

AD _____

Award Number: DAMD17-02-1-0729

TITLE: Use of Novel, Stable Gene Silencing Technology to
Determine the Contribution of the Receptor

PRINCIPAL INVESTIGATOR: Douglas S. Conklin, Ph.D.

CONTRACTING ORGANIZATION: Cold Spring Harbor Laboratory
Cold Spring Harbor, New York 11724

REPORT DATE: October 2003

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGEForm Approved
OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE October 2003	3. REPORT TYPE AND DATES COVERED Annual (30 Sep 2002 - 29 Sep 2003)
--	---------------------------------------	---

4. TITLE AND SUBTITLE Use of Novel, Stable Gene Silencing Technology to Determine the Contribution of the Receptor	5. FUNDING NUMBERS DAMD17-02-1-0729
--	---

6. AUTHOR(S) Douglas S. Conklin, Ph.D.	20040217 030
--	---------------------

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Cold Spring Harbor Laboratory Cold Spring Harbor, New York 11724 <i>E-Mail:</i> conklin@cshl.edu	8. PERFORMING ORGANIZATION REPORT NUMBER
---	---

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012	10. SPONSORING / MONITORING AGENCY REPORT NUMBER
--	---

11. SUPPLEMENTARY NOTES Original contains color plates: All DTIC reproductions will be in black and white.
--

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited	12b. DISTRIBUTION CODE
--	-------------------------------

13. ABSTRACT (Maximum 200 Words) An enormous amount of evidence has implicated members of the receptor protein tyrosine kinase (RPTK) family in many types of cancer. We have developed an extraordinarily powerful technique that uses small, DNA-encoded RNAs to suppress the expression of a desired gene in mammalian cells. For this project, short hairpin RNAs targeting each of the 58 human receptor protein tyrosine kinases will be created, introduced into MCF-7 cells and tested for effects on tumorigenicity in vivo and in vitro. At the end of the initial twelve months of the project there has been good progress. The goal of year one was to create a set of encoded hairpins targeted against the genomic complement of RPTKs. This goal has been met. By adopting high-throughput approaches to the construction of hairpins, we have created a set of silencing agents not only for the genomic complement of receptor tyrosine kinases, but also all tyrosine kinases and a large number of genes functionally related to the RPTKs. Other developments include the use of shRNAs in the production of transgenic animals and microarray-based shRNA delivery for phenotypic screening. Both of these are potentially useful in future studies of the tyrosine kinase targeting constructs.
--

14. SUBJECT TERMS Functional genomics, cell signaling, gene therapy, RNAI	15. NUMBER OF PAGES 41
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited
--	---	--	--

Table of Contents

Cover.....1

SF 298.....2

Table of Contents.....3

Introduction.....4

Body.....5

Key Research Accomplishments.....10

Reportable Outcomes.....11

Conclusions.....14

References.....15

Appendices.....16

Introduction

Gain-of-function mutations and the amplification of genes related to proliferation are universal components of cancer. Generally, these mutations account for the most profound phenotypic dissimilarities between cancer cells and the normal cells. As such the affected genes and gene products are the most useful possible targets for specific chemotherapeutics. Among the genes most often overexpressed in cancer are the receptor protein tyrosine kinases (RPTKs) (1). An enormous amount of evidence has implicated these proteins in many types of cancer. The HER-2/neu/ErbB-2 is an important target for a variety of second generation chemotherapeutics. A growing body of evidence has implicated other receptor protein tyrosine kinases in breast cancer. We have developed an extraordinarily powerful technique for suppressing gene expression in mammalian cells (for review see (2)). This technique uses small, DNA-encoded RNAs to suppress the expression of a desired gene, stably and inexpensively even in a diploid organism. DNA encoded short hairpin RNAs targeting each of the 58 human receptor protein tyrosine kinases will be cloned into mammalian expression vectors and introduced into MCF-7 cells. The effects of the constructs on RPTK expression, cell viability and proliferation, and estrogen dependence will be tested. Hairpin constructs that affect cellular correlates of tumorigenicity will be tested further for effects in three-dimensional culture models. Testing of these promising constructs in subcutaneous and orthotopic xenograft models will also be carried out. Hairpin RNAs that cause lethality or reversal of breast cancer-related phenotypes specifically in cancer cells will be identified as useful tools and potential therapeutics.

Body

There has been good progress on the proposed work for completion by the end of twelve months. The major goal of this time period was: **Creation of a set of encoded hairpins targeted against the genomic complement of RPTKs**. This goal has been met. As described below, by adopting high-throughput approaches to the construction of hairpins we have created a set of silencing agents not only for the genomic complement of receptor tyrosine kinases, but also all tyrosine kinases and a large number of genes functionally related to the RPTKs.

The year 1 aims fit in the overall objective of the proposal as filling a pool with shRNA constructs from which it will be possible to identify a small number that have inhibitory effects on breast cancer progression *in vivo* models. This would add to our understanding of the genes involved in breast cancer and identify their products as potential small molecule therapeutic targets. It would also pave the way for testing of the hairpins that we identify as gene therapeutics. It is not our intent to carry out a detailed investigation of the phenotypes that result from expression of a particular hairpin. Although undoubtedly interesting, it is beyond the scope of this proposal. Rather it is hoped that the hairpin constructs and resulting cell lines will succeed in identifying subjects for future study and serve as a resource for other investigators in the field. The first step of this was the generation of a large number silencing constructs to fill the pool.

Months 1-12 of the proposal were summarized in the original Statement of Work which is given below:

Creation of a set of encoded hairpins targeted against the genomic complement of RPTKs

1. Synthesis
2. Subcloning to recipient vector –sequence confirmation
3. Transfer to expression vector
4. Optimization of delivery to breast cancer cells
5. Immediate analysis of hairpins targeting RPTKs with known breast cancer phenotypes

Each task from the Statement of Work is discussed below:

1. Synthesis

The rate-limiting step in shRNA-based silencing has been the construction of the actual hairpin plasmids and their transfer into appropriate cell types. We had already made use of high throughput methodologies to significantly accelerate several of the steps involved. Perl-based hairpin oligo design scripts (available at the website: <http://katahdin.cshl.org:9331/RNAi>), and large scale oligo synthesis were already in place.

The first step in creating the set of DNA encoded short hairpin RNAs was to assemble all the accession numbers of the RPTKs (Table 1) . The list was processed by using the PERL scripts to generate potential shRNA sequences for each gene. At first three shRNA designs were obtained for each gene. Since little is currently understood regarding the variability in suppression strength for hairpins that target the same gene, we typically construct at least three hairpins for each gene that we are interested in silencing. Some hairpins cause a nearly 100% suppression of the cognate gene whereas others may only suppress 40% and others not at all. Although this results in more work, it also allows for the generation of hypomorphic alleles of varying severity which in some instances may be useful. For example complete suppression of an essential gene is likely to be lethal, whereas 40% suppression of an essential gene is likely to yield a milder phenotype.

Table 1. Accession numbers for the tyrosine kinases used in shRNA design.

NM_005157	NM_002110	NM_004431	NM_005211
NM_005158	NM_005365	NM_005233	NM_004119
NM_005781	NM_002350	NM_004438	NM_000222
NM_003985	NM_001721	L36644	NM_006206
NM_004383	NM_000061	NM_173655	NM_002609
NM_002378	NM_005546	NM_004440	NM_002821
NM_005607	NM_003215	AB040892	X12949
NM_004103	NM_003328	NM_004441	NM_005012
NM_005246	NM_003177	AF025304	NM_004560
NM_002005	L05148	NM_004443	NM_002944
NM_005975	NM_004304	NM_004444	S59184
NM_002031	NM_002344	NM_004445	NM_000459
NM_080823	NM_001699	M34641	NM_005424
NM_002227	NM_006343	NM_000141	NM_002529
NM_004972	NM_006293	NM_000142	NM_006180
NM_000215	NM_013993	NM_002011	NM_002530
NM_003331	NM_006182	NM_000875	NM_002019
NM_005248	NM_005228	NM_000208	NM_002020
NM_002037	NM_004448	J05046	NM_004920
NM_005417	NM_001982	NM_000245	NM_014916
NM_005433	NM_005235	NM_002447	NM_018423
NM_001715	NM_005232	NM_005592	

The shRNA targeting sequences for each gene were picked based upon a number of constraints. Only coding sequences were targeted, and each target site was chosen for maximum uniqueness in the human genome (>3* mismatches to any other sequence in the database) and, where possible, for sequence identity to the mouse ortholog of the targeted gene. Target sites with runs >3 T residues were eliminated to avoid premature termination by RNA polymerase III and 2-3 G:U base pairs were incorporated into the predicted stem to stabilize inverted DNA repeats in bacteria. Short hairpin RNA sequences were converted into a single 72nt primer sequence onto which are added 21nt of homology to the 3' end of the human U6 snRNA promoter.

Although hairpin primer sequences were directly ported and synthesized on a Mermade V oligonucleotide synthesizer (Bioautomation), it became apparent that for reasons related to reliability and production rate, that it was more effective to purchase oligos from commercial suppliers.

- 2. Subcloning to recipient vector –sequence confirmation*
- 3. Transfer to expression vector*

In the initial proposal, a PCR-based strategy was to be used to produce shRNA constructs. PCR reactions were carried out in 96-well format. The original plan called for the resulting fragments to be cloned directly into pENTR/D-topo (Invitrogen) and subsequently transferred to a recipient retroviral vector containing a Gateway acceptor site. The cost and poor reliability of the Gateway system made this an unreliable approach.

Instead we opted for a strategy in which the shRNA encoding PCR products were directly introduced into a recipient retroviral vector. Retrovirus-based vectors are generally favored for delivery of hairpins into mammalian cells (2). We find that mouse stem cell virus-based retroviral vectors are the most efficient method of gene transfer for the stable integration and expression of genetic constructs in mammalian cells. We also have significant experience with these types

of vectors (3).

shRNA's targeting the RPTKs have been constructed in a vector that contains a number of convenient design features (Figure 1). The vector is cable of producing self-inactivating Murine stem cell virus (MSCV) particles by transfection into commonly available retroviral packaging lines (eg Phoenix, LinX cells*). The vector also incorporates a convenient system for transferring shRNA expression cassettes to custom-designed or alternative vector systems without the need for *in vitro* manipulation. This relies on a new method called Mating Assisted Genetically Integrated Cloning, MAGIC designed in the laboratory of Steve Elledge, Harvard University. The MAGIC system consists of a donor vector in which the fragment of interest is flanked by two 50 bp sequences, termed homology region 1 (H1) and 2 (H2). Flanking H1 and H2 are sites for the rare restriction endonuclease I-SceI (4). The donor vector also includes an F' origin and a conditional origin of replication (RK6). The recipient vector, which contains compatible H1 and H2 sequences flanking two I-SceI sites surrounding a negative selectable marker (pheS) resides in a bacterial strain that contains an inducible I-SceI gene. Transfer of the fragment of interest into a recipient vector is accomplished by bacterial mating. Following F'-mediated transfer of the donor vector into the recipient host, induction of I-SceI cleaves both donor and recipient vectors, and these breaks are healed by homologous recombination via the H1 and H2 sequences in the donor and recipient plasmids. Selection against the unrecombined recipient containing pheS and I-SceI sites and for the capture of the appropriate insert (chloramphenicol resistance) give essentially 100% recovery of the desired plasmid. A Lentivirus MAGIC recipient vector based on the FUW vector has been developed (5).

A schematic of the vector housing the tyrosine kinase targeting clones is shown in Figure 1.

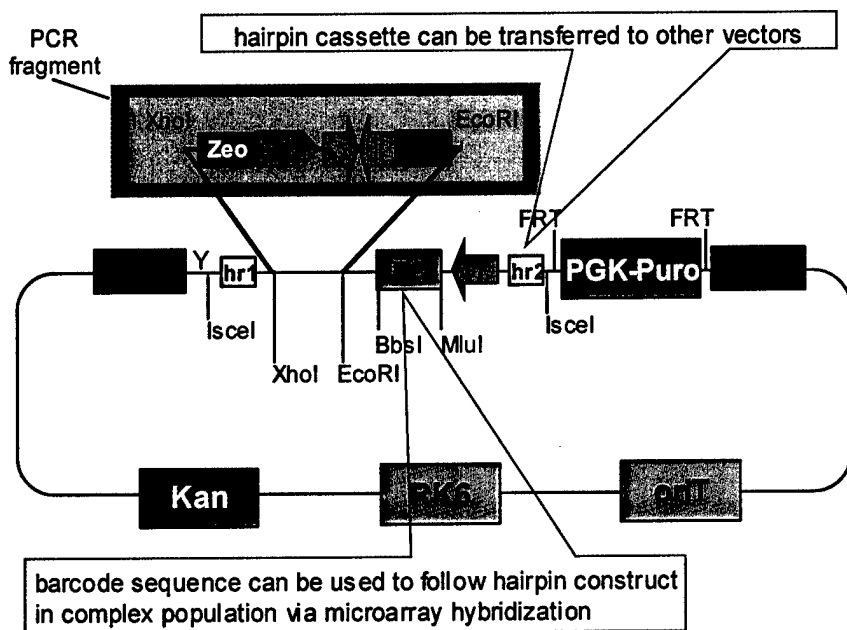


Figure 1. The revised vector for the construction of receptor tyrosine kinase targeting shRNAs. PCR fragments generated from oligos designed to encode shRNAs were cloned into the XhoI and EcoRI sites of the vector.

In this new strategy, shRNA encoding PCR products were directly introduced into the retroviral vector. Although this was fairly straightforward, it was apparent in pilot experiments that the frequency of obtaining successful clones in this manner was relatively low. For this reason we adopted a strategy in which every clone was subjected to DNA sequence analysis. This made possible the pooling of PCR fragments prior to ligation to vector. PCR fragments from four 96 well plates were pooled, digested with EcoRI and XhoI and ligated to vector. Several hundred clones were subjected to high throughput confirmatory sequencing carried out in the CSHL core

facility. Correct clones were identified by sequence analysis and isolated. In this way 295 shRNA vectors targeting tyrosine kinases were constructed. Bacteria harboring each plasmid were frozen in glycerol and stored in an archive. A sample of the clones that were produced is shown in Table 2. A complete list of clones generated to this point are given in the appendix. At least three clones have been produced for most of tyrosine kinases.

GENE	TARGETED SEQUENCE	PLATE	WELL
Homo sapiens EphB4 (EPHB4), mRNA.	CACCTTGATGTAGGGGTTCTCCATCCAGG	172840	G2
Homo sapiens EphB4 (EPHB4), mRNA.	CGGTAGTCCTGTTCATGGCATTGATCAC	172840	H2
Homo sapiens EphB4 (EPHB4), mRNA.	CGTCTTGACATTCACCTCCCGGTGGCCT	312940	C12
Homo sapiens feline sarcoma oncogene (FES), mRNA.	AAGCCTTGGTACTCAGCCTCAGGCTGGAT	173040	C2
Homo sapiens feline sarcoma oncogene (FES), mRNA.	ACAGCACCGACAGCACGTACTCCTGCTTG	316140	A11
Homo sapiens feline sarcoma oncogene (FES), mRNA.	TTCATGCCCTCCAGTAGACGAGCTCGG	317840	E7

Table 2. An example of information contained on clones constructed targeting the tyrosine kinases. Greater than three oligos were synthesized for each gene. The targeted sequence within each gene is shown. Plate and well refer to a bar-coded plate address in the hairpin archive of each clone.

4. Optimization of delivery to breast cancer cells

The infection of breast cancer cells required no optimization. Several manuscripts have appeared detailing efficient infection of MCF-7 cells with retroviruses (6, 7) which are identical to our typical protocols (3). On the second day after the transfection of packaging cells, the culture medium is replaced with fresh medium and the cells incubated overnight. The viral supernatants are harvested and then incubated with MCF-7 cells in the presence of 8 μ g/ml polybrene. After another overnight incubation, the viral supernatants are replaced with fresh culture medium and cells incubated for a further 48 h.

5. Immediate analysis of hairpins targeting RPTKs with known breast cancer phenotypes

The archival set of bacterial shRNA constructs targeting the tyrosine kinases is contained on four 96 well microplates. Cells from the entire set have been expanded and DNA has been prepared for transfection of each construct on an Eppendorf DNA Workstation.

We have not yet, however, progressed to the actual transfection of packaging cell lines. As described in the original proposal, we had hoped to start this in year one. That we have not progressed to this point is not due to an overly ambitious workplan but rather is an unfortunate consequence of the delay caused by my relocation to a new cancer institute (see Reportable Outcomes). This work is now underway.

Additional results

Year 1 funding also made possible other developments which are potentially useful in future studies of the tyrosine kinase targeting constructs.

shRNAs can be used to produce transgenic animals

We have shown that short hairpin RNAs can be used to generate transgenic knock-out animals (8). An shRNA construct that effectively targets the murine DNA glycosylase-encoding Neil1 gene in ES cells was used to generate transgenic mouse lines (Fig 3A). Animals that carried the Neil1-targeted shRNA expression vector all displayed reduced levels of Neil1 mRNA (Fig 3B) and contained siRNAs corresponding to a Dicer-processed form of the shRNA (Fig 3C). Consistent with a role for Neil1 in DNA repair, cells from these mice exhibited increased sensitivity to ionizing radiation. Together with our previous demonstration that plasmid-encoded shRNAs could induce a potent and specific RNAi response in adult mice (9), these studies indicate that RNAi is a powerful genetic tool that could be incorporated into conventional gene-

therapy strategies. Although outside the scope of the original proposal, this development allows for the direct testing of tyrosine kinase-targeting hairpins of interest in mouse cancer models.

High throughput microarray-based shRNA delivery for phenotypic screening

One of the challenges in functional genomics in mammalian cells regards the generation and phenotyping of large number of cell populations. Using microarray-based cell transfections, we have demonstrated an improved method for performing high throughput silencing screens in mammalian cells (10). This is also an effective method for identifying effective RNAi triggers for silencing genes in mammalian systems.

We tested the use of cell microarrays as a delivery approach for effective shRNAs. The microarray format has obvious organizational value for some cellular phenotypes. In the experiments, panels of 6 shRNAs directed against a test gene, MyoD, were arrayed together with plasmids encoding MyoD-EGFP and the internal control RFP and gelatin on glass slides. Cells plated on top of these microarray features internalized the arrayed plasmids and were in this way transfected. Quantitation of EGFP and the RFP internal control intensities in response to various shRNAs correlated with values obtained with conventional transfections.

These results suggest that the microarray based cell transfections ("RNAi microarrays") can be used in large scale RNAi screens. Using fully automated liquid-dispensing and plate handling robotic systems, it is possible to prepare mixtures containing silencing constructs which then can be printed at densities of up to 6,000 to 10,000 features per slide by modern microarrayers. Cellular phenotypes that can be assessed microscopically, (eg. cell shape, apoptosis, proliferation rate) or that can be linked to a visible reporter are perfect candidates for this approach. Although a potential application, the limited scope of the tyrosine kinase is probably best executed with conventional gene transfer and phenotype analysis methodologies.

Key Research Accomplishments

- Adopted a new vector and shRNA production methodology
- **Created a set of encoded hairpins targeted against each of the genomic complement of human receptor protein tyrosine kinase genes**
- Created a set of vectors targeting human tyrosine kinase-related genes
- Development of reverse transfection for high throughput gene transfer and phenotype analysis
- Development of shRNA-based transgenic knock-down mouse

Reportable Outcomes

Manuscripts:

Carmell, M.A., L. Zhang, D.S. Conklin, G.J. Hannon, and T.A. Rosenquist. 2003, Germline transmission of RNAi in mice. *Nat Struct Biol* 10:91-2

Conklin, D.S., 2003. RNA interference-based silencing of mammalian gene expression (*Chembiochem*. 4:1033-1039)

Kumar, R., D. S. Conklin and V. Mittal, 2003. High throughput selection of effective RNAi probes for gene silencing. (*Genome Res*. 13:2333-40)

Hannon, G. J. and D. S. Conklin 2003, RNAi by short hairpin RNAs expressed in vertebrate cells. In *mRNA Processing and Metabolism: Methods and Protocols, Methods in Molecular Biology Series, Humana Press*

McManus, M.T. and D. S. Conklin 2003, shRNA-mediated silencing of mammalian gene expression. In *RNAi: a guide to gene silencing, Cold Spring Harbor Press*

Paddison, PJ, Silva, JM, Conklin, DS, Schlabach, M., Li, M., Aruleba, S., Balija, V., O'Shaughnessy, A., Gnoj, L., Scobie, K., Chang, K., Westbrook, T., Sachidanandam, R., McCombie, WR, Elledge SJ and Hannon, GJ, 2003, A resource for large-scale RNAi based screens in mammals (*submitted*)

Abstracts:

High throughput RNAi in Mammals

Doug Conklin

High throughput mammalian RNAi, (invited presentation), IBC's 2nd International Conference on RNAi, Boston, 2003

Vectors that direct the expression of short hairpin RNAs (shRNAs) make possible the creation of continuous cell lines and transgenic animals in which suppression of a target gene is stably maintained by RNAi. High throughput methodologies, including perl-based hairpin oligo design scripts, large scale oligo synthesis, cloning and confirmatory sequencing, have been used to produce 23,000+ sequence-verified shRNA constructs. This hairpin library is carried on an MSCV-based retroviral vector that can be used to stably express hairpins in a cell type of interest or to create transgenic mice. Collectively, the hairpins target approximately 10,000 mammalian genes which encode the majority of protein classes that are viewed to be accessible to medicinal chemistry. It is expected that this resource will accelerate RNAi-based gene function analysis in mammals.

Mammalian applications of RNAi

Doug Conklin

6th Conference on Protein Expression in Animal Cells

RNA interference (RNAi) is now established as a general method to silence gene expression in a variety of organisms. Double-stranded RNA (dsRNA), when introduced to cells, interferes with the expression of homologous genes disrupting their normal function. In mammals, transient delivery of synthetic short interfering RNAs (siRNAs), which resemble the processed form of

standard double stranded RNAi triggers, is effective in silencing mammalian genes. Issues related to transfer efficiency and duration of the silencing effect, however, restrict the spectrum of the applications of siRNAs in mammals. These shortcomings of siRNAs have been solved by the cellular expression of short hairpin RNAs (shRNAs) from DNA vectors. shRNAs are indistinguishable from siRNAs in terms of efficacy and mechanism but can be produced within cells from standard mammalian expression vectors. In this way, shRNA expression makes possible the creation of continuous cell lines and transgenic animals in which suppression of a target gene is stably maintained by RNAi. As a result, the types of RNAi-based gene function analysis that can be carried out in mammals have been greatly expanded.

