nature genetics

Enzymatic production of RNAi libraries from cDNAs

Daisuke Shirane[1,2], Kohtaroh Sugao[1,2], Shigeyuki Namiki[1], Mao Tanabe[1], Masamitsu Iino[1] & Kenzo Hirose[1]

© 2004 Nature Publishing Group http://www.nature.com/naturegenetics

RNA interference (RNAi) induced by small interfering (siRNA) or short hairpin RNA (shRNA) is an important research approach in mammalian genetics. Here we describe a technology called enzymatic production of RNAi library (EPRIL) by which cDNAs are converted by a sequence of enzymatic treatments into an RNAi library consisting of a vast array of different shRNA expression constructs. We applied EPRIL to a single cDNA source and prepared an RNAi library consisting of shRNA constructs with various RNAi efficiencies. High-throughput screening allowed us to rapidly identify the best shRNA constructs from the library. We also describe a new selection scheme using the thymidine kinase gene for obtaining efficient shRNA constructs. Furthermore, we show that EPRIL can be applied to constructing an RNAi library from a cDNA library, providing a basis for future whole-genome phenotypic screening of genes.

Taking advantage of the huge quantity of genome data currently available[1–5], reverse genetics approaches that can determine the role of each gene by loss-of-function are important in determining gene-function relationships. Gene targeting[6] using homologous recombination is widely used, but its labor-intensive processes preclude its convenient and versatile application. Although antisense oligonucleotides can be more conveniently used, outcomes from this approach are often characterized by toxicity, instability and nonspecific effects[7].

RNAi, a gene suppression phenomenon triggered by double-stranded RNA[8], is a good alternative[7,9], because it is usually specific and can be achieved with unprecedented speed in a wide range of organisms. RNAi has been applied to genome-wide reverse genetics in *Caenorhabditis elegans*[10]. But the initial use of RNAi had been limited to invertebrates because long (>30 nucleotides, nt) double-stranded RNAs elicit interferon responses in higher vertebrates. This problem was overcome by the finding that a short (21–23 nt) double-stranded RNA (siRNA) directs RNAi in mammals without adverse effects[11]. The transient nature of siRNA effects has been overcome by the development of DNA-based vectors by which siRNA or shRNA is expressed intracellularly[12–17].

Despite this progress, there are still no general rules for designing siRNA or shRNA constructs with efficient gene-silencing activity; thus, it costs much time and money to identify suitable constructs. We developed a new technology called EPRIL to produce a library of shRNA expression constructs that can be systematically screened. We have also developed a technology for selecting the most efficient shRNA constructs from the library using the thymidine kinase gene. We furthermore showed that EPRIL can be used to produce an RNAi library from a cDNA library.

RESULTS

Production of an RNAi library

EPRIL comprises several steps of enzymatic treatments to produce an shRNA expression vector library from cDNAs of interest (**Fig. 1a**). First, double-stranded DNAs are quasi-randomly fragmented with DNase I (ref. 18). The fragments are then ligated to a hairpin-shaped adaptor containing the recognition sequence of *Mme*I. Although *Mme*I was reported to cut the top and bottom strands at 20 and 18 nt away from the recognition sequence, respectively[19], we found that *Mme*I cleaves DNA also at 21 and 19 nt away. Therefore, *Mme*I digestion liberates short 3′-protruding DNA fragments carrying sequences of either 20 or 21 nt from the target cDNAs. In PAGE analysis, *Mme*I-digested DNA appears as a band of ~40 nt (**Fig. 1b**). After *Mme*I digestion, the second adaptor is ligated to the liberated fragments. This adaptor carries two degenerate bases at its 3′ end on one strand so that the 3′ protruding end from the *Mme*I digest is filled in. After the second adaptor ligation, a primer extension reaction is carried out to convert the single-stranded hairpin DNA into a double-stranded DNA bearing inverted repeat sequences joined by a loop sequence (**Fig. 1**). Primer extension products are digested by appropriate restriction endonucleases to remove extra sequences outside the inverted repeats and are inserted into a plasmid vector described below. The extra sequence in a long loop flanked by inverted repeats is then removed with appropriate restriction endonucleases and religated to form a circle.

For library construction and shRNA expression, we used a retrovirus vector containing an RNA polymerase III–driven promoter from the mouse U6 gene to ensure stable RNAi expression[20].

RNAi library from cDNA encoding GFP

To evaluate the effectiveness of EPRIL, we produced a shRNA-expressing library from a cDNA encoding green fluorescent protein (GFP). Sequence analysis of individual clones from the library showed that

[1]Department of Pharmacology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. [2]These authors contributed equally to this work. Correspondence should be addressed to K.H. (kenzo@m.u-tokyo.ac.jp)

Published online 4 January 2004; doi:10.1038/ng1290

290 of 343 clones have inverted repeat sequences. Although 39 of these clones had imperfect inverted repeats, 251 clones had 19-nt or longer inverted repeats, and thus were candidate shRNA constructs for GFP silencing. The inverted repeat length of most clones was either 20 or 21 nt: 18 (7.2%), 139 (55.4%) and 94 (37.5%) of 251 clones were 19, 20 and 21 nt long, respectively. Because the first adaptor can be attached to both ends of DNase I–digested fragments, shRNA with two different orientations should be obtained. The guide sequence, which is complementary to mRNA sequences, was located on the 5′ or 3′ side of shRNA with almost the same frequency (54.6% versus 45.4%, $P > 0.1$). Analysis of target sequence positions showed that various shRNA constructs were generated from the target gene; 96.3% coverage was achieved across the entire 720-nt GFP-coding sequences with 157 nonredundant shRNA constructs from 251 independent clones. Thus, EPRIL allows high-throughput production of an array of shRNA constructs from a cDNA of interest.

