18

Isolation of Genetic Suppressor Elements (GSEs) from Random Fragment cDNA Libraries in Retroviral Vectors

Andrei V. Gudkov and Igor B. Roninson

1. Introduction

Cellular phenotypes resulting from decreased function of a specific gene are manifested as recessive, since they are suppressed in the presence of a normal allele of the corresponding gene. The powerful gene-transfer techniques, which have played a key role in the studies of dominant phenotypes, are not readily applicable to recessive traits, since expression of a recessive allele does not generally affect the cellular phenotype. In haploid organisms, random gene disruption can be used as a general method of cloning recessive genes, but this approach is limited to genes that are not essential for cell growth and is not, at least presently, applicable to diploid cells. This explains why studies of recessive genes of higher eukaryotes (e.g., tumor suppressor genes) are in many cases lagging far behind the analysis of dominant genes (e.g., oncogenes).

The identification and functional analysis of recessive genes in mammalian cells have been boosted by the ability to select genetic suppressor elements (GSEs) that induce the desired phenotype by suppression of specific genes *(1–7)*. GSEs are short (<500 bp) cDNA fragments that produce a phenotype when expressed in the cells; this phenotype is usually opposite to that of the full-length cDNA from which they are derived. GSEs inhibiting recessive genes behave as dominant selectable markers in gene-transfer protocols and can therefore serve as tools for studying recessive mechanisms.

There are two types of GSE: antisense-oriented GSEs encoding efficient inhibitory antisense RNA molecules and sense-oriented GSEs encoding functional protein domains that interfere with the protein function in a dominant

From *Methods in Molecular Biology, Vol 69 cDNA Library Protocols*
Edited by I G Cowell and C A Austin Humana Press Inc , Totowa, NJ

fashion. GSEs are isolated by preparing an expression library containing randomly fragmented DNA of the gene or genes targeted for suppression, introducing this library into the appropriate recipient cells, selecting cells with the desired phenotype, recovering the inserts from the expression vectors contained in the selected cells, and testing the recovered sequences for functional activity. Both sense- and antisense-oriented GSEs can be isolated through this procedure; the orientation, sequence, and mechanism of action of the GSEs are determined only after their biological activity has been established.

Principles and applications of the GSE methodology, as well as the results obtained with the use of this approach have been summarized elsewhere *(1)*. Briefly, the random fragment selection strategy was first applied in a prokaryotic system and used for the isolation of GSEs from the DNA of phage λ, by selecting GSEs for the ability to protect *Escherichia coli* from λ-induced lysis *(2)*. This approach was then extended to mammalian cells in a study where a set of GSEs were isolated from the cDNA encoding human topoisomerase II (topo II), by the ability of such GSEs to confer resistance to etoposide, a topo II-interacting drug *(3)*. In the latter study, a protocol was developed for preparing a random fragment library in the form of a mixture of recombinant retroviruses capable of highly efficient gene transfer into different mammalian cells. GSE selection was also carried out from random fragment libraries of several other single genes (e.g., p53 *[4]* and BCL-2 *[5]*), leading in all cases not only to the generation of new genetic tools, but also to new insights into the mechanisms of function and regulation of the studied genes.

Construction of GSE libraries from more complex targets, such as the mixture of multiple cDNA clones *(6)* or total cellular cDNA *(7)*, has extended the applications of GSE methodology from characterization of already known genes to identification of new genes whose inactivation is associated with various selectable phenotypes. GSEs isolated from such a library are most likely to induce their effect by inhibiting the gene from which they are derived. GSEs therefore can be used as probes for cloning full-length cDNA for the genes whose suppression caused the selected phenotypes. Several retroviral libraries carrying random fragments of normalized cellular mRNA sequences were constructed and used for isolation of GSEs inducing resistance to cytotoxic drugs or abnormal cell growth properties. This approach, for example, allowed us to demonstrate the involvement of the motor protein kinesin in drug sensitivity and senescence *(7)*.

Thus, the GSE strategy provides a general method for cloning genes involved in complex cellular phenotypes and for identifying functional domains of individual gene products. Isolated GSEs may be used as tools to study

the corresponding gene function not only in cell culture but, in perspective, in the whole organism, if introduced into the genome of transgenic animals. Phenotypic changes or developmental disorders associated with genetically introduced GSEs may indicate the normal function of the genes targeted by such GSEs.