Presentations:

High throughput RNAi in Mammals, (invited presentation), IBC's 2nd International Conference on RNAi, Boston, 2003

RNAi in mammalian functional genomics, (invited presentation), RNAi Symposium, David Axelrod Institute, 2003

Mammalian applications of RNAi, (invited presentation), 6th Annual PEACE Conference, Montreal, 2003

Theory and Applications of RNAi, Gordon Research Conference, 2003

Mammalian applications of RNAi, (invited presentation), University of Pennsylvania, 2003

Mammalian applications of RNAi, (invited presentation), University of Southern California, 2003

Mammalian applications of RNAi, (invited presentation), Dartmouth University, 2003

Mammalian applications of RNAi, (invited presentation), University at Albany, 2003

Mammalian applications of RNAi, (invited presentation), Aventis iLab Workshop, Wiesbaden, 2003

Mammalian applications of RNAi, (invited presentation), Drexel University, 2002

Mammalian applications of RNAi, (invited presentation), Ambion, 2002

Development of cell lines, tissue or serum repositories:

As was proposed, the collection of shRNA constructs targeting the tyrosine kinases will undoubtedly be useful to many investigators and will be made available as soon as testing is complete.

Funding applied for based on work supported by this award:

Department of Defense Breast Cancer Research Program

BC031982

FUNCTIONAL GENOMIC ANALYSIS OF BREAST CANCER CELL TUMORIGENICITY USING A NOVEL GENE SILENCING RESOURCE

Role: PI

Employment or research opportunities applied for and/or received based on experience/training supported by this award:

In the past year I was recruited to the State University of New York University at Albany's recently established Gen*NY*Sis Center for Excellence in Cancer Genomics. This is a recently developed institution that focuses on the type of work proposed in the present award and that I plan to pursue in my career.

This long term commitment to cancer genomics and improved proximity to a clinical department (Albany Medical College) was compelling. That I was carrying out the work on receptor tyrosine kinases and was funded by the DOD was instrumental in my being offered the job.

Conclusions

Year 1 of funding has resulted in the production of a large set of constructs targeting the tyrosine kinases. We have created a set of silencing agents not only for the genomic complement of receptor tyrosine kinases, but also all tyrosine kinases and a large number of genes functionally related to the RPTKs. This is an excellent start to our ultimate goal of identifying a small number of hairpins that have inhibitory effects on breast cancer progression in *in vivo* models.

Little has changed with respect to Year 2 Proposed Work. Hairpin constructs will be packaged in amphotropic retroviral packaging cells. Retroviral supernatants will be used to infect estrogen responsive MCF-7 cells in six well dishes. Infected populations will be monitored for phenotypic changes related to tumorigenic potential, e.g. cellular morphology, growth arrest, apoptosis, growth in soft agar, etc.

"So what?" one might ask.

Using high throughput methodologies, we have developed a resource for suppressing gene expression of one of the most important gene families in cancer. Although we are determining the effect that suppressing tyrosine kinase gene expression has on the tumorigenicity of breast cancer cells, it is hoped that the hairpin constructs will serve as a resource for other investigators studying a variety of cancers. This ultimately will add to our understanding of the genes encoding potential small molecule therapeutic targets in breast and other cancers. It should also pave the way for testing of the hairpins that we identify as gene therapeutics.

References

1. P. Blume-Jensen, T. Hunter, *Nature* **411**, 355-65. (2001).
2. D. S. Conklin, *ChemBiochem* **4**, 1033-9. (2003).
3. G. J. Hannon *et al.*, *Science* **283**, 1129-30 (1999).
4. P. Zhang, M. Z. Li, S. J. Elledge, *Nat Genet* **30**, 31-9. (2002).
5. C. Lois, E. J. Hong, S. Pease, E. J. Brown, D. Baltimore, *Science* **295**, 868-72. (2002).
6. J. M. Davis *et al.*, *Clin Cancer Res* **9**, 1161-70. 2: Liu D *et al.*
7. D. Liu *et al.*, *Eur J Cancer* **37**, 268-80. (2001).
8. M. A. Carmell, L. Zhang, D. S. Conklin, G. J. Hannon, T. A. Rosenquist, *Nat Struct Biol* **10**, 91-2. (2003).
9. A. P. McCaffrey *et al.*, *Nature* **418**, 38-9. (2002).
10. R. Kumar, D. S. Conklin, V. Mittal, *Genome Res* **13**, 2333-40.

Appendices

1. List of all shRNA constructs targeting human tyrosine kinases
2. Selected publications during support period

Conklin, D.S., 2003. RNA interference-based silencing of mammalian gene expression (*Chembiochem*. 4:1033-1039)

Kumar, R., D. S. Conklin and V. Mittal, 2003. High throughput selection of effective RNAi probes for gene silencing. (*Genome Res*. 13:2333-40)

McManus, M.T. and D. S. Conklin 2003, shRNA-mediated silencing of mammalian gene expression. In *RNAi: a guide to gene silencing*, Cold Spring Harbor Press (acknowledgement only)

Hannon, G. J. and D. S. Conklin 2003, RNAi by short hairpin RNAs expressed in vertebrate cells. In *mRNA Processing and Metabolism: Methods and Protocols*, Methods in Molecular Biology Series, Humana Press (acknowledgement only)

List of all shRNA constructs targeting human tyrosine kinases:

GENE	TARGETED SEQUENCE	PLATE	WELL
Homo sapiens v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene	GGAAGCCTTCCTTGATCATCTTGTAGAAC	333840	B3
Homo sapiens v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene	TGGACACTAGGAATGTGTAAGTGCCCTCCT	316140	E11
Homo sapiens discoidin domain receptor family, member 1 (DDR1),	GTTGCTCCATCCCACATAGTCATAGCCTG	320640	B4
Homo sapiens discoidin domain receptor family, member 1 (DDR1),	AAGCATCAGCTCATATAGGCCCTGCGGGC	353940	F7
Homo sapiens discoidin domain receptor family, member 1 (DDR1),	GGATAGTGTCCCCAGGGACAGAGAGGTGA	332940	H2
Homo sapiens discoidin domain receptor family, member 2 (DDR2),	CCTGCATGCCAGAGGATAGCGGCATAT	333840	C6
Homo sapiens discoidin domain receptor family, member 2 (DDR2),	ACAGCTCCATCATAGACAGAATCATTCAG	313440	G6
Homo sapiens EphA2 (EPHA2), mRNA.	CCACGGTGAAGGTGTAGTTCATGTGGGGC	173040	A2
Homo sapiens EphA2 (EPHA2), mRNA.	TTGATGGACTCCAGCCACTCGGACACCGT	314940	C4
Homo sapiens EphA2 (EPHA2), mRNA.	TTGGCCAGGTACTTTCATGCCAGCTGCGAT	316840	F1
Homo sapiens EphA4 (EPHA4), mRNA.	CCTTCCAGAGGGCTTGCTATCCACCCAAG	172940	F3
Homo sapiens EphA4 (EPHA4), mRNA.	ACAGGAGCCTCGAACTTCCACCAGGGAAG	173040	H7
Homo sapiens EphA4 (EPHA4), mRNA.	GCAATGGATGATGGTGCTGCTTGGTTGGT	172840	B2
Homo sapiens EphA7 (EPHA7), mRNA.	AACTGCCCCATGATGCTTGCTTCACACAA	312640	B4
Homo sapiens EphA7 (EPHA7), mRNA.	AGCTCCTTGCGGAATGATGGACAGCTCT	172940	E5
Homo sapiens EphA7 (EPHA7), mRNA.	CTTATATTCCTGCCAGTGTTCATAGTCTGT	322240	A3
Homo sapiens EphA8 (EPHA8), mRNA.	TCGATCCAGAAGGAGTAGTTCATGTGGGC	312640	E4
Homo sapiens EphA8 (EPHA8), mRNA.	CTTGCTGTAGCCACAGTGCCTCTTCTTGC	332440	H9
Homo sapiens EphA8 (EPHA8), mRNA.	ATGTACTCAGTCACAATCATTTGCCAGGCG	323840	G3
Homo sapiens EphB1 (EPHB1), mRNA.	TAATAACCGGTCCGACAGGTGCAGATGGG	307840	D2
Homo sapiens EphB1 (EPHB1), mRNA.	CTGGGACATTAGGGAGGCTGCTGCAGTCT	173040	G7
Homo sapiens EphB1 (EPHB1), mRNA.	AGGAGCAGTAGTAGATAATCCAGGGCCAT	315040	E6
Homo sapiens EphB2 (EPHB2), transcript variant 1, mRNA.	GTCTGCCAGGTAATTCATGCCAGTGCCTGA	332940	A6
Homo sapiens EphB2 (EPHB2), transcript variant 1, mRNA.	TCCACCTCTTCCGCATTGGCGATGCAGCT	314540	B11
Homo sapiens EphB3 (EPHB3), mRNA.	ACTTGAGCTCCACGTAGACCCGCTGCACA	172940	C5
Homo sapiens EphB3 (EPHB3), mRNA.	ATTGTACAGGAGGTTCATCCCGCCACCCA	351440	G6
Homo sapiens EphB3 (EPHB3), mRNA.	GCCCACCTTCAGCGTCTTGATGGCCACAA	322340	A7
Homo sapiens EphB4 (EPHB4), mRNA.	CACCTTGATGTAGGGGTTCTCCATCCAGG	172840	G2
Homo sapiens EphB4 (EPHB4), mRNA.	CGGTAGTCTCTTCAATGGCATTGATCAC	172840	H2
Homo sapiens EphB4 (EPHB4), mRNA.	CGTCTTGACATTACCTTCCCGGTGGCCT	312940	C12
Homo sapiens feline sarcoma oncogene (FES), mRNA.	AAGCCTTGGTACTCAGCCTCAGGCTGGAT	173040	C2
Homo sapiens feline sarcoma oncogene (FES), mRNA.	ACAGCACCGACAGCAGTACTCCTGCTTG	316140	A11
Homo sapiens feline sarcoma oncogene (FES), mRNA.	TCTCATGCCCTCCAGTAGACGAAGCTCGG	317840	E7
Homo sapiens fibroblast growth factor receptor 3 (achondroplasia,	CTTGTCTCCACGACGAGGTGTAGTTGC	321740	G11
Homo sapiens insulin receptor (INSR), mRNA.	CAGGATACGGGACCAGTCGATAGTGCCCA	173140	B2
Homo sapiens insulin receptor (INSR), mRNA.	TGAAGTTCACACAGCGCCAGTCTGGAAG	173140	F7
Homo sapiens insulin-like growth factor 1 receptor (IGF1R), mRNA.	TCTCTCCTTGTTATCCACTCTGCTCTCAA	315340	H4
Homo sapiens insulin-like growth factor 1 receptor (IGF1R), mRNA.	TTAGTCCCCGTCACCTTCTCCATGCGGTA	173040	F5
Homo sapiens Janus kinase 2 (a protein tyrosine kinase) (JAK2),	GCCTGGTGTACTCATCTATATGGAAGACA	173040	F2
Homo sapiens Janus kinase 2 (a protein tyrosine kinase) (JAK2),	TTCTCTGTGTAGAAGGCAGACTGCAGAGT	332840	D2
Homo sapiens Janus kinase 3 (a protein tyrosine kinase, leukocyte)	CCAGGTACTCCATGCCCTTGCAGATCTGC	173040	F6
Homo sapiens Janus kinase 3 (a protein tyrosine kinase, leukocyte)	AGAAGAAGTGTGGGAGTCCGTGGTCAGC	173040	B5
Homo sapiens Janus kinase 3 (a protein tyrosine kinase, leukocyte)	GGTGTCTCTCCTCGAAGATCGTGGGGTCT	307340	G1
Homo sapiens muscle, skeletal, receptor tyrosine kinase (MUSK),	AGGAACTCATTGAGGTCCACATAGGCCAT	332940	G3
Homo sapiens muscle, skeletal, receptor tyrosine kinase (MUSK),	CCCAGATTCAAGAAGTCAATTCGGGAAT	307440	A11
Homo sapiens neurotrophic tyrosine kinase, receptor, type 2	TTGGCCAAGTTCTGAAGGAGGGTATGGAT	322140	C7
Homo sapiens PTK2 protein tyrosine kinase 2 (PTK2), mRNA.	AGAACATTCGGAGCAGCAATGTCCCTGTG	351340	B1
Homo sapiens PTK2 protein tyrosine kinase 2 (PTK2), mRNA.	CATCAGTATCTCCACATACACACCAA	172840	F3
Homo sapiens protein tyrosine kinase 2 beta (PTK2B), mRNA.	AATGCCAGCTGAAGTACTGCCTGGCCCT	173140	C3

Homo sapiens protein tyrosine kinase 2 beta (PTK2B), mRNA.	AGGTCTGCTGGATCATCTTCCGGAAGTGT	333840	B9
Homo sapiens PTK7 protein tyrosine kinase 7 (PTK7), mRNA.	GGCCTCCTCATACCTCGCTACTACCACGT	322940	B5
Homo sapiens PTK7 protein tyrosine kinase 7 (PTK7), mRNA.	GTTGAGGCATTCATCTCTGGCTCCTCGC	322940	G5
Homo sapiens ret proto-oncogene (multiple endocrine neoplasia and	TGTTCTCATGCAGCCGTGTGCGGTACGTG	318840	B3
Homo sapiens ret proto-oncogene (multiple endocrine neoplasia and	TAGTCATGTACGGTCGCCCCGACGAAGCT	307840	A2
Homo sapiens tyrosine kinase with immunoglobulin and epidermal	CATTGACCTTGAAGCGCCGGCTGTCTTGG	172940	F8
Homo sapiens tyrosine kinase with immunoglobulin and epidermal	AGAGCTGCACATCCAGCTGGTAGTGGGTG	335240	E4
Homo sapiens v-src sarcoma (Schmidt-Ruppin A-2) viral oncogene	ACGGTGTCCGAGGAGTTGAAGCCTCCGAA	172940	H11
Homo sapiens v-src sarcoma (Schmidt-Ruppin A-2) viral oncogene	TACCACCTCCTCAGCCTGGATGGAGTCGGA	314640	A2
Homo sapiens v-src sarcoma (Schmidt-Ruppin A-2) viral oncogene	GCCGCTGTCCAGCTTGCGGATCTTGTAGT	307340	E8
Homo sapiens v-yes-1 Yamaguchi sarcoma viral related oncogene	GGAGCCGTCCACTTAATAGGGAAGCTTAGC	333940	B11
Homo sapiens v-yes-1 Yamaguchi sarcoma viral related oncogene	AATCATCCAGGACGCTCTGTAAGTAGTCA	173140	C11
Homo sapiens epidermal growth factor receptor (erythroblastic	ATACTCGGGGTGCCCACTGCAGTGTGT	353940	B7
Homo sapiens epidermal growth factor receptor (erythroblastic	TGAGCTGCGTGATGAGTTGCACGGTGGAG	312940	A11
Homo sapiens epidermal growth factor receptor (erythroblastic	CTTGCGGACGCCGTCTTCCCTCCATCTCAT	335340	D4
Homo sapiens v-abl Abelson murine leukemia viral oncogene homolog 1	CATAGACCTTCTCTGGGCAGCCTTCTGGG	332840	H6
Homo sapiens v-abl Abelson murine leukemia viral oncogene homolog 1	ACTTCTTCCACAGGTGGGGAGACCGGAAG	332840	H10
Homo sapiens v-abl Abelson murine leukemia viral oncogene homolog 1	AAGGCCGAGGATGCTGATGGCAACGTGGA	313040	F10
Homo sapiens v-abl Abelson murine leukemia viral oncogene homolog 2	TCCCCAGAGGGTGTCTGAGAGGAAGAGC	331440	E12
Homo sapiens v-abl Abelson murine leukemia viral oncogene homolog 2	AACAGATGACGAGGAGCGGCTCTCCCAA	314140	C5
Homo sapiens v-abl Abelson murine leukemia viral oncogene homolog 2	GGTGCCAGCTCTGTACCATCTTCATGGG	352140	F2
Homo sapiens activated p21cdc42Hs kinase (ACK1), mRNA.	CATCCTCATCCTGGGCCACGTCGTCATAG	314440	A12
Homo sapiens activated p21cdc42Hs kinase (ACK1), mRNA.	TTGGCTGTCTGAGAGCAGCAGAGATACTG	315140	H8
Homo sapiens Bruton agammaglobulinemia tyrosine kinase (BTK), mRNA.	TAGTTAATGAGCTCAGGGATGGTGTGAA	173040	C8
Homo sapiens Bruton agammaglobulinemia tyrosine kinase (BTK), mRNA.	CATACTGCACCAGCTTCTCATGGGAA	333840	E12
Homo sapiens Bruton agammaglobulinemia tyrosine kinase (BTK), mRNA.	GTCTCAGCAGTCCGAGCCTGGTAATCAA	335340	D2
Homo sapiens fyn-related kinase (FRK), mRNA.	ATAGCGGTACCCAGAGCTCTGGTTCAGGC	173040	E8
Homo sapiens FYN oncogene related to SRC, FGR, YES (FYN),	TGGTGACCAGCTCTGTGAGTAAGATTCCA	172940	B8
Homo sapiens FYN oncogene related to SRC, FGR, YES (FYN),	CAGGTAGGCCATGCCCTCACACACATCCA	172840	C9
Homo sapiens IL2-inducible T-cell kinase (ITK), mRNA.	AACTGCATCTTCTTCATCATGGCATAAGT	173140	G4
Homo sapiens Janus kinase 1 (a protein tyrosine kinase) (JAK1),	TCGGTGCAGCTCCACCTCAGCAGTACAT	335340	A4
Homo sapiens Janus kinase 1 (a protein tyrosine kinase) (JAK1),	TGGATGTAATTCTGCGACTCCAGTAACA	331640	F5
Homo sapiens PTK6 protein tyrosine kinase 6 (PTK6), mRNA.	TCCACTTGTAGGGGATATTGTGGTCATGG	324140	H11
Homo sapiens PTK6 protein tyrosine kinase 6 (PTK6), mRNA.	CTCTCCTCCAAGTACTTTCATGCCCATGGA	316240	G6
Homo sapiens spleen tyrosine kinase (SYK), mRNA.	CTTGACCAGCCTGTCTGTGTCAGCAGCA	332340	E12
Homo sapiens tyrosine kinase 2 (TYK2), mRNA.	TGGGTATGAGGTTCTCGAAGGTTGGGCGA	313540	H8
Homo sapiens tyrosine kinase 2 (TYK2), mRNA.	GCCATATCAACCAGCTGTGGAAGCTTCAA	335340	G10
Homo sapiens v-yes-1 Yamaguchi sarcoma viral oncogene homolog 1	GATCCTCCCCTGTAGATCCAGCAGTCA	317440	F11
Homo sapiens zeta-chain (TCR) associated protein kinase (70 kD)	TCCTGTGAGAAGCAGACAGCTGAGCAGCG	319440	H6
Homo sapiens fms-related tyrosine kinase 1 (vascular endothelial	TCTGGAAGACAGGAAGCTCCATGCCTCTGG	324340	C12
Homo sapiens fms-related tyrosine kinase 1 (vascular endothelial	CCCAGGTGAGCCCAGATCCATGGTTCACA	353640	H9
Homo sapiens fms-related tyrosine kinase 3 (FLT3), mRNA.	ACCCCAGAGACCACACATCGCTCTTGCT	322340	F9
Homo sapiens AXL receptor tyrosine kinase (AXL), transcript variant	TGGCCCATGATAGTTGCCTCTCGAAGGAA	324040	B12
Homo sapiens EphA1 (EPHA1), mRNA.	TGTATGTGGGTCAACATAAGTCTGAGAC	322340	A9
Homo sapiens EphA3 (EPHA3), mRNA.	CTGGGAGTCCATGGGTACCGTGTCTGGAA	333840	H1
Homo sapiens EphA3 (EPHA3), mRNA.	GCTACTGATGATCTTCTTCTGTGGCCCAA	315040	G5
Homo sapiens EphA3 (EPHA3), mRNA.	TTGTCCAAGTTCATGGTCTTCACTGCCAC	313440	G5
Homo sapiens c-mer proto-oncogene tyrosine kinase (MERTK), mRNA.	TCATCGGGCAATGAGCTTCCCAAGGGCAG	324140	A1
Homo sapiens c-mer proto-oncogene tyrosine kinase (MERTK), mRNA.	CCAGGAGGTATTGGGTATCCATTGATGGG	331640	A1
Homo sapiens receptor tyrosine kinase-like orphan receptor 1	TCCTGATTGGATCCTGATGAGGGCAAGCT	334840	A6
Homo sapiens receptor tyrosine kinase-like orphan receptor 1	CCCCAGCAGCTTGTAGTAATCGGCGGCAT	319140	E10
Homo sapiens receptor tyrosine kinase-like orphan receptor 2	GCTGGCACTAATGCTGTATCCAGTTCAA	323940	G12
Homo sapiens RYK receptor-like tyrosine kinase (RYK), mRNA.	TCTGTGATGAAGACCTCGCAGCTTACAAC	335240	B12

