Published in final edited form as:

Mol Biol. 2006 May 1; 40(3): 396–405. doi:10.1134/S002689330603006X.

Optimization of a Genome-Wide Disordered Lentivector-Based Short Hairpin RNA Library

O.A. Guryanova^{a,b}, M. Makhanov^c, A.A. Chenchik^c, P.M. Chumakov^{b,d}, and E.I. Frolova^a

O.A. Guryanova: ; M. Makhanov: ; A.A. Chenchik: ; P.M. Chumakov: ; E.I. Frolova: eifr@ibch.ru

^a Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, 117997 Russia

^b Lerner Research Institute, Cleveland Clinic Foundation, Cleveland, Ohio 44195, United States

^c System Biosciences (SBI), Mountain View, California 94041, United States

^d Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia

Abstract

To obtain a whole genome library that suppresses the total diversity of human mRNAs, lentiviral vector constructs and a short hairpin RNA (shRNA) expression cassette were optimized. The optimization of the vector increased the virus titer in preparations by 15-20 times. A simple shRNA structure with a 21-bp stem proved to be the most effective. Lentivector-based shRNA expression constructs were obtained by using *puro^R*, *copGFP*, or *H-2K^k* as a selectable marker. The efficiency of the optimized library was demonstrated when screening for shRNAs reactivating the tumor suppressor p53 in HeLa cells. Cells carried a reporter construct ensuring p53-responsive synthesis of a fluorescent protein, which allowed selection of cells with reactivated p53 by flow cytometry.

Keywords

RNA interference; functional genomics; short hairpin RNA; *copGFP*; *H-2K^k*; flow cytometry; pseudovirus particle packing; p53; genome-wide library; efficient screening

Introduction

Since the human genome structure has been deciphered and a large body of data on gene expression has been obtained by hybridization with cDNA microarrays, it is now possible to establish the functional relationships of genes, as well as arrange their products into functional chains. It is especially important in connection with this to develop new and efficient methods for functional genomics, which are aimed at identifying the functions of individual genes.

Gene silencing via RNA interference is a common and promising tool of reverse genetics and is used to study the functions of genes whose structures are known. RNA interference is the inhibition of gene expression by homologous double-stranded RNA (dsRNA) [1]. Fire et al. [2] were the first to describe RNA interference in nematode *Caenorhabditis elegans*. It has been found that dsRNA is far more effective in gene silencing as compared with single-stranded antisense RNA, just a few dsRNA copies per cell are already sufficient to exert a considerable effect. More recently RNA interference has been observed in a variety of organisms from invertebrates to mammals, including human [3-5]. In animals an additional mechanism of RNA interference utilizes endogenous microRNAs, which are short noncoding RNAs that contain imperfect inverted repeats and form hairpins with a stem of about 70 bp and incomplete homology [6,7].

An important role in RNA interference is played by the enzyme Dicer, which is similar to RNase III [8]. Short dsRNAs and microRNAs are recognized and processed by Dicer to yield small interfering RNAs (siRNAs), which are 21-22 bp in size, are phosphorylated at the 5' end, and contain two protruding non-paired nucleotides at each 3' end [9]. To be efficiently recognized by Dicer, microRNA is preliminarily processed in the nucleus by Drosha, which is another RNase III-like enzyme [10]. Then, microRNA is transferred into the cytoplasm and binds with Dicer. The resulting siRNAs are incorporated into a multi-component effector RNA-induced silencing complex (RISC) [11]. The sense strand of siRNA is displaced and lost owing to an RISC component acting as helicase, while the antisense strand is used to scan the structure of an mRNA passing through RISC. As a homologous region is found, the mRNA is degraded by an Argonaute-family endonuclease [12], which is also a component of RISC [13,14].

Soon after its discovery, RNA interference started to be used as an efficient, convenient, and, specific tool for inhibiting the expression of particular genes [15,16]. Two fundamentally different techniques are employed to induce RNA interference: exogenous dsRNA is introduced into the cell via transfection or is expressed from an artificial construct by the transcription system of the cell. In the latter case, the expression product is a short hairpin RNA (shRNA) that has a stem of 19-21 bp and is processed by Dicer to yield functional siRNA. Target cells containing siRNA are then tested for phenotypic changes [17].

As this approach is extended to the whole genome by constructing siRNA or shRNA libraries covering all known and unknown genes, functional genomics will receive additional impetus [18], and targets will be identified for developing new drugs [19,20]. Such ordered whole-genome libraries have already been obtained in plates, with each well containing a particular siRNA or an siRNA-expressing vector [21]. The libraries are screened gene by gene, with each specimen being introduced into cells in a separate transfection experiment. It is clear that experiments with these libraries are highly laborious and mostly feasible with special robots [22]. A reverse transfection technique has been developed to simplify the procedure; transfection mixtures in the form of DNA- or RNA-lipid complexes are applied to wells and dried, and cells are then seeded in the wells to be transfected in situ [23]. Modification of the reverse transfection technique has yielded libraries consisting of cell microarrays: cells are seeded on a glass dotted with transfection mixtures containing individual components of a library [24]. The drawbacks are that such libraries are exhaustible and expensive; experiments with them require expensive robots and are laborious, and their application is limited. Because of these drawbacks, such libraries are virtually unavailable for non-commercial labs.