© 2004 Nature Publishing Group http://www.nature.com/naturegenetics

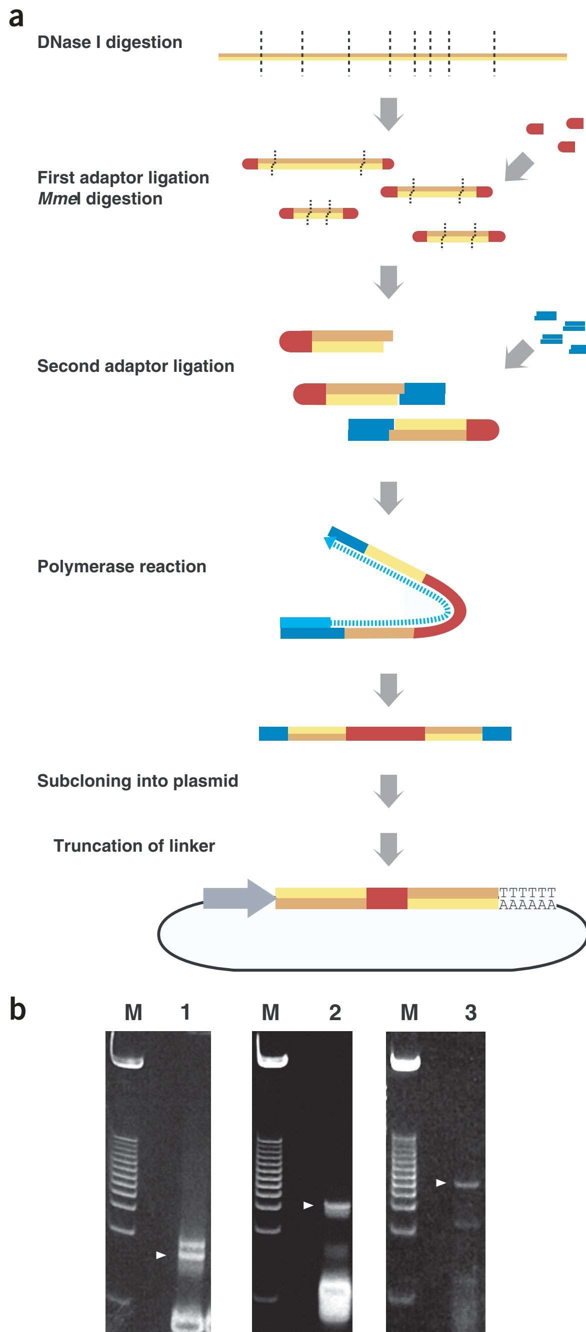

Figure 1 Outline of EPRIL. (**a**) Schematic of protocol for enzymatic derivation of cDNA into shRNA-expressing library. (**b**) PAGE of the DNAs in the course of enzymatic derivation. M, 25-nt ladder marker. Lanes 1, 2 and 3 show the products after first adaptor ligation and *Mme*I digestion, adaptor 2–ligated fragments and fragments after polymerase reaction, respectively. In each lane, the products are indicated by arrowheads.

We produced retroviruses harboring shRNA-expressing constructs from individual plasmids in 96-well plates, allowing us to obtain a large array of independent viruses. We then infected Jurkat T cells expressing GFP with the viruses and determined the level of expression of GFP by flow cytometry. We analyzed the distribution of RNAi efficiency from the compiled results of 262 nonredundant constructs (**Fig. 2a**). About 56% of constructs had low RNAi efficiency (reduction in GFP expression by a factor of less than 1.5). To estimate the probability of finding shRNA constructs with an RNAi efficiency causing at least a given relative reduction (x) in GFP expression, we plotted the fractional abundance of shRNA constructs with a relative reduction greater than x against x (**Fig. 2b**). Approximately 30% of shRNA constructs reduced expression by a factor of at least 2, whereas only a small percentage reduced expression by a factor of 8 or more. The overall probability was proportional to the relative reduction value with approximately −1.7 power-law scaling, which means, for example, that to find RNAi constructs that are 5 times more efficient, 15 times ($5^{1.7}$) more candidate constructs must be screened.

We then analyzed the regional dependence of RNAi efficiency. Efficient and inefficient shRNA constructs were distributed throughout the entire region (**Fig. 2c**). For quantitative analysis, we grouped the data into ten equal subregions (**Fig. 2d**) and carried out variance analysis. We found no significant regional dependency (Kruskal Wallis test, $P > 0.4$). Notably, the results indicated that even the region within 72 nt of the start codon can serve as a good target, despite the previous presumption that the region is a poor substrate for RNAi[9]. As to the orientation of shRNA constructs, there was the overall tendency for guide sequences residing in the 3′ side to be more effective.

The patterns of variability in RNAi efficiency are cell-type-independent, as the ten representative constructs tested showed similar efficiency profiles in HEK293 or HeLa cells (data not shown). Direct transfection of the plasmid or PCR-amplified minimal shRNA expression cassettes had profiles similar to those with retroviral transduction (**Fig. 3a**), excluding the effect of differences in viral titer. Furthermore, RNAi profiles with direct transfection of *in vitro*–transcribed shRNA correlated well to that of DNA-based expression (**Fig. 3b**). We obtained similar results with *in vitro*–transcribed shRNAs predigested with Dicer. These results indicate that the factors determining the RNAi efficiency profile are downstream of the transcription and Dicer-processing step.