As a methodology, the GSE approach is still young and open to improvements. Some stages of library preparation, delivery, or selection are flexible and may be done in several alternative ways (we have tried to describe some of the alternatives in Section 4.). Thus, the protocols presented herein reflect our as yet limited experience in GSE development and may not necessarily provide the optimal experimental design.

2. Materials

2.1. GSE Libraries from Individual cDNAs

2.1.1. Random Fragmentation of a cDNA Insert by DNase I

1. GeneClean kit for DNA purification from agarose gels (BIO 101, Midwest Scientific, St. Louis, MO)
2. DNase I (Sigma, St. Louis, MO): 1 mg/mL in 0.01*N* HCl; keep frozen at –70°C in small aliquots.
3. 10X DNase I buffer: 500 m*M* Tris-HCl, pH 7.5, 100 m*M* $MnCl_2$, 1 mg/mL bovine serum albumin (BSA); prepare fresh and keep on ice.

2.1.2. Filling the Ends of DNase I-Generated Fragments

1. T4 DNA polymerase (New England Biolabs, Beverly, MA)
2. Klenow fragment of *E. coli* DNA polymerase (New England Biolabs).
3. 10X T4 polymerase buffer: 500 m*M* NaCl, 100 m*M* Tris-HCl, pH 7 9, 100 m*M* $MgCl_2$, 50 mg/mL BSA, 10 m*M* dithiothreitol (DTT)
4. Stock solution of dNTPs, 2.5 m*M* each.

2.1.3. Adapter Ligation, Size Fractionation and Polymerase Chain Reaction (PCR) Amplification

1. Two oligonucleotides annealed to form a double-stranded adapter (Fig. 1): store frozen in annealing buffer: 10 m*M* Tris-HCl, pH 7 5, 7 m*M* $MgCl_2$, 100 m*M* NaCl.
2. T4 DNA ligase (New England Biolabs).
3. 5X Ligation buffer: 250 m*M* Tris-HCl, pH 7.8, 100 m*M* $MgCl_2$, 50 m*M* DTT, 5 m*M* adenosine triphosphate (ATP), 125 mg/mL BSA.
4. *Taq* DNA polymerase (Promega, Madison, WI) or equivalent enzyme
5. 10X PCR buffer: 500 m*M* KCl, 100 m*M* Tris-HCl (pH 9.0 at 25°C).
6. dNTP 10X stock solution, 2.5 m*M* each.
7. Sense-oriented adaptor oligonucleotide (Fig. 1), water solution, 1 mg/mL; keep frozen at –70°C in small aliquots.

Fig. 1. Adapter sequences used for different GSE libraries. **(A)** Adapter used for topo II random fragment library *(3)*, contains *Cla*I cloning site, single ATG codon in sense, and three stop codons in all reading frames in antisense orientation. Forms relatively long inverted repeats at the ends of inserts. **(B)** Adapter with *Cla*I cloning site and three ATG codons in all reading frames; was used for large GSE library from normalized NIH 3T3 cDNA *(7)*; does not contain stop codons, which are provided by vector sequences. **(C)** Asymmetrical adapters used for p53 library *(4)*, one contains *Eco*RI cloning site and three ATG codons, and another one contains three stop codons and *Bam*HI site; allows for oriented cloning. In this and in the previous adapters, ATG contexts do not perfectly match Kozak's initiator codon consensus. **(D)** Optimized, though never experimentally tested, adapter sequence in which ATG codons are put in perfect Kozak's consensus for initiator codons.

2.1.4. Cloning and Amplification of the Library

1. Centricon-100 filters (Amicon, Beverly, MA, catalog no. 4212).
2. Purified vector plasmid DNA (pLNCX, pLXSN *[8]* or other retroviral vectors).
3. Restriction enzymes and their 10X reaction buffers.
4. Calf intestinal alkaline phosphatase (New England Biolabs).
5. 10X Alkaline phosphatase buffer: 500 m*M* NaCl, 100 m*M* Tris-HCl, pH 7.9, 100 m*M* $MgCl_2$, 10 m*M* DTT.
6. L-broth (per liter): 10 g tryptone, 5 g yeast extract, 10 g sodium chloride Autoclave.