To overcome the above problems, we cooperated with System Biosciences (SBI) to develop a method for constructing whole-genome shRNA libraries on the basis of a lentiviral vector. The use of lentiviruses makes it possible to infect even non-dividing cells [25], and the integration of such constructs into the genome of target cells ensures stable expression of shRNA through many passages [26]. Moreover, it is possible with this approach to obviate the construction of an ordered library and to obtain a complete set of shRNAs in one tube, which makes the subsequent functional screening simpler and cheaper. After transduction of a library, cells expressing the phenotype of interest are selected and the fragment responsible for the phenotypic changes is isolated from the resulting cell sub-population, is then amplified, and is identified by sequencing or hybridization with oligonucleotide components of an siRNA library ordered in a microarray. It is clear that a lentiviral library should meet several requirements in order to produce reliable results. For instance, a high titer is necessary for pseudovirus particles expressing siRNA with an optimal structure of the RNA hairpin, efficiently inhibiting the expression of the target. Cell selection from the transfected population should be quite rapid.

The objective of this work was to solve the above problems. As a result, the titer of packed virus particles increased 15- to 20-fold. The highest efficiency was observed for shRNA having

a simple structure with a 21-bp stem. To improve the selection we used lentiviral constructs expressing both shRNA and a selectable marker: *puro^R*, *copGFP*, or *H-2K^k*. The efficiency of the optimized library was verified in screening experiments, using a cell line with fluorescence intensity that depended on the p53 activity.

Experimental

Cell lines, culture conditions, and lentivirus preparations

Cells of the lines 293T, H1299, HeLa, and HeLa LC5 with the reporter β -galactosidase gene (*lacZ*) controlled by a p53-responsive promoter were grown in DMEM supplemented with 10% fetal bovine serum, penicillin, and streptomycin. To allow the packing of virus particles, 293T cells were seeded so that the culture density was not more than 70% at the moment of transfection. Cells were transfected with lentiviral constructs and helper plasmids, using Lipofectamine and Plus Reagent as recommended by Invitrogen. The medium containing virus particles was collected 24, 36, and 48 h after transfection; filtered through a 0.45- μ m filter; and then frozen at -70°C. The virus titer was assayed with H1299 cells. Cells were seeded in a 12-well plate at 1×10^5 cells per well and infected with a virus preparation for 4 h. To improve the efficiency of infection, virus preparations were supplemented with the polycationic reagent Polybrene (Sigma) to 5 mg/ml. The fluorescence of the green fluorescent protein (GFP) was assayed by flow cytometry 48 h after infection. Only viable cells were included in the analysis; 10,000 events were recorded in each experiment.

Constructs expressing shRNA

Oligonucleotides were selected so that their annealing yielded sticky ends complementary to the *Bam*HI and *Eco*RI restriction sites, which were utilized for cloning in pSIH-Puro (Fig. 1a). To obtain constructs expressing *copGFP* [27] or *H-2K^k* [28] as a selectable marker, the protein-coding region of a gene was amplified with highly selective *Taq* DNA polymerase (a component of a High Fidelity PCR Supermix, Invitrogen) and cloned into the *Xba*I-*Sal*I sites of pSIH in place of the puromycin resistance gene. The recognition sites for the necessary restriction enzymes were introduced with PCR primers.

Cell staining with ONPG for β -galactosidase

After infection HeLa LC5 cells were incubated for 72 h and plated on a 96-well plate at 5×10^4 cells per well. Cells were stained with an ONPG solution (1 mM MgCl₂, 250-mM Tris-HCl, pH 7.4, 0.02% NP-40, 2 mg/ml ONPG in PBS) at 37°C for 4h, until a yellow color developed. Absorption was measured using a plate reader at 405 nm. To standardize the cell amount in samples, cells were stained with 1% Methylene Blue in 50% methanol, the excess dye was washed out with water, an equal volume of 1% SDS was added, and 600 nm was measured. All measurements were performed in duplicate.

PCR on genomic DNA and reverse transcription (RT)-PCR

DNA and RNA were simultaneously isolated from cell cultures with the use of TRIzol as recommended by Invitrogen. To estimate the number of provirus copies integrated in the cell genome, the WPRE region of the lentiviral construct was amplified in PCR with 1 μ g of genomic DNA and the specific primers WPRE-forward (5'-TCATGCTATTGCTTC-CCGTATGGCTTTC-3') and WPRE-reverse (5'-GGATTGAGGGCCGAAGGGACGTAG-3'). The first cDNA strand was synthesized using SuperScript II reverse transcriptase (Invitrogen), an oligo-dT primer, and 5 μ g of total RNA. PCR was carried out using thermostable Titanium *Taq* polymerase (BD Clontech); 250 ng of cDNA; and specific primers to *GAPDH* (*GAPDH*-forward 5'-CTCGTCTCTGGAAGATGG-3' and *GAPDH*-reverse 5'-

CAATGACCCCTTCATTGAC-3'), *p53* (*p53*-forward 5'-GACCTCAAAGCTGTTC-CGTCC-3' and *p53*-reverse 5'-CGTCTGGGCTTCTTGCATTC-3'), or HPV18 *E6* (*E6*-forward 5'-GCTGCATGCCATAAATGTATAGAT-3' and *E6*-reverse 5'-CTATGTTGTGAAATCGTCGTT-3'). Amplification involved 20-28 cycles of 94°C for 30 s, 68°C for 30 s, and 72°C for 1 min.