RNAi library from cDNA encoding type 1 IP3R

To determine whether our strategy is applicable to endogenous genes, we used a DNA encoding the type 1 inositol 1,4,5-trisphosphate receptor[21] (IP3R) as a target. As with GFP, we obtained various shRNA constructs for IP3R. Of the 256 clones with sequence data, 214 carried the shRNA constructs, which included 199 nonredundant constructs with different target positions. For the rapid screening of effective shRNA constructs, we developed Jurkat T cells with a reporter construct that expresses head-to-tail linked mRNAs of GFP and IP3R. The degradation of target mRNA by shRNA constructs can be evaluated by monitoring the reduction in GFP fluorescence[22].

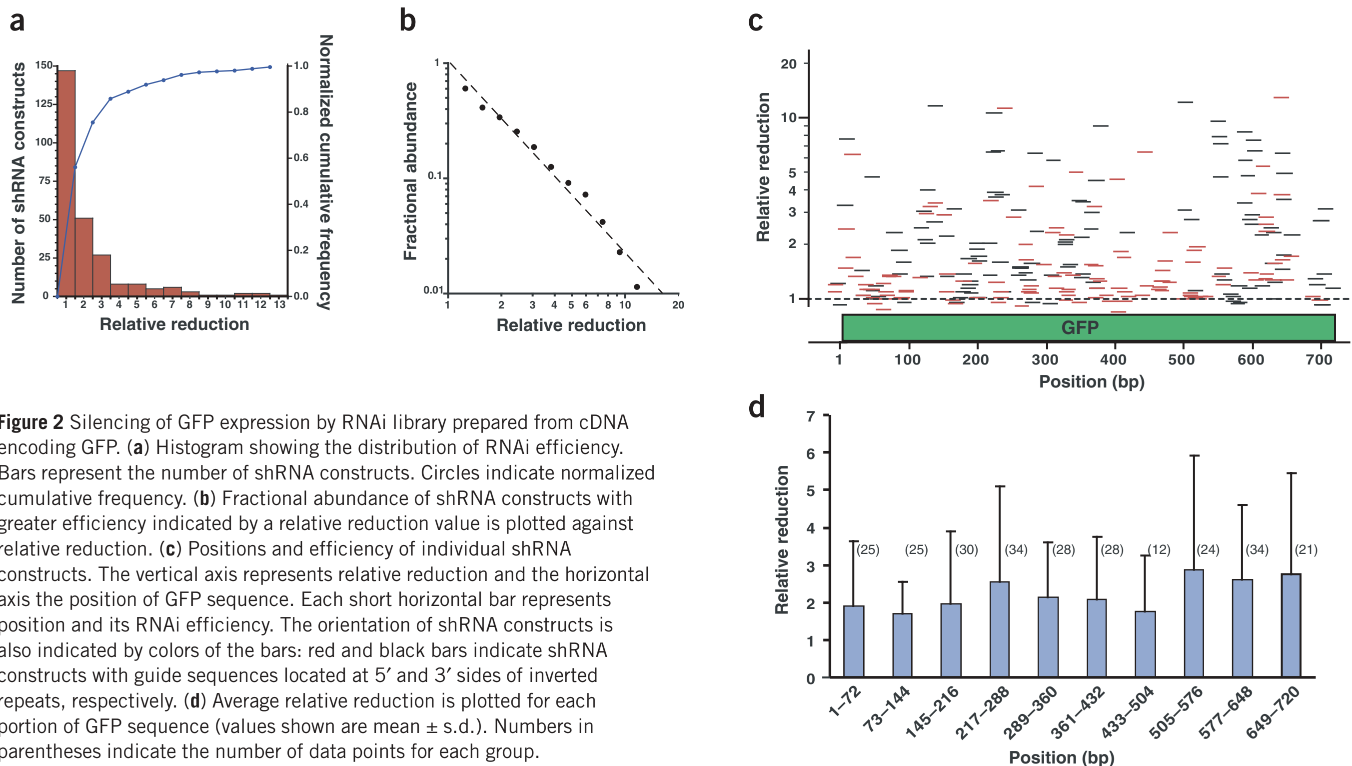

Figure 2 Silencing of GFP expression by RNAi library prepared from cDNA encoding GFP. (**a**) Histogram showing the distribution of RNAi efficiency. Bars represent the number of shRNA constructs. Circles indicate normalized cumulative frequency. (**b**) Fractional abundance of shRNA constructs with greater efficiency indicated by a relative reduction value is plotted against relative reduction. (**c**) Positions and efficiency of individual shRNA constructs. The vertical axis represents relative reduction and the horizontal axis the position of GFP sequence. Each short horizontal bar represents position and its RNAi efficiency. The orientation of shRNA constructs is also indicated by colors of the bars: red and black bars indicate shRNA constructs with guide sequences located at 5′ and 3′ sides of inverted repeats, respectively. (**d**) Average relative reduction is plotted for each portion of GFP sequence (values shown are mean ± s.d.). Numbers in parentheses indicate the number of data points for each group.

