other vectors are not necessarily inferior.

To verify the sequence of inserted oligonucleotides, pSuper derivates can be sequenced with T3, T7 and M13 primers. Refer to the exact map of a pSuper derivate to select a suitable primer. For sequencing inserts in pTMP/pLMP vectors, we use the following primers, which are localized upstream of the hairpin insertion site: pTMP: 5′-TTGACCTCCATAGAAGACACCG-3′, pLMP: 5′-CCTCATCACCCAGGTTAAGAT-3′.

2.5. Reagents

1. Restriction enzymes (Fermentas, New England Biolabs, 10 U/μL) with buffers: *Bgl*II and *Hin*dIII for pSuper derivates or *Xho*I and *Eco*RI for pTMP and pLMP.
2. Agarose (e.g., Invitrogen), and electrophoresis running buffer. We use SB buffer: 10 mM NaOH, 36 mM boric acid.
3. T4 DNA ligase with 10× ligation buffer (Fermentas).
4. Chemically competent *Escherichia coli* cells (e.g., *DH5*α strain). shRNA-expressing vectors are usually readily propagated in normal lab strains. However, inverted repeats in plasmids occasionally cause complications. In such a case, one can use strains, which are able to maintain DNA with potentially highly structured sequences, such as Sure (Stratagene) or Stbl4 cells (Invitrogen).
5. Luria–Bertani (LB) medium [possibly terrific broth (TB) medium for more yield).
6. LB agar plates: 1.5% Agar in LB medium with 100 μg/mL ampicillin.
7. TE buffer: 10 mM Tris–HCl (pH 7.5), 1 mM EDTA.
8. Ampicillin: Stock solution 100 mg/mL in water, working concentration 100 μg/mL.
9. 60% Glycerol, sterile.
10. Gel extraction kit (e.g., QIAquick Gel Extraction Kit).
11. Miniprep kit (e.g., QIAprep Spin Miniprep kit).
12. MIDI or MAXIprep kit (e.g., Qiagen HiSpeed Plasmid Midi or Maxi kit).
13. Transfection reagents for plasmid transfection: Turbofect (Fermentas).
14. Transfection reagent for siRNA: Turbofect (Fermentas) or Oligofectamine + OptiMEM (Invitrogen).
15. 30% Fetal calf serum (FCS): A stock made of 50 mL FCS, 1.7 mL 200 mM glutamine, 1.7 mL penicilin (10,000 U/mL)/streptomycin (10 mg/mL), and 113.6 mL DMEM.

16. PolyI:C (Sigma), 5 mg/mL stock in water.
17. Plasmid expressing immunostimulatory shRNA, 200 ng/μL.

3. Methods

The following protocol is a basic protocol we use for RNAi experiments. Analysis of the interferon response is a compilation of data found in the literature, our experience with occasional appearance of interferon response in RNAi experiments (28), and ongoing analysis of effects of long dsRNA expression in mammalian somatic cells.

3.1. siRNA Sequence Design

Time consideration: 2 h

Proper siRNA design is crucial for conducting a successful RNAi experiment. Incorrect siRNA sequence will result in decreased siRNA efficiency and/or specificity of the silencing effect. Own siRNA design carries a risk of a failure and it is not unusual when only one of three, or even four siRNAs induces good repression of the cognate gene. Ideally, one should obtain two siRNAs of different sequences targeting the same gene, which sometimes makes siRNA design a painstaking process. If only a single siRNA is available, one of the controls should include a rescue experiment (see also Subheading 1.5 and Note 2). Therefore, prior to making own siRNA design, it is very useful to search PubMed and WWW (particularly siRNA vendor sites) to find whether suitable functional siRNA sequences are available already. [AU1]

Several important criteria for siRNA design have been identified and they were built into a number of freely available Web-based design tools (summarized for example in ref. 45). Preference of one siRNA design tool over another is to some extent a matter of personal choice (reviewed in ref. 46). We combine two design tools: BIOPREDsi (47) and RNAxs (48) and we subsequently verify the specificity of siRNAs using the Specificity Server (49). BIOPREDsi is a neural network-based algorithm that has been trained on a large set of siRNAs and has been used for a genome-wide design of siRNA. BIOPREDsi was a top-scoring approach in a comparative study of several common siRNA design tools (50) and it is routinely used by us or our collaborators. BIOPREDsi is a representative of tools that design siRNAs according to optimized parameters of an siRNA sequence but not taking into an account the interaction of an siRNA with its cognate mRNA. Therefore, we complement the BIOPREDsi siRNA prediction with RNAxs tool, which introduces the analysis of the cognate sequence accessibility. RNAxs predicts secondary structures within the siRNA binding site and evaluates probability of efficient recognition of the binding site by the RISC complex, as it has been shown in biochemical studies (48, 51).

1. Obtain the target mRNA sequence from NCBI (http://www.ncbi.nlm.nih.gov) or ENSEMBL (http://www.ensembl.org). For a simple knockdown, it is often recommended to use the coding region (CDS) of target mRNA to design siRNAs because the unique coding sequence reduces off-target risk. The downside of using CDS for siRNA design is that a rescue expression construct must carry specifically positioned mutations in the cognate sequence (see the section controls). When targeting a 3′ UTR, one can just use a different 3′ UTR to make a rescue expression plasmid.
2. Search for suitable siRNAs at the BIOPREDsi page (http://www.BIOPREDsi.org). Paste the target mRNA sequence in FASTA format (use Readseq to convert the sequence into FASTA format. Readseq online can be found at different sites, we typically use the one available in Sequence utilities at the BCM Search Launcher site (http://searchlauncher.bcm.tmc.edu/)). Make sure to have the "Input Type" set to RNA Sequence. Set "# of predicted siRNAs" to 20, click "Design siRNA sequences" and save the results. (Note: the siRNA design using BIOPREDsi might be a little more time-consuming). If the Web site is not available, you can use another similar siRNA designer Web site, e.g., http://www.dharmacon.com/designcenter/.
3. Repeat the same procedure with the RNAxs (http://rna.tbi.univie.ac.at/cgi-bin/RNAxs). Set "Maximal number of siRNAs" in the Output option to 20 again and click "REPRESS IT".
4. Pick the sequences ranked from best to worst in RNAxs that also have high scores from the BIOPREDsi prediction. Ideally, at least four siRNAs with high scores should be selected. Verify that these sequences are specific using the Specificity Server (http://informatics-eskitis.griffith.edu.au/SpecificityServer (49)). Paste the RefSeq code (NM_XXXXXX) of the target mRNA and the siRNA sequence to the Option B "Database search for matches" window and press "Search". The server returns Automatic recommendation "OK" when the inspected siRNA passes specificity criteria. Similarly, test the specificity of all selected sequences.
5. Optional: To test if siRNAs will be designed with high scores in other siRNA design tools, one can use, for example DSIR (http://cbio.ensmp.fr/dsir/) (52). An extensive list of other siRNA design tools can be found elsewhere (45).

3.2. Production of Vectors Expressing Desired shRNAs

Time consideration: 1–2 weeks

3.2.1. pTer/pSuper shRNA Vector Cloning

1. Design and obtain the sense and antisense oligonucleotides as described in Fig. 3a.
2. Oligonucleotide annealing: Mix 5 μg of sense oligonucleotide with 5 μg of antisense oligonucleotide in TE buffer in a total volume of 100 μL. Place the tube with oligonucleotide mixture