Immunofluorescent cell staining for flow cytometry

The portion of transfected cells was assayed by the expression of H-2K^k 48 h after infection. Cells were treated with trypsin, washed three times with 0.1% BSA in PBS, incubated with 0.1% rat serum in PBS on ice for 1 h (to block nonspecific binding), stained with FITC-conjugated antibodies against H-2K^k (1:50, Miltenyi Biotech) in the blocking buffer on ice for 1 h, and washed three times as above. Green fluorescence was assayed by flow cytometry; only viable cells were analyzed and 10,000 events were registered.

A *p53*-responsive reporter construct

The *p53*-responsive cassette WafConA-mCMV-EGFP [29] was cloned in the improved lentiviral vector pLU. The nucleotide sequence coding for destabilized GFP (copGFPds) was amplified using highly specific *Taq* polymerase from a High Fidelity PCR Supermix (Invitrogen) and specific primers containing the *Sall* and *BamHI* sites; the product was cloned into the *XhoI*-*BamHI* sites of the reporter vector instead of the EGFP gene.

Screening of the 8.5K shRNA library

HeLa cells (1×10^7) carrying the construct that determined their fluorescence as dependent on the *p53* activity were transfected with the packed 8.5K shRNA library for 8 h and were passed (1:3) 48 h after infection. After 24 h cells were treated with trypsin and washed. Two fractions of 2×10^5 cells each and one fraction of 1×10^6 cells were selected from 5% of cells with the most intense fluorescence by flow sorting. As a control two fractions of 2×10^5 cells each were selected from non-sorted cells. The resulting fractions were used to isolate DNA with TRIzol (Invitrogen). Amplification of the integration sites of shRNA-expressing constructs, hybridization with a GeneChip 8.5K Human Genome Focused Array (Affymetrix), and analysis of the results were performed by System Biosciences.

Results

Optimization of the Packing of Pseudovirus Particles

Lentiviral vectors are advantageous for delivering genetic material into the cell, because even non-dividing cells can be infected and transduction efficiency can reach 100%. However, the main steps of transduction should be optimized. With a high-titer lentivirus preparation, it is possible to avoid its concentration and to use highly diluted supernatants, thereby minimizing the nonspecific effects exerted by the toxic metabolites of producer cells. In addition, experiments are cheaper with high-titer preparations.

To optimize the packing of pseudovirus particles, we used the lentiviral construct pLU-CMV-EGFP with *GFP* controlled by the cytomegalovirus (CMV) promoter and two variants of helper plasmids, pCMVAR8.2 (I.M. Verma, Salk Institute, La Jolla, United States) or pGag1 and pRev2 (J.G. Sodroski, Dana Farber Cancer Institute, Boston, United States). Virus particles were pseudotyped with the coat protein G of the vesicular stomatitis virus expressed from pVSV-G. All plasmids were used in equal amounts for transfection. To estimate the titer of pseudovirus particles, H1299 cells were infected with a virus preparation diluted eightfold.

When pCMVAR8.2 was used as a helper, only 5% of cells started producing GFP. With pGag1 and pRev2, fluorescence was observed in up to 41% of cells. Thus, pGag1 and pRev2 ensured a virus titer an order of magnitude higher than with pCMVAR8.2.

To further optimize the packing of virus particles, we examined transfection mixtures with various proportions of the lentiviral construct and the helper plasmids. The recombinant virus was used in a tenfold dilution to infect cells. The virus titer dramatically decreased with an increasing pVSV-G portion (Table 1, sample A4), which agrees with the data that protein G is toxic for eukaryotic cells [30]. An increase in pRev2 considerably improved the efficiency of virus packing, especially when the pGag1 amount was not reduced (Table 1, samples A3 and A6).

Based on the results, we further optimized the plasmid ratio (Table 1, samples B1-B5). Cells were infected with a virus preparation diluted tenfold. The optimal plasmid ratio was pLU-GFP:pGag1:pRev2:pVSV-G = 1:2:4:1. With this ratio the virus titer can be increased 15- to 20-fold as compared with the titer observed with the plasmids used in equal amounts.

Optimization of the shRNA Structure

When a shRNA library is obtained in one tube it is important that no more than one or two copies of a lentiviral construct are integrated in the cell genome; with a high multiplicity of infection, it is difficult to identify the shRNA responsible for the phenotype observed. Consequently, it is necessary that shRNA is highly efficient so that a single copy of the construct suffices to strongly silence the target gene. The extent of target gene silencing by shRNA depends not only on the choice of the mRNA target site, but also on the efficiency of shRNA transport from the nucleus into the cytoplasm, where mRNA is degraded, and on the affinity for processing enzymes. An important role in these processes is played by the hairpin structure, formed by complementary regions of siRNA. We tested four variants of hairpin structures, which were based on three 22-nt sequences homologous to the *p53* mRNA: p53-1 (5'-GATCTGGATCCACCAAGACTTG-3'), p53-2 (5'-TACATTCTGCAAGCATCTGC-3'), and p53-3 (5'-ACTGGAAGACTC-CAGTGGTAAT-3').