Homo sapiens TEK tyrosine kinase, endothelial (venous	CCAGCCACTGTGTTACACTGCAGACCCA	319440	H5
Homo sapiens TEK tyrosine kinase, endothelial (venous	ATCCATCCGTAACCCATCCTTCTTGATGC	314940	H8
Homo sapiens TYRO3 protein tyrosine kinase (TYRO3), mRNA.	ACCGGATCAGGGTCTGGAGGGGTAGGTTA	314940	H3
Homo sapiens TYRO3 protein tyrosine kinase (TYRO3), mRNA.	CACAGTGGGTCCTCAGCTCTCTCGATGT	317140	D9
Homo sapiens TYRO3 protein tyrosine kinase (TYRO3), mRNA.	AAGGAGCCATCCTCTTGCTTCAGCTGGGC	314640	F10
Homo sapiens v-erb-b2 erythroblastic leukemia viral oncogene	ACAGGAGTGGGTGCAGTTGATGGGGCAAG	322640	B1
Homo sapiens v-erb-b2 erythroblastic leukemia viral oncogene	CTGCTGCCGTCGCTTGATGAGGATCCCAA	319340	G7
Homo sapiens v-erb-b2 erythroblastic leukemia viral oncogene	CATTGTCTAGCACGGCCAGGGCATAAGTTG	332840	D7
Homo sapiens fibroblast growth factor receptor 1 (fms-related	TCTTCCAGGGCTTCCAGAACGGTCAACCA	313740	B3
Homo sapiens fibroblast growth factor receptor 4 (FGFR4),	AGACCACACGTCACCTCTGGTGTGTGTACA	173140	G5
Homo sapiens fibroblast growth factor receptor 4 (FGFR4),	AAGACAGAATCGCTGGAGGAGCAGGTGCT	335340	E10
Homo sapiens fibroblast growth factor receptor 4 (FGFR4),	AGGATGATGTCCGTATACCTGGCCTCGGG	307140	C4
Homo sapiens colony stimulating factor 1 receptor, formerly	TAGCAAGCTTGGGCTCAGGCTGGTGGTCA	306940	H4
Homo sapiens platelet-derived growth factor receptor, beta	CACAGAGACGTTGATGGATGACACCTGGA	331540	E2
Homo sapiens platelet-derived growth factor receptor, beta	ATGATCTCATAGATCTCGTCGGAGGCATG	307240	A4
Homo sapiens v-ros UR2 sarcoma virus oncogene homolog 1 (avian)	GAGGTTAACATTCAATGGAGCCTTCCAAT	319440	A3
Homo sapiens TXK tyrosine kinase (TXK), mRNA.	GAGACGAGTCATGAGACCGGCTGCATTGT	320640	C11
Homo sapiens TXK tyrosine kinase (TXK), mRNA.	GAGGGCGATATAGCCTGAAGCCTTCAGAA	313740	C5
Homo sapiens BMX non-receptor tyrosine kinase (BMX), mRNA.	ACACCTCTGGAGCTGACCACTTGACTGGA	319440	A8
Homo sapiens BMX non-receptor tyrosine kinase (BMX), mRNA.	TAGACTGGTACTTGAAGATGGTGGCTGGG	317540	G8
Homo sapiens BMX non-receptor tyrosine kinase (BMX), mRNA.	GATCATGCCTGCTGAATTGTGTTGATGAT	321040	C9
Homo sapiens fer (fps/fes related) tyrosine kinase (phosphoprotein	TGCCCTCGAATCGATACATGTTATCAACA	315040	C9
Homo sapiens fer (fps/fes related) tyrosine kinase (phosphoprotein	GCCAGAAGATGAATACACTCCACCATCCT	315040	C7
Homo sapiens fer (fps/fes related) tyrosine kinase (phosphoprotein	TGCACCATGGTACCAGTCTGTCTGCCA	313740	C3
Homo sapiens B lymphoid tyrosine kinase (BLK), mRNA.	TTCTCTTCTGTGACGAGTGACCTGGCCA	318440	D6
Homo sapiens B lymphoid tyrosine kinase (BLK), mRNA.	ATTCAGTGTGATGATTCGAGCCAAGCCA	314940	D11
Homo sapiens anaplastic lymphoma kinase (K1-1) (ALK), mRNA.	AAGATCTGGAAGAGAAGGCGGGCTGGGA	321040	F7
Homo sapiens platelet-derived growth factor receptor, alpha	TGTGGCCCCAGGTGAGTCATTATCTTTCAG	321040	C7
Homo sapiens platelet-derived growth factor receptor, alpha	GTCTCGTCTCTCTTGTGATGAAGGTGGA	315340	A4
Homo sapiens neurotrophic tyrosine kinase, receptor, type 3	CATAGTCCAGCCAGACGCTTCCCAGCAAG	315040	A7
Homo sapiens neurotrophic tyrosine kinase, receptor, type 3	GGCGTTCACATGCTGTTCCTTGAATCCT	331440	C1
Homo sapiens KIAA1079 protein (KIAA1079), mRNA.	AATTCTGGAGCAGTCCATCGCAGAGGGAA	316940	F6
Homo sapiens KIAA1079 protein (KIAA1079), mRNA.	TGGCAATGTTGACGGGAGAGAAGGCTTGG	333840	F7
Homo sapiens tyrosine kinase, non-receptor, 1 (TNK1), mRNA.	CTATGCTTCTGGGTGTTGATCTCCCCGG	313440	E7
Homo sapiens tyrosine kinase, non-receptor, 1 (TNK1), mRNA.	ACCAGCCCGAAGTCAGCCACCTTGATGGT	324240	A7
Homo sapiens EphB6 (EPHB6), mRNA.	AGGCTGAGCTGAGCCACATCACTGAAGGT	313740	E1
Homo sapiens EphB6 (EPHB6), mRNA.	AGGAGCAGCACCCATAGGCTACACCCAT	310340	G3
Homo sapiens similar to Ephrin type-A receptor 5 precursor	GCCATTCACCTACTGATCTGTAGGCCCA	332340	A10
Homo sapiens insulin receptor-related receptor (INSRR), mRNA.	GAAGTAGAGGTGGCGGCAAGCTACACAGG	316140	E2
Homo sapiens insulin receptor-related receptor (INSRR), mRNA.	AAGCTACACAGGCACGAGGGTCTTCTGGC	318440	A3
Homo sapiens ribosomal protein S4, Y-linked (RPS4Y), mRNA.	CCAATCACCATACACAAATTGCCTGTATC	334740	E7
Homo sapiens colony stimulating factor 1 receptor, formerly	ACACGACCACCTCTGTGCTAGCACGTTTC	352340	D8
Homo sapiens colony stimulating factor 1 receptor, formerly	GTTCTTATAGTCGACGCCCTCCCTCGGGGT	334740	C12
Homo sapiens RYK receptor-like tyrosine kinase (RYK), mRNA.	TGAAGTGTGTCATCAATGACACAGTTCCT	314840	H8
Homo sapiens RYK receptor-like tyrosine kinase (RYK), mRNA.	CACTTCTCAAGTCGTTCTTCTCTATCCGC	312940	A9
Homo sapiens IL2-inducible T-cell kinase (ITK), mRNA.	CAGCCAGTTGACGCAGCAGTCTGGAGAAG	323940	C11
Homo sapiens IL2-inducible T-cell kinase (ITK), mRNA.	ACCAGATAACTGCTTGGTACATATCCTTC	351340	E11
Homo sapiens PTK6 protein tyrosine kinase 6 (PTK6), mRNA.	TGAGCTCCGTGATGATGTACACGGGGTCC	352440	D1
Homo sapiens PTK6 protein tyrosine kinase 6 (PTK6), mRNA.	AAGGCCTCATGGTTGGACATGCCTGGGTA	313740	C11
Homo sapiens TYRO3 protein tyrosine kinase (TYRO3), mRNA.	ATTGGCACAGCGCACCCCTGAGGCTGTAGT	314140	H11
Homo sapiens TYRO3 protein tyrosine kinase (TYRO3), mRNA.	CTCCACACTGCAGTTGAGCTTACCAGGCT	320240	C9
Homo sapiens EphA8 (EPHA8), mRNA.	GCAGGGCACTTCTTATAGTAGATGCGGAG	307140	A3
Homo sapiens EphA1 (EPHA1), mRNA.	CTGGCTGTGGTGAAGATCCGATGGGCAAT	313740	E11

Homo sapiens EphA1 (EPHA1), mRNA.	CAGCTCCAGCACACACTCCATGGTGTCCA	307040	C4
Homo sapiens EphA2 (EPHA2), mRNA.	ATGCTGACGATGTCAGCGAACTTGGGGCG	318440	E12
Homo sapiens EphA2 (EPHA2), mRNA.	CTTCAGCATGCCCTTGACACCTCCCCAA	352240	D1
Homo sapiens EphA3 (EPHA3), mRNA.	CACACTTCATATTACCATCCAATGCCTTG	352540	A2
Homo sapiens EphA3 (EPHA3), mRNA.	CTCGTAGAGTGAACCTTGAGCTCCACATAA	314740	A4
Homo sapiens EphA3 (EPHA3), mRNA.	AGCAGCTTCTGGGTCATCCTCCAGGACAC	314740	F5
Homo sapiens insulin-like growth factor 1 receptor (IGF1R), mRNA.	CACTCTGGCCCCACGGATATGGTCGTTC	315240	A1
Homo sapiens insulin-like growth factor 1 receptor (IGF1R), mRNA.	GCAGCAAGTACTCGGTAATGACCGTGAGC	315340	H2
Homo sapiens insulin-like growth factor 1 receptor (IGF1R), mRNA.	CACCCCTGGCAACTCCTTCATAGACCATC	316440	C9
Homo sapiens PTK7 protein tyrosine kinase 7 (PTK7), mRNA.	ATGCAGCGGTAGATCCCTGCATTGCGTGG	313040	C12
Homo sapiens PTK7 protein tyrosine kinase 7 (PTK7), mRNA.	AGATGTGCATCCTGGGTCCCAGCTGGTG	308040	C5
Homo sapiens PTK7 protein tyrosine kinase 7 (PTK7), mRNA.	AAGACAATGGCTGTCTGGGTACCGCCAG	351340	A12
Homo sapiens EphA7 (EPHA7), mRNA.	GCTTACACAATATTCTGTAGGTCACATCG	314140	H6
Homo sapiens EphA7 (EPHA7), mRNA.	ATCCAGAAGAGGGCTTATTGGCCTACTAC	351240	G4
Homo sapiens EphA7 (EPHA7), mRNA.	AATAGTCATCCTGGCTACTGATTCAAGGG	331940	H10
Homo sapiens v-erb-a erythroblastic leukemia viral oncogene homolog	GGCAGCCATGGGGCATAAGTTGAGTAACC	314140	E6
Homo sapiens epidermal growth factor receptor (erythroblastic	TCTCTTAATTCCTTGATAGCGACGGGAAT	308040	C4
Homo sapiens epidermal growth factor receptor (erythroblastic	TTCTGGAAGTCCATCGACATGTTGCTGAG	307340	G7
Homo sapiens epidermal growth factor receptor (erythroblastic	ATTCCTGCCTCGGCTGACATTCGGCAA	318440	C12
Homo sapiens Janus kinase 1 (a protein tyrosine kinase) (JAK1),	TCAGCAGCCACACTCAGGTTCTTGAGTTC	352240	B1
Homo sapiens Janus kinase 1 (a protein tyrosine kinase) (JAK1),	AGAGGAGAGATACGGCATGCCTGTGCAGC	314740	D4
Homo sapiens EphB1 (EPHB1), mRNA.	ATAGCCAGGCTTGCAGGTGCATCGCCAA	173040	A11
Homo sapiens EphB1 (EPHB1), mRNA.	TGGAGACTCCATTGATGGCCTGGATGTCA	352440	E1
Homo sapiens muscle, skeletal, receptor tyrosine kinase (MUSK),	ACCAGTGAATGTTGACGAGCTCTCTCAT	307240	F1
Homo sapiens muscle, skeletal, receptor tyrosine kinase (MUSK),	CCCACACCATTGTTGGCCGTGCAGCAGTA	314740	B1
Homo sapiens muscle, skeletal, receptor tyrosine kinase (MUSK),	TCTGCTGAGTAGATGTTCTCGGAGAGGCC	331940	A10
Homo sapiens platelet-derived growth factor receptor, alpha	CGAATCATGCATGATGCTCTGGCCAGGC	307340	E4
Homo sapiens platelet-derived growth factor receptor, alpha	GATTCTCCACAATCTCACTCAGGTGGTAA	352440	F1
Homo sapiens v-abl Abelson murine leukemia viral oncogene homolog 1	TTACTGTGCCTCGGGTCACCTGGTCAGAC	332040	B6
Homo sapiens v-abl Abelson murine leukemia viral oncogene homolog 1	AAGAAGCTGCCATTGATCCCGCTGCTCAG	313040	C5
Homo sapiens zeta-chain (TCR) associated protein kinase (70 kD)	TAGAACTCGCAGAGCTCTGCCGGTCCACA	351240	B11
Homo sapiens zeta-chain (TCR) associated protein kinase (70 kD)	TATCCATCTGAGTTGAGGGTGTGATTCG	332040	F6
Homo sapiens zeta-chain (TCR) associated protein kinase (70 kD)	TATCGCGCTTCAGGAAGAGCTTCTTGTC	334940	E6
Homo sapiens EphB3 (EPHB3), mRNA.	CTTCATCTCGTAGTCCAGGATGACTCCGT	353740	B2
Homo sapiens EphB3 (EPHB3), mRNA.	ATGTCCACGCCAACTCAGATGTTACCCAT	307940	E7
Homo sapiens v-abl Abelson murine leukemia viral oncogene homolog 2	TTCTCTTCTCATGAAGGAGCTGAAGAAG	314740	A1
Homo sapiens TEK tyrosine kinase, endothelial (venous	TCATCATCACAGTTGAGGGCTTCTCCAG	307140	C1
Homo sapiens TEK tyrosine kinase, endothelial (venous	TTCTTGCCCTTGAACTTGTAAACGGATAGT	334940	B5
Homo sapiens EphB4 (EPHB4), mRNA.	GATCGGAAGAGTTCTCCTCCAGGAATCGG	323440	H3
Homo sapiens fer (fps/fes related) tyrosine kinase (phosphoprotein	GCCTCTATCTGCTGATGAACACCTATGTA	307240	C9
Homo sapiens fer (fps/fes related) tyrosine kinase (phosphoprotein	CCATAGATGTAACCTGATCGTGCATCTTCT	314440	F2
Homo sapiens fer (fps/fes related) tyrosine kinase (phosphoprotein	GCCAAGACTTGGATACGTTGCTGACATAA	317840	H10
Homo sapiens protein tyrosine kinase 2 beta (PTK2B), mRNA.	GTFCTTGTCCGCTCCAGGTAGTGGCCCA	331940	B3
Homo sapiens EphA4 (EPHA4), mRNA.	CACCGGATAGGAATCTTGCCACCCCTGGT	172940	B10
Homo sapiens EphA4 (EPHA4), mRNA.	GTATACCCCTGGAACCTGTGACAGCGTCGC	319340	B4
Homo sapiens tec protein tyrosine kinase (TEC), mRNA.	CAACACTGATAACTCCATCTGTCCAGAA	351940	C4
Homo sapiens tec protein tyrosine kinase (TEC), mRNA.	ACACATGCTCAGCAGTACGTCTCTACTGA	321840	B4
Homo sapiens tec protein tyrosine kinase (TEC), mRNA.	TCAAAGTATTCATCTATTGTGCGCAGCAG	316240	H10
Homo sapiens fibroblast growth factor receptor 2	TCCAAGTATTCTCATTTGGTTGTGAGAGT	352040	B7
Homo sapiens Bruton agammaglobulinemia tyrosine kinase (BTK), mRNA.	TCAAGATCTGGTCCCTCAGCGGTTGGG	314140	E7
Homo sapiens Bruton agammaglobulinemia tyrosine kinase (BTK), mRNA.	GGTAGTTCAGGAGGCGCCATTGGCCATG	321940	E6
Homo sapiens Bruton agammaglobulinemia tyrosine kinase (BTK), mRNA.	CTTCTTACTGCCTCTTCTCCCACGTTCAA	307540	H5
Homo sapiens Gardner-Rasheed feline sarcoma viral (v-fgr) oncogene	ACCAGTCACCTTCAGTATTGTTCCAGGATG	321940	F7

Homo sapiens Gardner-Rasheed feline sarcoma viral (v-fgr) oncogene	TCTGGTCTGATCCCAGTCCCGGATGGACA	316340	G8
Homo sapiens Gardner-Rasheed feline sarcoma viral (v-fgr) oncogene	CCTGTCCACCTGTTCACACTTCCCGT	317940	D10
Homo sapiens AXL receptor tyrosine kinase (AXL), transcript variant	TTGTAGATCTTCTTGGAGAGCCCGAAGTC	320240	F7
Homo sapiens AXL receptor tyrosine kinase (AXL), transcript variant	CTTCGGGCTCCTCCAGGAAGTAAGGCAAG	314740	F3
Homo sapiens v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene	ATCATTCTTGATGTCTCTGGCTAGACCAA	324240	A6
Homo sapiens fibroblast growth factor receptor 3 (achondroplasia,	TCTAGCTCCTTGTTCGGTGGTGTAGCGCC	321840	D7
Homo sapiens fibroblast growth factor receptor 3 (achondroplasia,	TTGGACTCCAGGGACACCTGTGCGTTGAG	334940	H2
Homo sapiens fibroblast growth factor receptor 3 (achondroplasia,	CATCACGTTGTCTCGGTCCAGCACAT	173040	H10
Homo sapiens Janus kinase 3 (a protein tyrosine kinase, leukocyte)	TGTACTTGGCCATGAGCGAGTGCCGGTCT	315340	E7
Homo sapiens Janus kinase 3 (a protein tyrosine kinase, leukocyte)	TAGCTGACAGTCTTCCAGCAGCTCTCCCGG	352040	G3
Homo sapiens insulin receptor-related receptor (INSRR), mRNA.	AGTACACGATGAAGCTGAGCAGGTCCGCGG	316340	H10
Homo sapiens insulin receptor-related receptor (INSRR), mRNA.	AGATGTGTTCCAAGCAGAGCGCGGGTTG	312640	D7
Homo sapiens insulin receptor-related receptor (INSRR), mRNA.	AGATCCTGTGCCGCTGGATGGAGTCGAT	317840	G8
Homo sapiens insulin receptor (INSR), mRNA.	AGGAAGGTTGGCCTCATCTGGGGTTGAA	351840	H3
Homo sapiens insulin receptor (INSR), mRNA.	ACGGCTCCCATCTCAGCAAGATCTTGTCA	332040	D5
Homo sapiens tyrosine kinase with immunoglobulin and epidermal	GGCTCCAGGGCTGTTGTGCACGTAGATGA	351840	F3
Homo sapiens tyrosine kinase with immunoglobulin and epidermal	GTGCTTCCATGTCAGCTGGATCTCGGAGT	315240	E11
Homo sapiens tyrosine kinase with immunoglobulin and epidermal	GAGAAGCCAAGAAGAGGATGGGGAGCAAG	314440	H2
Homo sapiens v-src sarcoma (Schmidt-Ruppin A-2) viral oncogene	AAGGTCCCGGTGGACGTAGTTCATCCGCT	321940	A4
Homo sapiens v-erb-b2 erythroblastic leukemia viral oncogene	GTAACCTGTGATCTCTCCAGAGTCTCAA	321840	A4
Homo sapiens v-yes-1 Yamaguchi sarcoma viral related oncogene	CTTAATAACATCACCATGCACAGGGTCAA	324840	B3
Homo sapiens v-yes-1 Yamaguchi sarcoma viral related oncogene	TGTTGTTATAGTAACCCATCCAGACTTCC	352140	E7
Homo sapiens c-src tyrosine kinase (CSK), mRNA.	GGTAATCGCCCAGCATCACGTCTCCGAAC	315240	B10
Homo sapiens c-src tyrosine kinase (CSK), mRNA.	TGGCCATGTACTCAGTGACGATGTAGAGC	313640	B12
Homo sapiens leukocyte tyrosine kinase (LTK), mRNA.	AGCCCAGAGGTTGTCAGTCTCTGAAGCGT	316340	D10
Homo sapiens leukocyte tyrosine kinase (LTK), mRNA.	TGAGTGCAGTACTGCAGACGCTCCAAGAT	323440	C3
Homo sapiens leukocyte tyrosine kinase (LTK), mRNA.	AGGCTCAGTGTGAGGTTGCCACCACAGC	314140	E9
Homo sapiens activated p21cdc42Hs kinase (ACK1), mRNA.	CCAGGTTCCAGTCAACATCTCCAGCACT	318940	A4
Homo sapiens Janus kinase 2 (a protein tyrosine kinase) (JAK2),	CATACCCTTGATATCTGAGATGTGTACT	316040	C6
Homo sapiens v-yes-1 Yamaguchi sarcoma viral oncogene homolog 1	GTTCCATCCATGTTCCCATCCCACTTC	317840	C10
Homo sapiens v-yes-1 Yamaguchi sarcoma viral oncogene homolog 1	CCTGTATCCTCGCTCCACTTGTCTAGTA	331740	D2
Homo sapiens v-yes-1 Yamaguchi sarcoma viral oncogene homolog 1	CCAGCAGGATATGAACTGGCACCCTGA	310540	F9
Homo sapiens receptor tyrosine kinase-like orphan receptor 1	CCAGCTGCAATCTGAATTGCAATGTGCAG	324440	E12
Homo sapiens receptor tyrosine kinase-like orphan receptor 1	TCGGGAACATGTAATTAGGATATCTGGGG	334740	A2
Homo sapiens KIAA1079 protein (KIAA1079), mRNA.	TGTGCAGGAAGTGCAGCTTGTGCATGGCG	320240	D2
Homo sapiens KIAA1079 protein (KIAA1079), mRNA.	TGATGCACCTGCTCAGGTAGGGAGAGCCT	317640	A12
Homo sapiens fibroblast growth factor receptor 4 (FGFR4),	ATGCTGCCACAGCGTTCTCTACCAGGCA	312640	B2
Homo sapiens feline sarcoma oncogene (FES), mRNA.	GTCGCAGGAAGCCTTGGTACTCAGCCTCA	319140	G11
Homo sapiens feline sarcoma oncogene (FES), mRNA.	ACACCACTCTTCTTGGTGAGGGGCTGCTG	307640	H7
Homo sapiens receptor tyrosine kinase-like orphan receptor 2	CAGGTGCGTAGACGTGCCGATCATGGTGA	324040	B7
Homo sapiens receptor tyrosine kinase-like orphan receptor 2	ATCCTGGTGGAGTACCCGAGTAGGGCT	314740	A2
Homo sapiens frns-related tyrosine kinase 4 (FLT4), mRNA.	AACACCTTGCAAGTAGGGCTGGCGTCTGT	324040	E6
Homo sapiens similar to Ephrin type-A receptor 5 precursor	ACAAGCTGAATCACAGTGAAGTGCATC	314840	E11
Homo sapiens similar to Ephrin type-A receptor 5 precursor	GCCAGTTATTCTGATTCTGTCCATCACT	324640	C1
Homo sapiens similar to Ephrin type-A receptor 5 precursor	ATTGGGACGATCTGGTCTTGCCAAGACA	319140	A11
Homo sapiens tyrosine kinase 2 (TYK2), mRNA.	CTCGTCATCCATCTTGCCCTCCTCAGGGT	307740	C7
Homo sapiens tyrosine kinase 2 (TYK2), mRNA.	TCTACTGGCCACCCTGGATGCCACCAGT	317640	D3
Homo sapiens tyrosine kinase 2 (TYK2), mRNA.	ACCCTGAGCAATGCCATGAGCTCAAGGA	307440	G8
Homo sapiens frns-related tyrosine kinase 3 (FLT3), mRNA.	CATCACTCTTAATGGTGTAGATGCCTTCA	316040	C2
Homo sapiens frns-related tyrosine kinase 3 (FLT3), mRNA.	GGCCAGGCTCTGTGAACACAGACTTAA	314740	H4
Homo sapiens frns-related tyrosine kinase 3 (FLT3), mRNA.	TCTGGCACAGCACCTATGTCCGTCCTCAA	319140	F11
Homo sapiens neurotrophic tyrosine kinase, receptor, type 2	CTTGGAGTGTCTTGCCAACCTAAGCAGAA	324040	F11
Homo sapiens neurotrophic tyrosine kinase, receptor, type 2	GGTCCATGCCACCTTATCCAGGACGACAT	316140	D8