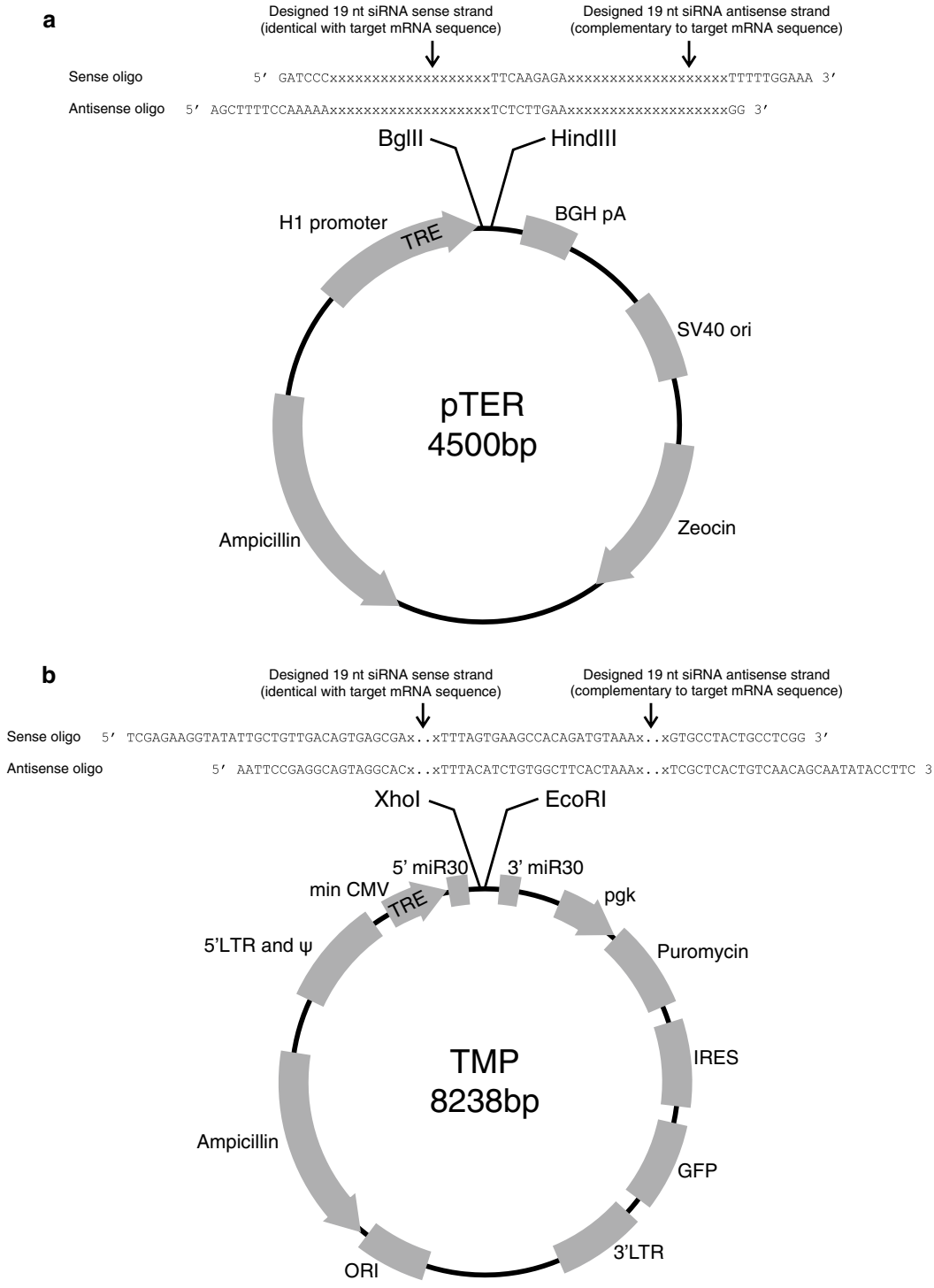


Fig. 3. Schematic overview of pTer (**a**) and TMP (**b**) vectors. Restriction sites for insertion of oligonucleotides carrying shRNA sequences are visualized. The oligonucleotides should carry sense and antisense strands of the *in silico* designed 19-nt siRNA as indicated in (**a**) for pTer and in (**b**) for TMP.

in a beaker with a large volume of boiling water (500–1,000 mL) and incubate for 2 min. Turn heating off and leave the oligonucleotide mix in hot water to cool down slowly. Dilute annealed oligonucleotide 100× with TE buffer.

3. pTer vector digestion: Mix 2 μg of pTer plasmid DNA with 2 μL of 10× restriction buffer and add water to a final volume of 18 μL. Add 1 μL of *Bgl*II enzyme and 1 μL *Hin*dIII enzyme and incubate for 2 h at 37°C.
4. Resolve the digested plasmid in a 1.0% agarose gel in SB buffer and extract DNA from the gel using a commercial gel extraction kit according to manufacturer's recommendations (e.g., QIAquick Gel Extraction Kit). Elute the extracted DNA in 30 μL of elution buffer.
5. Ligation: Mix 2 μL of digested pTer DNA with 5 μL of annealed diluted oligonucleotide (1 ng/μL); add 2 μL of 10× ligation buffer and nuclease-free water to a final volume of 19 μL. Add 1 μL of T4 DNA ligase and incubate for 2 h at room temperature (RT) or overnight at 16°C.
6. Transformation: Add 2 μL of the ligation mixture to 50 μL of chemically competent *E. coli* cells (e.g., *DH5α* strain) and incubate for 20 min on ice. Submit cells to a 30 s heat-shock at 42°C and immediately add 1 mL of LB medium. Incubate for 1 h at 37°C with shaking.
7. Selection: Plate 100 μL of transformed cells on LB agar plate supplemented with 100 μg/mL ampicillin and incubate overnight at 37°C.
8. Vector preparation: Pick several colonies from the selection plate and incubate for 8 h to overnight at 37°C in 5 mL LB media (containing 100 μg/mL ampicillin) each. Isolate the vector DNA using a commercial Miniprep kit (e.g., QIAprep Miniprep kit). Make a glycerol stock of the vector by mixing 750 μL of the remaining culture and 250 μL of sterile 60% glycerol in a cryotube and store in −80°C for later use. Verify the insert sequence by sequencing (it is an important step because oligonucleotides often carry mutations). Once the vector sequence is verified, prepare sufficient amount of DNA for intended experiments.

   Note: Incompatible ends in the vector allow for direct insertion of annealed oligonucleotides without a need for dephosphorylation of the vector and phosphorylation of oligonucleotides. However, an incomplete restriction digest can produce a high background of empty vectors, so we recommend verifying the presence of the insert by restriction digest prior to sequencing. A typical cloning procedure yields up to a hundred of colonies. If several hundred colonies and more are obtained, this usually indicates high background of empty clones.

3.2.2. TMP/LMP shRNA Vector Cloning

1. Design and obtain the sense and antisense oligonucleotides as described in Fig. 3b.
2. Oligonucleotide annealing: Mix 5 μg of sense oligonucleotide with 5 μg of antisense oligonucleotide in TE buffer in a total volume of 100 μL. Place the tube with oligonucleotide mixture in a beaker with boiling water and incubate for 2 min. Turn heating off and leave the oligonucleotide mix in hot water to cool down slowly. Dilute annealed oligonucleotide 100× with TE buffer.
3. TMP vector digestion: Mix 1 μg of TMP plasmid DNA with 2 μL of 10× restriction buffer and add water to a final volume of 18 μL. Add 1 μL of *Xho*I enzyme and 1 μL *Eco*RI enzyme and incubate for 2 h at 37°C.
4. Separate the digested plasmid in a 1.0% agarose gel in SB buffer (sodium borate buffer: 10 mM sodium hydroxide, pH adjusted to 8.5 with boric acid.) and extract from gel using commercial gel extraction kit (e.g., QIAquick Gel Extraction Kit). Elute the extracted DNA in 30 μL of elution buffer.
5. Ligation: Mix 2 μL of digested TMP DNA with 5 μL of annealed oligonucleotide (1 ng/μL), add 2 μL of 10× ligation buffer and nuclease-free water to a final volume of 19 μL. Add 1 μL of T4 DNA ligase and incubate for 2 h at RT or overnight at 16°C.
6. Transformation: Add 2 μL of the ligation mixture to 50 μL of chemically competent *E. coli* cells (e.g., *DH5*α strain) and incubate for 20 min on ice. Submit cells to a 30 s heat-shock at 42°C and immediately add 1 mL of LB medium. Incubate for 1 h at 37°C with shaking.
7. Selection: Plate 100 μL of transformed cells on LB agar supplemented with 100 μg/mL ampicillin and incubate overnight at 37°C.
8. Vector preparation: Pick several colonies from the selection plate and incubate for 8 h to overnight at 37°C in 5 mL LB media (containing 100 μg/mL ampicillin) each. Isolate the vector DNA using a commercial Miniprep kit (e.g., QIAprep Miniprep kit). Make a glycerol stock of the vector by mixing 750 μL of the remaining culture and 250 μL of sterile 60% glycerol in a cryotube and store in −80°C for later use. Verify the insert sequence by sequencing (it is an important step because oligonucleotides often carry mutations). Once verified the vector sequence is verified, prepare sufficient amount of DNA for intended experiments.