According to the published data, features characteristic of microRNA (an extended stem with non-complementary nucleotides) may improve the affinity of an exogenously expressed shRNA for the RNA-processing enzyme Dicer and increase the efficiency of its association with RISC [31]. When a microRNA-like hairpin is combined with the regions flanking the human mir-30 hairpin [32], the resulting structure is recognized by Drosha, simulating natural endogenous silencing [33]. Proceeding from these findings, we obtained the following lentivector-based constructs (Fig. 1b): (i) shRNA, which was a basic construct expressing a short hairpin with a 22-bp stem from under the control of the H1 promoter; (ii) microRNA, in which the basic 22-bp hairpin was combined with a short stem with non-complementary nucleotides; (iii) mir-30-flank, in which the microRNA structure was combined with the 50-nt regions flanking the human mir-30 hairpin; and (iv) U6-leader, in which structure (iii) was combined with the leader sequence of the U6 promoter.

The efficiency of *p53* silencing was assayed using the reporter HeLa LC5 line, producing P-galactosidase under the control of a *p53*-responsive promoter. The highest inhibition of *p53* expression was observed with the simplest (basic) variant of the shRNA structure (a hairpin with a 22-bp stem) (Fig. 1c). To verify, the level of the target *p53* mRNA was assayed by RT-PCR and the copy number of the integrated provirus was estimated by amplifying the WPRE segment from genomic DNA of transfected cells (Fig. 1d). The results confirmed that the simplest RNA hairpin was most effective in silencing the target gene. It is of interest to note that our constructs inhibited *p53* expression by more than 90% in spite of the considerable variation of the copy number of integrated provirus DNA.

Construction of Vectors for shRNA Expression with Various Selectable Markers

An initial whole-genome shRNA library was constructed on the basis of the lentiviral vector pSIH-Puro, with shRNA being expressed from the H1 promoter and the CMV promoter controlling the expression of the puromycin resistance gene (*puroR*). However, selection by antibiotic resistance has its drawbacks: the test is time-consuming (taking 7-14 days), antibiotic sensitivity varies among cell lines and should be estimated preliminarily, representation of some shRNAs in a cell culture may change in the case of their effect on cell division, and some active shRNAs may even be lost if exerting a toxic effect on the cell. In addition, virus titer assays are laborious and time consuming. To overcome these drawbacks, special vectors were designed to allow rapid estimation of the virus titer and rapid selection of infected cells. Convenient markers are provided by copGFP and H-2K^k, a transmembrane receptor devoid of the cytoplasmic domain. The advantage of copGFP as a selectable marker is that copGFP can be directly detected in the cell under a fluorescence microscope. Moreover, it is possible with flow sorting to quantify the copGFP expression in a cell population without a preliminary sample preparation and to select cells from a living transfected culture. A drawback of copGFP is that it is unfeasible to simultaneously monitor other cell parameters in the green range (e.g., to estimate the level of reactive oxygen species by DCF fluorescence) or to use an additional transcription reporter construct utilizing *GFP*. In addition, flow sorting impairs cell viability, and a considerable fraction of the cell population is lost. These drawbacks can be overcome by using H-2K^k as a selectable marker. With H-2K^k, the portion of transfected cells can be estimated by flow sorting after vital immunofluorescent staining with FITC- or TRITC-conjugated antibodies against H-2K^k. Pure populations of transfected cells can be rapidly selected (within 30 min) and under mild conditions via binding with anti-H-2K^k antibodies fixed on magnetic beads. This procedure does not require expensive instruments, which is an additional advantage. The drawbacks of selection with H-2K^k are that a preliminary treatment (immunofluorescent staining) of cells is necessary for flow sorting, and that the antibodies and the selection kits are rather expensive.

To test the above selection variants, we constructed the lentiviral vectors pSIH-copGFP and pSIH-H-2K^k for shRNA expression. The efficiency of packing into virus particles was estimated for the above basic constructs and for the vectors expressing p53-specific shRNAs. In either variant the constructs utilizing copGFP ensured transfection of more than 97% of cells in a population (data not shown). In the case of H-2K^k, the efficiency of transfection was 91% with the empty vector and 80% in the case of pSIH-H-2K^k-shp53-1 (Fig. 2a).

These findings suggest that pSIH-H-2K^k is packed in virus particles somewhat less efficiently as compared with pSIH-copGFP. The average copy number of integrated constructs was estimated by PCR. The results showed that the packing efficiency was 20% lower with constructs expressing H-2K^k than with similar constructs expressing copGFP (Fig. 2b). The difference is rather low, demonstrating the applicability of both selectable markers.

Testing of a Fluorescent p53-Responsive Construct

Attempts have already been made to obtain a fluorescent reporter construct based on a lentiviral vector. For instance, the p53-responsive construct pLV-Waf-ConA-mCMV-EGFP has been designed in our lab [29]. However, the high stability of EGFP masks the effect of p53 activation. We optimized this reporter construct (Fig. 3a). For this purpose the p53-responsive cassette was cloned in the improved lentiviral vector pLU, and *EGFP* was replaced by the gene coding for copGFPs, whose half life is about 2 h.

The sensitivity of the new p53-responsive construct was studied with HeLa cells. These cells synthesize the wild-type p53, which is rapidly degraded in proteasomes, because of the binding with HPV18 E6. Suppression of *E6* in HeLa cells was expected to allow the accumulation of

active p53 and, consequently, activation of the reporter gene. To silence HPV18 *E6* with shRNA, we chose the sequence 5'-ctaactcggttatacaa-3' (Fig. 3b).