© 2004 Nature Publishing Group http://www.nature.com/naturegenetics

The shRNA constructs targeting IP3R also had varied RNAi efficiencies (**Fig. 4a**). We selected two of the most effective shRNA constructs, SI2A5 and SI3G6, and tested whether they could direct effective RNAi for endogenously expressed IP3R in the vascular smooth muscle cell line A7r5 (ref. 23). Western-blot analysis confirmed the reduction in IP3R expression levels (**Fig. 4b**). The loss of function was also confirmed by measuring vasopressin-elicited intracellular Ca^{2+} responses (**Fig. 4c**). These results indicate that EPRIL is useful for generating effective shRNA constructs targeting endogenous genes.

Selection of efficient shRNA constructs in cells

To facilitate high-throughput screening, we developed a new selection scheme that uses a specialized selection marker gene for positive selection of efficient shRNA constructs in cells (**Fig. 5a**). The marker gene consists of two parts linked head-to-tail: the first part encodes a fusion protein consisting of thymidine kinase and puromycin N-acetyltransferase[24] and the second encodes the target mRNA. Normally, ganciclovir kills cells carrying this marker gene owing to an intracellular accumulation of a toxic derivative of ganciclovir by thymidine kinase action[24]. Cells transduced with shRNA that have efficient RNAi activity for the target gene will escape death by silencing thymidine kinase expression. We constructed such a marker gene using GFP as the target, introduced it into Jurkat T cells and selected them with puromycin. We then produced retroviruses expressing shRNAs as a mixture from the shRNA library targeting GFP and used them to infect cells expressing the marker gene. We challenged the infected cells by treatment with ganciclovir for 48 h and cultured the ganciclovir-resistant cells. We recovered shRNA constructs by PCR amplification from surviving cells and reconstituted them into retrovirus expression vectors to obtain the selected library.

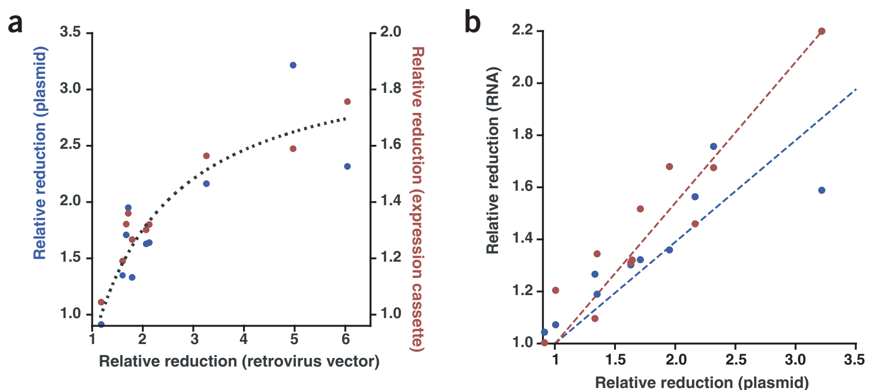

Figure 3 RNAi efficiency profiles with various transduction protocols. (**a**) Efficiency of GFP silencing by transfection of plasmid (left vertical axis, blue circles) or PCR-amplified expression cassette (right vertical axis, red circles) is compared with that by retrovirus transduction (horizontal axis). (**b**) Efficiency of GFP silencing by transfection of *in vitro*–transcribed shRNA with (red circles) and without (blue circles) Dicer digestion are plotted against efficiency (horizontal axis).

Figure 4 Silencing of type 1 IP3R expression by RNAi library derived from IP3R cDNA. (a) RNAi efficiency estimated on the basis of GFP fluorescence intensity. Relative reduction values in GFP fluorescence intensity are shown. Box represents the IP3R coding region, and dots indicate the target position of each shRNA construct. Data for shRNA constructs SI2A5 and SI3G6 are marked. (b) Western blotting of A7r5 cells infected with retroviruses carrying no shRNA constructs (Control), GFP-silencing shRNA expressing vector (shGFP) or shRNA constructs targeting IP3R (SI2A5 and SI3G6). Actin serves as loading control. Compiled normalized IP3R expression levels are shown in the lower panel (values shown are mean ± s.e.m., $n = 3$). (c) Intracellular calcium responses of A7r5 cells elicited by vasopressin with or without (black line) transduction of IP3R-targeting shRNA constructs (red and green lines denote SI2A5 and SI3G6, respectively). Bar indicates application of arginine vasopressin (AVP).

© 2004 Nature Publishing Group http://www.nature.com/naturegenetics

To test whether the ganciclovir selection concentrated efficient shRNA constructs, we infected Jurkat T cells expressing GFP with a mixture of viruses prepared from the reconstituted library. Flow cytometry showed that GFP expression was more attenuated in cells transduced with retroviruses from the ganciclovir-selected library than in cells transduced with retroviruses from the untreated or parental library, indicating that ganciclovir selection concentrated the more efficient shRNA constructs (**Fig. 5b**). We confirmed this by analyzing individual clones and observed a marked increase in the number of highly efficient shRNA constructs in the ganciclovir-selected library (**Fig. 5c**). These results indicate that the new selection scheme is useful for obtaining efficient shRNA constructs from the library in cells.

RNAi library from a cDNA library

EPRIL provides an opportunity for constructing a shRNA library from a complex mixture of cDNAs, such as a cDNA library, rather than a single cDNA source. We tested the applicability of the technology in a comprehensive search for genes responsible for particular cellular functions. We prepared a cDNA library from mRNA expressed in mouse myeloid precursor FL5.12 cells and subjected it to EPRIL to make a shRNA library. We sequenced randomly selected clones and found that 215 of the 240 sequences obtained contained inverted repeats (**Table 1**). Among these, inverted repeats from 35 clones contained sequences consisting of more than 10-nt stretches of poly(A) or poly(T), which presumably originated from poly(A) tails of mRNAs.