3.3. RNAi with Transient Transfection of siRNA with Oligofectamine

Time consideration: 4 days

Before performing RNAi experiments, we recommend to test different concentrations of each siRNA (ideally 5, 10, 20, 50, and 100 nM) and subsequently use the lowest concentration required for

efficient silencing. The following protocol is based on transfection of 50 nM siRNA into HEK293 and HeLa cells grown in 6-well plates. Values/volumes indicate amounts per well, values in parentheses indicate amounts per well for 24-well plates.

1. One day before transfection, plate cells to achieve optimal confluency. Suggested confluency of HeLa cells for transfection with Oligofectamine is ~50%. For HEK293 cells, we recommend a little bit higher confluency (60–70%) as they easily detach. For other cell types, optimization of ideal confluency may be needed. When using media with antibiotics, remove the medium with antibiotics and replace it with a medium without antibiotics before transfection.
2. Prepare siRNA-oligofectamine complexes. For each well, mix in a 1.5-mL tube (Tube A) 2.5 μL (1.25 μL) of siRNA (20 μM stock) and 180 μL (190 μL) of Opti-MEM media, mix gently by tapping on the tube, and incubate for 7 min at RT. To another tube (tube B), add 2 μL (1 μL) of Oligofectamine and 16 μL (8 μL) of Opti-MEM, mix gently by tapping on the tube, and incubate for 7 min at RT. After 7-min incubation, add content of tube B into tube A, mix gently by tapping, and incubate for 20 min at RT.
3. Transfect cells. Wash cells with Opti-MEM, add 800 μL (400 μL) of Opti-MEM and transfection mixture from above, and incubate for 4 h (37°C). Then add 500 μL (250 μL) of 30% FCS, and incubate for 48 h. It is also possible to replace the transfection medium with DMEM containing 10% FCS instead of adding 30% FCS to Opti-MEM. Cells can be cultured from 24 to 72 h after transfection but it should be kept in mind that shorter incubation may not be sufficient for detecting an RNAi effect and longer incubation increases the risk of interferon response.

We also use Turbofect (Fermentas) for siRNA transfection with good results. siRNA transfection with Turbofect is performed as described in Subheading 3.4, except siRNA is used instead of DNA to achieve final concentration of 40 nM during transfection.

3.4. RNAi with Transient Transfection of shRNA-Expressing Vector with Turbofect

Time consideration: 1 week

Plate cells to achieve density required for transfection. Knowing optimal density is critical for efficient transfection and may differ from one transfection reagent to another. The following protocol is for transfection of HEK293 and HeLa cells grown in 6-well plates, values in parentheses indicate amounts for 24-well plates.

1. For transfection of HeLa or HEK293 cells, dilute cells in culture medium to a density 60,000 cell/mL and plate 3 mL (0.5 mL) of suspension per well, 24 h prior to transfection. For optimal transfection conditions, cells should be 50–70% confluent.

2. Change culture media just before transfection and add 1 mL (0.5 mL) of fresh media per well.
3. Transfect cells using Turbofect transfection reagent according to the manufacturer's protocol. Briefly, combine the desired amount of shRNA-containing plasmid (usually 1.5 μg (0.5 μg)) and pBluescript plasmid up to the total DNA amount of 2.5 μg (1.0 μg) for each well and mix DNA with 500 μL (100 μL) DMEM.
4. Mix by pipetting and incubate for 30 min at RT.
5. Add the mixture dropwise into each well.
6. 6 h after transfection, add 1.5 mL (0.4 mL) of fresh DMEM medium supplemented with 10% FCS and antibiotics.
7. Harvest the cells 24–72 h (typically 48 h) after transfection and assay for RNAi effects.

Transfection efficiency of pTMP or pLMP can be easily monitored because these plasmids carry a EGFP reporter. To monitor transfection efficiency in pSuper or pTer, one can use an EGFP-expressing plasmid for transfections instead of pBluescript.

3.5. Detection of the Interferon Response in RNAi Experiment

Table 1 lists several genes transcriptionally activated by the interferon response found in the literature, which are suitable markers. Particularly *IFIT1* (also known as p56) seems to be a suitable, sensitive marker for ISG activation (27, 29).

3.5.1. By RT-PCR or Quantitative Real-Time PCR

If an analysis of knockdown effects does not include RT-PCR already, isolate total RNA from cells in 6-well plates 48 h after transfection using Trizol (Invitrogen) according to the manufacturer's protocol. Remove DNA contamination by DNAse I treatment (Fermentas), followed by phenol/chloroform extraction (53). Prepare cDNA using Superscript III Reverse Transcriptase (Invitrogen) primed with random hexamer primers according to the manufacturer's instructions. Perform quantitative or semi-quantitative PCR using the primers in Table 2.

3.5.2. In Culture

For the detection of interferon IFN-α or β by a conventional sandwich enzyme immunoassay (ELISA), some of the commercially available kits can be used (see Table 3). The IFN concentration in cell supernatants is measured 24–48 h after transfection.

3.5.3. On Arrays

Induction of the interferon pathway in RNAi experiments is generally monitored using RT-PCR or western blotting on interferon response marker genes or proteins, respectively. However, in experiments where effects of RNAi are studied on transcriptome level, the interferon stimulation can be assessed directly from microarray data. In such a case, the overall interferon stimulated gene (ISG) profile may vary between different cell types. We explored publicly available microarray data dealing with interferon induction in

Table 1
Primers for RT-PCR detection of the interferon activation

Primers for human samples

Gene name	Primer		Quantification	References
IFIT1 (p56, ISG56-K)	F R	CTAAGCAAAACCCTGCAGAACG GGAATTCAATCTGATCCAAGACTC	Real-time	This work
IFIT2	F R	GCCACAAAAAATCACAAGCCA CCATTGTCTGGATTTAAGCGG	Real-time	(57)
RIG-I (DDX58)	F R	CATGTCCACCTTCAGAAGTGTCTG GGTTTTTCCACAACCTGTAGGAGC	Real-time	This work
OAS1	F R	CTTTGATGCCCTGGGTCAGTTG CTCTGTAGTTCTGTGAAGCAGGTG	Real-time	This work
STAT1	F R	TGGGTTTGACAAGGTTCTT TATGCAGTGCCACGGAAAG	Semi-quantitative	(58)
IFN-α2	F R	GGATGAGACCCTCCTAGACAAAT ATGATTTCTGCTCTGACAACCTC	Real-time	(59)
IFN-β	F R	ATGAGTGGTGGTTGCAGGC AAGCATCAGAGGCGGACTCTGGGA	Real-time	(60)

Primers for murine samples

Gene name	Primer		Quantification	References
Ifit1 (Isg56)	F R	AGAGAGTCAAGGCAGGTTTCTGAG TCTCACTTCCAAATCAGGTATGTCA	Real-time	This work
Ifit2 (Isg54)	F R	ATGAAGCAGGTGCTGAATACTAGTGA TGGTGAGGGCTTTCTTTTTCC	Real-time	(61)
Rig-I (Ddx58)	F R	AGCTTACTCGGAGGTTTGAAGAAA CAGTCAGTATGCCAGGCTTTAGAA	Real-time	This work
Oas1b	F R	AGACGTTGTGGAGTGAAGTTTGAG T CCCAGCTTCTCCTTACACAGTTG	Semi-quantitative	This work
Stat1	F R	CACATTCACATGGGTGGAAC TCTGGTGCTTCCTTTGGTCT	Real-time	(62)
IFN-α2	F R	TCTGTGCTTTCCTCGTGATG TTGAGCCTTCTGGATCTGCT	Real-time	(62)
IFN-β	F R	GGAGATGACGGAGAAGATGC CCCAGTGCTGGAGAAATTGT	Real-time	(63)

Table 2
TaqMan probes for markers for interferon activation in human and mouse cells

Gene name	TaqMan probe	
Ifit1 (ISG56)	Hs Mm	Hs03027069_s1 Mm00515153_m1
Ifit2 (ISG54)	Hs Mm	Hs01922738_s1 Mm00492606_m1
Rig-I (Ddx58)	Hs Mm	Hs01061434_m1 Mm01216860_m1
Oas1	Hs Mm	Hs00973635_m1 Mm01198570_m1
Stat1	Hs Mm	Hs01014000_m1 Mm01257291_m1
IFN-α2	Hs Mm	Hs00999940_s1 Mm00833961_s1
IFN-β	Hs Mm	Hs01077958_s1 Mm00439552_s1

Table 3
Selection of ELISA detection kits for assaying interferons

Product	Manufacturer	Reference
Human IFN-α ELISA kit	R&D Systems, USA	(64)
Human IFN-β ELISA kit	R&D Systems, USA	(64)
Mouse IFN-α ELISA kit	PBL Biomedical Laboratories, USA	(65)
Human (mouse) IFN-β ELISA kit	BioSource International, USA	(66)

RNAi experiments in various human cell types and generated a list of ISGs that were induced (with one exception – *OAS2*) in at least two studies. These human ISGs with corresponding probe IDs for 3 widely used microarray systems are summarized in Table 4 and their mouse counterparts in Table 5. The expression of the listed genes should be verified in all microarray data obtained from RNAi experiment to exclude nonspecific RNAi mediated interferon stimulation. Eight particularly important known ISG marker genes are highlighted in bold.