7 L-agar: L-broth containing 1 5% bacto-agar. Autoclave to melt agar and sterilize
8 Ampicillin.

2.2. GSE Libraries from Normalized Total Cellular cDNA

2.2.1. cDNA Synthesis

1 cDNA synthesis system, such as Gibco BRL (Grand Island, NY; catalog no 8267SA) or equivalent kit.
2. PolyA$^+$ RNA, 2–10 µg, water solution 0.2–1 mg/mL (keep at –70°C)
3. 5'-Phosphorylated random hexanucleotides, water solution, 1 mg/mL.
4. α-[^{32}P]-dCTP at 5000 Ci/mmol (Amersham, Arlington Heights, IL)

2.2.2. Filling the Ends, Ligation with Adapter, and PCR Amplification

Refer to the materials listed in Section 2.1.3.

2.2.3. Normalization of cDNA

1 2X Hybridization solution: 0 6*M* Na-phosphate, pH 7 5, 2 m*M* EDTA, 0.2% SDS
2. Hydroxylapatite for nucleic acid chromatography (Bio-Rad, Richmond, CA).
3 0.6*M* Phosphate buffer, pH 7 2.
4 Centricon-100 (Amicon)

2.2.4. Cloning

1 Highly efficient competent bacterial cells for electroporation (Invitrogen, San Diego, CA, INVaF' or TOP10 Library Size Electroporation Kits)
2. L-agar, ampicillin

2.3. Screening of GSE Libraries

2.3.1. GSE Library Delivery

1. Packaging cell lines: BOSC23 *(9)* (ecotropic) and BING *(10)* (amphotropic)
2. Dulbecco's modified Eagle's medium (DMEM) tissue-culture media with 10% newborn or fetal calf serum.
3. Target cells and appropriate culture media.
4 Solutions for calcium-phosphate transfection *(9)*: 2X HBS· 50 m*M* HEPES, pH 7.05, 10 m*M* KCl, 12 m*M* dextrose, 280 m*M* NaCl, 1 5 m*M* Na_2HPO_4; 1*M* $CaCl_2$, 25 m*M* chloroquine.
5 CsCl-quality purified retroviral vector plasmids
6 Polybrene (hexadimethtine bromide, Sigma), 4 mg/mL sterile water stock solution.
7. Geneticin (G418) (Gibco BRL) or other appropriate selective agent

2.3.2. Virus Rescue from Selected Cells

1. PCR solutions (*see* Section 2.1.3.).
2. PCR primers.
3 Standard cloning reagents (purified vectors, restriction enzymes, etc.) *(11)*

2.3.3. Control of Library Representativity

1 PCR solutions.
2. Primers.
3 Reagents for standard Southern blot hybridization *(11)*.

3. Methods

3.1. GSE Libraries from Individual cDNAs

GSE libraries from individual cDNAs are useful for structure/functional analysis of specific genes whose suppression leads to selectable changes in cellular phenotype. GSEs serve as efficient genetic tools allowing suppression of the gene's function in various cellular systems. In addition, the identification of sense-oriented GSEs is a method of mapping functional protein domains. Examples of successful GSE isolation from single cDNA libraries include topo II *(3)*, p53 *(4)*, and BCL-2 *(5)*.