Figure 5 Thymidine kinase–based system for selection of efficient RNAi constructs. (a) Schema of selection strategy. Upper drawing shows the selection marker gene construct. TK-Puro indicates an mRNA encoding a fusion protein of thymidine kinase and puromycin N-acetyltransferase. Ganciclovir (GCV) is converted to a toxic derivative of ganciclovir (GCV-PPP) leading to cell death. (b) RNAi effects by a mixture of reconstituted shRNA constructs in RNAi library after ganciclovir (GCV) selection (left panel, red area) or without selection (right panel, blue area). Continuous curves in right panel represent data for the parental shRNA library. Dotted curves represent GFP fluorescence distribution without shRNA construct transduction. (c) Analysis of individual RNAi library clones with or without ganciclovir (GCV) selection. Individual clones were classified into four categories according to their RNAi efficiency: weak (relative reduction by a factor of <1.5), moderate (1.5–2.5), high (2.5–4.5) and very high (>4.5). Normalized population represents the number of clones in each category normalized to the total number of clones ($n = 121$ and 101 for clones with and without ganciclovir selection, respectively).

© 2004 Nature Publishing Group http://www.nature.com/naturegenetics

We subjected the remaining 180 clones to BLAST search; the sequences of 165 clones matched cDNA sequences or ESTs (**Supplementary Table 1** online). We grouped these clones originating from gene transcripts according to UniGene clusters (Build 126) and found that 146 of the 165 clones belong to at least one UniGene cluster. Of these, some clones corresponded to the same clusters, suggesting that these clones originated from the same gene. These include genes encoding heat shock protein 8, ubiquitin B and ribosomal protein L41, all of which are highly expressed in many organs. As the size of a UniGene cluster is a rough estimate of the relative expression level, we compared cluster size with the number of appearances of the same gene. The cluster size with repeats of more than one was significantly greater than that without repeats (Mann-Whitney U-test, $P < 0.003$). These features suggest that the RNAi library represents the original expression profiles. We further analyzed the target positions of the transcripts and found that 53% corresponded to the coding region, 44% to the 3′ untranslated region and 3% to the 5′ untranslated region. Thus, shRNA constructs were obtained from various portions of the transcripts. Collectively, the results indicate that EPRIL can be applied to a cDNA library consisting of a complex mixture of various genes.

DISCUSSION

We established a technology called EPRIL by which a shRNA-expressing library can be enzymatically derived from cDNA templates. When applied to single cDNA sources, the library provides an array of candidate shRNA constructs from various regions of the cDNA. The combination of EPRIL with high-throughput screening and the selection scheme in cells provides a general platform to produce shRNA constructs for every gene. Such collections will greatly contribute to RNAi-based high-throughput reverse genetics in mammals.

We attempted to identify sequence preferences for efficient RNAi. Although we found a weak but common tendency for shRNA constructs having uracil at the fourth nucleotide from the 5′ end of the guide sequence to be efficient for silencing GFP and IP3R, preferences at other positions were not always consistent. For example, guanine at the second nucleotide tended to be favorable for silencing IP3R but not GFP. Clearly, more data from various genes are required to determine general sequence preferences.

EPRIL enables the production of a large shRNA library from the cDNA library consisting of a complex mixture of cDNAs. We estimated that the library contains ~3–40 10^5 independent cDNA-derived shRNA constructs. Therefore, the present method may enable RNAi-based forward genetics in which one can identify genes by selection based on absolute phenotypic criteria, such as cell morphology, adhesion, cell death, expression level of a key molecule or other functional indexes. shRNA sequences can serve as reliable tags for gene identification, because the probabilities of 9.1 10^{-13} and 2.3 10^{-13} that tags of 20 nt and 21 nt, respectively, with random sequences can be perfectly matched[25] are sufficiently small, considering the size of the mouse or human genome (~3 10^9 nt).

We postulate that forward genetics using the RNAi library may be applied at the whole-animal level. Because lentivirus-mediated transduction enables the production of mice carrying shRNA constructs[26], we can efficiently produce an array of mutant mice with shRNA constructs targeting different genes. Once these are systematically screened based on phenotypes, the gene responsible for a certain phenotype can be readily identified. This feature contrasts with an ongoing large-scale mutagenesis project in mice using ethylnitrosourea[27]. Although ethylnitrosourea mutagenesis efficiently produces mutant mice, the processes used to determine loci are tedious. Furthermore, the variability in RNAi efficiency of shRNA constructs may lead to discovering unique phenotypes owing to the variability in the degree of gene silencing. EPRIL will thus contribute to the elucidation of the gene-function relationship at the whole-genome level through both forward and reverse genetics.

Table 1 Summary of sequencing of shRNA library clones prepared from FL5.12 cDNA library using EPRIL

Successful sequences	240
No inverted repeats	25
Inverted repeat–containing clones	215
Sources of inverted repeats	
Poly(A)[a]	35
cDNA/EST match	165
Mitochondrial gene	6
Genomic DNA[b]	3
Primer	3
Unidentified	3

[a]Inverted repeats consisting of more than 10-nt stretches of adenines or thymines.
[b]Sequences that matched genomic DNA sequences but not cDNA or EST sequences.