Table 4
List of human ISGs induced by dsRNA or in RNAi experiments inferred from published microarray data

Gene name	RefSeq ID	Ensembl ID	Affymetrix HU133 ID	Illumina Human_WG-6v3 ID	Agilent Human Genome ID	References
ADAR	**NM_001111**	**ENSG00000160710**	**201786_s_at**	**ILMN_1776777**	**A_23_P200439**	(24, 28, 67)
CCL3	NM_002983	ENSG00000006075	205114_s_at	ILMN_1671509	A_23_P373017	(67, 68)
CEBPD	NM_005195	ENSG00000180733	203973_s_at	ILMN_1782050	A_23_P31810	(54, 68)
CXCL10	NM_001565	ENSG00000169245	204533_at	ILMN_1791759	A_24_P303091	(27, 67)
CXCL11	NM_005409	ENSG00000169248	210163_at	ILMN_2067895	A_24_P20607	(27, 68)
G1P2	NM_005101	ENSG00000187608	205483_s_at	ILMN_2054019	A_23_P819	(26, 28, 54, 67)
G1P3	NM_022873	ENSG00000126709	204415_at	ILMN_1687384	A_23_P201459	(26, 28, 54)
GBP1	NM_002053	ENSG00000117228	231577_s_at	ILMN_2148785	A_32_P107372	(24, 27, 28)
GMPR	NM_006877	ENSG00000137198	204187_at	ILMN_1729487	A_24_P277657	(27, 68)
IFI16	NM_005531	ENSG00000163565	206332_s_at	ILMN_1710937	A_23_P217866	(54, 67)
IFI27	NM_005532	ENSG00000165949	202411_at	ILMN_2058782	A_23_P48513	(27, 28, 54)
IFI35	NM_005533	ENSG00000068079	209417_s_at	ILMN_1745374	A_23_P152782	(27, 67)
IFI44	NM_006417	ENSG00000137965	214453_s_at	ILMN_1760062	A_23_P23074	(27, 54)
IFIT1	**NM_001548**	**ENSG00000185745**	**203153_at**	**ILMN_1707695**	**A_23_P52266**	(24, 26–28, 54, 67)
IFIT2	NM_001547	ENSG00000119922	217502_at	ILMN_1739428	A_23_P24004	(26, 28, 54)
IFIT3	NM_001549	ENSG00000119917	204747_at	ILMN_1701789	A_23_P35412	(28, 54, 67)
IFIT5	NM_012420	ENSG00000152778	203595_s_at	ILMN_1696654	A_24_P30194	(27, 54, 67)
IFITM1	**NM_003641**	**ENSG00000185885**	**214022_s_at**	**ILMN_1801246**	**A_23_P72737**	(26–28, 54, 67)
IFITM2	NM_006435	ENSG00000185201	201315_x_at	ILMN_1673352	A_24_P287043	(26, 54)
IRF2	NM_002199	ENSG00000168310	203275_at	ILMN_2090607	A_23_P136478	(24, 67)
IRF7	NM_004029	ENSG00000185507	208436_s_at	ILMN_1798181	A_24_P378019	(27, 28, 67)

(continued)

Table 4 (continued)

Gene name	RefSeq ID	Ensembl ID	Affymetrix HU133 ID	Illumina Human_WG-6v3 ID	Agilent Human Genome ID	References
ISG20	**NM_002201**	**ENSG00000172183**	**33304_at**	**ILMN_1659913**	**A_23_P32404**	(27, 28, 54, 67, 68)
ISGF3G	NM_006084	ENSG00000213928	203882_at	ILMN_1745471	A_23_P65442	(27, 54, 67)
MDA5	**NM_022168**	**ENSG00000115267**	**219209_at**	**ILMN_1781373**	**A_23_P68155**	(54, 67, 68)
MX1	NM_002462	ENSG00000157601	202086_at	ILMN_1662358	A_23_P17663	(27, 28, 54)
MX2	NM_002463	ENSG00000183486	204994_at	ILMN_2231928	A_23_P6263	(27, 67)
NMI	NM_004688	ENSG00000123609	203964_at	ILMN_1739541	A_23_P154235	(24, 54, 67)
OAS1	**NM_016816**	**ENSG00000089127**	**205552_s_at**	**ILMN_1672606**	**A_23_P64828**	(26–28, 67)
OAS2	NM_016817	ENSG00000111335	204972_at	ILMN_1674063	A_24_P343929	(26)
OAS3	NM_006187	ENSG00000111331	218400_at	ILMN_2184262	A_23_P47955	(26, 27, 54, 67)
PKR	NM_002759	ENSG00000055332	204211_x_at	ILMN_1706502	A_23_P142750	(54, 67)
PLSCR1	NM_021105	ENSG00000188313	202446_s_at	ILMN_1745242	A_23_P69109	(27, 54, 67)
RIG-I	**NM_014314**	**ENSG00000107201**	**218943_s_at**	**ILMN_1797001**	**A_23_P20814**	(54, 67)
SP100	NM_003113	ENSG00000067066	210218_s_at	ILMN_2284998	A_23_P349928	(27, 67)
SP110	NM_004509	ENSG00000135899	209761_s_at	ILMN_2415144	A_23_P120002	(27, 54)
STAT1	**NM_007315**	**ENSG00000115415**	**200887_s_at**	**ILMN_1777325**	**A_24_P274270**	(26, 27, 54, 67)
STAT2	NM_005419	ENSG00000170581	205170_at	ILMN_1690921	A_23_P76090	(54, 67)
TAP1	NM_000593	ENSG00000168394	202307_s_at	ILMN_1751079	A_23_P59005	(24, 54)
TLR3	NM_003265	ENSG00000164342	206271_at	ILMN_2155708	A_23_P29922	(28, 54)
UBE2L6	NM_198183	ENSG00000156587	201649_at	ILMN_1769520	A_23_P75741	(26, 27)
USP18	NM_017414	ENSG00000184979	219211_at	ILMN_1740200	A_23_P132159	(54, 67)

There are probe IDs of 3 commonly used microarray systems from Affymetrix, Illumina, and Agilent included. Eight well known ISG marker genes are highlighted in bold

Table 5
List of mouse counterparts to human ISGs induced in microarrays from RNAi experiments with probe IDs of 3 commonly used microarray systems from Affymetrix, Illumina, and Agilent