Homo sapiens v-ros UR2 sarcoma virus oncogene homolog 1 (avian)	CACAGACACATACCAGTTCATCCATGATG	315340	A2
Homo sapiens v-ros UR2 sarcoma virus oncogene homolog 1 (avian)	CTATTGTGAATCTTCACCTCTCTGGGATA	308040	H6
Homo sapiens v-erb-b2 erythroblastic leukemia viral oncogene	CAGCCAATCGTAGCCCTGCATAGGGCTCT	314640	F9
Homo sapiens v-erb-b2 erythroblastic leukemia viral oncogene	TGGCATTCCGGGTGGCAGGAGAAGCATTTC	323840	B1
Homo sapiens neurotrophic tyrosine kinase, receptor, type 3	AGCCCAGTGACTATCCAGTCCACATCAGG	307640	G8
Homo sapiens neurotrophic tyrosine kinase, receptor, type 3	GCCCACCACGTTCTCTGCAATGCACGTC	320640	D1
Homo sapiens neurotrophic tyrosine kinase, receptor, type 3	CCATTGATGGTCTGGTTGGCTGTGCCAG	307340	G4
Homo sapiens BMX non-receptor tyrosine kinase (BMX), mRNA.	TCCAAGAAGGCCATGCCTTCACAGACATC	317640	C3
Homo sapiens BMX non-receptor tyrosine kinase (BMX), mRNA.	CAGCTGTACATGATCTGGTAGATGGTGTC	317540	H10
Homo sapiens hemopoietic cell kinase (HCK), mRNA.	AAGTCATCCAGCACACTCTGGATGTATTC	323940	D7
Homo sapiens hemopoietic cell kinase (HCK), mRNA.	TGGAGGAACTTGGACTTCATCGACCCCAT	317940	A8
Homo sapiens hemopoietic cell kinase (HCK), mRNA.	TGGCAATGGCTGCTTGCTGCCCTCATCAC	331840	A7
Homo sapiens discoidin domain receptor family, member 2 (DDR2),	TTGATGGCACTGGCCATTCGGTGGTGAG	324440	C12
Homo sapiens discoidin domain receptor family, member 2 (DDR2),	CAGGACAGGCTGGTTGGCACTGACATCTA	331840	A2
Homo sapiens discoidin domain receptor family, member 2 (DDR2),	AGCAGCTGAGCATCAGCTTATACACAGAG	312640	F2
Homo sapiens met proto-oncogene (hepatocyte growth factor receptor)	ATCAGGCTACTGGGCCCAATCACTACATG	334640	B6
Homo sapiens fyn-related kinase (FRK), mRNA.	TCCAGAGCCTCTGACAGATGTTGCTCAT	307140	F9
Homo sapiens fyn-related kinase (FRK), mRNA.	GCCATGTCTACCTGTTGAGTCAGATGGAT	313640	E11
Homo sapiens c-mer proto-oncogene tyrosine kinase (MERTK), mRNA.	TATCCGGTAGCCACCAGTTCTCCATCCT	313640	G10
Homo sapiens c-mer proto-oncogene tyrosine kinase (MERTK), mRNA.	TCATGACTGAGCCATTACCCAGCGGATCA	315340	D6
Homo sapiens c-mer proto-oncogene tyrosine kinase (MERTK), mRNA.	CTGGTGACGGCTGCAATCCTCACTGTGCA	352240	C10
Homo sapiens PTK2 protein tyrosine kinase 2 (PTK2), mRNA.	CCAAGAGCACACTTGAAGCATTCCCTTATC	314540	C3
Homo sapiens PTK2 protein tyrosine kinase 2 (PTK2), mRNA.	AGATGTGTAATGGCAGGCTTCTTGAAGAA	352140	D6
Homo sapiens TXK tyrosine kinase (TXK), mRNA.	CTTCTGCACTGAACAGCAACAGCAGCAAC	314140	A6
Homo sapiens macrophage stimulating 1 receptor (c-met-related)	CTTAGTGACTTGATGGCACATTGGATTCCG	351240	D3
Homo sapiens macrophage stimulating 1 receptor (c-met-related)	TTAGGCAACGCCACAGGTCAGGAAGTGG	353940	G11
Homo sapiens fms-related tyrosine kinase 1 (vascular endothelial	CAACAGAGTGCTGCTGTCGCCCTGGTAGT	351940	F5
Homo sapiens fms-related tyrosine kinase 1 (vascular endothelial	GCATTGTTCTGTTATTAAGTGTCCGCAGT	320640	B10
Homo sapiens spleen tyrosine kinase (SYK), mRNA.	CTGGAGAAGTTGTAGTAGTTGATGCATTTC	324040	H6
Homo sapiens spleen tyrosine kinase (SYK), mRNA.	TTGTCTGCAGCCAGGGTGCAGTTCTGG	352440	F10
Homo sapiens platelet-derived growth factor receptor, beta	CACAGATCTTGACCAGCTTGCCCTTCACAG	319340	E2
Homo sapiens neurotrophic tyrosine kinase, receptor, type 1	GAGATTCAGGCGACTGAGCCGAGGAGTGA	307040	D12
Homo sapiens melanoma antigen, family A, 9 (MAGEA9), mRNA.	TCTTCATCAGGCTTGCACTGCGGACTCCT	334640	A2

RNA-Interference-Based Silencing of Mammalian Gene Expression

Douglas S. Conklin*[a]

1. Introduction

In the five years since its initial description, RNA interference (RNAi) has become a remarkably powerful method for the suppression of gene expression in a variety of organisms. As a technique that only requires short stretches of sequence information to generate loss-of-function alleles for any gene, it holds enormous promise in the functional analysis of genes found in the recently completed genome sequences. It has already become a standard tool in those organisms in which traditional gene-knockout analysis is not feasible. In mammals, in which traditional gene knockouts are labor-intensive, the advent of RNAi has generated a great deal of excitement in its application to drug-target identification and in its development as a potential therapeutic.

RNAi was first described as a startling phenomenon in which exogenously supplied double-stranded RNA (dsRNA) had potent and specific effects in reducing the expression of homologous endogenous genes.^[1] Although it was immediately put to work as a genetic tool, several groups set out to uncover the novel molecular mechanism behind the phenomenon. The results of these studies, which include genetic evidence from *C. elegans* and biochemical evidence from a number of systems,^[2] have provided a model in which RNAi occurs through a two-step mechanism (Figure 1). In the first step, a sequence-

specific silencing factor composed of a 21–25-nucleotide (nt) short dsRNA is produced from the longer input dsRNA.^[3, 4] Short dsRNAs of this size were originally isolated from *Arabidopsis* plants undergoing posttranscriptional silencing and were found to be complementary to both strands of the silenced gene.^[5, 6] That similar short dsRNAs were also involved in RNAi was confirmed when, after isolation from cells undergoing silencing, they were sufficient to suppress gene expression in *Drosophila* S2 cells.^[7–9] Since they appear to be a requisite component and are sufficient to trigger RNA interference, these small RNAs have been termed short interfering RNAs (siRNAs).

The structure of siRNAs gave clues to the mechanism by which they were produced. Sequence analysis indicated that the siRNAs contained two nucleotide 3'-overhangs.^[8] This structure suggested that the dsRNA was converted into siRNAs by an RNase III family nuclease. Such an enzyme was first identified in *Drosophila* cells and subsequently found to exist in a number of eukaryotes capable of RNAi. In keeping with its function it was named Dicer.^[10] Structurally, Dicer enzymes contain an amino-terminal DExH/DEAH RNA helicase domain, tandemly repeated RNase III catalytic domains, and a carboxy-terminal dsRNA-binding domain.^[11, 12] They also contain a *piwi-argonaute-zwille* (PAZ) domain, which is specific to proteins involved in the RNAi pathway.

The second step of the RNAi mechanism was elucidated largely on the basis of biochemical evidence obtained in *Drosophila* cells.^[8, 13–15] This work established that siRNA products of Dicer were incorporated into a multicomponent nuclease complex, termed RISC (RNA-induced silencing complex). This complex uses the sequence information contained within the siRNA as a specificity determinant in the identification and nucleolytic destruction within the region of homology of cognate mRNAs. Although the mechanism of this complex is still poorly understood, a number of its conserved protein components are now known. These include homologues of the Argonaute protein of *Arabidopsis*,^[16] the fragile X mental-retardation protein and the Vasa intronic gene (VIG) protein.^[17] Each of these has been implicated in some form of posttran-

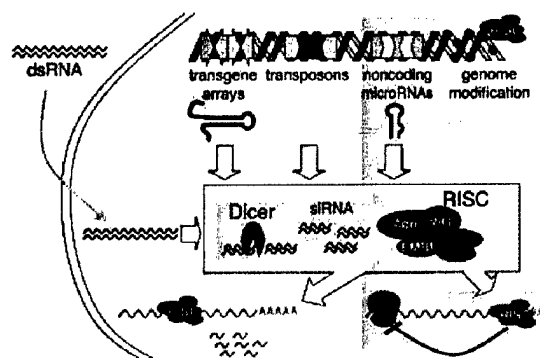


Figure 1. Cellular roles of the RNA-interference (RNAi) machinery. In addition to processing exogenous dsRNA, the RNAi machinery is involved in the processing of other double-stranded RNA (dsRNA) species shown. Exogenous dsRNA and transgene-derived RNAs are processed into small interfering RNAs (siRNAs), which silence mRNA by degradation. In the case of the microRNA (miRNA) precursors, cognate targets are silenced by translational repression. The detailed mechanisms of transposon and chromatin silencing are unknown at present but involve transcription of repeat structures as an early step.

[a] Current address: D. S. Conklin
Cancer Genome Center
Cold Spring Harbor Laboratory
500 Sunnyside Blvd, Woodbury, NY 11797 (USA)
Fax: (+1) 516-422-4134
E-mail: conklin@cshl.org

scriptional gene regulation; however, their roles within RISC are unknown. In general, the study of RISC is difficult. A number of *argonaute* gene family members are found in most species, which raises the possibility that different types of RISC complexes may exist within a cell, and the complex appears to play a part in multiple cellular processes (Figure 1).^[18]

That dsRNA-induced silencing phenomena exist in a variety of evolutionarily diverse organisms^[11, 19] suggested that the components of this system played a basic role in cellular biology. Since double-stranded RNA viruses and mobile genetic elements have the potential to form dsRNA structures and are virtually ubiquitous, it was thought that the RNAi pathway may have evolved early in eukaryotes as a cell-based immunity against genetic parasites such as viruses and transposable elements. In this light, the dsRNA-dependent silencing of transgenes in plants and endogenous genes in *C. elegans* were viewed as reprogrammed antiviral responses. It is now clear that this is only part of the story. The RNAi pathway is a central player in a variety of cellular processes related to the regulation of gene expression (Figure 1).

Endogenous small hairpin-shaped RNAs, which contain regions of dsRNA that are processed by the RNAi pathway, are now believed to be ubiquitous regulators of gene expression. The first examples of these, the *C. elegans lin-4* and *let-7* RNAs, were dubbed small temporal RNAs (stRNAs) owing to their role in the timing of developmental events.^[20–23] Transcribed as short (~70 nt) hairpins, these RNAs are processed into a 21–22-nt mature form by the first step in the RNAi pathway and subsequently guide RISC complexes to the 3' untranslated region (UTR) of target mRNAs through imperfect base-pairing interactions.^[24–28] In the case of *lin-4* and *let-7*, expression of the targeted genes is inhibited at the level of protein synthesis, not mRNA destruction. This is now believed to be a common regulatory mechanism in eukaryotes, since hundreds of short hairpins collectively termed microRNAs (miRNAs) have been identified in *C. elegans*, *Drosophila*, mice, and humans.^[29–33]

The RNAi machinery has also been shown to be involved in gene silencing that occurs in heterochromatin.^[34] The large number of repeats and transposons found in eukaryotic genomes are frequently associated with centromeres within large regions of silent chromatin. In *Schizosaccharomyces pombe*, deletion of the genes encoding Argonaute, Dicer, or RNA-dependent RNA polymerase, which is responsible for the amplification of dsRNA in some species,^[35, 36] abrogated the normal silencing of reporter constructs integrated within regions of centromeric heterochromatin. Deletion of these genes also impaired centromere function.^[37] The precise mechanism of this process is a mystery. dsRNA transcribed from repeated sequences at the centromeres is processed by the RNAi machinery and in some way directs the methylation of histone 3 (H3) on lysine 9. Methylated K9 H3 binds heterochromatin protein 1, which in turn inhibits local transcription. Although surprising, this work echoes links between dsRNAs and silencing of plant transgenes by chromatin methylation of the homologous DNA region.^[38, 39] These results are important in that they establish that the RNAi machinery is required for the pretranscriptional silencing of genes and the proper function of centromeric DNA. They have

also sparked enormous interest in latent, noncoding dsRNA transcripts.

2. siRNAs in Mammalian Cells

From the outset, RNAi proved to be a useful genetic tool. It virtually revolutionized the genetics of *C. elegans*, enabling genome-scale RNAi-mediated gene-function analysis.^[40–43] As mammals incorporate many of the same components of the RNAi pathway as *C. elegans*, it was hoped that it would have a similar effect on the genetics of mammals. Despite the mechanistic similarities, however, several hurdles related to the biology of mammals needed to be cleared before RNAi was proven to be an effective genetic technique in mammals.

The largest impediment to the implementation of RNAi in mammals is the physiological responses to dsRNA by these cells, which, when triggered, result in cell death. These include the induction of type I interferon (IFN) and the activation of two classes of IFN-induced enzymes: PKR, the dsRNA-dependent protein kinase, and 2',5'-oligoadenylate synthetases, the products of which activate RNase L. As little as one molecule of dsRNA longer than 30 nt is sufficient to trigger these responses, eventually resulting in global inhibition of translation and apoptosis.^[44–47] Although these responses are absent in some cell types, such as murine F9 and P19 embryonic carcinoma cell lines, which can therefore be subjected to long dsRNA-based gene silencing,^[48–50] a more general solution to this problem was needed for RNAi to become universally useful in mammalian cells.

The key development in methodology for dsRNA-based silencing in mammalian cells was to employ dsRNAs that would fail to trigger the nonspecific dsRNA responses and yet still induce RNAi-type silencing.^[9, 51] By using dsRNAs that mimicked siRNA duplexes produced by Dicer processing of long dsRNAs, Tuschl and co-workers were able to demonstrate specific gene silencing in a variety of mammalian cells.^[9] Since these molecules are easily produced by chemical synthesis, this has rapidly become a standard technique for gene manipulation in mammalian cells. Applications of siRNAs have been reviewed extensively^[52–56] and a variety of up-to-the-minute information is generally available online (e.g. <http://www.dharmacon.com/>, <http://www.ambion.com/techlib/resources/RNAi/index.html>)

One interesting concept that has emerged from the extensive use of siRNAs in gene silencing is the variable susceptibility of target genes to siRNAs. The central event in RNAi-mediated gene silencing is the interaction of the siRNA contained in RISC with its complementary sequence within an mRNA. However, it appears that all interactions between siRNAs and target sequences are not equal. Not only do different genes respond differently to silencing, but considerable variability in the degree of suppression exists between target sequences within a single gene.^[57, 58] It is thought that unknown intrinsic factors related to mRNA abundance, structure, translation rate, or other features of the RNAi mechanism are responsible. Further study of the mechanism of RNAi is required before we can accurately predict the suitability of a specific target sequence.

The target RNA-cleavage reaction guided by siRNAs is generally regarded as highly sequence-specific, requiring near identity between the siRNA and its cognate mRNA. For example, mismatches near the center of an siRNA duplex are most critical to target recognition and essentially abolish target RNA cleavage.^[59] Mismatched bases near the ends of an siRNA contribute little to the specificity of target recognition. Some researchers have taken advantage of this to design target-allele-specific siRNAs.^[60, 61] Despite these indications of the specificity, it is now clear that siRNAs can have effects on nontargeted sequences. Microarray studies have revealed that a single siRNA can affect the levels of a variety of messages within a cell that are not targeted.^[62] These effects include the silencing of nontargeted genes containing as few as 11 contiguous nucleotides of identity to the siRNA. In a practical sense, it indicates that it is advisable to test several different siRNAs when silencing a gene of interest.

3. Stable Silencing

The development of siRNAs for use in cell-culture genetics is an enormous advance over the available technologies. Nevertheless, when compared to the power of RNAi in *C. elegans*, for example, mammalian siRNA-mediated silencing is somewhat limited. The first significant difference between RNAi in *C. elegans* and mammalian cells is that mammalian cells do not take up exogenously applied dsRNA efficiently. Mammalian cells must be subjected to cationic lipids or electroporation for siRNAs to be effective. Neither of these methods is particularly effective in vivo. Another major distinction between *C. elegans* and mammalian cells is that the RNAi response does not persist in mammalian cells. Mammalian cells lack the ability to amplify the RNAi response that worms have, and thus RNAi is limited to approximately 6–8 cell doublings.^[63]

The issues of transfer efficiency and persistence of siRNAs in mammalian cells served as a catalyst for the development of stable RNAi-based silencing. In lower organisms, RNAi analysis of gene function was greatly improved by the in vivo expression of long dsRNA hairpins 500–1000 nt in length.^[64–67] In each case, the production of siRNAs by in vivo transcription and endogenous Dicer cleavage improved both the delivery and duration of the silencing effect as compared to transient, dsRNA-based methods. Vector systems based on these ideas are now available for use in *C. elegans*,^[65] *Drosophila*,^[66] and plants.^[68]

The development of stable long dsRNA hairpin-based expression systems in mammals, however, was not so straightforward. Long dsRNA is an effective approach for silencing genes in mammalian embryos,^[69] and vectors have been constructed for

this purpose (see Figure 2).^[70] This approach is not universally applicable to cell lines. Although silencing by the use of long dsRNA has been accomplished in mouse embryonic stem (ES) cells and embryonic carcinoma cells,^[70–73] the nonspecific effects

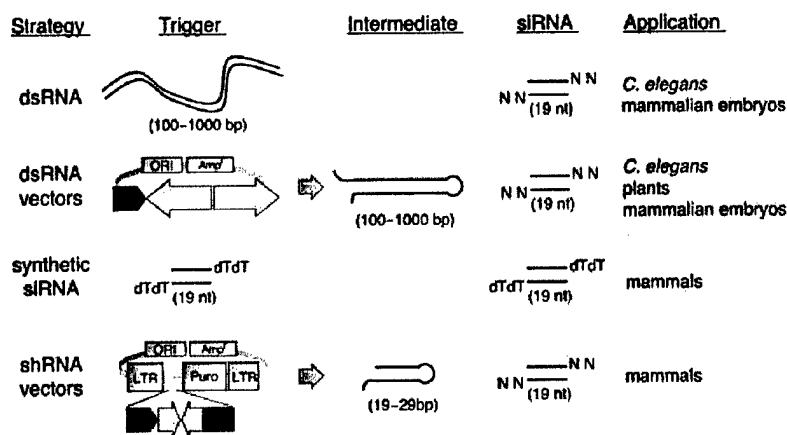


Figure 2. RNA-interference delivery strategies. Four major methods of introducing RNAi triggers are listed with schematics and typical applications of each. A representative long dsRNA and short hairpin RNA (shRNA) vector are shown. The elements depicted in the shRNA vector are typically included to improve the gene-transfer efficiency and stability of the silencing effect. Note that a siRNA or siRNA-like molecule is generated in each approach.

of long (> 30 nt) dsRNA expression in eliciting the interferon response in normal, differentiated somatic cell types^{[44][46]} required a novel strategy for stable dsRNA-based silencing in these cells.

To overcome these problems, many groups turned to the endogenous short hairpin miRNAs as a model for expressing dsRNAs for silencing in cells.^[58, 74–76] With short double-stranded stems, they were unlikely to trigger the nonspecific responses and were already known to be substrates of Dicer. Although the overall structure of miRNAs was retained, sequences within the stems of the encoded RNAs were engineered to be homologous to targeted gene sequences. In this way, silencing could be reprogrammed to specifically target any gene of choice. The similarity to miRNAs was somewhat superficial, however, as simple stem structures that were perfectly identical to the targeted gene and caused its degradation rather than translational arrest were found to be most effective.^[76] Nevertheless, this approach largely solved the problems related to RNAi persistence and transfer efficiency in mammalian cells (see Figure 2).