METHODS

Cell culture. We cultured Jurkat T cells in RPMI1640 medium (Invitrogen) containing 10% fetal calf serum, penicillin and streptomycin. We maintained GP293, HEK293, A7r5 and HeLa cells in Dulbecco's modified Eagle's medium (Sigma or Invitrogen) containing 10% fetal calf serum. We maintained FL5.12 cells (a gift from T. Inaba, University of Hiroshima) in RPMI1640 medium (Sigma) containing 10% fetal calf serum and 1 ng ml^{-1} IL-3 (Wako, Japan).

Plasmid construction. We constructed all plasmids, including that carrying the shRNA-expressing retrovirus vector, pNAMA-U6, using standard molecular biological techniques. Briefly, we obtained the U6 promoter and termination signal–encoding DNA by PCR amplification from pSilencer1.0-U6 (Ambion) and inserted them into the *Nhe*I site of the plasmid carrying the 3′ long terminal repeat (LTR) from the pMX retrovirus vector (a gift from T. Kitamura, University of Tokyo). We subjected the plasmid to PCR amplification (primer sequences available on request) to form *Bbs*I-*Bsm*I sites for inserting shRNA-encoding DNA fragments to obtain pBsk-U63-3LTR. To construct pNAMA-U6, we excised a *Hin*dIII-*Sal*I fragment bearing the 3′ LTR containing the U6 promoter at the *Nhe*I site from pBsk-U63-3LTR and inserted it into the *Hin*dIII-*Sal*I site of pda5LTR-DsRed2-M4, which carries the pMX vector backbone lacking the 3′ LTR with the gene DsRed2 originating from pDsRed2-N1 (Clontech).

pMS240-PNS encodes the retrovirus vector expressing the thymidine kinase–puromycin–N-acetyltransferase fusion protein[24] under the control of the SV40 promoter. We constructed it from PCR-amplified fragments encoding thymidine kinase and puromycin acetyltransferase. We subcloned DNA fragments containing the thymidine kinase and puromycin N-acetyltransferase genes into a self-inactivating retrovirus vector, pMS240, that we developed. pMS240-PNS has *Bam*HI-*Not*I sites for cloning DNA fragments encoding target genes. To select effective shRNA constructs targeting GFP, we inserted GFP-encoding fragments from pd2EGFP-1 (Clontech) into the sites.

To construct pMX-d2EGFP-IP3R, a plasmid encoding the retrovirus vector for monitoring IP3R silencing, we prepared DNA encoding destabilized GFP variants from pd2EGFP-1 and inserted it into pMX. We then inserted DNA encoding IP3R from pBS-IP3R into a site downstream of d2EGFP.

EPRIL. For shRNA constructs targeting GFP, we obtained the *Bam*HI-*Not*I DNA fragment encoding GFP from pEGFP-1 (Clontech). For IP3R, we prepared a fragment encoding the entire coding region of rat type 1 IP3R by *Eco*RI-*Not*I digestion of pBS-IP3R1. To produce cDNA library–derived shRNA constructs, we generated a double-stranded cDNA library from mRNA prepared from FL5.12 cells using a Super SMART PCR cDNA synthesis kit (Clontech) and used this directly for further derivation without subcloning. We then applied EPRIL using these DNA fragments according to the following six-step protocol.

© 2004 Nature Publishing Group http://www.nature.com/naturegenetics

Step 1. Preparation of randomly fragmented DNA. We partially digested the DNA fragment with DNase I in the presence of 1 mM $MnCl_2$, 0.1 mg ml^{-1} bovine serum albumin and 50 mM Tris-HCl (pH 7.5) and repaired its ends with 0.1 U l^{-1} T4 DNA polymerase (Takara). We empirically determined DNase I (Takara) concentration for each preparation to obtain digests with an average size of ~100–200 nt on PAGE.

Step 2. First adaptor ligation and *Mme*I cleavage. We ligated the digests to a hairpin-shaped oligonucleotide, adaptor 1 (sequence available on request), with T4 DNA ligase (DNA ligation kit version 2, Takara). The hairpin-shaped adaptor 1 was purified from a chemically synthesized oligonucleotide by native PAGE before use. We repaired a nick between the 5′ end of adaptor 1 and the 3′ end of the digested DNA with 0.006 U l^{-1} *E. coli* ligase (Takara) in the presence of 0.1 mM NAD, 1.2 mM EDTA, 10 mM $(NH_4)_2SO_4$, 4 mM $MgCl_2$ and 30 mM Tris-HCl (pH 8.0) after phosphorylation with 1.0 or 1.25 U l^{-1} T4 polynucleotide kinase. The adaptor-linked short DNA fragments were liberated by cleavage with *Mme*I (NEB). We detected a liberated DNA fragment as a single band migrating at ~40 nt on native PAGE. We excised the band in the gel and obtained the *Mme*I-cleaved product with phenol-chloroform extraction and ethanol precipitation.

Step 3. Second adaptor ligation. We prepared adaptor 2 by annealing an oligonucleotide pair (sequences available on request) and purification using PAGE. The latter oligonucleotide contains two degenerate bases at the 3′ end, denoted as N, which represents A, C, G or T. We ligated adaptor 2 to the *Mme*I-cleaved DNA fragment from step 2 using T4 DNA ligase. After purification by PAGE, we repaired a nick on the adaptor 2–ligated fragment by treatment with T4 polynucleotide kinase (Takara) and T4 DNA ligase.