HeLa cells carrying the fluorescent p53-responsive construct were transfected with the lentiviral vector expressing an E6-specific shRNA, or with the 8.5K shRNA library used as a control (Fig. 3c). Transfection with the shRNA library only slightly shifted the peak of copGFPs fluorescence to the right, which probably reflected the effect of p53-activating components of the library. The suppression of *E6* in the reporter cell line caused a dramatic rightward shift in the cell distribution by copGFPs fluorescence, suggesting considerable activation of the reporter gene.

Screening of the 8.5K shRNA Library

The whole-genome shRNA library was constructed on the basis of the pSIH lentiviral vector, with an optimal hairpin structure. To construct the library, we used 85,000 human expressed nucleotide sequences of a GeneChip 8.5K Human Genome Focused Array (Affymetrix). Three to five shRNAs were designed for each sequence; target sites were selected to ensure optimal hybridization with the microarray components. The resulting library includes about 4×10^4 individual constructs. Preparative synthesis of oligonucleotides for the library was performed simultaneously on one microarray. In each oligonucleotide a restriction site corresponding to one of the cloning sites of the vector was introduced in the basal region attached to the membrane. The distal region of each oligonucleotide contained another restriction site and a common sequence serving for primer annealing in synthesis of the second DNA strand. After the second strand was synthesized on the microarray, double-stranded oligonucleotides were detached from the membrane with a mixture of restriction enzymes, purified by PAGE, and cloned in the lentiviral vector under the control of the H1 RNA gene promoter. This procedure yielded a disordered lentiviral library for shRNA expression.

The library was screened using the above p53-responsive HeLa cell line. Cells were transformed with the library. Then, flow sorting was used to select cells with the most intense fluorescence, suggesting activation of the p53-responsive construct. Genomic DNA was isolated from the selected cells. The regions coding for shRNAs presumably responsible for this phenotype were amplified and then identified by hybridization with a microarray (see Experimental). The genes showing enrichment in comparison with a non-sorted cell population are listed in Table 2. It is noteworthy that the set included the gene for the YY1 transcription factor, which is involved in activating transcription from the regulatory region of the HPV genome [34]. Possibly, YY1 gene silencing decreases the production of viral proteins, including E6, and thereby reactivates p53.

Thus, we identified several dozens of shRNAs that show enrichment in the cell fraction with the activated p53-responsive construct. In some cases several different shRNAs corresponded to one gene, suggesting a higher reliability of the result. It remains unclear what role is played by these genes in regulating the p53 activity in HeLa cells and other HPV-containing human cervical carcinomas. Relevant work is in progress in our lab.

Discussion

RNA interference is widely used as a convenient and highly specific tool in reverse genetics. This technique is especially promising for functional genomics, with increasing importance as the genome structure is deciphered and there is rapid accumulation of gene expression data. In view of this, many labs have focused on constructing siRNA or shRNA whole-genome libraries [21-24]. Several attempts have been made to utilize viruses for delivering the library material into cells [35-38]. We described a whole-genome shRNA library that is based on a lentiviral vector and is suitable for transforming virtually all cell populations. The intention is that the

library is screened and individual integrated constructs are isolated and sequenced in order to identify the gene whose silencing is responsible for the observed phenotype. To simplify the procedure, a bar code has been proposed for identifying individual clones by hybridization on microarrays; an arbitrary unique 60-bp sequence is introduced in a shRNA-expressing construct and is used as an identifier [36,37]. However, such arbitrary sequences are not all optimal for hybridization, as they may introduce an additional uncontrollable error in the results. Another modification of this technique takes advantage of the unique nucleotide sequence of each component of a library, utilizing the shRNA-coding regions for identification [38]. It is essential for both modifications that special individual microarrays are obtained, which is rather expensive.

We used an optimized technique of internal molecular bar codes. The structure of each RNA hairpin was selected to ensure optimal hybridization with the microarray in order to minimize individual variation and nonspecific noise associated with this step of screening. When designing shRNAs we took into account the presence of such nucleotide sequences in common microarrays produced by Affymetrix. The 8.5K lentivirus-based human shRNA library is fully compatible with the commercially available Gene-Chip 8.5K Human-Genome Focused Array (Affymetrix), which made the screening far cheaper.

The selection of a pure cell population is another important problem in works with disordered libraries. Various antibiotic resistance genes are commonly introduced in constructs for this purpose. We constructed vectors with alternative selectable markers. For instance, the expression of H-2K^k on the cell surface makes it possible to obtain a population consisting of transfected cells to about 100% in less than 1 h. Our vector, based on GFP, allows library screening not only with cell cultures, but also with whole organisms *in vivo*, e.g., for studying the tissue-specific processes or the mechanisms of metastasis.

Further progress in large-scale gene silencing via RNA interference will greatly improve the efficiency of searching for new biological activities and drug targets.

Acknowledgments

This work was supported by the Russian Foundation for Basic Research and the National Institutes of Health grants CA104903 and AG025276.