Gene name	RefSeq	Ensembl ID	Affymetrix 430_v2 ID	Illumina MouseWG-6_v2 ID	Agilent Mouse Genome ID
Adar	NM_001038587	ENSMUSG00000027951	1434268_at	ILMN_2489167	A_52_P183181
Ccl3	NM_011337	ENSMUSG00000000982	1419561_at	ILMN_1253919	A_51_P140710
Cebpd	NM_007679	ENSMUSG00000071637	1423233_at	ILMN_2588570	A_51_P444447
Cxcl10	NM_021274	ENSMUSG00000034855	1418930_at	ILMN_1214419	A_51_P432641
Cxcl11	NM_019494	ENSMUSG00000060183	1419698_at	ILMN_1247446	A_52_P676403
G1p2	NM_015783	ENSMUSG00000035692	1431591_s_at	ILMN_1256257	A_52_P463936
Gbp1	NM_010259	ENSMUSG00000028269	1420549_at	ILMN_1233293	A_51_P398766
Gmpr	NM_025508	ENSMUSG00000000253	1448530_at	ILMN_2602581	A_51_P495986
Ifi16	NM_008329	ENSMUSG00000073489	1452348_s_at	ILMN_3010089	A_51_P408343
Ifi27	NM_029803	ENSMUSG00000079017	1426278_at	ILMN_2762944	A_52_P90363
Ifi35	NM_027320	ENSMUSG00000010358	1445897_s_at	ILMN_2625290	A_51_P414889
Ifi44	NM_133871	ENSMUSG00000028037	1423555_a_at	ILMN_2680136	A_51_P487690
Ifit1	NM_008331	ENSMUSG00000034459	1450783_at	ILMN_2774340	A_51_P327751
Ifit2	NM_008332	ENSMUSG00000045932	1418293_at	ILMN_2981167	A_52_P542388
Ifit3	NM_010501	ENSMUSG00000074896	1449025_at	ILMN_2944666	A_51_P359570
Ifitm1	NM_026820	ENSMUSG00000025491	1424254_at	ILMN_2640765	A_52_P541802
Ifitm2	NM_030694	ENSMUSG00000060591	1417460_at	ILMN_1232667	A_51_P168459
Irf2	NM_008391	ENSMUSG00000031627	1418265_s_at	ILMN_1251696	A_51_P316523
Irf7	NM_016850	ENSMUSG00000025498	1417244_a_at	ILMN_1227573	A_51_P421876

(continued)

Table 5 (continued)

Gene name	RefSeq	Ensembl ID	Affymetrix 430_v2 ID	Illumina MouseWG-6_v2 ID	Agilent Mouse Genome ID
Isg20	**NM_020583**	**ENSMUSG00000039236**	**1419569_a_at**	**ILMN_2735615**	**A_51_P510713**
Isgf3g	NM_008394	ENSMUSG00000002325	1421322_a_at	ILMN_1233461	A_52_P176013
Mda5	**NM_027835**	**ENSMUSG00000026896**	**1426276_at**	**ILMN_2648913**	**A_52_P121468**
Mx1	NM_010846	ENSMUSG00000000386	1451905_a_at	ILMN_2707870	A_52_P446431
Mx2	NM_013606	ENSMUSG00000023341	1419676_at	ILMN_1239219	A_51_P514085
Nmi	NM_019401	ENSMUSG00000026946	1425719_a_at	ILMN_2755958	A_51_P125067
Oas1	**NM_001083925**	**ENSMUSG00000029605**	**1425119_at**	**ILMN_2613140**	**A_52_P110877**
Oas2	NM_145227	ENSMUSG00000032690	1425065_at	ILMN_2670150	A_52_P587071
Oas3	NM_145226	ENSMUSG00000032661	1425374_at	ILMN_1216020	A_51_P472867
Pkr	NM_011163	ENSMUSG00000024079	1440866_at	ILMN_1250410	A_52_P559919
Plscr1	NM_011636	ENSMUSG00000032369	1429527_a_at	ILMN_2911344	A_52_P654841
Rig-i	**NM_172689**	**ENSMUSG00000040296**	**1456890_at**	**ILMN_2717127**	**A_52_P523946**
Sp100	NM_013673	ENSMUSG00000026222	1451821_a_at	ILMN_2846812	A_52_P127720
Sp110	NM_175397	ENSMUSG00000070034	1456493_at	ILMN_1214911	A_52_P512201
Stat1	**NM_009283**	**ENSMUSG00000026104**	**1420915_at**	**ILMN_2510233**	**A_52_P496503**
Stat2	NM_019963	ENSMUSG00000040033	1421911_at	ILMN_2657822	A_51_P225808
Tap1	NM_013683	ENSMUSG00000037321	1416016_at	ILMN_1250409	A_51_P100327
Tlr3	NM_126166	ENSMUSG00000031639	1422782_s_at	ILMN_2697002	A_52_P85174
Ube2l6	NM_019949	ENSMUSG00000027078	1417172_at	ILMN_2431619	A_52_P214740
Usp18	NM_011909	ENSMUSG00000030107	1418191_at	ILMN_2433990	A_51_P164219

Eight well-known ISG marker genes are highlighted in bold

Table 6
siRNA sequences and oligos for shRNA expression for deliberate induction of interferon response by small RNAs

siRNA (sense strand)	Reference
GUCCGGGCAGGUCUACUUUTT	(36)
AGCUUAACCUGUCCUUCAAdTdT	(35)
UGUCCUUCAAUGUCCUUCAA	(35)
CUACACAAAUCAGCGAUUU	(34)

shRNA (designed for TMP vector)	Reference
TGCTGTTGACAGTGAGCGCTACACAAATCAGCG ATTTTTTTAGTGAAGCCACAGATGTAAAAAAAT CGCTGATTTGTGTAGTGCCTACTGCCTCGGA	This work
TGCTGTTGACAGTGAGCGTGTCCTTCAATGTCCT TCAATTTAGTGAAGCCACAGATGTAAATTGAAGG ACATTGAAGGACATGCCTACTGCCTCGGA	This work

3.6. Deliberate Induction of Interferon Response

3.6.1. With Poly I:C

The production of interferons in a positive control sample can be induced by direct addition of the polyinosinic:polycytidylic acid (polyI:C) to the cultured cells. The suggested final concentration in culture media is 50 μg/mL of polyI:C. Add polyI:C to cells 6 h after transfection.

3.6.2. With Interferon-Inducing shRNA/siRNAs

As discussed above, some siRNAs may stimulate the interferon response in immune cells (see Table 6). Some siRNAs contain specific immunostimulatory motifs within single-stranded RNAs (ssRNA) from siRNAs, such as UGUGU and GUCCUUCAA, (34, 35), while others do not have defined sequence motifs (36). As positive control for interferon induction, we use two of these siRNAs expressed from pTMP vector (Fig. 4). Oligonucleotides used for production of pTMP vectors are listed in Table 6. For transfection, use the same protocol as above.

4. Notes

A good experiment should include appropriate controls, which would help to pinpoint the cause of a problem. The most common problems and their solutions are discussed below.

1. *Troubleshooting poor RNAi knockdown.* It can be caused by problems with the vector, its delivery, or inefficient siRNA.

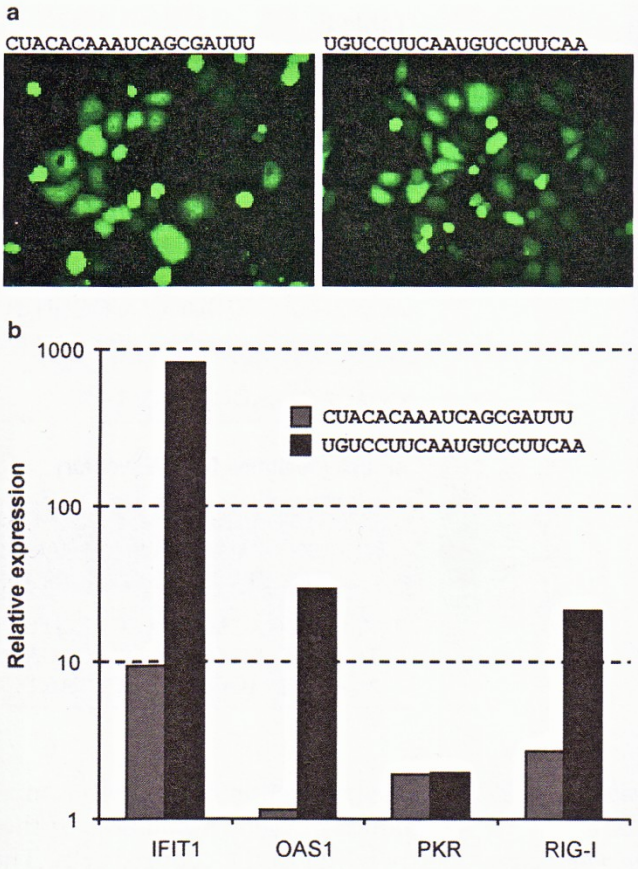


Fig. 4. Interferon response to immunostimulatory shRNA expressed from pTMP vectors. (**a**) EGFP expression from pTMP EGFP reporter in HeLa cells transfected with pTMP vectors expressing immunostimulatory shRNA (0.5 μg per well in a 6-well plate). (**b**) Induction of interferon response markers by expressing immunostimulatory shRNAs. HeLa cells were transfected with 0.5 μg of each plasmid pTMP plasmid derivate per well in a 6-well plate and relative interferon response marker expression was estimated by quantitative Real-Time RT-PCR. Expression of the markers in cells transfected with empty pTMP vector was set to 1.