Step 4. Primer extension. We then subjected the product from step 3 to a primer extension reaction, incubated it with the primer oligonucleotide (sequence available on request) at 94 °C for 135 s in the reaction buffer containing 0.1% Triton X-100, 0.2 mM dNTPs, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$ and 20 mM Tris-HCl (pH 8.8) and then cooled it to 62 °C. We added 0.02 U l^{-1} *Bst* DNA polymerase large fragment (NEB) to the reaction mixture to initiate the primer extension reaction. The reaction was carried out for 300 s at 65 °C and stopped by cooling to 4 °C. We purified the reaction product by PAGE.

Step 5. Subcloning into plasmid. We digested the product from step 4 with *Bpm*I, blunt-ended it with the Klenow fragment (Takara) and then digested it with *Bbs*I. We digested pNAMA-U6 with *Bsm*I, similarly blunt-ended it, digested it with *Bbs*I and treated it with bacterial alkaline phosphatase (Takara). We gel-purified the digested fragments and ligated them at a molar ratio of about 3:1. We then electroporated the ligation mixture into ElectroMAX DH5α-E competent cells (Invitrogen). We plated the transformed cells on 500-cm^2 LB-agar plates containing 100 g ml^{-1} carbenicillin. After overnight incubation, we collected the bacterial lawn using a scraper and prepared plasmid DNA using a plasmid purification kit (Plasmid MIDI kit, Qiagen).

Step 6. Truncation of excessive linker. We digested the plasmid purified from step 5 with *Bcg*I, blunt-ended it with T4 DNA polymerase and recircularized it by self-ligation. This treatment removes most of the sequences from adaptor 1 but retains the short linker sequence. To eliminate as much contamination by incompletely digested plasmids as possible, we digested the DNA with *Mfe*I, whose recognition sequence resides in the *Bcg*I-truncated portion of sequences from adaptor 1. We transformed ElectroMAX DH5α-E competent cells by the recircularized plasmids and selected them on LB-agar plates containing 100 g ml^{-1} carbenicillin. After overnight cultivation on the plate, we obtained a plasmid library stock.

Recovery of shRNA expression constructs from genome. We recovered shRNA expression constructs by PCR amplification from 100 ng of genomic DNA prepared from the transduced FL5.12 cells. We used *Vent* DNA polymerase (NEB) for PCR (primer sequences available on request). We digested the PCR-amplified DNA with *Not*I and *Afl*II and subcloned it into pBsk-3LTR. After transformation and purification, we excised the DNA encoding recovered shRNA constructs containing the 3′ LTR with *Hin*dIII and *Not*I and subcloned it into pda5LTR-DsRed2-M4. The resultant plasmid is structurally the same as pNAMA-U6 and can be subjected to packaging to produce retroviruses carrying the recovered shRNA constructs.

shRNA preparation *in vitro*. We synthesized shRNAs by T7 promoter–based *in vitro* transcription. To prepare templates for *in vitro* transcription, we annealed two chemically synthesized oligonucleotides, a T7 promoter-encoding oligonucleotide and a shRNA-encoding oligonucleotide (sequences available on request) at 45 °C for 30 s after incubating for 135 s at 94°C. We then converted the annealed oligonucleotides to double-stranded DNA templates by extension reaction using the *Bst* DNA polymerase large fragment (0.08 U l^{-1}) at 50 °C for 600 s. We produced shRNA from the purified templates using a CUGA7 *in vitro* transcription kit (Nippon Gene) according to the manufacturer's protocol. We purified the transcribed shRNA using a gel filtration spin column (MicroSpin G-25, Amersham Biosciences). To produce Dicer-digested shRNA, we treated shRNA with recombinant human Dicer (Gene Therapy Systems) according to the manufacturer's protocol. Oligonucleotide sequences of the 11 different shRNA-encoding oligonucleotides are available on request.

Retrovirus production and transduction. For high-throughput screening, we produced and transduced retrovirus carrying shRNA constructs in 96-well plates. We prepared plasmids in 96-well plates using a QIAwell 96 Ultra Plasmid kit (Qiagen) according to the manufacturer's protocol. We transfected GP293 packaging cell lines (Clontech) with ~200 ng of the retrovirus vector plasmid and 17 ng of the VSV-G-encoding plasmid using Lipofectamine 2000 (Invitrogen) for each well in the 96-well plates. We obtained the culture medium containing retrovirus 2 d after transfection. We transduced retrovirus by adding 50 l of the medium containing retrovirus particles to 50 l of Jurkat T cell suspension (1.0 10^5 cells per ml). For the middle-scale production of retrovirus, we purified DNA with a Plasmid MIDI kit and packaged retrovirus in GP293 grown on 10-cm dishes by transfection with 24 g of the retrovirus vector plasmid and 2 g of the VSV-G-encoding plasmid. When necessary, we concentrated the retrovirus particles by overnight centrifugation followed by resuspension in an appropriate culture medium.

Flow cytometry and evaluation of RNAi efficiency. We estimated relative GFP expression levels based on fluorescence intensity as analyzed by a FACScan flow cytometer (BD). We used relative reduction in GFP fluorescence (control fluorescence intensity divided by that in cells with shRNA transduction) as a measure of RNAi efficiency. To correct batch-to-batch differences in RNAi efficiency due to variation in retrovirus titer, we included a series of internal control shRNA constructs in each 96-well plate.

Western-blot analysis. We infected A7r5 cells with retrovirus harboring shRNA constructs. Four days later, we collected the cells by trypsinization, solubilized them and subjected them to SDS-PAGE. After transferring the SDS-PAGE products to PVDF membranes, we probed the protein corresponding to IP3R with the rabbit IgG antibody to type1 IP3R (Alomone) and with horseradish peroxidase–conjugated antibody to rabbit IgG (MBL) as the primary and secondary antibodies, respectively, and detected it by chemiluminescence. We quantified immunoblot signal intensity using IPLab spectrum software (Scanalytics). We normalized the IP3R signal with the immunoblot signal of actin probed using an antibody to actin (Santa Cruz).