To date, several groups have developed workable strategies for stable gene silencing in mammalian cells.^[58, 74–77] In most cases, mammalian promoters are used to drive the expression of an interfering short hairpin RNA (shRNA) from DNA vectors introduced to cells by commonly used gene-transfer methods such as transfection or infection. Promoters using RNA polymerase III are employed in most cases, since this enzyme precisely initiates and terminates small, highly structured RNA transcripts^[78] and is active in most, if not all, embryonic and somatic cell types. These include the mouse^[79] and human U6-sRNA

promoters^[76, 80, 81], the human RNase P (H1) RNA promoter^[58, 74] and the human Val-tRNA promoter.^[61] Although the production of shRNAs by using pol III is the most common approach, other strategies exist. These include expressing miRNA-like constructs from an RNA polymerase II (cytomegalovirus) CMV immediate-early promoter^[75] and using separate U6 promoters to produce single-stranded 21-nt RNAs, which presumably anneal within the cell to form structures identical to siRNAs.^[77]

For shRNAs, the available evidence suggests that construct design is relatively flexible. Double-stranded stems between 18 and 29 nt in length are roughly equivalent in efficacy.^[82] Sequences of this length are long enough to serve as substrates for Dicer and contain unique silencing information, yet are still small enough to evade the PKR and interferon pathways. One of the strands of the stem structure should be complementary to the sense strand of the targeted mRNA. Whether it is the 5' stem strand or the 3' stem strand is not important.^[58, 74, 76, 79–81] The sequence of the loop is also fairly unimportant. Loop structures between 3 and 9 nt in length work well, longer loops may be deleterious.^[82] Target selection is poorly understood as is the case with siRNAs. The somewhat imprecise published guidelines with which to select hairpin target sites suggest a target sequence near the 5' end of the gene with a GC content of approximately 50%. Many target sites that do not share these criteria are highly effective, including several cases in which the 3' end of the gene was the best choice.^[83, 84]

The many vagaries of short hairpin construct design have been incorporated in online design tools that simplify the entire process (<http://katahdin.cshl.org:9331/RNAi/>, <http://jura.wi.mit.edu/bioc/siRNA/>, <http://www.dharmacon.com/>, http://www.ambion.com/techlib/misc/siRNA_finder.html). Once a target gene is selected, DNA oligonucleotides that encode shRNA corresponding to the gene are cloned downstream of a promoter into a vector suitable for gene transfer into mammalian cells. The power of shRNAs lies in their ability to continually silence targets after being stably transferred to host cells. For this reason, vectors that make use of gene-transfer methodologies that are inherently stable, such as retroviral integration, lentiviral integration, and adenoviral expression, are better choices for shRNA expression.

4. shRNA-Mediated Silencing in Animals

The stable suppression afforded by shRNAs has also been harnessed to affect the phenotypes of animals. The function of some genes, such as those that function in developmental, behavioral, and other complex processes, can only be studied in terms of the effect that a mutation has at the level of the organism. For this reason, classical mouse knockouts have been invaluable in investigating the function of many genes. Expression of an shRNA generates what is effectively a dominant loss-of-function mutation. That it is dominant avoids a major impediment of conventional knockout technologies, which require a significant investment of time and resources in the production of homozygous disruption alleles.

That stable RNAi could be used effectively to create loss-of-function alleles at the organism level in mammals was proven by

using a modified hydrodynamic transfection method to deliver siRNAs and a luciferase reporter gene to the livers of adult mice.^[85] Monitoring luciferase activity with quantitative whole-body imaging demonstrated that expression of a targeted luciferase reporter was specifically silenced to levels less than 5% of controls by a co-injected siRNA- or an shRNA-expression plasmid.

Recently, it has been demonstrated that short hairpin RNAs can be used to generate transgenic knockout animals.^[86, 87] ES cells carrying shRNA vectors that targeted a ubiquitously expressed GFP reporter were used to generate transgenic mice that exhibited silencing in all tissues.^[86] Suppression levels of up to 4% of untreated controls were observed and lasted for several weeks in F1 transgenic animals after introduction of the transgene. This approach has also been shown to be effective with endogenous targets.^[87] An shRNA construct that effectively targets the murine DNA glycosylase-encoding *neil1* gene in ES cells was used to generate transgenic mouse lines. Animals that carried the *neil1*-targeted shRNA expression vector all displayed reduced levels of *neil1* mRNA and contained siRNAs corresponding to a Dicer-processed form of the shRNA. Consistent with a role for *neil1* in DNA repair, cells from these mice exhibited increased sensitivity to ionizing radiation.

Several groups have now reported improved vector systems for the construction of shRNA-based knockdown animals.^[88–90] These vectors are based on self-inactivating lentiviruses, which are less susceptible to silencing during mouse development and are ideally suited for the generation of transgenic animals through infection of embryonic stem cells or single-cell embryos.^[91] In addition to an shRNA expression cassette, most express EGFP as a reporter that allows infected cells to be selected by flow cytometry. A variety of reporter and endogenous genes have been silenced in mice using these systems.

A general issue with RNAi-based approaches is whether the reduction of expression, which typically approaches but does not reach 100% reduction in expression, is sufficient to generate an obvious phenotype. Although it was originally hoped that RNAi in transgenic animals would substitute for the classical generation of knockouts, it seems likely that in some instances RNAi will prove to be a poor substitute for the complete disruption of a gene. Nevertheless, it is now clear that hairpins that are less than completely effective may be equally valuable as an investigation tool. Incomplete loss-of-function mutations in genes essential for viability are, in general, much more useful than nulls. This is true of genes involved in development where classical gene knockouts are of limited use as they eliminate gene function universally in the embryo. Even in genes that are not lethal, hairpins that do not completely silence a targeted gene are useful since they can give rise to subtle phenotypes that aid in determining the function of a gene in a given process. In a study with retroviral shRNA constructs that had intrinsic differences in their ability to silence the targeted mouse p53 gene, different hairpins produced phenotypes of varying severity at both the cellular and organism levels.^[92] p53 shRNAs that drastically reduced levels of the tumor-suppressor protein led to the formation of aggressive tumors and premature death of mice. Relatively ineffective p53 shRNAs had minimal effects on

p53 levels and resulted in animals with a mild form of the disease with no effects on mortality. The utility of shRNAs that do not completely silence their target, coupled with the increased ease of generating mutant animals represent significant advances in the methodologies for elucidating gene function *in vivo*.

5. Summary and Outlook

Owing to the robust nature of the process, RNAi-based gene silencing is certain to be a valuable technique for the foreseeable future. That it is effective *in vitro* and *in vivo* suggests that the limits to its utility are only in the types of phenotypes and assays that apply to mammalian cells. The rugged nature of the approach derives in large part from the fact that RNAi, as opposed to antisense, makes use of a collection of cellular enzymes that has been honed by millions of years of evolution for the express purpose of gene silencing.

Although highly useful already, a number of improvements to the current technologies are likely to emerge in the near future (see Figure 3). Inducible hairpin promoters that express shRNAs only in response to small molecule inducers or in specific tissue

Perhaps the most pertinent application of RNAi in mammalian cells is directly related to the fact that its discovery has coincided with the appearance of thousands of genes of unknown function in the mammalian genomic sequencing project databases. Since RNAi requires only short DNA sequences to manufacture potent reagents to knockdown gene expression, the sequence databases provide all the information required for genome-wide gene-function analysis projects. Arrayed libraries of siRNAs or hairpins that target each ORF in a mammalian genome are being constructed for use in phenotype-based, mid- to high-throughput screens *in vitro* and *in vivo*.^[2, 93, 94] The hope is that the role of any gene product in a given biological process can be tested without extensively developed reagents or presuppositions of its function.

Finally, RNAi may have a future in the direct treatment of disease.^[94–104] In contrast to standard gene therapy, which normally relies on the ectopic expression of proteins, RNAi can diminish the effects of deleterious gain-of-function mutant genes or the genes of infectious pathogens. Evidence for the potential of RNAi as a therapeutic has been demonstrated by a number of groups.^[105–107] As with any gene-therapy approach, however, issues related to delivery and safety are likely to pose major obstacles. Until these are overcome; RNAi will remain one of the most important tools for target discovery and validation in the development of small-molecule therapeutics.

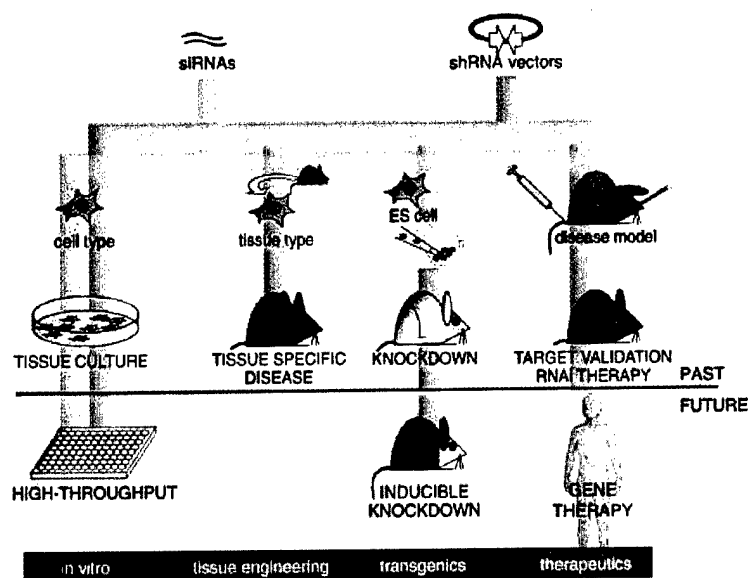


Figure 3. Applications of RNA-interference-based silencing in mammals. Schematic summary of current (above red line) and probable future (below red line) uses of RNA interference. siRNAs are typically used in cell cultures, but have also been used in some therapeutic studies in mice. shRNA vectors are especially useful in the generation of stable, hypomorphic alleles of genes in tissue culture cells, in specific cell types within mice, and in entire animals. High-throughput, inducible knockouts and gene therapy are under development at the time of this writing.

types would be valuable in the analysis of genes that are essential, have multiple roles in development, or that are related to behavior. Vectors that express several silencing constructs simultaneously would improve combinatorial silencing.^[82] Since RNAi acts in a dominant fashion, multiply mutant cells and animals can be generated in the time that it takes to generate a single heterozygous founder by conventional methods.

Acknowledgements

Shola Aruleba, Greg Hannon, David Kupper-smith, Ravi Sachidanandam, Kim Scobie, and the rest of the Cold Spring Harbor RNAi Group are thanked for comments and helpful discussions. D.S.C. is supported by the US Army Breast Cancer Research Program.

Keywords: antisense agents · gene expression · genomics · RNA · RNA interference

- [1] A. Fire, S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, C. C. Mello, *Nature* **1998**, *391*, 806–811.
- [2] G. J. Hannon, *Nature* **2002**, *418*, 244–251.
- [3] A. Grishok, H. Tabara, C. C. Mello, *Science* **2000**, *287*, 2494–2497.
- [4] H. Tabara, M. Sarkissian, W. G. Kelly, J. Fleenor, A. Grishok, L. Timmons, A. Fire, C. C. Mello, *Cell* **1999**, *99*, 123–132.
- [5] A. J. Hamilton, D. C. Baulcombe, *Science* **1999**, *286*, 950–952.
- [6] L. Jones, A. J. Hamilton, O. Voinnet, C. L. Thomas, A. J. Maule, D. C. Baulcombe, *Plant Cell* **1999**, *11*, 2291–2302.
- [7] D. Yang, H. Lu, J. W. Erickson, *Curr. Biol.* **2000**, *10*, 1191–1200.
- [8] P. D. Zamore, T. Tuschl, P. A. Sharp, D. P. Bartel, *Cell* **2000**, *101*, 25–33.
- [9] S. M. Elbashir, J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, T. Tuschl, *Nature* **2001**, *411*, 494–498.
- [10] E. Bernstein, A. A. Caudy, S. M. Hammond, G. J. Hannon, *Nature* **2001**, *409*, 363–366.
- [11] E. Bernstein, A. M. Denli, G. J. Hannon, *RNA* **2001**, *7*, 1509–1521.

- [12] R. H. Nicholson, A. W. Nicholson, *Mamm. Genome* **2002**, *13*, 67–73.
- [13] A. Nykanen, B. Haley, P. D. Zamore, *Cell* **2001**, *107*, 309–321.
- [14] T. Tuschl, P. D. Zamore, R. Lehmann, D. P. Bartel, P. A. Sharp, *Genes. Dev.* **1999**, *13*, 3191–3197.
- [15] S. M. Hammond, E. Bernstein, D. Beach, G. J. Hannon, *Nature* **2000**, *404*, 293–296.
- [16] S. M. Hammond, S. Boettcher, A. A. Caudy, R. Kobayashi, G. J. Hannon, *Science* **2001**, *293*, 1146–1150.
- [17] A. A. Caudy, M. Myers, G. J. Hannon, S. M. Hammond, *Genes Dev.* **2002**, *16*, 2491–2496.
- [18] M. A. Carmell, Z. Xuan, M. Q. Zhang, G. J. Hannon, *Genes Dev.* **2002**, *16*, 2733–2742.
- [19] S. M. Hammond, A. A. Caudy, G. J. Hannon, *Nat. Rev. Genet.* **2001**, *2*, 110–119.
- [20] I. Ha, B. Wightman, G. Ruvkun, *Genes Dev.* **1996**, *10*, 3041–3050.
- [21] R. C. Lee, R. L. Feinbaum, V. Ambros, *Cell* **1993**, *75*, 843–854.
- [22] F. J. Slack, M. Basson, Z. Liu, V. Ambros, H. R. Horvitz, G. Ruvkun, *Mol. Cell* **2000**, *5*, 659–669.
- [23] B. Wightman, I. Ha, G. Ruvkun, *Cell* **1993**, *75*, 855–862.
- [24] A. Grishok, A. E. Pasquinelli, D. Conte, N. Li, S. Parrish, I. Ha, D. L. Baillie, A. Fire, G. Ruvkun, C. C. Mello, *Cell* **2001**, *106*, 23–34.
- [25] R. F. Ketting, S. E. Fischer, E. Bernstein, T. Sijen, G. J. Hannon, R. H. Plasterk, *Genes Dev.* **2001**, *15*, 2654–2659.
- [26] S. W. Knight, B. L. Bass, *Science* **2001**, *293*, 2269–2271.
- [27] G. Hutvagner, J. McLachlan, A. E. Pasquinelli, E. Balint, T. Tuschl, P. D. Zamore, *Science* **2001**, *293*, 834–838.
- [28] B. J. Reinhart, F. J. Slack, M. Basson, A. E. Pasquinelli, J. C. Bettinger, A. E. Rougvie, H. R. Horvitz, G. Ruvkun, *Nature* **2000**, *403*, 901–906.
- [29] M. Lagos-Quintana, R. Rauhut, W. Lendeckel, T. Tuschl, *Science* **2001**, *294*, 853–858.
- [30] N. C. Lau, L. P. Lim, E. G. Weinstein, D. P. Bartel, *Science* **2001**, *294*, 858–862.
- [31] R. C. Lee, V. Ambros, *Science* **2001**, *294*, 862–864.
- [32] Z. Mourelatos, J. Dostie, S. Paushkin, A. Sharma, B. Charroux, L. Abel, J. Rappsilber, M. Mann, G. Dreyfuss, *Genes Dev.* **2002**, *16*, 720–728.
- [33] A. E. Pasquinelli, B. J. Reinhart, F. Slack, M. Q. Martindale, M. I. Kuroda, B. Maller, D. C. Hayward, E. E. Ball, B. Degnan, P. Muller, J. Spring, A. Srinivasan, M. Fishman, J. Finnerty, J. Corbo, M. Levine, P. Leahy, E. Davidson, G. Ruvkun, *Nature* **2000**, *408*, 86–89.
- [34] T. A. Volpe, C. Kidner, I. M. Hall, G. Teng, S. I. Grewal, R. A. Martienssen, *Science* **2002**, *297*, 1833–1837.
- [35] P. Stein, P. Svoboda, M. Anger, R. M. Schultz, *Rna* **2003**, *9*, 187–192.
- [36] T. Sijen, J. Fleenor, F. Simmer, K. L. Thijssen, S. Parrish, L. Timmons, R. H. Plasterk, A. Fire, *Cell* **2001**, *107*, 465–476.
- [37] T. Volpe, V. Schramke, G. L. Hamilton, S. A. White, G. Teng, R. A. Martienssen, R. C. Allshire, *Chromosome Res.* **2003**, *11*, 137–146.
- [38] M. F. Mette, W. Aufsatz, J. van der Winden, M. A. Matzke, A. J. Matzke, *EMBO J.* **2000**, *19*, 5194–5201.
- [39] L. Jones, F. Ratcliff, D. C. Baulcombe, *Curr. Biol.* **2001**, *11*, 747–757.
- [40] N. J. O'Neil, R. L. Martin, M. L. Tomlinson, M. R. Jones, A. Coulson, P. E. Kuwabara, *Am. J. Pharmacogenomics* **2001**, *1*, 45–53.
- [41] P. Gonczy, C. Echeverri, K. Oegema, A. Coulson, S. J. Jones, R. R. Copley, J. Duperon, J. Oegema, M. Brehm, E. Cassin, E. Hannak, M. Kirkham, S. Pichler, K. Flohrs, A. Goessen, S. Leidel, A. M. Alleaume, C. Martin, N. Ozlu, P. Bork, A. A. Hyman, *Nature* **2000**, *408*, 331–336.
- [42] A. G. Fraser, R. S. Kamath, P. Zipperlen, M. Martinez-Campos, M. Sohrmann, J. Ahringer, *Nature* **2000**, *408*, 325–330.
- [43] A. J. Walhout, J. Reboul, O. Shtanko, N. Bertin, P. Vaglio, H. Ge, H. Lee, L. Doucette-Stamm, K. C. Gunsalus, A. J. Schetter, D. G. Morton, K. J. Kemphues, V. Reinke, S. K. Kim, F. Piano, M. Vidal, *Curr. Biol.* **2002**, *12*, 1952–1958.
- [44] P. I. Marcus, M. J. Sekellick, *Virology* **1985**, *142*, 411–415.
- [45] B. R. Williams, *Oncogene* **1999**, *18*, 6112–6120.
- [46] J. Gil, M. Esteban, *Apoptosis* **2000**, *5*, 107–114.
- [47] M. J. Clemens, A. Elia, *J. Interferon Cytokine Res.* **1997**, *17*, 503–524.
- [48] G. R. Stark, I. M. Kerr, B. R. Williams, R. H. Silverman, R. D. Schreiber, *Annu. Rev. Biochem.* **1998**, *67*, 227–264.
- [49] H. Harada, K. Willison, J. Sakakibara, M. Miyamoto, T. Fujita, T. Taniguchi, *Cell* **1990**, *63*, 303–312.
- [50] M. K. Francis, J. M. Lehman, *Mol. Cell. Biol.* **1989**, *9*, 3553–3556.
- [51] V. Bitko, S. Barik, *BMC Microbiol.* **2001**, *1*, 34.
- [52] M. T. McManus, P. A. Sharp, *Nat. Rev. Genet.* **2002**, *3*, 737–747.
- [53] M. Scherr, M. A. Morgan, M. Eder, *Curr. Med. Chem.* **2003**, *10*, 245–256.
- [54] Y. Shi, *Trends Genet.* **2003**, *19*, 9–12.
- [55] M. Famulok, S. Verma, *Trends Biotechnol.* **2002**, *20*, 462–466.
- [56] L. Timmons, *Mol. Cell* **2002**, *10*, 435–437.
- [57] A. B. Kapadia, A. Bideau-Andersen, F. V. Chisari, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 2014–2018.
- [58] M. T. McManus, C. P. Petersen, B. B. Haines, J. Chen, P. A. Sharp, *Rna* **2002**, *8*, 842–850.
- [59] S. M. Elbashir, J. Martinez, A. Patkaniowska, W. Lendeckel, T. Tuschl, *Embo J.* **2001**, *20*, 6877–6888.
- [60] P. Lassus, J. Rodriguez, Y. Lazebnik, *Sci. STKE* **2002**, *2002*, PL13.
- [61] H. Kawasaki, K. Taira, *Nucleic Acids Res.* **2003**, *31*, 700–707.
- [62] A. L. Jackson, S. R. Bartz, J. Schelter, S. V. Kobayashi, J. Burchard, M. Mao, B. Li, G. Cavet, P. S. Linsley, *Nat. Biotechnol.* **2003**, *21*, 635–637.
- [63] P. Stein, P. Svoboda, M. Anger, R. M. Schultz, *RNA* **2003**, *9*, 187–192.
- [64] N. A. Smith, S. P. Singh, M. B. Wang, P. A. Stoutjesdijk, A. G. Green, P. M. Waterhouse, *Nature* **2000**, *407*, 319–320.
- [65] N. Tavemarakis, S. L. Wang, M. Dorovkov, A. Ryazanov, M. Driscoll, *Nat. Genet.* **2000**, *24*, 180–183.
- [66] J. R. Kennerdell, R. W. Carthew, *Nat. Biotechnol.* **2000**, *18*, 896–898.
- [67] C. F. Chuang, E. M. Meyerowitz, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 4985–4990.
- [68] S. V. Wesley, C. A. Helliwell, N. A. Smith, M. B. Wang, D. T. Rouse, Q. Liu, P. S. Gooding, S. P. Singh, D. Abbott, P. A. Stoutjesdijk, S. P. Robinson, A. P. Gleave, A. G. Green, P. M. Waterhouse, *Plant J.* **2001**, *27*, 581–590.
- [69] F. Wianny, M. Zernicka-Goetz, *Nat. Cell. Biol.* **2000**, *2*, 70–75.
- [70] P. Svoboda, P. Stein, R. M. Schultz, *Biochem. Biophys. Res. Commun.* **2001**, *287*, 1099–1104.
- [71] E. Billy, V. Brondani, H. Zhang, U. Muller, W. Filipowicz, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 14428–14433.
- [72] S. Yang, S. Tutton, E. Pierce, K. Yoon, *Mol. Cell Biol.* **2001**, *21*, 7807–7816.
- [73] P. J. Paddison, A. A. Caudy, G. J. Hannon, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 1443–1448.
- [74] T. R. Brummelkamp, R. Bernards, R. Agami, *Science* **2002**, *296*, 550–553.
- [75] Y. Zeng, E. J. Wagner, B. R. Cullen, *Mol. Cell* **2002**, *9*, 1327–1333.
- [76] P. J. Paddison, A. A. Caudy, E. Bernstein, G. J. Hannon, D. S. Conklin, *Genes Dev.* **2002**, *16*, 948–958.
- [77] M. Miyagishi, K. Taira, *Nat. Biotechnol.* **2002**, *20*, 497–500.
- [78] R. S. Goomer, G. R. Kunkel, *Nucleic Acids Res.* **1992**, *20*, 4903–4912.
- [79] J. Y. Yu, S. L. DeRuiter, D. L. Turner, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 6047–6052.
- [80] C. P. Paul, P. D. Good, I. Winer, D. R. Engelke, *Nat. Biotechnol.* **2002**, *20*, 505–508.
- [81] G. Sui, C. Soohoo, B. Affar el, F. Gay, Y. Shi, W. C. Forrester, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5515–5520.
- [82] J. Y. Yu, J. Taylor, S. L. DeRuiter, A. B. Vojtek, D. L. Turner, *Mol. Ther.* **2003**, *7*, 228–236.
- [83] M. T. McManus, B. B. Haines, C. P. Dillon, C. E. Whitehurst, L. van Parijs, J. Chen, P. A. Sharp, *J. Immunol.* **2002**, *169*, 5754–5760.
- [84] Q. Ge, M. T. McManus, T. Nguyen, C. H. Shen, P. A. Sharp, H. N. Eisen, J. Chen, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 2718–2723.
- [85] A. P. McCaffrey, L. Meuse, T. T. Pham, D. S. Conklin, G. J. Hannon, M. A. Kay, *Nature* **2002**, *418*, 38–39.
- [86] H. Hasuwa, K. Kaseda, T. Einarsdottir, M. Okabe, *FEBS Lett.* **2002**, *532*, 227–230.
- [87] M. A. Carmell, L. Zhang, D. S. Conklin, G. J. Hannon, T. A. Rosenquist, *Nat. Struct. Biol.* **2003**, *10*, 91–92.
- [88] S. A. Stewart, D. M. Dykxhoorn, D. Palliser, H. Mizuno, E. Y. Yu, D. S. An, D. M. Sabatini, I. S. Chen, W. C. Hahn, P. A. Sharp, R. A. Weinberg, C. D. Novina, *RNA* **2003**, *9*, 493–501.
- [89] D. A. Rubinson, C. P. Dillon, A. V. Kwiatkowski, C. Sievers, L. Yang, J. Kopinja, M. Zhang, M. T. McManus, F. B. Gertler, M. L. Scott, L. Van Parijs, *Nat. Genet.* **2003**, *33*, 401–406.
- [90] G. Tiscornia, O. Singer, M. Ikawa, I. M. Verma, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1844–1848.
- [91] C. Lois, E. J. Hong, S. Pease, E. J. Brown, D. Baltimore, *Science* **2002**, *295*, 868–872.
- [92] M. T. Hemann, J. S. Fridman, J. T. Zilfou, E. Hernando, P. J. Paddison, C. Cordon-Cardo, G. J. Hannon, S. W. Lowe, *Nat. Genet.* **2003**, *3*, 3.
- [93] P. J. Paddison, G. J. Hannon, *Cancer Cell* **2002**, *2*, 17–23.