- Check the quality of the vector DNA (electrophoresis and sequencing).
- Include a positive transfection control.
- Increase the amount of the shRNA vector in transfection.
- Assay at different time points (protein is downregulated later than mRNA).
- Try another siRNA sequence.

2. *Handling interferon response.* If the interferon response is detected in an RNAi experiment, the following troubleshooting procedure may help. Verify that the RNAi trigger and not the delivery or induction method cause the effect since delivery methods themselves may induce the interferon response.

A typical experiment should include appropriate controls, which would help to identify the cause. Changing the amount of the transfection reagent, using another transfection reagent, producing stable inducible cell lines, and titration of inducing and selecting agents to lower doses may help if the interferon response is caused by other factors than the RNAi trigger itself.

If the interferon response is detected in an RNAi experiment, the following troubleshooting procedure may help. Verify that the RNAi trigger and not the delivery or induction method cause the effect since delivery methods themselves may induce the interferon response. A typical experiment should include appropriate controls, which would help to identify the cause. Changing the amount of the transfection reagent, using another transfection reagent, producing stable inducible cell lines, and titration of inducing and selecting agents to lower doses may help if the interferon response is caused by other factors than the RNAi trigger itself.

If the RNAi trigger seems to be causing the effect and it is an siRNA, the following solutions may work. If the siRNA is from a commercial source, titrate the amount of the siRNA to the lowest effective dose first and/or use an earlier time-point for analysis. For example, the interferon response may be observed at 72 h after transfection but not at 24 or 48 h post-transfection (28). If one needs rather late time-points, it is worth considering the production of inducible stable cell lines. The purity of the siRNA is also very important and changing the manufacturer may help in some cases. Poor quality of chemically synthesized siRNAs may induce the interferon response (29). Also, avoid motifs that stimulate the interferon response discussed above (see Subheading 1.3). Sometimes, switching to another cell line can help, as different cell lines have different sensitivity to the interferon activation (54). As mentioned above, if the siRNAs are home-made by T7 polymerase, treatment of siRNAs with RNAse T1 and alkaline phosphatase to remove 5′ triphosphate GTP should help (30). If the problem still persists, it is probably more economical to purchase siRNAs from one of the established sources rather than spending time and financial resources to solve the problem. If the problem appears to be linked to an shRNA expressed from a vector controlled by the U6 or H1 promoter, it is advisable to verify that the vector sequence does not contain the critical AA dinucleotide motif near the transcription start site (33). If that is not the case, one should try using another siRNA sequence and/or expression vector (e.g., switching from the type I to the type II hairpin).

3. *Dealing with off-targeting.* A commentary by Echeverri et al. (42) summarizing the guidelines for appropriate controls in RNAi experiments offers two ways to address off-target effects in RNAi experiments named "the two R's": rescue or redundancy.

Rescue experiments are performed by delivering expression of a functional version of a targeted gene, which is mutated such that the base-pairing with a short RNA is disrupted. If the short RNA is targeting the 3′UTR, it should not be a problem to appropriately replace the 3′UTR. If the short RNA is targeting the coding sequence, the situation is little bit more complicated because one has to mutate/degenerate numerous positions within codons of the target sequence. It is important to consider codon usage so as not to introduce more rare codons than necessary. One or two nucleotides should be mutated in the middle of the sequence to impair Ago2-mediated cleavage. In addition, it is useful to mutate one nucleotide in a position complementary to nucleotides 2–4 of the binding siRNA sequence to interfere with recognition of the sequence by RISC.

Redundancy is another way to address off-target effects. Two or more RNA triggers with distinct sequences producing the same phenotype decrease the probability that the phenotype is caused by off-targeting. However, some phenotypes are fairly common (e.g., slower growth, apoptosis, developmental arrest), so using two or even more different RNAi inducers may not be enough to decipher if the phenotype is specific to the downregulated gene or not. In such a case, a rescue experiment is highly advisable. When microarray analysis is being performed on knockdown cells, one can estimate the extent of off-targeting on the cellular transcriptome for each short RNA used. Similarly to analysis of hexamer and heptamer seed enrichment among upregulated transcripts upon repression of RNA silencing (55), one can test if the seeding motif corresponding to the 5′ end of an active siRNA strand is significantly enriched in 3′UTRs of transcripts downregulated in knockdown cells.

It must be pointed out that "nontargeting" controls (scrambled siRNAs or shRNAs, or short RNAs against nonexpressed genes such as EGFP, RL-luciferase, etc.) are not appropriate controls for off-targeting for reasons mentioned above. It is a common misconception that ignores the fact that off-targeting is individual to each RNAi trigger because it is sequence-specific. "Nontargeting" RNAi triggers rather serve as controls for the sequence-independent effects, such as interferon response and saturation of RNA silencing with an excess of exogenous short RNAs.

Acknowledgements

We thank Witold Filipowicz group at the FMI for sharing their experience and protocols and Daniela Schmitter, Radek Malik, and Lenka Sarnova for help with preparation of the manuscript.

Protocols are based on research supported by the GAAV grant IAA 501110701, EMBO SDIG grant 2006–1483, GACR grant 204/09/0085, and the Purkynje Fellowship.

References

1. Sontheimer, E. J., Carthew, R. W. (2005) Silence from within: endogenous siRNAs and miRNAs. *Cell* **122:** 9–12.
2. Zamore, P. D., Haley, B. (2005) Ribo-gnome: the big world of small RNAs. *Science* **309:** 1519–1524.
3. Filipowicz, W. (2005) RNAi: the nuts and bolts of the RISC machine. *Cell* **122:** 17–20.
4. Tolia, N. H., Joshua-Tor, L. (2007) Slicer and the argonautes. *Nat. Chem. Biol.* **3:** 36–43.
5. Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., Tuschl, T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411:** 494–498.
6. Manche, L., Green, S. R., Schmedt, C., Mathews, M. B. (1992) Interactions between double-stranded RNA regulators and the protein kinase DAI. *Mol. Cell. Biol.* **12:** 5238–5248.
7. Svoboda, P. (2008) RNA silencing in mammalian oocytes and early embryos. *Curr. Top. Microbiol. Immunol.* **320:** 225–256.
8. Mineno, J., Okamoto, S., Ando, T., Sato, M., Chono, H., Izu, H., Takayama, M., Asada, K., Mirochnitchenko, O., Inouye, M., Kato, I. (2006) The expression profile of microRNAs in mouse embryos. *Nucleic Acids Res.* **34:** 1765–1771.
9. Cummins, J. M., He, Y., Leary, R. J., Pagliarini, R., Diaz, L. A., Jr., Sjoblom, T., Barad, O., Bentwich, Z., Szafranska, A. E., Labourier, E., Raymond, C. K., Roberts, B. S., Juhl, H., Kinzler, K. W., Vogelstein, B., Velculescu, V. E. (2006) The colorectal microRNAome. *Proc. Natl. Acad. Sci. U S A* **103:** 3687–3692.
10. Liu, J., Carmell, M. A., Rivas, F. V., Marsden, C. G., Thomson, J. M., Song, J. J., Hammond, S. M., Joshua-Tor, L., Hannon, G. J. (2004) Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305:** 1437–1441.
11. Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., Tuschl, T. (2004) Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cells* **15:** 185–197.
12. Pillai, R. S., Artus, C. G., Filipowicz, W. (2004) Tethering of human AGO proteins to mRNA mimics the miRNA-mediated repression of protein synthesis. *RNA - Publ. RNA Soc.* **10:** 1518–1525.
13. Song, J. J., Smith, S. K., Hannon, G. J., Joshua-Tor, L. (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305:** 1434–1437.
14. Doench, J. G., Petersen, C. P., Sharp, P. A. (2003) siRNAs can function as miRNAs. *Genes Dev.* **17:** 438–442.
15. Hutvagner, G., Zamore, P. D. (2002) A microRNA in a multiple-turnover RNAi enzyme complex. *Science* **297:** 2056–2060.
16. Yekta, S., Shih, I. H., Bartel, D. P. (2004) MicroRNA-directed cleavage of HOXB8 mRNA. *Science* **304:** 594–596.
17. Brummelkamp, T. R., Bernards, R., Agami, R. (2002) A system for stable expression of short interfering RNAs in mammalian cells. *Science* **296:** 550–553.
18. McManus, M. T., Petersen, C. P., Haines, B. B., Chen, J., Sharp, P. A. (2002) Gene silencing using micro-RNA designed hairpins. *RNA - Publ. RNA Soc.* **8:** 842–850.
19. Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J., Conklin, D. S. (2002) Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev.* **16:** 948–958.
20. van de Wetering, M., Oving, I., Muncan, V., Pon Fong, M. T., Brantjes, H., van Leenen, D., Holstege, F. C., Brummelkamp, T. R., Agami, R., Clevers, H. (2003) Specific inhibition of gene expression using a stably integrated, inducible small-interfering-RNA vector. *EMBO Rep.* **4:** 609–615.
21. Chung, K. H., Hart, C. C., Al-Bassam, S., Avery, A., Taylor, J., Patel, P. D., Vojtek, A. B., Turner, D. L. (2006) Polycistronic RNA polymerase II expression vectors for RNA interference based on BIC/miR-155. *Nucleic Acids Res.* **34**: e53.
22. Zeng, Y., Wagner, E. J., Cullen, B. R. (2002) Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol. Cells* **9:** 1327–1333.
23. de Veer, M. J., Sledz, C. A., Williams, B. R. (2005) Detection of foreign RNA: implications for RNAi. *Immunol. Cell Biol.* **83:** 224–228.
24. Geiss, G., Jin, G., Guo, J., Bumgarner, R., Katze, M. G., Sen, G. C. (2001) A comprehensive view of regulation of gene expression by