3.1 1. Random Fragmentation of a cDNA Insert by DNase I

1. Excise cDNA insert from the plasmid using appropriate restriction enzyme(s), and purify the insert from agarose gel by GeneClean procedure or any other alternative method (e.g , phenol extraction from low-melting agarose or electroelution *[11]*)
2 Keep the isolated DNA in TE buffer at a concentration of 0 1–1 mg/mL As a sufficient amount for further procedures, we recommend preparing at least 10 μg of the insert
3 To estimate the conditions of DNase I digestion, first treat several 0.5-μg aliquots of isolated DNA inserts with DNase I (freshly prepared 1.1000 dilution of the stock solution in reaction buffer) at 16°C in freshly prepared DNase I buffer (*see* Note 1) for different time periods (from 5 s to 1 min). Stop the reactions by adding 0.5*M* EDTA to the final concentration of 25 m*M* and estimate the degree of DNA degradation by electrophoresis in 1.5% agarose gel next to a 123-bp ladder (Gibco BRL) or other appropriate size marker
4. Use the reaction conditions leading to the desired size range (*see* Note 2) to scale up the reaction proportionally to the amount of DNA. Remove DNase I by phenol–chloroform extraction and ethanol precipitation (here and in Section 3.1.2 where this procedure is mentioned, do two phenol–chloroform extractions followed by one chloroform extraction, add EDTA to 20 m*M*, precipitate with 2.5 vol of ethanol, and wash DNA precipitate twice with 70% ethanol to remove traces of organic solvents).

3.1.2. Filling the Ends of DNase I-Generated Fragments

1. Dissolve DNA in T4 DNA polymerase buffer containing 0.25 m*M* of each of dNTPs (approx 10 μL for 1 μg DNA).
2. Add T4 DNA polymerase and Klenow fragment of *E coli* DNA polymerase (5 U/μg each), and incubate at 37°C for 1 h.

3. Stop reaction by adding EDTA to 25 m*M*.
4. Purify DNA by phenol–chloroform extraction.

3.1.3. Adapter Ligation, Size Fractionation, and PCR Amplification

1. Mix equimolar amounts of sense and antisense strands of the adaptor (5–10 μg total) in 50 μL of annealing buffer. Put the mixture into a water bath preheated to 90°C, and allow to cool down slowly to room temperature.
2. Mix 1 μg of DNase-treated cDNA insert with 1 μg of annealed adapter (Fig 1; *see* Note 3) in 10 μL reaction buffer with ATP and 1 μL of T4 ligase.
3. Incubate the reaction at room temperature for at least 1 h.
4. Load the ligation mixture on 1.5% agarose gel. After electrophoresis, cut out DNA fragments of the desired size range (*see* Note 2), and elute them from agarose using any appropriate method.
5. If necessary, amplify the resulting mixture of fragments by PCR using "sense" strand of the adapter as a primer for both ends. Before PCR, incubate the mixture with *Taq* DNA polymerase for 10 min at 55°C to eliminate nicks between the adapter and cDNA fragments. Before doing large-scale PCR, estimate in a pilot set of reactions the minimal number of cycles necessary to reach the plateau (under our conditions, this number varied from 8–16 cycles). To avoid artifacts, do not exceed this value in large-scale PCR (*see* Note 4)

3.1.4. Cloning and Amplification of the Library

1. Purify the PCR product by phenol–chloroform extraction, and centrifuge through Centricon-100 filters to remove primers and short artifactual PCR products.
2. Estimate the quantity of your cDNA preparation, and treat it with the desired restriction enzyme(s) using ≥50 U of enzyme/μg of DNA for 2–16 h.
3. Purify cDNA fragments by phenol-chloroform extraction and centrifugation through Centricon-100
4. Digest vector DNA (*see* Note 5) with the appropriate restriction enzyme(s), and (in case of using a single restriction enzyme) treat it with alkaline phosphatase according to manufacturer's recommendations.
5. Before large-scale ligation, determine optimal ligation conditions using different ratios of insert and vector DNAs, and subsequently testing transformation efficiency and the proportion of recombinant clones among the transformants (our standard ligation conditions: molar ratio 5:1, DNA concentration 10–50 ng/mL).
6. Once optimal conditions are selected, do a large-scale ligation and transformation of competent bacterial cells, and plate them on L-agar with an appropriate antibiotic. Based on our experience, an extensive set of GSEs can be derived from single-gene cDNA libraries containing about 5000–10,000 recombinant clones/kb of cDNA.
7. Wash off the bacterial colonies from agar by thoroughly scraping the surface with a spatula, using several milliliters of L-broth/plate.

8 To amplify the library, allow the resulting bacterial suspension to grow for 3–4 h in L-broth with the appropriate antibiotic (about 1 L of broth/30,000 colonies)
9. Isolate and purify plasmid DNA. For purification, use either CsCl density centrifugation or any other alternative method that allows highly purified plasmid DNA to be used for calcium-phosphate transfection.