- [94] R. Agami, *Curr. Opin. Chem. Biol.* **2002**, *6*, 829–834.
- [95] T. Abbas-Terki, W. Blanco-Bose, N. Deglon, W. Pralong, P. Aebischer, *Hum. Gene Ther.* **2002**, *13*, 2197–2201.
- [96] A. Borkhardt, *Cancer Cell* **2002**, *2*, 167–168.
- [97] D. Lawrence, *Lancet* **2002**, *359*, 2007.
- [98] H. Xia, Q. Mao, H. L. Paulson, B. L. Davidson, *Nat. Biotechnol.* **2002**, *20*, 1006–1010.
- [99] L. Zhang, N. Yang, A. Mohamed-Hadley, S. C. Rubin, G. Coukos, *Biochem. Biophys. Res. Commun.* **2003**, *303*, 1169–1178.
- [100] N. J. Caplen, J. P. Taylor, V. S. Statham, F. Tanaka, A. Fire, R. A. Morgan, *Hum. Mol. Genet.* **2002**, *11*, 175–184.
- [101] M. Wilda, U. Fuchs, W. Wossmann, A. Borkhardt, *Oncogene* **2002**, *21*, 5716–5724.
- [102] M. A. Martinez, B. Clotet, J. A. Este, *Trends Immunol.* **2002**, *23*, 559–561.
- [103] D. P. Cioca, Y. Aoki, K. Kiyosawa, *Cancer Gene Ther.* **2003**, *10*, 125–133.
- [104] Y. Aoki, D. P. Cioca, H. Oidaira, J. Kamiya, K. Kiyosawa, *Clin. Exp. Pharmacol. Physiol.* **2003**, *30*, 96–102.
- [105] T. Yokota, N. Sakamoto, N. Enomoto, Y. Tanabe, M. Miyagishi, S. Maekawa, L. Yi, M. Kurosaki, K. Taira, M. Watanabe, H. Mizusawa, *EMBO Rep.* **2003**, *4*, 602–608.
- [106] L. Zhang, N. Yang, A. Mohamed-Hadley, S. C. Rubin, G. Coukos, *Biochem. Biophys. Res. Commun.* **2003**, *303*, 1169–1178.
- [107] E. Song, S. K. Lee, J. Wang, N. Ince, N. Ouyang, J. Min, J. Chen, P. Shankar, J. Lieberman, *Nat. Med.* **2003**, *9*, 347–351.

Received: June 11, 2003 [M686]

High-Throughput Selection of Effective RNAi Probes for Gene Silencing

Rajeev Kumar, Douglas S. Conklin, and Vivek Mittal¹

Cancer Genome Research Center, Cold Spring Harbor Laboratory, Woodbury, New York 11797, USA

RNA Interference (RNAi) is a process of sequence-specific posttranscriptional gene silencing mediated by double-stranded RNA. RNAi has recently emerged as a powerful genetic tool to analyze gene function in mammalian cells. The power of this method is limited however, by the uncertainty in predicting the efficacy of small interfering RNAs (siRNAs) in silencing a gene. This has imposed serious limitations not only for small-scale but also for high-throughput RNAi screening initiatives in mammalian systems. We have developed a reliable and quantitative approach for the rapid and efficient identification of the most effective siRNA against any gene. The efficacy of siRNA sequences is monitored by their ability to reduce the expression of cognate target-reporter fusions with easily quantified readouts. Finally, using microarray-based cell transfections, we demonstrate an unlimited potential of this approach in high-throughput screens for identifying effective siRNA probes for silencing genes in mammalian systems. This approach is likely to have implications in the use of RNAi as a reverse genetic tool for analyzing mammalian gene function on a genome-wide scale.

[Supplemental material is available online at www.genome.org. The following individuals kindly provided reagents, samples, or unpublished information as indicated in the paper: G. Hannon, J. Silva, R. Benzera, K. LaVine, S. Gupta, D. Helfman, E. Hatchwell, S. Kantaraci, and A. Lassar.]

RNA interference (RNAi) is a process by which double-stranded RNA (dsRNA) homologous to the target locus can specifically inactivate gene function in plants, invertebrates, and mammalian systems (Sharp 1999; Hammond et al. 2001). The endogenous mediators of sequence-specific mRNA degradation are 21- and 22-nucleotide (nt) siRNAs generated from longer dsRNAs by the ribonuclease III activity of the evolutionarily conserved dicer enzyme (Bernstein et al. 2001; Elbashir et al. 2001b). Gene-specific long dsRNAs have been successfully used in *Caenorhabditis elegans* and *Drosophila melanogaster* for RNAi-mediated gene silencing (Hannon 2002). In mammalian cells, however, dsRNAs longer than 30 base pairs (bp) trigger the antiviral/interferon pathways, which results in global shut-down of protein synthesis (Marcus and Sekellick 1985; Gil and Esteban 2000). It was recently demonstrated that RNAi-mediated gene silencing can be obtained in cultured mammalian cells by delivery of either chemically synthesized short (<30-nt) double-stranded siRNA molecules (Elbashir et al. 2001a) or by endogenous expression of short hairpin RNAs (shRNAs) bearing fold-back stem-loop structure (Brummelkamp et al. 2002; Elbashir et al. 2002; Paddison et al. 2002; Paul et al. 2002; Sul et al. 2002).

General guidelines for designing siRNA oligonucleotides are available (Elbashir et al. 2002) which require avoiding regions of mRNA such as 5' and 3' untranslated regions (UTRs) and regions within 75 bases of a start codon, and sequences having a GC content of >70% or <30%. Although the existing rules for siRNA selection are a reliable general guide, they do not ensure that each selected siRNA will work. For example, even though targeting the UTR is not generally recommended, successful gene inhibition has been reported for siRNAs targeting 3' UTRs of the Myc oncogene (Bryom et al. 2002) and CD8 α receptor (McManus

et al. 2002). Thus it has been observed that in many cases a majority of siRNAs or shRNAs designed against a gene are not effective (Elbashir et al. 2001a; Holen et al. 2002; Lee et al. 2002; Yu et al. 2002; Kapadia et al. 2003). On average, only 1 in 5 of the siRNAs/shRNAs selected for targeting a region of a gene show efficient gene silencing (McManus et al. 2002; Kapadia et al. 2003; R. Kumar and V. Mittal, unpubl.). Instability of an siRNA probe in vivo, inability to interact with components of the RNAi machinery, and inaccessibility of the target mRNA due to local secondary structural constraints have all been proposed as possible causes for the failure of most tested siRNAs. An analysis of nucleotide sequences, melting temperatures, and secondary structures has thus far not revealed any obvious differences between effective and noneffective siRNA (Hohjoh 2002), thus preventing the development of more effective siRNA design methods. Moreover, empirical approaches such as RNase H susceptibility assays of siRNA/target duplexes (Lee et al. 2002) and shotgun approaches in which a mixture of siRNA produced by RNase III-mediated hydrolysis of long double-stranded RNA is used (Calegari et al. 2002; Yang et al. 2002) do not easily translate to the design of effective RNAi probes. At present, the lack of a reliable target-site selection method has imposed serious limitations on the efficiency of RNAi as a general tool for gene function analysis.

We have developed a quantitative procedure for rapid identification of an effective siRNA/shRNA for inhibition of target gene expression. Effective RNAi probes are identified based on their ability to inactivate cognate sequences in an ectopically expressed target gene-reporter chimeric mRNA. Using either a fluorescent or enzymatic reporter, the siRNA effect is monitored quantitatively. By examining a variety of genes with diverse biological functions in various mammalian cell lines, we have shown a strong correlation in the suppression of ectopically expressed and cognate endogenous genes by siRNA probes. Furthermore, using microarray-based cell transfections, we demonstrate that this approach can be tailored to high-throughput screens for identifying effective siRNA probes in mammalian systems. Such

¹Corresponding author.

E-MAIL mittal@cshl.org; FAX (516) 422-4109.

Article and publication are at <http://www.genome.org/cgi/doi/10.1101/gr.1575003>.

a screen would have unlimited potential for analyzing mammalian gene function on a genome-wide scale.

RESULTS AND DISCUSSION

Experimental Design for Selecting Effective siRNA Probes

The crucial event in RNAi-mediated gene silencing is the interaction of a 21-nt siRNA probe with its complementary sequence within an mRNA. This led us to hypothesize that the expression of an endogenous gene will be as efficiently silenced as that of

ectopically expressed mRNA by an effective siRNA. We designed constructs that express target protein fused at either the N- or C-terminus to a fluorescent (such as enhanced green fluorescent protein, EGFP or red fluorescent protein, RFP) or enzymatic (*Renilla luciferase*) reporter polypeptide (Fig. 1). These constructs were transfected into mammalian cells together with individual siRNAs designed against various regions of the target gene (Fig. 1A,B). If any siRNA effectively targets and inactivates expression of its cognate mRNA, a marked reduction in reporter expression (EGFP/RFP fluorescence or enzymatic activity) is expected. Conversely if it fails to efficiently target its cognate mRNA, no significant change in reporter expression should be observed. Both the fluorescence and enzymatic activities can be quantitated.

The feasibility of the experimental design was tested by evaluating critical parameters associated with the target-reporter fusion products such as stability of fusion proteins, accessibility of target site in the chimeric mRNA, and specificity of siRNA probes in suppressing cognate mRNA expression as reflected by changes in reporter expression. EGFP and RFP were used as a target and reporter respectively, and an effective siRNA specific to EGFP mRNA was used (Donze and Picard 2002). A plasmid expressing EGFP was cotransfected with an effective siRNA targeting specifically EGFP mRNA (Donze and Picard 2002) with a control plasmid expressing RFP. A significant reduction in EGFP, but not in RFP, expression was observed (Fig. 1C, upper panel), demonstrating specificity of the siRNA for the target gene. Transfection of the cells with a nonspecific (NON-SP) siRNA affected neither EGFP nor RFP expression (Fig. 1C, lower panel). We then tested both an N-terminal and C-terminal target-reporter fusion (EGFP-RFP and RFP-EGFP) to determine whether siRNA against the target (EGFP) would result in the abrogation of reporter (RFP) gene expression. The results showed reduced expression of EGFP and RFP for both N- and C-terminal fusions (Fig. 1D,E, upper panel) as compared to nonspecific siRNA (Fig. 1D,E, lower panel), suggesting that the siRNA-mediated suppression of the target gene (EGFP) expression is faithfully reported by the reporter (RFP) portion of the fusion constructs. Taken together the data suggest that the target-reporter products are stable and that the reporter portion of the fusion construct accurately reflects silencing of the target when fused to either its N- or C-terminus. The latter property is particularly attractive, because it allows for flexibility in the construction of chimeric fusions.

To provide flexibility in the choice of reporter systems, we also explored the use of an enzymatic rather than a fluo-

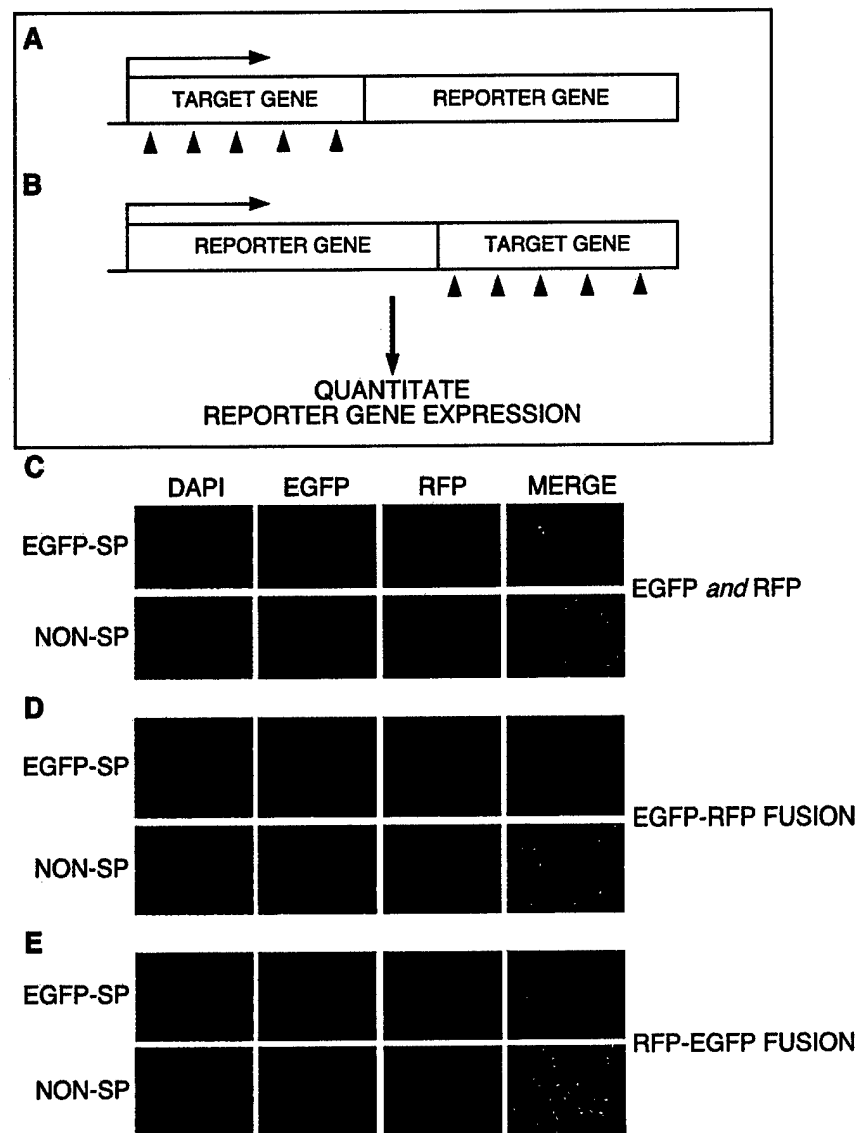


Figure 1 Strategy and experimental verification for screening effective siRNA probes using target-reporter fusion. (A) Reporter gene fused at either C-terminal or (B) N-terminal of target gene. A panel of siRNAs or shRNAs against target gene is shown (▲). Efficacy of siRNA-mediated target gene silencing was measured by quantitation of reporter gene expression. C2C12 cells were cotransfected with either EGFP-specific siRNA (upper panel) or nonspecific siRNA (lower panel) with (C) two independent plasmids expressing EGFP and RFP, (D) a plasmid expressing EGFP-RFP fusion protein, and (E) a plasmid expressing RFP-EGFP fusion protein. DAPI, 4', 6'-diamidino-2-phenylindole hydrochloride; EGFP, enhanced green fluorescent protein; RFP, red fluorescent protein; NON-SP, nonspecific siRNA; SP, specific siRNA. A scrambled siRNA served as a nonspecific control.

rescent reporter. Plasmids encoding an EGFP-*Renilla* luciferase fusion protein and internal control firefly luciferase, respectively, were transfected with varying concentrations of either EGFP-specific or nonspecific siRNAs. The internal control was used to normalize experimental variation associated with transfection efficiency. The results showed a specific dose-dependent decrease in EGFP expression with 50–250 ng of siRNA (Fig. 2A; upper panel). In contrast, nonspecific siRNA had no detectable effect on EGFP expression in this range (Fig. 2A, lower panel). Moreover, normalized quantitation of *Renilla* and firefly luciferase activity from lysates of aliquots of the same transfected cells showed that EGFP siRNA specifically decreases *Renilla* luciferase activity in a dose-dependent manner (Fig. 2B), consistent with the decrease in EGFP fluorescence (Fig. 2A). An approximately fivefold reduction in *Renilla* luciferase by specific siRNA relative to a control nonspecific siRNA was observed. Thus, both a fluorescent and an enzymatic reporter can be used effectively to monitor RNAi-mediated gene silencing. In alternative embodiments, a reporter gene which encodes a protein sequence such as His-tags, immunoglobulin domains, myc tags, poly-glycine tags, FLAG tags, or HA-tags may be used. These reporters can be conveniently detected by immunoassay methods, such as Western blotting, immunohistochemistry, ELISA, and/or immunoprecipitation.

Application of the Method in Identifying Effective siRNA Probes

To evaluate the usefulness of the screening method for identifying effective siRNA probes, we designed five siRNAs targeting various regions spanning the MyoD coding sequence (Fig. 3A). The *MyoD* gene was used for the first screen because of its robust expression in muscle precursor cells and the availability of reliable antibodies to monitor levels of the protein. Cells were cotransfected with a construct expressing a MyoD-EGFP fusion protein, the dSRed2-N1 plasmid as an internal control for transfection, and individual siRNA probes. siRNA 25 showed dramatic reduction in the number of EGFP-positive cells compared to cells transfected with nonspecific siRNA (NON-SP, Fig. 3B), as determined by fluorescence microscopy. The normalized fluorescence intensity ratio of target (MyoD-EGFP) to internal control (RFP) was obtained from lysates of aliquots of the same transfected cells. The results show that siRNA 25 was the most effective (almost 50% reduction in GFP expression relative to nonspecific

siRNA) in the suppression of ectopic *MyoD* gene expression (Fig. 3C), in agreement with the microscopic imaging (Fig. 3B). Partial reduction in normalized GFP fluorescence was observed with siRNA 294 (20%) and siRNA 538 (10%). To determine whether there is a correlation between the ability of the siRNA to inhibit expression of both ectopically expressed MyoD-EGFP and endogenous MyoD, cells were transfected with only siRNAs specific to MyoD. The level of endogenous MyoD was then examined by Western blot analysis. We observed a strong correlation between suppression of ectopic MyoD-EGFP and endogenous MyoD gene expression by the same panel of siRNAs (Fig. 3D, see percent reduction in MyoD expression), with siRNA 25 being the most effective (80% reduction) in silencing endogenous MyoD expression (Fig. 3D, upper panel) followed by siRNA 294 (35%) and siRNA 538 (10%).

This strategy was then assayed for selecting effective plasmid-encoded shRNA. Plasmid-encoded shRNAs have distinct advantages over their chemically synthesized siRNA counterparts. They are cost-effective and provide a stable and continuous expression of hairpin RNA that is useful for analysis of phenotypes that develop over extended periods of time as a result of loss of gene function. We chose five different target sequences for MyoD hairpins, as shown in Figure 3E. shRNA 708 showed a dramatic reduction in the number of EGFP-positive cells (Fig. 3F, upper right panel) compared to nonspecific shRNA (Fig. 3F, upper left panel) or other shRNAs (data not shown). The normalized fluorescence intensity ratio of target (MyoD-EGFP) to internal control RFP was quantitated and confirmed the effectiveness of shRNA 708 (almost 65% reduction in GFP expression relative to nonspecific siRNA), consistent with the microscopy imaging analysis (Fig. 3F). To monitor the ability of the various shRNAs to inhibit expression of the endogenous MyoD gene, cells were transfected only with shRNAs. Western blot experiments showed that shRNA 708 (85% reduction in MyoD protein levels) was the most effective in suppressing endogenous MyoD gene expression (Fig. 3G, upper panel) followed by shRNA 312 and shRNA 897 (50% reduction in each case), in agreement with the fluorescent intensity results (Fig. 3F).

We next determined the general applicability of the siRNA screening strategy in identifying effective RNAi candidates for additional genes. The silencing of Lamin A/C by a panel of siRNAs was analyzed (Fig. 4A). siRNA 608, which was previously shown to be effective against Lamin A/C (Harborth et al. 2001),

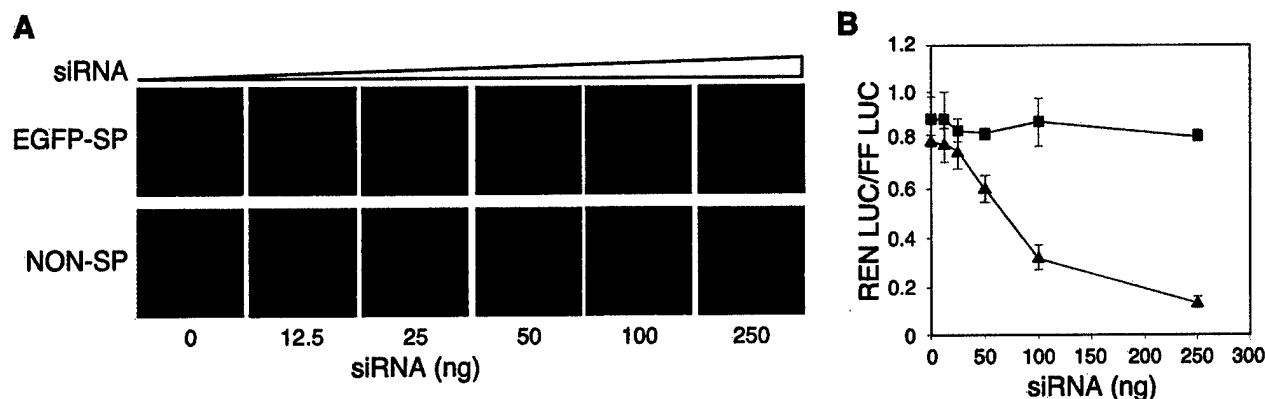


Figure 2 Quantitation of siRNA-mediated target gene silencing. (A) C2C12 cells were transfected with a constant amount of EGFP-Rluc (expressing EGFP-*Renilla* luciferase fusion protein), pGL3-Control (expressing firefly luciferase as internal control), and increasing concentrations of EGFP-specific siRNA (upper panel) and nonspecific siRNA (lower panel). At 24 h posttransfection cells were imaged for EGFP expression. (B) Normalized relative amount of *Renilla* and firefly luciferase (n=3) as a function of treatment with nonspecific siRNA (■) or EGFP-specific siRNA (▲) were plotted. Cotransfection without siRNA served as a control. A scrambled siRNA served as a nonspecific control.