double-stranded RNA-mediated cell signaling. *J. Biol. Chem.* **276:** 30178–30182.

25. Schlee, M., Hornung, V., Hartmann, G. (2006) siRNA and isRNA: two edges of one sword. *Mol. Ther.* **14:** 463–470.
26. Sledz, C. A., Holko, M., de Veer, M. J., Silverman, R. H., Williams, B. R. (2003) Activation of the interferon system by short-interfering RNAs. *Nat. Cell Biol.* **5:** 834–839.
27. Bridge, A. J., Pebernard, S., Ducraux, A., Nicoulaz, A. L., Iggo, R. (2003) Induction of an interferon response by RNAi vectors in mammalian cells. *Nature Genet.* **34:** 263–264.
28. Svoboda, P. (2007) Off-targeting and other non-specific effects of RNAi experiments in mammalian cells. *Curr. Opin. Mol. Ther.* **9:** 248–257.
29. Marques, J. T., Devosse, T., Wang, D., Zamanian-Daryoush, M., Serbinowski, P., Hartmann, R., Fujita, T., Behlke, M. A., Williams, B. R. (2006) A structural basis for discriminating between self and nonself double-stranded RNAs in mammalian cells. *Nat. Biotechnol.* **24:** 559–565.
30. Kim, D. H., Longo, M., Han, Y., Lundberg, P., Cantin, E., Rossi, J. J. (2004) Interferon induction by siRNAs and ssRNAs synthesized by phage polymerase. *Nat. Biotechnol.* **22:** 321–325.
31. Hornung, V., Ellegast, J., Kim, S., Brzozka, K., Jung, A., Kato, H., Poeck, H., Akira, S., Conzelmann, K. K., Schlee, M., Endres, S., Hartmann, G. (2006) 5'-Triphosphate RNA is the ligand for RIG-I. *Science* **314:** 994–997.
32. Pichlmair, A., Schulz, O., Tan, C. P., Naslund, T. I., Liljestrom, P., Weber, F., Reis e Sousa, C. (2006) RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates. *Science* **314:** 997–1001.
33. Pebernard, S., Iggo, R. D. (2004) Determinants of interferon-stimulated gene induction by RNAi vectors. *Differentiation* **72:** 103–111.
34. Judge, A. D., Sood, V., Shaw, J. R., Fang, D., McClintock, K., MacLachlan, I. (2005) Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA. *Nat. Biotechnol.* **23:** 457–462.
35. Hornung, V., Guenthner-Biller, M., Bourquin, C., Ablasser, A., Schlee, M., Uematsu, S., Noronha, A., Manoharan, M., Akira, S., de Fougerolles, A., Endres, S., Hartmann, G. (2005) Sequence-specific potent induction of IFN-alpha by short interfering RNA in plasmacytoid dendritic cells through TLR7. *Nat. Med.* **11:** 263–270.
36. Sioud, M. (2005) Induction of inflammatory cytokines and interferon responses by double-stranded and single-stranded siRNAs is sequence-dependent and requires endosomal localization. *J. Mol. Biol.* **348:** 1079–1090.
37. Marques, J. T., Williams, B. R. (2005) Activation of the mammalian immune system by siRNAs. *Nat. Biotechnol.* **23:** 1399–1405.
38. Jackson, A. L., Bartz, S. R., Schelter, J., Kobayashi, S. V., Burchard, J., Mao, M., Li, B., Cavet, G., Linsley, P. S. (2003) Expression profiling reveals off-target gene regulation by RNAi. *Nat. Biotechnol.* **21:** 635–637.
39. Jackson, A. L., Burchard, J., Schelter, J., Chau, B. N., Cleary, M., Lim, L., Linsley, P. S. (2006) Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA - Publ. RNA Soc.* **12:** 1179–1187.
40. Scacheri, P. C., Rozenblatt-Rosen, O., Caplen, N. J., Wolfsberg, T. G., Umayam, L., Lee, J. C., Hughes, C. M., Shanmugam, K. S., Bhattacharjee, A., Meyerson, M., Collins, F. S. (2004) Short interfering RNAs can induce unexpected and divergent changes in the levels of untargeted proteins in mammalian cells. *Proc. Natl. Acad. Sci. U S A* **101:** 1892–1897.
41. Brown, K., Samarsky, D. (2005) RNAi off-targeting: Light at the end of the tunnel. *J. RNAi Gene Silencing* **2:** 175–177.
42. Echeverri, C. J., Beachy, P. A., Baum, B., Boutros, M., Buchholz, F., Chanda, S. K., Downward, J., Ellenberg, J., Fraser, A. G., Hacohen, N., Hahn, W. C., Jackson, A. L., Kiger, A., Linsley, P. S., Lum, L., Ma, Y., Mathey-Prevot, B., Root, D. E., Sabatini, D. M., Taipale, J., Perrimon, N., Bernards, R. (2006) Minimizing the risk of reporting false positives in large-scale RNAi screens. *Nat. Methods* **3:** 777–779.
43. Snove, O., Jr., Holen, T. (2004) Many commonly used siRNAs risk off-target activity. *Biochem. Biophys. Res. Commun.* **319:** 256–263.
44. Jackson, A. L., Burchard, J., Leake, D., Reynolds, A., Schelter, J., Guo, J., Johnson, J. M., Lim, L., Karpilow, J., Nichols, K., Marshall, W., Khvorova, A., Linsley, P. S. (2006) Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing. *RNA - Publ. RNA Soc.* **12:** 1197–1205.
45. Birmingham, A., Anderson, E., Sullivan, K., Reynolds, A., Boese, Q., Leake, D., Karpilow, J., Khvorova, A. (2007) A protocol for designing siRNAs with high functionality and specificity. *Nat. Protoc.* **2:** 2068–2078.
46. Peek, A. S., Behlke, M. A. (2007) Design of active small interfering RNAs. *Curr. Opin. Mol. Ther.* **9:** 110–118.
47. Huesken, D., Lange, J., Mickanin, C., Weiler, J., Asselbergs, F., Warner, J., Meloon, B., Engel, S., Rosenberg, A., Cohen, D., Labow, M., Reinhardt, M., Natt, F., Hall, J. (2005) Design of a genome-wide siRNA library using an artificial neural network. *Nat. Biotechnol.* **23:** 995–1001.