References

1. Dykxhoorn DM, Novina CD, Sharp PA. Killing the messenger: Short RNAs that silence gene expression. *Nature Rev Mol Cell Biol* 2003;4:457–467. [PubMed: 12778125]
2. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Cae-norhabditis elegans*. *Nature* 1998;391:806–811. [PubMed: 9486653]
3. Ngo H, Tschudi C, Gull K, Ullu E. Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proc Natl Acad Sci USA* 1998;95:14687–14692. [PubMed: 9843950]
4. Yang S, Tutton S, Pierce E, Yoon K. Specific double-stranded RNA interference in undifferentiated mouse embryonic stem cells. *Mol Cell Biol* 2001;21:7807–7816. [PubMed: 11604515]
5. Hutvagner G, McLachlan J, Pasquinelli AE, Balint Ę, Tuschl T, Zamore PD. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 2001;293:834–838. [PubMed: 11452083]
6. Moss EG. MicroRNAs: Hidden in the genome. *Curr Biol* 2002;12:R138–R140. [PubMed: 11864587]
7. Bartel DP. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* 2004;116:281–297. [PubMed: 14744438]
8. Bernstein E, Caudy AA, Hammond SM, Hannon GJ. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 2001;409:363–366. [PubMed: 11201747]

9. Caplen NJ, Parrish S, Imani F, Fire A, Morgan RA. Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. *Proc Natl Acad Sci USA* 2001;98:9742–9747. [PubMed: 11481446]
10. Lee Y, Ahn C, Choi H, Kim J, Yim J, Lee J, Provost P, Radmark O, Kim S, Kim VN. The nuclear RNase III Drosha initiates microRNA processing. *Nature* 2003;425:415–U9. [PubMed: 14508493]
11. Hammond SM, Bernstein E, Beach D, Hannon GJ. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* 2000;404:293–296. [PubMed: 10749213]
12. Carmell MA, Xuan Z, Zhang MQ, Hannon GJ. The Argonaute family: Tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev* 2002;16:2733–2742. [PubMed: 12414724]
13. Sledz CA, Williams BRG. RNA interference in biology and disease. *Blood* 2005;106:787–794. [PubMed: 15827131]
14. Hammond SM. Dicing and slicing: The core machinery of the RNA interference pathway. *FEBS Lett* 2005;579:5822–5829. [PubMed: 16214139]
15. Sui G, Soohoo C, El Bashir A, Gay F, Shi Y, Forrester WC, Shi Y. A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. *Proc Natl Acad Sci USA* 2002;99:5515–5520. [PubMed: 11960009]
16. Paddison PJ, Caudy AA, Bernstein E, Hannon GJ, Conklin DS. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev* 2002;16:948–958. [PubMed: 11959843]
17. Cullen BR. Induction of stable RNA interference in mammalian cells. *Gene Therapy* 2005;1–6. [PubMed: 15510174]
18. Silva J, Chang K, Hannon GJ, Rivas FV. RNA-interference-based functional genomics in mammalian cells: Reverse genetics coming of age. *Oncogene* 2004;23:8401–8409. [PubMed: 15517022]
19. Heinonen JE, Smith CIE, Nore BF. Silencing of Bruton's tyrosine kinase (Btk) using short interfering RNA duplexes (siRNA). *FEBS Lett* 2002;527:274–278. [PubMed: 12220673]
20. Van Es HHG, Arts GJ. Biology calls the targets: Combining RNAi and disease biology. *Drug Discovery Today* 2005;10:1385–1391. [PubMed: 16253877]
21. Boutros M, Kiger AA, Armknecht S, Kerr K, Hild M, Koch B, Haas SA, Paro R, Perrimon N. Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science* 2004;303:832–835. [PubMed: 14764878]
22. Wheeler DB, Carpenter AE, Sabatinin DM. Cell microarrays and RNA interference chip away at gene function. *Nature Genet* 2005;37:S25–S30. [PubMed: 15920526]
23. Ziauddin J, Sabatini DM. Microarrays of cells expressing defined cDNAs. *Nature* 2001;411:107–110. [PubMed: 11333987]
24. Silva JM, Mizuno H, Brady A, Lucito R, Hannon GJ. RNA interference microarrays: High-throughput loss-of-function genetics in mammalian cells. *Proc Natl Acad Sci USA* 2004;101:6548–6552. [PubMed: 15084744]
25. Nguyen TH, Oberholzer J, Birraux J, Majno P, Morel P, Trono D. Highly efficient lentiviral vector-mediated transduction of non-dividing, fully reimplantable primary hepatocytes. *Mol Therapy* 2002;6:199–209.
26. Pfeifer A, Ikawa M, Dayn Y, Verma IM. Trans-genesis by lentiviral vectors: Lack of gene silencing in mammalian embryonic stem cells and preimplantation embryos. *Proc Natl Acad Sci USA* 2002;99:2140–2145. [PubMed: 11854510]
27. Lukyanov KA, Chudakov DM, Fradkov AF, Labas YA, Matz MV, Lukyanov S. Discovery and properties of GFP-like proteins from non-bioluminescent anthozoa. *Methods Biochem Anal* 2006;47:121–138. [PubMed: 16335712]
28. Fitchen JH, Hays EF, Ferrone S. Expression of the H-2 antigenic complex on murine haematopoietic stem cells. *Scand J Immunol* 1982;15:547–551. [PubMed: 6981838]
29. Razorenova OV, Agapova LS, Budanov AV, Iva-nov AV, Strunina SM, Chumakov PM. Retroviral reporter systems for assessing the activity of stress-inducible signal transduction pathways controlled by the p53, HIF-1, and HSF-1 transcription factors. *Mol Biol* 2005;39:286–293.
30. Chen ST, Iida A, Guo L, Friedmann T, Yee JK. Generation of packaging cell lines for pseudo-typed retroviral vectors of the G protein of vesicular stomatitis virus by using a modified tetracycline inducible system. *Proc Natl Acad Sci USA* 1996;93:10057–10062. [PubMed: 8816750]