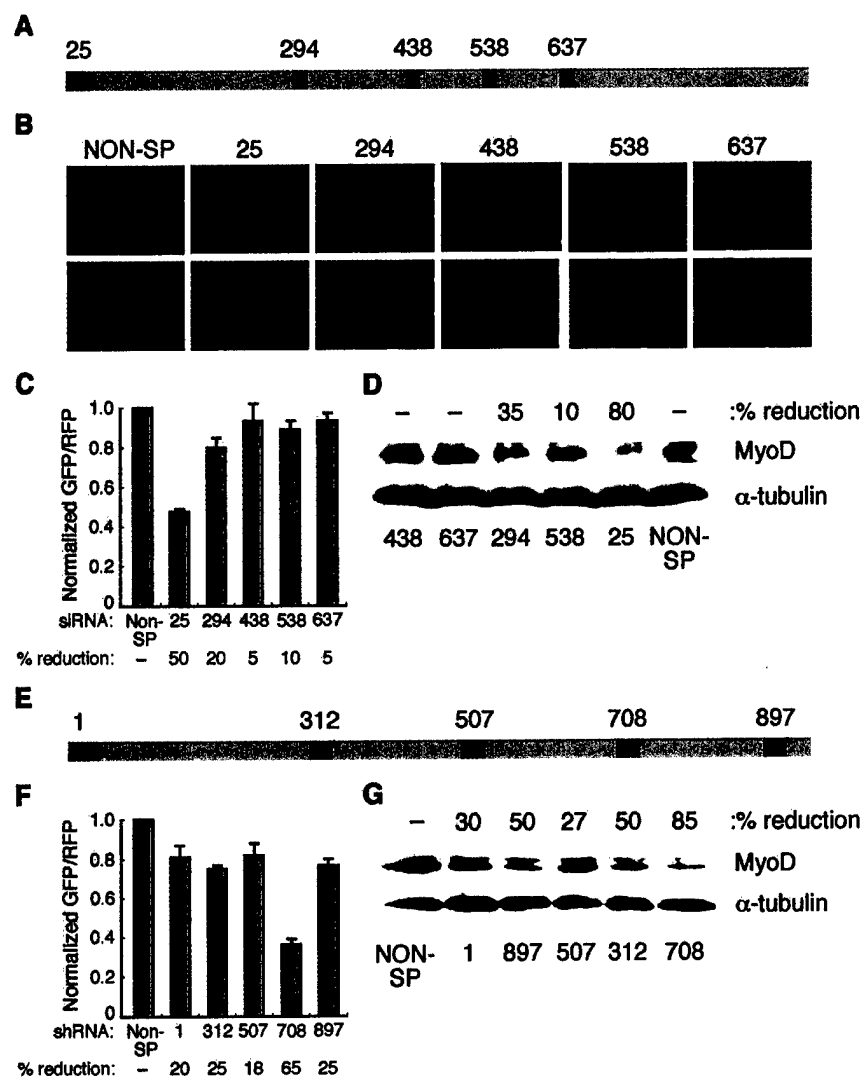


Figure 3 Correlation between siRNA/shRNA screening results and suppression of endogenous MyoD. (A) Schematic showing location of various siRNA probes against MyoD gene. The location of siRNA probes on the target sequences were assigned with respect to the translation start site (nucleotide position 1). (B) C2C12 cells were transfected with MyoD-EGFP, pDsRed2-N1 (internal control), and specific siRNAs against MyoD or a nonspecific siRNA as indicated. Upper and lower panels show EGFP and RFP expression, respectively, after 24 h of transfection. (C) Normalized relative amounts of EGFP and RFP for each siRNA. Percentage (%) reduction in GFP fluorescence is indicated below each siRNA. (D) C2C12 cells were transfected with specific MyoD siRNAs or a nonspecific siRNA and subjected to Western blot analysis after 48 h of transfection for MyoD (upper panel) and α -tubulin (lower panel) protein. The percent reduction in band intensity relative to nonspecific control is shown. (E) Schematic showing location of the shRNAs along the sequence of the MyoD gene. (F) Normalized relative amount of EGFP and RFP for each shRNA. % reduction in GFP fluorescence is indicated below each shRNA. (G) C2C12 cells were transfected with specific MyoD shRNAs or a nonspecific shRNA and subjected to Western blot analysis after 48 h of transfection for MyoD (upper panel) and α -tubulin (lower panel) protein. shRNA against firefly luciferase and scrambled siRNA served as a nonspecific control. All of the blots were reprobbed with anti- α -tubulin (Sigma) antibody to show equal loading.

was included. Of the five siRNAs tested, siRNA 608 was by far the most effective in suppressing both ectopic Lamin-GFP fusion (60% reduction in relative fluorescence intensity, Fig. 4B) and endogenous Lamin A/C (87% reduction in lamin A/C protein level, Fig. 4C), in agreement with the 90% reduction observed previously using the same siRNA (Elbashir et al. 2001a). Similarly, in a screen of a panel of shRNA probes against the human tumor suppressor p53, a sequence known to be effective (Brummelkamp

et al. 2002) performed most efficiently in our screen (Table 1). Taken collectively, these results validate the reliability of this approach for RNAi screening.

We have screened panels of siRNAs and shRNA probes against genes with diverse biological functions in both murine and human cell lines. Table 1 summarizes the screening results described above obtained with MyoD, Lamin A/C, and p53 as well as results obtained with the human EF-hand calcium binding protein S100 α -subunit and members of the Id family of genes. These results further underscore that only a minority of RNAi probes are effective in gene silencing, and that the effective RNAi probes can be identified from this screen. Thus far we have always observed a correlation between the efficiency of an siRNA in the screen and the suppression of the cognate endogenous gene. However, the generality of this correlation can only be determined by performing a comprehensive analysis involving a broad spectrum of genes. It is possible that factors such as enhanced transcriptional rates, protein stability, and density of RNA binding proteins associated with some genes may compromise the in vivo effectiveness of siRNA. This may result in a small proportion of false-negatives in this screening methodology.

Our assay can be used to screen for effective siRNAs directed against targets that cannot be monitored directly, either because of lack of reagents such as antibodies, or because of very low expression levels. Another strength of the screening method is its ability to identify the most robust siRNA candidate within 24 h of transfection irrespective of the status of the endogenous protein. This is particularly attractive compared to determining the efficacy of siRNA probes by directly monitoring their ability to suppress cognate endogenous genes, which may involve time-consuming optimization of siRNA dose and incubation time (Elbashir et al. 2001a; Harborth et al. 2001; Mendez et al. 2002). In several cases we performed the screening using a variety of human and murine cell lines (HeLa, NIH3T3, murine endothelial, HEK293, IMR90, C2C12); in all cases the results were independent of the cell lines used (data not shown; also note that siRNA 25 and shRNA 708 against MyoD were identified using both murine C2C12 (Fig. 3F) and human HeLa cells [Fig. 5F,G]).

The screening method allows not only the identification of the most effective siRNAs, but also that of siRNAs that display partial suppression of target gene expression. We predict that such siRNAs would be useful in studies where partial downregulation of gene expression is desired, particularly in cases where partial downregulation might result in distinct phenotypes. For

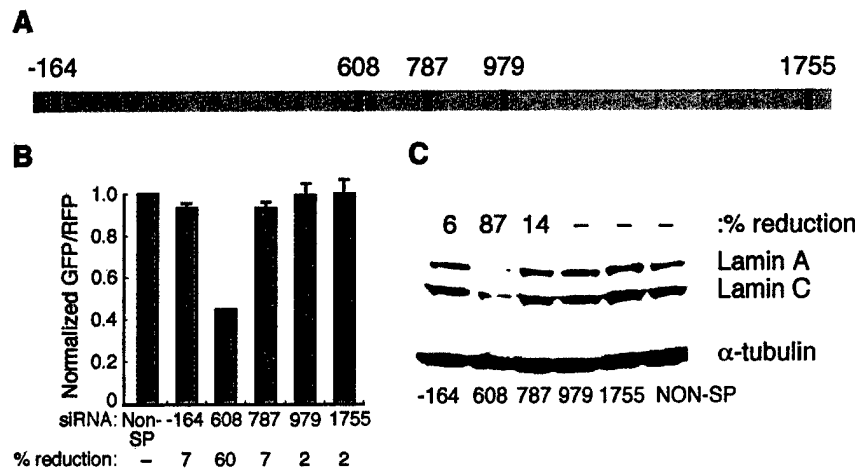


Figure 4 Correlation between siRNA screening results and suppression of endogenous Lamin A/C. (A) The location of the siRNAs against the *Lamin A/C* gene. (B) HeLa cells were transfected with a constant amount of EGFP-Lamin A/C, pDsRed2-N1, and specific Lamin A/C siRNA and nonspecific siRNA. Normalized relative amounts of EGFP and RFP for each siRNA and % reduction in GFP fluorescence for each siRNA are shown. (C) Western blot analysis using α -Lamin A/C antibody.

example, shRNAs showing varying levels of p53 suppression generated distinct tumor phenotypes in vivo (Hemann et al. 2003). They would also be useful where lethality associated with complete suppression of critical genes is of concern.

A High-Throughput Selection of siRNA Probes by RNAi Microarrays

Currently a major effort is geared towards using RNAi as a reverse genetic approach for analyzing mammalian gene function on a genome-wide scale as demonstrated in *C. elegans* (Kamath et al. 2003). However unlike *C. elegans*, where long dsRNAs work well in sequence-specific gene silencing due to lack of an antiviral/interferon response, successful applications using RNAi in somatic mammalian cells would require reliable RNAi screens to identify effective siRNA probes for every gene. We were interested in determining whether our method for selecting effective siRNA probes would work in a highly parallel assay. We made use of a DNA microarray platform where two plasmids, one encoding a target gene-EGFP fusion and another encoding RFP (as an internal control) were arrayed with either specific siRNA or nonspecific siRNA on a glass slide and then incubated with mammalian cell culture (Fig. 5A). Only the cells growing in close proximity to the DNA spots were transfected, resulting in spatially distinct groups of transfected cells within a lawn of untransfected cells, as observed previously (Ziauddin and Sabatini 2001). A laser scanner was used to monitor EGFP and RFP fluorescence intensities in each individual cluster. Figure 5C is a laser scan image of spots expressing EGFP (green) and RFP (red) expression, and Figure 5D is a microscopy image of a representative spot. Typically, each cluster was comprised of 300–500 fluorescent cells. We next sought to determine whether the microarray-based cell transfection format would recapitulate siRNA-mediated suppression of ectopic gene expression as a function of siRNA concentration and specificity, as observed in conventional transfections (Fig. 2). Quantitation of normalized mean intensities (EGFP/RFP) showed a dose-dependent suppression of EGFP expression by its cognate siRNA (Fig. 5E), with 300 ng of siRNA providing maximal suppression. Moreover, an effective siRNA directed to EGFP showed specific suppression of EGFP and not RFP expression (Fig. 5B). All of these results are also

consistent with those obtained in a conventional transfection assay (Fig. 2).

We then tested the use of cell microarrays to screen effective siRNA probes. First the concentrations of the target-reporter fusions were optimized for detection of fluorescence by a laser scanner (data not shown). Next, the panels of six siRNAs and six shRNAs directed against MyoD described above (Fig. 3A,E) were arrayed together with the plasmids encoding MyoD-EGFP and the internal control RFP. The quantitation of EGFP and the RFP internal control intensities in response to various shRNAs and siRNAs (Fig. 5F,G) identified shRNA 708 and siRNA 25 as the most effective in suppression of MyoD-EGFP expression, in agreement with our observations with conventional transfections (Fig. 3). Although by using conventional transfections we observed a partial suppression of MyoD gene expression by siRNA 294 (Fig. 3C,D) and shRNA 312 (Fig. 3G), the same probes did not show

similar trends in the microarray-based screen (Fig. 5F,G). This is likely due to differences in the sensitivity and specificity of detection of the two screening platforms. Nevertheless, the two types of screening are in perfect agreement in their ability to identify the most effective RNAi probe.

These results suggest that the microarray-based cell transfections (“RNAi microarrays”) can be used in large-scale RNAi screens. Using fully automated liquid-dispensing and plate-handling robotic systems, it is possible to prepare mixtures containing constructs expressing target-reporter fusions, internal controls, various shRNAs, and synthetic siRNAs which then can be printed at densities of up to 6000 to 10,000 features per slide by modern microarrayers. Although similar screens can be carried out in 96- or 384-well plates, the microarray platform provides distinct advantages. The arrays are cost-effective because they require minimal materials (plasmids, shRNA, or siRNA) and expensive transfection reagents. Mass transfection on arrays allows homogenous distribution of cell population in comparison to variability often encountered in well-to-well transfections. Besides, several slides can be generated and the screens performed in duplicate or triplicate to obtain statistically significant results. However, a 96- or 384-well plate format is more suitable for enzymatic assays, such as for luciferase, β -galactosidase, etc.

Table 1. siRNA/shRNA-Induced Gene Silencing for Ectopically Expressed Target-Reporter Fusions

Gene	Accession	siRNA*	shRNA*
MyoD	M84918	5 (1)	5 (1)
Lamina/C	NM_005572	5 (1)	0
S-100	NM_002961.2	5 (0)	0
Id1	AK008264	5 (1)	8 (1)
Id2	AF077860	5 (0)	3 (1)
Id3	AK002820	5 (1)	8 (1)
Id4	AF077859	5 (1)	3 (0)
p53	X02469	5 (1)	4 (1)

*Of the total siRNA/shRNA tested in the screens, the number of the most effective probes are mentioned in parentheses.

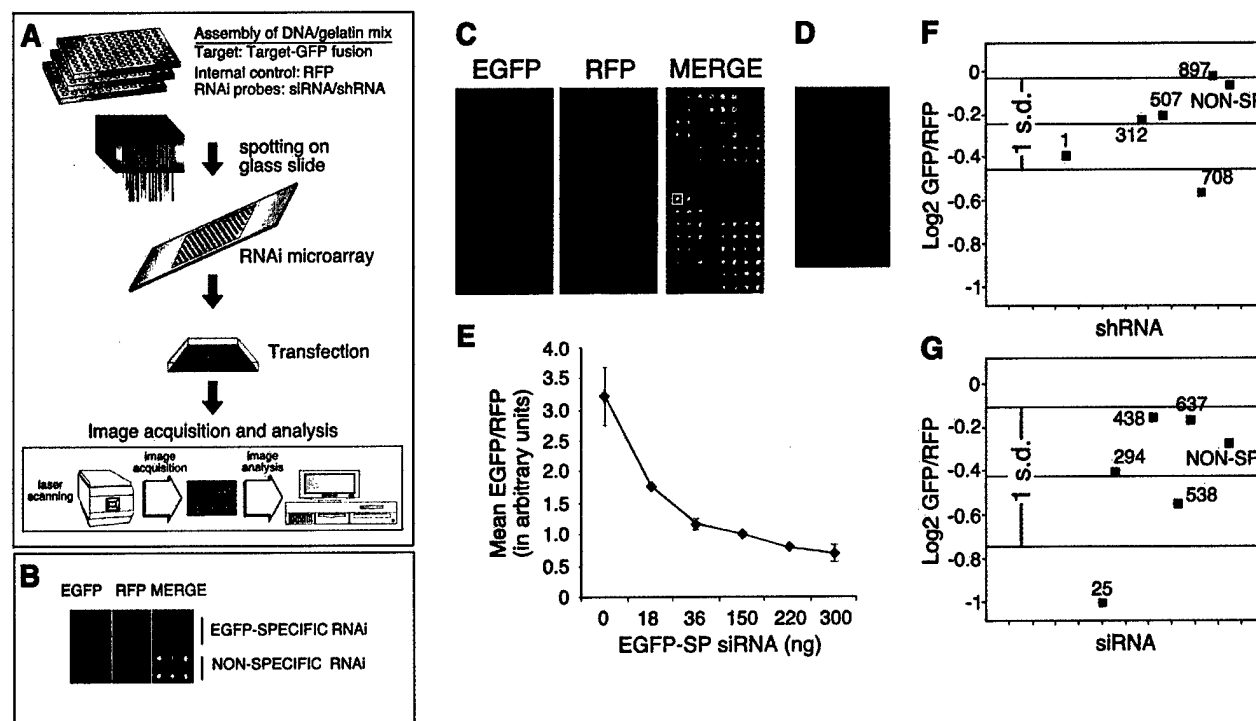


Figure 5 siRNA screens using microarray-based cell transfections. (A) Schematic showing manufacture and analyses of an RNAi microarray. (B) RNAi microarray showing that an effective EGFP specific siRNA results in the suppression of EGFP but not RFP expression. A nonspecific siRNA did not affect EGFP or RFP expression. (C) A laser scan image of EGFP- and RFP-expressing microarray. Each feature was 500 μ m in diameter with a pitch of 750 μ m. (D) Magnified image of a feature represented by a boxed area of an image from A. (E) Dose-dependent response of EGFP-specific siRNA in suppression of EGFP expression as quantitated by mean trimmed normalized values of EGFP/RFP mean intensities ($n=4$). (F) Screen for identifying effective shRNA against *MyoD* gene. (G) Screen for identifying effective siRNA against *MyoD*. Normalized EGFP/RFP median intensities (y-axis) were log transformed, normalized ($n=4$), and plotted against individual siRNA/shRNA probes (x-axis). In each case probes within and outside 1 standard deviation (s.d.) from the mean value were considered noneffective (marked in black) and effective (marked in red), respectively.

Although a major time constraint is the generation of constructs expressing the target-reporter fusion proteins, a novel cloning strategy to rapidly generate both N- and C-terminus GFP fusions in a high-throughput format is available (Simpson et al. 2000). This method allows for cloning of amplified open reading frames (ORFs) by recombination and therefore circumvents the need for screening for restriction sites in the ORF in question. Another possibility is the use of an internal ribosomal entry site (IRES) between the target and the reporter. This would bypass the stringent requirement of maintaining ORFs and simplify the cloning of target-reporter fusion constructs considerably. We observed weak IRES-mediated expression of GFP that was unsuitable for quantitation in our screening assays (data not shown).

We envisage that RNAi microarrays can also be directly exploited for scoring RNAi-generated phenotypes in mammalian cells that result from the loss of gene function, such as NuMa, GAS41, Lamina/C/B1/B2, β/γ actin, zyxin, Eg5, CENP-E, and Cdk1 shown previously in conventional transfection assays (Harborth et al. 2001).

In summary, we have designed a rapid, inexpensive, powerful, reliable, and quantitative method for screening the most effective siRNA or shRNA for suppressing gene expression in mammalian cells. This method is likely to be useful until detailed biochemical mechanisms of RNAi pathways are uncovered that will provide more rational strategies for efficiently targeting suppression of any desired gene. The ability to quickly identify effective siRNA for silencing any gene is likely to have significant implications not only in basic research, but also in RNAi-based therapeutics (Agami 2002; Cottrell and Doering 2003; Shi 2003)

and in the generation of genetically modified animal models (Hasuwa et al. 2002; Kim et al. 2002; Carmell et al. 2003).

METHODS

Generation of siRNAs, shRNAs, and Target-Reporter Fusion Constructs

Synthetic siRNAs were either purchased from Dharmacon Research or synthesized by *in vitro* transcription (Donze and Picard 2002). Specific siRNA against EGFP (Donze and Picard 2002) and scrambled siRNA (sense 5'-ATGATACTCGAGGGCATGTCTC-TATAGTGAGTCGTATTA-3' and antisense 5'-CGGAGACATGC-CCTCGAGTATCTATAGTGAGTCGTATTA-3') were synthesized. siRNA and shRNA probes (see Supplemental Fig. 1, available online at www.genome.org) against target genes were designed using computer software (<http://www.cshl.org/public/SCIENCE/hannon.html>). The location of the siRNAs and shRNAs on the target sequences were assigned with respect to the translational start site, ATG (where nucleotide A was designated position 1). shRNA sequences (two complementary ~83-nt DNA oligos) were annealed and cloned directly into a U6 promoter-containing vector (Paddison et al. 2002) using a ligation-independent cloning method (S. Aruleba, K. Scobie, and D.S. Conklin, unpubl.). shRNA against firefly luciferase (Paddison et al. 2002) and a scrambled siRNA served as a nonspecific control. The MyoD-EGFP fusion construct was prepared by amplifying MyoD cDNA (M 84918) from a pCMV-MyoDs vector and inserted into XhoI and BamHI sites of pEGFP-N2 (Clontech). EGFP-*Renilla* luciferase (EGFP-Rluc) fusion construct was prepared by cloning PCR-amplified EGFP product from pEGFP-N2 (Clontech) into the

BamHI site of pRLuc-N3 (PerkinElmer).

EGFP-Lamin A/C fusion construct was prepared by blunt end ligation (filled in by T4 DNA polymerase) of *Bsr*GI and *Not*I-digested pEGFP-N2 vector with a *Not*I and *Sall*-digested Lamin A/C (NP_005563) insert from Lamin A/C-pSPORT I vector (Research Genetics). In all cases, the integrity and orientation of the target gene with respect to the reporter gene were confirmed by restriction enzyme digestion and DNA sequencing.