3.2. GSE Libraries from Normalized Total Cellular cDNA

Highly complex GSE libraries of random fragments of total cellular mRNA are expected to contain suppressors of practically all cellular genes. Such libraries can therefore be used for the identification and cloning of new genes whose suppression alters the cell phenotype. A total polyA$^+$ RNA preparation is a mixture of unequally represented sequences, some very abundant and some extremely rare. In order to increase the probability of isolating GSEs from rare cDNAs, a polyA$^+$ RNA-derived cDNA preparation is subjected to a normalization procedure, providing for uniform abundance of different DNA sequences Assuming that the mammalian genome contains at least 10,000 genes, and that the incidence of GSE-containing clones in a single-gene cDNA library is on the order of 1/100–1/500 (in our experience), the minimal complexity of a mammalian GSE library has to be at least 10^7 independent recombinant clones.

3.2.1. cDNA Synthesis

To achieve equal representation of 5'- and 3'-mRNA sequences in a random fragment cDNA preparation, cDNA synthesis is carried out on fragmented mRNA using random oligonucleotides as primers (*see* Note 6).

1. To determine the conditions for RNA fragmentation, treat 3–5 aliquots of the polyA$^+$ RNA (each aliquot contains 50 ng of RNA in 10 µL) in a boiling water bath for 2–15 min.
2. Use treated RNA samples for pilot cDNA synthesis using a commercially available cDNA synthesis kit according to the manufacturer's protocol for double-stranded cDNA synthesis, using random hexanucleotides at a concentration of 500 ng/mL as primers. Monitor the size of the first strand of cDNA by labeling it with α-[^{32}P]-dCTP and subjecting it to electrophoresis under denaturing conditions *(11)*
3. For large-scale cDNA synthesis, use the conditions of mRNA degradation under which most of the product ranges between 200 and 1,000 bp (*see* Note 7). Synthesize at least 100 ng of cDNA

3.2.2. Filling the Ends, Ligation with Adapter, and PCR Amplification

Proceed as described for the single-gene library protocol (*see* Section 3.2.1.) PCR amplification is an essential step in a complex cDNA library preparation, since it is necessary to generate enough cDNA for normalization (20 µg) and cloning (10–20 µg).

3.2.3. Normalization of cDNA

This is a modification of the procedure of Patanjali et al. *(12)*.

1. Denature 20 µg of cDNA by boiling for 5 min in 25 µL of TE buffer, followed by immediate cooling on ice.
2. Add 25 µL of 2X hybridization solution, and divide the mixture into four aliquots in Eppendorf tubes, 12.5 µL each.
3. Add one to two drops of mineral oil to each sample to avoid evaporation, and place the tubes into a 68°C water bath for annealing.
4. Freeze down one tube every 12 h.
5. After the last time-point, dilute each of the annealing mixtures with water to a final volume of 500 µL and subject DNA to hydroxylapatite (HAP) chromatography (*see* Note 8)
6. Put into Eppendorf tubes HAP suspension equilibrated with 0.01*M* phosphate-buffered saline (PBS) so that the volume of HAP pellet is approx 100 µL.
7. Preheat tubes with HAP and all the solutions used below, and keep them at 65°C. Remove the excess of PBS, and add diluted annealing solution Mix thoroughly, but gently, by shaking in a 65°C water bath
8. Leave the tubes in the water bath until HAP pellet is formed (a 15-s centrifugation to collect the pellet is okay; do not exceed 1000*g* in the microcentrifuge to avoid damage of HAP crystals).
9. Carefully replace the supernatant with 1 mL of preheated 0.01*M* PBS, and repeat the process.
10. To elute single-stranded DNA (ssDNA), suspend the HAP pellet in 500 µL of PBS at the single-strand elution concentration determined as described in Note 8 (e.g., 0.16*M*), collect the supernatant, and repeat the process. Combine supernatants, and remove traces of HAP by centrifugation.
11. Concentrate ssDNA by centrifugation, and wash it three times using 1 mL of water on Centricon-100.
12. Amplify isolated ssDNA sequences by PCR with the sense primer from the adapter, using a minimal number of cycles to obtain 10 µg of the product (it usually requires about 20 PCR cycles to synthesize 10 µg of PCR product in 12–20 50-µL reactions using approx 1/10 of the isolated ssDNA).
13. Check that the size of the PCR product remains within the desired range (200–500 bp) (*see* Notes 2, 3, and 7).
14. Test the normalization quality by Southern or slot-blot hybridization with ^{32}P-labeled probes for high, moderate- and low-expressing genes using 0.3–1.0 µg of normalized cDNA/lane (slot) (Fig. 2) We use β-actin and β-tubulin cDNAs as probes for high-expressing genes, c-*myc* and topo II cDNAs for moderate, and c-*fos* cDNA for low-expressing genes.
15. Compare cDNA isolated after different annealing times with the original unnormalized cDNA. Make sure that the probes are of a similar size and specific activity. Use the best-normalized ssDNA fraction, which produces the most uniform signal intensity with different probes, for large-scale PCR amplification to