48. Tafer, H., Ameres, S. L., Obernosterer, G., Gebeshuber, C. A., Schroeder, R., Martinez, J., Hofacker, I. L. (2008) The impact of target site accessibility on the design of effective siRNAs. *Nat. Biotechnol.* **26:** 578–583.
49. Chalk, A. M., Sonnhammer, E. L. L. (2008) siRNA specificity searching incorporating mismatch tolerance data. *Bioinformatics* **24:** 1316–1317.
50. Matveeva, O., Nechipurenko, Y., Rossi, L., Moore, B., Saetrom, P., Ogurtsov, A. Y., Atkins, J. F., Shabalina, S. A. (2007) Comparison of approaches for rational siRNA design leading to a new efficient and transparent method. *Nucleic Acids Res.* **35:** e63.
51. Ameres, S. L., Martinez, J., Schroeder, R. (2007) Molecular basis for target RNA recognition and cleavage by human RISC. *Cell* **130:** 101–112.
52. Vert, J. P., Foveau, N., Lajaunie, C., Vandenbrouck, Y. (2006) An accurate and interpretable model for siRNA efficacy prediction. *BMC Bioinformatics* **7:** 520.
53. Sambrook, J., Fritsch, E. F., Maniatis, T. (1989) *Molecular cloning: a laboratory manual.* 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
54. Reynolds, A., Anderson, E. M., Vermeulen, A., Fedorov, Y., Robinson, K., Leake, D., Karpilow, J., Marshall, W. S., Khvorova, A. (2006) Induction of the interferon response by siRNA is cell type- and duplex length-dependent. *RNA - Publ. RNA Soc.***12:** 988–993.
55. Schmitter, D., Filkowski, J., Sewer, A., Pillai, R. S., Oakeley, E. J., Zavolan, M., Svoboda, P., Filipowicz, W. (2006) Effects of Dicer and Argonaute down-regulation on mRNA levels in human HEK293 cells. *Nucleic Acids Res.* **34:** 4801–4815.
56. Zeng, Y., Cai, X., Cullen, B. R. (2005) Use of RNA polymerase II to transcribe artificial microRNAs. *Methods Enzymol.* **392:** 371–380.
57. Abe, Y., Sinozaki, H., Takagi, T., Minegishi, T., Kokame, K., Kangawa, K., Uesaka, M., Miyamoto, K. (2006) Identification of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD)-inducible genes in human amniotic epithelial cells. *Reprod. Biol. Endocrinol.* **4:** 27.
58. Hu, X., Herrero, C., Li, W. P., Antoniv, T. T., Falck-Pedersen, E., Koch, A. E., Woods, J. M., Haines, G. K., Ivashkiv, L. B. (2002) Sensitization of IFN-gamma Jak-STAT signaling during macrophage activation. *Nat. Immunol.* **3:** 859–866.
59. Kato, A., Homma, T., Batchelor, J., Hashimoto, N., Imai, S., Wakiguchi, H., Saito, H., Matsumoto, K. (2003) Interferon-alpha/beta receptor-mediated selective induction of a gene cluster by CpG oligodeoxynucleotide 2006. *BMC Immunol.* **4:** 8.
60. Kamezaki, K., Shimoda, K., Numata, A., Matsuda, T., Nakayama, K., Harada, M. (2004) The role of Tyk2, Stat1 and Stat4 in LPS-induced endotoxin signals. *Int. Immunol.* **16:** 1173–1179.
61. Sakamoto, S., Qin, J., Navarro, A., Gamero, A., Potla, R., Yi, T., Zhu, W., Baker, D. P., Feldman, G., Larner, A. C. (2004) Cells Previously Desensitized to Type 1 Interferons Display Different Mechanisms of Activation of Stat-dependent Gene Expression from Naive Cells. *J. Biol. Chem.* **279:** 3245–3253.
62. Wang, Z. W., Sarmento, L., Wang, Y., Li, X.-q., Dhingra, V., Tseggai, T., Jiang, B., Fu, Z. F. (2005) Attenuated Rabies Virus Activates, while Pathogenic Rabies Virus Evades, the Host Innate Immune Responses in the Central Nervous System. *J. Virol.* **79:** 12554–12565.
63. Stewart, M. J., Smoak, K., Blum, M. A., Sherry, B. (2005) Basal and Reovirus-Induced Beta Interferon (IFN-{beta}) and IFN-{beta}-Stimulated Gene Expression Are Cell Type Specific in the Cardiac Protective Response. *J. Virol.* **79:** 2979–2987.
64. Pauls, E., Senserrich, J., Bofill, M., Clotet, B., Este, J. A. (2007) Induction of interleukins IL-6 and IL-8 by siRNA. *Clin. Exp. Immunol.* **147:** 189–196.
65. Yang, R., Murillo, F. M., Cui, H., Blosser, R., Uematsu, S., Takeda, K., Akira, S., Viscidi, R. P., Roden, R. B. S. (2004) Papillomavirus-Like Particles Stimulate Murine Bone Marrow-Derived Dendritic Cells To Produce Alpha Interferon and Th1 Immune Responses via MyD88. *J. Virol.* **78:** 11152–11160.
66. Bolcato-Bellemin, A.-L., Bonnet, M.-E., Creusat, G., Erbacher, P., Behr, J.-P. (2007) Sticky overhangs enhance siRNA-mediated gene silencing. *Proc. Natl. Acad. Sci. USA* **104:** 16050–16055.
67. Cekaite, L., Furset, G., Hovig, E., Sioud, M. (2007) Gene expression analysis in blood cells in response to unmodified and 2'-modified siRNAs reveals TLR-dependent and independent effects. *J. Mol. Biol.* **365:** 90–108.
68. Persengiev, S. P., Zhu, X., Green, M. R. (2004) Nonspecific, concentration-dependent stimulation and repression of mammalian gene expression by small interfering RNAs (siRNAs). *RNA - Publ. RNA Soc.* **10:** 12–18.

Author's contribution

Supplement 1:

I performed the literature search regarding articles utilizing dsRNA and participated in the writing of the manuscript.

Supplement 2:

I performed the complete analysis of previously generated MosIR mouse transgene and Mos3 x MosIR hybrids. I analyzed pCagEGFP-MosIR-transfected cells, except microarray analysis and western blotting. I was involved in deep sequencing data analysis and I participated in the writing of the manuscript.

Supplement 3:

I generated most of transfection and flow cytometry data and I was involved in the analysis of deep sequencing data. I participated in the writing of the manuscript.

Supplement 4:

I generated most of the data with a shared contribution to cloning, generation of stable lines, and analysis of deep sequencing data. I participated in the writing of the manuscript.

Supplement 5:

I participated in the writing of the manuscript, especially sections 5.2.5. and 5.3.

Supplement 6:

I participated in the writing of the manuscript and I generated a large part of data in the methodical section.

List of publications

Nejepinska J, Malik R, Filkowski J, Flemr M, Filipowicz W, Svoboda P. dsRNA expression in the mouse elicits RNAi in oocytes and low adenosine deamination in somatic cells. Nucleic Acids Res. 2012 Jan; 40(1):399-413.
(IF = 8.026)

Nejepinska J, Malik R, Moravec M, Svoboda P. Deep sequencing reveals complex spurious transcription from transiently transfected plasmids. PLoS One. 2012; 7(8):e43283.
(IF = 4.092)

Chalupnikova K, Nejepinska J, Svoboda P. Production and application of long dsRNA in mammalian cells. Methods Mol Biol. 2013; 942:291-314.
(Book chapter)

Nejepinska J, Flemr, Svoboda P. The canonical RNA interference pathway in animals. Regulatory RNAs. 2012; 111-150.
(Book chapter)

Nejepinska J, Flemr M, Svoboda P. Control of the interferon response in RNAi experiments. Methods Mol Biol. 2012; 820:133-61.
(Book chapter)

Nejepinska J, Malik R, Svoboda P. Expressed dsRNA provokes selective sequence-independent PKR response in transient co-transfection experiments.
(Manuscript attached)

Flemr M, Malik R, Franke V, Nejepinska J, Beck IM, Vlahovicek K, Svoboda P. A retrotransposon-driven variant of mouse Dicer fuels endogenous siRNA production.
(Submitted)