Cell Culture, Transfection, and Western Blot Analysis

Murine C2C12 and human HeLa cells (ATCC) were cultured in DMEM supplemented with 10% FBS, 100 U/mL penicillin, and 100 µg/mL streptomycin (Life Technologies). Cells were transfected in 6-well plates using Lipofectamine PLUS (Life Technology) and 2 µg of siRNAs or shRNAs. For Western blot analysis, cells were harvested at 48 h posttransfection and washed with TBS (50 mM Tris, pH 8.0, 150 mM NaCl) and lysed in 100 µL of RIPA lysis buffer (TBS supplemented with 1% NP-40 and complete protease inhibitors, Roche Applied Science). Equal amounts of lysate were subjected to Western blot analysis (Harlow and Lane 1988) using α -MyoD or Lamin A/C primary antibody (Santa Cruz) and horseradish peroxidase-conjugated secondary antibody. The blots were stripped twice by stripping buffer (100mM β -mercaptoethanol, 2% SDS, 62.5 mM Tris-HCl, pH 6.7) at 55°C and reprobed with anti- α -tubulin (Sigma) primary antibody and horseradish peroxidase-conjugated secondary antibody. The blots were developed using an ECL Plus Western Blotting Detection System (Amersham Biosciences). Signal intensities were determined after background corrections by using Alpha-Imager 2000 Documentation and Analysis Software (Alpha Innotech). The percentage reduction in band intensity for each siRNA/shRNA relative to nonspecific control (assumed to be unchanged) was calculated.

Screening Assay

For each well of a 24-well plate, 150 ng of siRNAs or shRNAs, 100 ng of target-reporter fusion construct, and 50 ng of pSRed2-N1 (internal control) were used in transfections. Twenty-four h posttransfection, EGFP and RFP images were captured using a Zeiss AxioCam HRm camera at equal exposure time for all panels. Cleared cell lysates were prepared as described above for the Western blot analysis, and EGFP and RFP intensities were quantitated using a Multilabel Counter (PerkinElmer) with Wallac 1420 software. The EGFP/RFP ratio was calculated and normalized against nonspecific control as described (Chiu and Rana 2002).

Dual Luciferase Assay

In each well of a 24-well plate, C2C12 cells were cotransfected with 300 ng EGFP-Rluc, 200 ng pGL3-Control (Promega), and increasing concentrations (12.5–250 ng) of specific EGFP siRNA or nonspecific siRNA. After 24 h of transfection, the relative amounts of *Renilla* and firefly luciferase were analyzed by a dual luciferase assay (Promega) using a luminometer (Model 3010, Analytical Scientific Instruments). The *Renilla*/firefly luciferase ratio was calculated and normalized against the control.

Microarray-Based Cell Transfections (RNAi Microarray): Printing, Transfection, Laser Scanning, and Quantitation

Essentially a mixture containing plasmid DNA and various siRNA or shRNA in 0.18% gelatin solution was prepared and arrayed in a 384-well plate. These were printed on CMT GAPS glass slides (Corning) by an arrayer (VP478A, V & P Scientific) at 4°C. The final plasmid concentrations for EGFP fusion construct or nuc-EGFP (Clontech) and pSRed2-N1 (internal control) were 150 ng/µL and 50 ng/µL, respectively. The shRNA or siRNA concentration was kept constant at 300ng/µL or as mentioned. For transfections, 24 µL of Lipofectamine 2000 (Invitrogen) was mixed with 300 µL of OPTI-MEM I media (GibcoBRL) and pipetted onto a 40 × 20 × 0.2 mm cover well (PC200; Grace Bio-

Labs). A slide with the printed side down was placed on the cover well and incubated for 45 min. The cover well was removed and the slide was incubated with 1×10^6 HeLa cells per mL for 24 h with a media change at 6 h. The cells were fixed and mounted with a coverslip. More details on printing and transfection can be found at (Ziauddin and Sabatini 2001). The slides were scanned using a laser scanner (ScanArray 5000; PerkinElmer) at 20 µm resolution to measure EGFP and RFP fluorescence. After scanning, the EGFP and RFP intensities of each spot were quantitated using GenePix 4.0 software (Axon Instruments). In all analyses, features showing obvious blemishes and morphological defects were eliminated. Features with low intensities (<100 units) in the red channel (RFP fluorescence) were considered inefficient transfections and were removed from further analysis. Data were normalized by trimmed mean with the trim value set to 5% (highest 5% of the values and the lowest 5% of the values are excluded from the calculated mean) to reduce the effects of outliers. Each spot was represented in quadruplet, and average values were used for final quantitation. Data files are available as Supplemental Figures 2 and 3.

ACKNOWLEDGMENTS

We thank Nouria Hernandez, Eli Hatchwell, James Egan, and Sunita Gupta for comments on the manuscript and helpful discussions, and Greg Hannon, Jose Silva, Robert Benezra, Kimberly LaVine, Sunita Gupta, David Helfman, Eli Hatchwell, Sibel Kantaraci, and Andrew Lassar for reagents.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

- Agami, R. 2002. RNAi and related mechanisms and their potential use for therapy. *Curr. Opin. Chem. Biol.* **6**: 829–834.
- 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.
- Brummelkamp, T.R., Bernards, R., and Agami, R. 2002. A system for stable expression of short interfering RNAs in mammalian cells. *Science* **296**: 550–553.
- Bryom, M., Pallota, V., Brown, D., and Ford, L. 2002. Visualizing siRNA in mammalian cells: Fluorescence analysis of the RNAi effect. *Ambion TechNotes* **9**: 3–5. www.ambion.com/tech/ib/tn/101/4.html.
- Calegari, F., Haubensak, W., Yang, D., Huttner, W.B., and Buchholz, F. 2002. Tissue-specific RNA interference in postimplantation mouse embryos with endoribonuclease-prepared short interfering RNA. *Proc. Natl. Acad. Sci.* **99**: 14236–14240.
- Carmell, M.A., Zhang, L., Conklin, D.S., Hannon, G.J., and Rosenquist, T.A. 2003. Germline transmission of RNAi in mice. *Nat. Struct. Biol.* **10**: 91–92.
- Chiu, Y.L. and Rana, T.M. 2002. RNAi in human cells: Basic structural and functional features of small interfering RNA. *Mol. Cell* **10**: 549–561.
- Cottrell, T.R. and Doering, T.L. 2003. Silence of the strands: RNA interference in eukaryotic pathogens. *Trends Microbiol.* **11**: 37–43.
- Donze, O. and Picard, D. 2002. RNA interference in mammalian cells using siRNAs synthesized with T7 RNA polymerase. *Nucleic Acids Res.* **30**: e46.
- Eibashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. 2001a. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**: 494–498.
- Eibashir, S.M., Lendeckel, W., and Tuschl, T. 2001b. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes & Dev.* **15**: 188–200.
- Eibashir, S.M., Harborth, J., Weber, K., and Tuschl, T. 2002. Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods* **26**: 199–213.
- Gil, J. and Esteban, M. 2000. Induction of apoptosis by the dsRNA-dependent protein kinase (PKR): Mechanism of action. *Apoptosis* **5**: 107–114.
- Hammond, S.M., Caudy, A.A., and Hannon, G.J. 2001. Posttranscriptional gene silencing by double-stranded RNA. *Nat. Rev. Genet.* **2**: 110–119.
- Hannon, G.J. 2002. RNA interference. *Nature* **418**: 244–251.
- Harborth, J., Eibashir, S.M., Beichert, K., Tuschl, T., and Weber, K. 2001. Identification of essential genes in cultured mammalian cells using

- small interfering RNAs. *J. Cell Sci.* **114**: 4557–4565.
- Harlow, E. and Lane, D. 1988. *Antibodies: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Hasuwa, H., Kaseda, K., Elnarsdottir, T., and Okabe, M. 2002. Small interfering RNA and gene silencing in transgenic mice and rats. *FEBS Lett.* **532**: 227–230.
- Hemann, M.T., Fridman, J.S., Zilfou, J.T., Hernando, E., Paddison, P.J., Cordon-Cardo, C., Hannon, G.J., and Lowe, S.W. 2003. An epiallelic series of p53 hypomorphs created by stable RNAi produces distinct tumor phenotypes in vivo. *Nat. Genet.* **3**: 3.
- Hohjoh, H. 2002. RNA interference (RNAi) induction with various types of synthetic oligonucleotide duplexes in cultured human cells. *FEBS Lett.* **521**: 195–199.
- Holen, T., Amarzguloul, M., Wiiger, M.T., Babale, E., and Prydz, H. 2002. Positional effects of short interfering RNAs targeting the human coagulation trigger tissue factor. *Nucleic Acids Res.* **30**: 1757–1766.
- Kamath, R.S., Fraser, A.G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., et al. 2003. Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* **421**: 231–237.
- Kapadia, S.B., Brideau-Andersen, A., and Chisari, F.V. 2003. Interference of hepatitis C virus RNA replication by short interfering RNAs. *Proc. Natl. Acad. Sci.* **100**: 2014–2018.
- Kim, M.H., Yuan, X., Okumura, S., and Ishikawa, F. 2002. Successful inactivation of endogenous Oct-3/4 and c-mos genes in mouse preimplantation embryos and oocytes using short interfering RNAs. *Biochem. Biophys. Res. Commun.* **296**: 1372–1377.
- Lee, N.S., Dohjima, T., Bauer, G., Li, H., Li, M.J., Ehsani, A., Salvaterra, P., and Rossi, J. 2002. Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat. Biotechnol.* **20**: 500–505.
- Marcus, P.I. and Sekellick, M.J. 1985. Interferon induction by viruses. XIII. Detection and assay of interferon induction-suppressing particles. *Virology* **142**: 411–415.
- McManus, M.T., Petersen, C.P., Haines, B.B., Chen, J., and Sharp, P.A. 2002. Gene silencing using micro-RNA designed hairpins. *RNA* **8**: 842–850.
- Mendez, J., Zou-Yang, X.H., Kim, S.Y., Hidaka, M., Tansey, W.P., and Stillman, B. 2002. Human origin recognition complex large subunit is degraded by ubiquitin-mediated proteolysis after initiation of DNA replication. *Mol. Cell* **9**: 481–491.
- Paddison, P.J., Caudy, A.A., Bernstein, E., Hannon, G.J., and Conklin, D.S. 2002. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes & Dev.* **16**: 948–958.
- Paul, C.P., Good, P.D., Winer, I., and Engelke, D.R. 2002. Effective expression of small interfering RNA in human cells. *Nat. Biotechnol.* **20**: 505–508.
- Sharp, P.A. 1999. RNAi and double-strand RNA. *Genes & Dev.* **13**: 139–141.
- Shi, Y. 2003. Mammalian RNAi for the masses. *Trends Genet.* **19**: 9–12.
- Simpson, J.C., Wellenreuther, R., Poustka, A., Pepperkok, R., and Wiemann, S. 2000. Systematic subcellular localization of novel proteins identified by large-scale cDNA sequencing. *EMBO Rep.* **1**: 287–292.
- Sui, G., Soohoo, C., Affar el, B., Gay, F., Shi, Y., and Forrester, W.C. 2002. A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. *Proc. Natl. Acad. Sci.* **99**: 5515–5520.
- Yang, D., Buchholz, F., Huang, Z., Goga, A., Chen, C.Y., Brodsky, F.M., and Bishop, J.M. 2002. Short RNA duplexes produced by hydrolysis with *Escherichia coli* RNase III mediate effective RNA interference in mammalian cells. *Proc. Natl. Acad. Sci.* **99**: 9942–9947.
- Yu, J.Y., DeRulter, S.L., and Turner, D.L. 2002. RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells. *Proc. Natl. Acad. Sci.* **99**: 6047–6052.
- Ziauddin, J. and Sabatini, D.M. 2001. Microarrays of cells expressing defined cDNAs. *Nature* **411**: 107–110.

WEB SITE REFERENCES

<http://www.cshl.org/public/SCIENCE/hannon.html>; RNAi OligoRetriever.

Received May 22, 2003; accepted in revised form August 1, 2003.

shRNA-mediated Silencing of Mammalian Gene Expression

Michael T. McManus* and Douglas S. Conklin†

*Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139;

†Cancer Genome Research Center, Cold Spring Harbor Laboratory, Woodbury, New York 11797

THE RNA HAIRPIN STRUCTURE IS PERHAPS THE SIMPLEST of all RNA secondary structures. It is formed when a single-stranded RNA folds back onto itself, making a double-stranded RNA (dsRNA) stem joined at one end by a single-stranded loop. The hairpin, also referred to as a "stem-loop," is found ubiquitously throughout RNAs of all classes. Its fundamental architectural roles in RNAs such as tRNAs or rRNAs are well recognized, but hairpin RNAs also function in a diversity of other cellular roles. They define nucleation sites for folding, determine tertiary interactions in RNA enzymes, protect mRNAs from degradation, and are recognized by RNA-binding proteins. Recently, a special class of hairpin RNAs has been shown to be responsible for the posttranscriptional regulation of gene expression. These RNAs are central to hairpin-mediated stable gene silencing, a powerful technology for the genetic manipulation of mammalian cells *in vitro* and *in vivo*.

RNA HAIRPINS THAT REGULATE GENE EXPRESSION

A novel strategy to regulate gene expression was discovered when *lin-4* and *let-7* hairpin RNAs were identified in *Caenorhabditis elegans* (Lec et al. 1993; Pasquinelli et al. 2000; Reinhart et al. 2000). The *C. elegans lin-4* regulatory gene, a 22-nucleotide RNA processed from a precursor hairpin RNA (Figure 6.1), was identified in a screen for mutations that

C. elegans lin-4 hairpin precursor

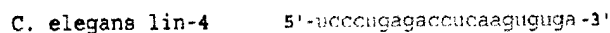
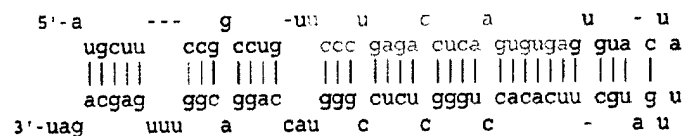


FIGURE 6.1. *lin-4* is processed from a hairpin RNA precursor. Although the exact structure of the precursor has not been determined, the *lin-4* precursor RNA is predicted to form a hairpin precursor RNA containing multiple bulges. The RNase III enzyme Dicer is believed to process the stable *lin-4* small temporal RNA (shown in red) from the hairpin.

number of potential applications. The limitations appear to be only in the nature of the phenotypes and the relative ease of their assay, since the robust nature of this process is likely to make stable silencing a useful technique for the foreseeable future.

ACKNOWLEDGMENTS

We thank Phillip A. Sharp, Chris Petersen, Carl Novina, Shola Aruleba, Greg Hannon, Patrick Paddison, Ravi Sachidanadam, Kim Scobie, and the rest of the Cold Spring Harbor RNAi Group for comments and helpful discussions. M.T.M. is partially supported by the Cancer Research Institute and by the U.S. Health Service MERIT Award (R37-GM34277) from the National Institutes of Health and PO1-CA42063 from the National Cancer Institute. D.S.C. is supported by the U.S. Army Breast Cancer Research Program.

REFERENCES

- Abbas-Terki T., Blanco-Bose W., Deglon N., Pralong W., and Aebischer P. 2002. Lentiviral-mediated RNA interference. *Hum. Gene Ther.* **13**: 2197–201.
- Agami R. 2002. RNAi and related mechanisms and their potential use for therapy. *Curr. Opin. Chem. Biol.* **6**: 829–834.
- Ambros V., Bartel B., Bartel D.P., Burge C.B., Carrington J.C., Chen X., Dreyfuss G., Eddy S.R., Griffiths-Jones S., Marshall M., Matzke M., Ruvkun G., and Tuschl T. 2003. A uniform system for microRNA annotation. *RNA* **9**: 277–279.
- Aoki Y., Cioca D.P., Oidaira H., Kamiya J., and Kiyosawa K. 2003. RNA interference may be more potent than antisense RNA in human cancer cell lines. *Clin. Exp. Pharmacol. Physiol.* **30**: 96–102.
- 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.
- Billy E., Brondani V., Zhang H., Muller U., and Filipowicz W. 2001. Specific interference with gene expression induced by long, double-stranded RNA in mouse embryonal teratocarcinoma cell lines. *Proc. Natl. Acad. Sci.* **98**: 14428–14433.
- Bitko V. and Barik S. 2001. Phenotypic silencing of cytoplasmic genes using sequence-specific double-stranded short interfering RNA and its application in the reverse genetics of wild type negative-strand RNA viruses. *BMC Microbiol.* **1**: 34.
- Borkhardt A. 2002. Blocking oncogenes in malignant cells by RNA interference—New hope for a highly specific cancer treatment? *Cancer Cell* **2**: 167–168.
- Brennecke J., Hipfner D.R., Stark A., Russell R.B., and Cohen S.M. 2003. bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene *hid* in *Drosophila*. *Cell* **113**: 25–36.
- Brummelkamp T.R., Bernards R., and Agami R. 2002. A system for stable expression of short interfering RNAs in mammalian cells. *Science* **296**: 550–553.
- Caplen N.J., Taylor J.P., Statham V.S., Tanaka E., Fire A., and Morgan R.A. 2002. Rescue of polyglutamine-mediated cytotoxicity by double-stranded RNA-mediated RNA interference. *Hum. Mol. Genet.* **11**: 175–184.
- Carmell M.A., Zhang L., Conklin D.S., Hannon G.J., and Rosenquist T.A. 2003. Germline transmission of RNAi in mice. *Nat. Struct. Biol.* **10**: 91–92.
- Castanotto D., Li H., and Rossi J.J. 2002. Functional siRNA expression from transfected PCR products. *RNA* **8**: 1454–1460.
- Chanfreau G., Buckle M., and Jacquier A. 2000. Recognition of a conserved class of RNA tetraloops by *Saccharomyces cerevisiae* RNase III. *Proc. Natl. Acad. Sci.* **97**: 3142–3147.
- Chuang C.F. and Meyerowitz E.M. 2000. Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci.* **97**: 4985–4890.
- Cioca D.P., Aoki Y., and Kiyosawa K. 2003. RNA interference is a functional pathway with therapeutic potential in human myeloid leukemia cell lines. *Cancer Gene Ther.* **10**: 125–133.
- Clemens M.J. and Elia A. 1997. The double-stranded RNA-dependent protein kinase PKR: Structure and function. *J. Interferon Cytokine Res.* **17**: 503–524.

Uncorrected Proof Copy

20

RNA Interference by Short Hairpin RNAs Expressed in Vertebrate Cells

Gregory J. Hannon and Douglas S. Conklin

Abstract

RNA interference (RNAi) is now established as a general method to silence gene expression in a variety of organisms. Double-stranded RNA (dsRNA), when introduced to cells, interferes with the expression of homologous genes, disrupting their normal function. In mammals, transient delivery of synthetic short interfering RNAs (siRNAs), which resemble the processed form of standard double-stranded RNAi triggers, is effective in silencing mammalian genes. Issues related to transfer efficiency and duration of the silencing effect, however, restrict the spectrum of the applications of siRNAs in mammals. These shortcomings of siRNAs have been solved by the cellular expression of short hairpin RNAs (shRNAs) from DNA vectors. shRNAs are indistinguishable from siRNAs in terms of efficacy and mechanism but can be produced within cells from standard mammalian expression vectors. In this way, shRNA expression makes possible the creation of continuous cell lines and transgenic animals in which suppression of a target gene is stably maintained by RNAi. As a result, the types of RNAi-based gene function analysis that can be carried out in mammals have been greatly expanded. We describe methods for the construction and transfer of stable shRNA expressing vectors suitable for generating loss of function alleles in mammalian cells in vitro or in vivo.

Key Words

RNAi; gene silencing; retrovirus; knock-outs; mammalian genetics.

1. Introduction

The recent completion of the human and mouse genomes has brought mammalian gene function analysis to the forefront of biology (1-3). Until recently, the techniques available for mammalian gene function analysis were far less effective than those commonly used in model genetic organisms. This was especially true in the generation of loss-of-function mutations. Rapid methods for directed gene ablation in a number of model organisms have existed for 20 yr.

From: *Methods in Molecular Biology*, vol. 257: *mRNA Processing and Metabolism*
Edited by: D. R. Schoenberg © Humana Press Inc., Totowa, NJ

Uncorrected Proof Copy

Uncorrected Proof Copy

6. Oligonucleotides can be obtained from any commercial supplier. A 0.05- μ mol synthesis scale is sufficient. Resuspend the oligonucleotides at 100 μ M before annealing.
7. A thermal cycler can also be used to anneal oligonucleotides. Add 10- μ L aliquots of mixed oligonucleotides into polymerase chain reaction tubes (500- μ L size). Place the tubes in a thermal cycler and set up a program to perform the following profile: step 1: heat to 95°C and remain at 95°C for 2 min, step 2: ramp cool to 25°C over a period of 10 min, step 3: proceed to a storage temperature of 4°C.
8. The completeness of digestion can be monitored by agarose gel electrophoresis. Although not generally necessary, digested vector can also be purified following agarose gel electrophoresis.
9. White colonies result from the displacement of the *lacZ α* gene, which drastically reduces the vector background in the cloning reaction. The inclusion of the *Hind*III site in the loop of the hairpin allows for rapid identification of clones containing a hairpin. Digestion of successful hairpin constructs with *Hind*III produces a 570-bp band. A few colonies for each hairpin construct should be sequenced.
10. Retroviruses produced by these methods are potentially hazardous! Caution should be exercised in their production and use. Although a powerful technique for gene transfer, appropriate NIH and other regional guidelines should be followed to ensure the safety of those working in the laboratory.
11. Retroviruses are more stable at 32°C than 37°C. Although incubation at 32°C during production and infection is not necessary, it will improve titers.
12. The volume of virus will depend on several factors including the titer of virus, the susceptibility of the strain to infection, etc. A good starting point is to infect the host cells using a 0.5X dilution of the retroviral media supernatant collected in Subheading 7, step 10.
13. Reverse transcription and integration of viral genomes generally take place within the first 24–36 h. Although cells can be assayed for the phenotype of interest at this time, selection of infected cells with puromycin greatly improves the penetrance of most phenotypes. For each host cell line, the minimum puromycin concentration required to kill uninfected cells should be determined in titration experiments before infection. One mg/mL final concentration is a good starting point for titration. It should be noted that puromycin typically kills most cells after only 2 d exposure.

Acknowledgments

Shola Aruleba, David Koppersmith, Ravi Sachidanandam, Kim Scobie, and the rest of the Cold Spring Harbor RNAi Group are thanked for comments and helpful discussions. GJH is a Rita Allen Foundation Scholar and is supported by the US Army Breast Cancer Research Program and the NIH. DSC is supported by the US Army Breast Cancer Research Program.

Uncorrected Proof Copy