31. Zeng Y, Wagner EJ, Cullen BR. Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Mol Cell* 2002;9:1327–1333. [PubMed: 12086629]
32. Lagos-Quintana M, Rauhit R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. *Science* 2001;294:853–858. [PubMed: 11679670]
33. Cullen BR. RNAi the natural way. *Nature Genet* 2005;37:1163–1165. [PubMed: 16254559]
34. Bauknecht T, See RH, Shi Y. A novel C/E BP beta-YY1 complex controls the cell-type-specific activity of the human papillomavirus type 18 upstream regulatory region. *J Virol* 1996;70:7695–7705. [PubMed: 8892890]
35. Kolfschoten IGM, van Leeuwen B, Berns K, Mullenders J, Beijersbergen RL, Bernards R, Voorhoeve PM, Agami R. A genetic screen identifies PITX1 as a suppressor of RAS activity and tumorigenicity. *Cell* 2005;121:849–858. [PubMed: 15960973]
36. Westbrook TF, Martin ES, Schlabach MR, Leng Y, Liang AC, Feng B, Zhao JJ, Roberts TM, Mandel G, Hannon GJ, DePinho RA, Chin L, Elledge SJ. A genetic screen for candidate tumor suppressors identifies REST. *Cell* 2005;121:837–848. [PubMed: 15960972]
37. Silva JM, Li MZ, Chang K, Ge W, Golding MC, Rickles RJ, Siolas D, Hu G, Paddison PJ, Schlabach MJ, Sheth N, Bradshaw J, Burchard J, Kul-karni A, Cavet G, Sachidanandam R, McCombie WR, Cleary MA, Elledge SJ, Hannon GJ. Second-generation shRNA libraries covering the mouse and human genomes. *Nature Genet* 2005;37:1281–1288. [PubMed: 16200065]
38. Berns K, Hijmans EM, Mullenders J, Brummelkamp TR, Velds A, Heimerlikx M, Kerkhoven RM, Madiredjo M, Nijkamp W, Weigelt B, Agami R, Ge W, Cavet G, Linsley PS, Beijersbergen RL, Bernards R. A large-scale RNAi screen in human cells identifies new components of the p53 pathway. *Nature* 2004;428:431–437. [PubMed: 15042092]

Figure 1.

(a) Map of the shRNA expression vector pSIH-Puro. The vectors pSIH-copGFP and pSIH-H-2K^k are similar but have, respectively, *copGFP* and *H-2K^k* in place of the puromycin resistance gene (*puroR*). (b) Structure of RNA hairpins expressed from the corresponding constructs. Black arrows show the cleavage sites for Dicer and gray arrows, for Drosha. (c) Efficiency of *p53* silencing with various RNA hairpins in reporter HeLa LC5 cells expressing the P-galactosidase gene under the control of a *p53*-responsive promoter. The P-galactosidase activity in control (nontreated) cells was taken as 100%. (d) RT-PCR analysis of the *p53* mRNA expression in the same cells. The copy number of integrated constructs was estimated by PCR amplification of the WPRE region from cell genomic DNA.

Figure 2.

(a) Expression of the marker *H-2K^k* gene in H1299 cells transfected with pSIH-H-2K^k-shp53-1 or with the empty vector as revealed by flow cytometry after immunofluorescent staining. The range of fluorescent cells is shown. (b) Comparison of the copy number in HeLa cells for similar constructs based on pSIH with *copGFP* or *H-2K^k*. The number of integrated copies was estimated by PCR-amplification of the WPRE region from cell genomic DNA. The signal intensity of cells transfected with pSIH-*copGFP*-shp53-1 was taken as unity. Constructs that did not express shRNA were used as controls.

Figure 3.

(a) Structure of the fluorescent p53-responsive cassette: p53RE, the WafConA cluster of p53-responsive elements [29]; mCMV, the minimal CMV promoter. (b) Silencing of the HPV18 *E6* gene in HeLa cells by shRNA. The mRNA level was estimated by RT-PCR. (c) Activation of the fluorescent p53-responsive construct in HeLa cells in response to *E6* silencing by shRNA as revealed by flow cytometry: 1, copGFP is undetectable; 2, fluorescent cells; 3, region of the bright fluorescence of copGFP.

Table 1
Optimization of the packing of lentivirus particles by varying the helper plasmid ratio in lipofection of 293T producer cells*

Sample	Plasmid ratio in lipofection				Fluorescent cells, %
	pLU-GFP	pGag1	pRev2	pVSV-G	
A1	0.25	0.25	0.25	0.25	33.31
A2	0.2	0.4	0.2	0.2	26.20
A3	0.2	0.2	0.4	0.2	46.55
A4	0.2	0.2	0.2	0.4	17.77
A5	0.25	0.35	0.15	0.25	19.33
A6	0.35	0.15	0.35	0.15	39.85
B1	0.2	0.3	0.3	0.2	45.88
B2	0.2	0.2	0.4	0.2	60.77
B3	0.222	0.222	0.444	0.111	64.38
B4	0.125	0.25	0.5	0.125	72.41
B5	0.133	0.267	0.533	0.067	50.44

* Virus preparations diluted 1:10 were used to infect H1299 cells; the portion of transfected cells was estimated by flow cytometry 48 h after infection.