Intracellular Ca^{2+} measurements. For intracellular Ca^{2+} measurements, we seeded A7r5 cells seeded on coverslips, infected them with shRNA-expressing retroviruses and loaded them with the Ca^{2+} indicator Fura-2 4 d after infection. We evaluated changes in intracellular Ca^{2+} concentration by ratiometric fluorescence measurements on an inverted microscope equipped with a CCD camera (Photometrics), as described previously[28]. We stimulated the cells by applying 1 nM arginine vasopressin in the presence of an L-type Ca^{2+} channel blocker, nicardipine (10 M).

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS
We thank A. Hashimoto for technical advice regarding FL5.12 cell culture and K. Serizawa for help in data analysis. This work was supported in part by Grants-in-

Aid for Scientific Research and the Advanced and Innovational Research program in Life Sciences from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 27 October; accepted 15 December 2003
Published online at http://www.nature.com/naturegenetics/

1. The *C. elegans* Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012–2018 (1998).
2. Adams, M.D. *et al.* The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185–2195 (2000).
3. Waterston, R.H. *et al.* Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
4. Lander, E.S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
5. Venter, J.C. *et al.* The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
6. Capecchi, M.R. The new mouse genetics: altering the genome by gene targeting. *Trends Genet.* **5**, 70–76 (1989).
7. Opalinska, J.B. & Gewirtz, A.M. Nucleic-acid therapeutics: basic principles and recent applications. *Nat. Rev. Drug Discov.* **1**, 503–514 (2002).
8. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811 (1998).
9. Dykxhoorn, D.M., Novina, C.D. & Sharp, P.A. Killing the messenger: short RNAs that silence gene expression. *Nat. Rev. Mol. Cell Biol.* **4**, 457–467 (2003).
10. Kamath, R.S. & Ahringer, J. Genome-wide RNAi screening in *Caenorhabditis elegans*. *Methods* **30**, 313–321 (2003).
11. Elbashir, S.M. *et al.* Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494–498 (2001).
12. Brummelkamp, T.R., Bernards, R. & Agami, R. A system for stable expression of short interfering RNAs in mammalian cells. *Science* **296**, 550–553 (2002).
13. Miyagishi, M. & Taira, K. U6 promoter-driven siRNAs with four uridine 3′ overhangs efficiently suppress targeted gene expression in mammalian cells. *Nat. Biotechnol.* **20**, 497–500 (2002).
14. Yu, J.Y., DeRuiter, S.L. & Turner, D.L. RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells. *Proc. Natl. Acad. Sci. USA* **99**, 6047–6052 (2002).
15. Sui, G. *et al.* A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. *Proc. Natl. Acad. Sci. USA* **99**, 5515–5520 (2002).
16. Lee, N.S. *et al.* Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat. Biotechnol.* **20**, 500–505 (2002).
17. Paul, C.P., Good, P.D., Winer, I. & Engelke, D.R. Effective expression of small interfering RNA in human cells. *Nat. Biotechnol.* **20**, 505–508 (2002).
18. Anderson, S. Shotgun DNA sequencing using cloned DNase I-generated fragments. *Nucleic Acids Res.* **9**, 3015–3027 (1981).
19. Boyd, A.C., Charles, I.G., Keyte, J.W. & Brammar, W.J. Isolation and computer-aided characterization of MmeI, a type II restriction endonuclease from Methylophilus methylotrophus. *Nucleic Acids Res.* **14**, 5255–5274 (1986).
20. Paddison, P.J. & Hannon, G.J. RNA interference: the new somatic cell genetics? *Cancer Cell* **2**, 17–23 (2002).
21. Mignery, G.A., Newton, C.L., Archer, B.T. 3rd & Sudhof, T.C. Structure and expression of the rat inositol 1,4,5-trisphosphate receptor. *J. Biol. Chem.* **265**, 12679–12685 (1990).
22. Kumar, R., Conklin, D.S. & Mittal, V. High-throughput selection of effective RNAi probes for gene silencing. *Genome Res.* **13**, 2333–2340 (2003).
23. De Smedt, H. *et al.* Determination of relative amounts of inositol trisphosphate receptor mRNA isoforms by ratio polymerase chain reaction. *J. Biol. Chem.* **269**, 21691–21698 (1994).
24. Chen, Y.T. & Bradley, A. A new positive/negative selectable marker, puDeltatk, for use in embryonic stem cells. *Genesis* **28**, 31–35 (2000).
25. Saha, S. *et al.* Using the transcriptome to annotate the genome. *Nat. Biotechnol.* **20**, 508–512 (2002).
26. Rubinson, D.A. *et al.* A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference. *Nat. Genet.* **33**, 401–406 (2003).
27. Kile, B.T. *et al.* Functional genetic analysis of mouse chromosome 11. *Nature* **425**, 81–86 (2003).
28. Hirose, K., Kadowaki, S., Tanabe, M., Takeshima, H. & Iino, M. Spatiotemporal dynamics of inositol 1,4,5-trisphosphate that underlies complex Ca^{2+} mobilization patterns. *Science* **284**, 1527–1530 (1999).

© 2004 Nature Publishing Group http://www.nature.com/naturegenetics