Table 2

Genes potentially involved in regulating the p53 activity on evidence of screening the 8.5K whole-genome lentivector-based shRNA library*

GenBank	Total shRNAs	Description
Z48481	3	<i>Homo sapiens</i> mRNA for membrane-type matrix metalloproteinase 1
NM_003746.2	1	<i>Homo sapiens</i> dynein, light chain, LC8-type 1 (DYNLL1), transcript variant 3, mRNA
NM_003022.1	1	<i>Homo sapiens</i> SH3 domain binding glutamic acid-rich protein like (SH3BGRL), mRNA
NM_021978.3	2	<i>Homo sapiens</i> suppression of tumorigenicity 14 (colon carcinoma) (ST14), mRNA
NM_002717.2	2	<i>Homo sapiens</i> protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), alpha isoform (PPP2R2A), mRNA
NM_012289.3	1	<i>Homo sapiens</i> kelch-like ECH-associated protein 1 (KEAP1), transcript variant 2, mRNA
NM_003403.3	3	<i>Homo sapiens</i> YY1 transcription factor (YY1), mRNA
NM_002199.2	2	<i>Homo sapiens</i> interferon regulatory factor 2 (IRF2), mRNA
NM_001833.1	3	<i>Homo sapiens</i> clathrin, light polypeptide (Lca) (CLTA), transcript variant nonbrain, mRNA
N22903		yx66e04.s1 Soares melanocyte 2 NbHM <i>Homo sapiens</i> cDNA clone IMAGE:266718 3, mRNA sequence
NM_014936.3		<i>Homo sapiens</i> ectonucleotide pyrophosphatase/phosphodiesterase 4 (putative function) (ENPP4), mRNA
BG251266		602365892F1 NIH_MGC_90 <i>Homo sapiens</i> cDNA clone IMAGE:4474055 5-, mRNA sequence
NM_000965.2		<i>Homo sapiens</i> retinoic acid receptor, beta (RARβ), transcript variant 1, mRNA
NM_003476.2		<i>Homo sapiens</i> cysteine and glycine-rich protein 3 (cardiac LIM protein) (CSRP3), mRNA
NM_001874.3		<i>Homo sapiens</i> carboxypeptidase M (CPM), transcript variant 1, mRNA
NM_002202.1		<i>Homo sapiens</i> ISL1 transcription factor, LIM/homeodomain, (islet-1) (ISL1), mRNA
NM_005097.1		<i>Homo sapiens</i> leucine-rich, glioma inactivated 1 (LGI1), mRNA
NM_031311.2		<i>Homo sapiens</i> carboxypeptidase, vitellogenic-like (CPVL), transcript variant 1, mRNA
NM_006538.2		<i>Homo sapiens</i> BCL2-like 11 (apoptosis facilitator) (BCL2L11), transcript variant 6, mRNA
Z97056		Human DNA sequence from clone RP3-434P1 on chromosome 22, complete sequence
U93239.1		Human Sec62 (Sec62) mRNA, complete cds
L47125.1		<i>Homo sapiens</i> (chromosome X) glypican (GPC3) mRNA, complete cds
AB002559.1		<i>Homo sapiens</i> mRNA for hunc18b2, complete cds
BC002552.2		<i>Homo sapiens</i> high-mobility group 20B, mRNA (cDNA clone MGC:1965 IMAGE:3138919), complete cds
AW080549		xc33a08.x1 NCI_CGAP_Co18 <i>Homo sapiens</i> cDNA clone IMAGE:2586038 3-similar to contains <i>Alu</i> repetitive element; contains element MSR1 repetitive element, mRNA sequence
NM_012081.2		<i>Homo sapiens</i> ELL-related RNA polymerase II, elongation factor (ELL2), mRNA
NM_000549.2		<i>Homo sapiens</i> thyroid stimulating hormone, beta (TSHB), mRNA
NM_000922		<i>Homo sapiens</i> phosphodiesterase 3B, cGMP-inhibited (PDE3B), mRNA
BF184274		601843339F1 NIH_MGC_54 <i>Homo sapiens</i> cDNA clone IMAGE:4064072 5-, mRNA sequence
X98405.1		<i>Homo sapiens</i> mRNA for myelin associated glycoprotein, S-MAG
NM_013374.3		<i>Homo sapiens</i> programmed cell death 6 interacting protein (PDCD6IP), mRNA
NM_016339.1		<i>Homo sapiens</i> Rap guanine nucleotide exchange factor (GEF)-like 1 (RAPGEFL1), mRNA
NM_022158.2		<i>Homo sapiens</i> fructosamine 3 kinase (FN3K), mRNA
NM_015899.1		<i>Homo sapiens</i> pleckstrin homology domain containing, family A (phosphoinositide binding specific) member 9 (PLEKHA9), mRNA

GenBank	Total shRNAs	Description
NM_020369.1	1	<i>Homo sapiens</i> fascin homolog 3, actin-bundling protein, testicular (Strongylocentrotus purpuratus) (FSCN3), mRNA
NM_019063.2		<i>Homo sapiens</i> echinoderm microtubule associated protein like 4 (EML4), mRNA
AA775177		ac79a06.s1 Stratagene lung (#937210) <i>Homo sapiens</i> cDNA clone IMAGE:868786 3-, mRNA sequence
AK023133.1		<i>Homo sapiens</i> cDNA FLJ13071 fis, clone NT2RP3001792, moderately similar to HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN M

* Experiments were performed with the HeLa cell line carrying the fluorescent p53-responsive construct.