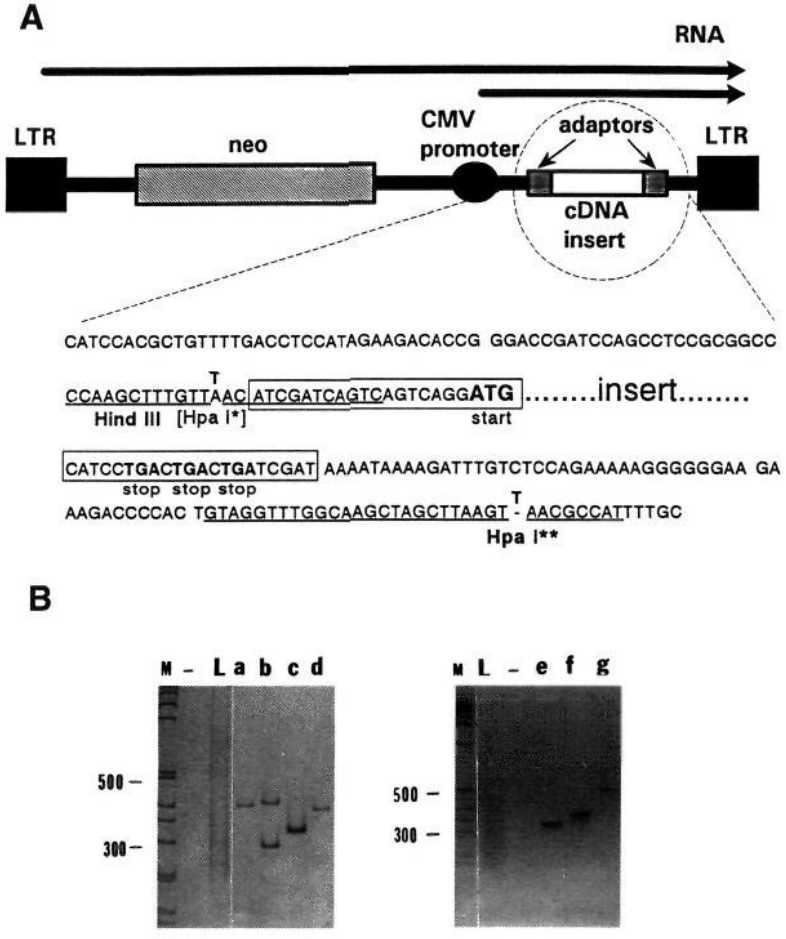

Fig. 2. (A) Structure of an integrated provirus containing an insert of topo II cDNA insert. Adapter-derived sequences are bracketed, with translation initiation and termination codons shown in boldface. Sequences used as PCR primers for isolation and recloning of the inserts are underlined; changes in the primer sequences leading to inactivation (*) or creation (**) of the indicated restriction sites are shown above the vector sequence. **(B)** PCR amplification of proviral inserts from DNA of HeLa cells after infection with the retroviral library (lane L) and from DNA of HeLa cell clones obtained after etoposide selection (lanes a–g). The sense strand of the adapter shown in Fig. 1A was used as the PCR primer. Too many PCR cycles used for amplification of the library cDNA inserts (lane L) in the left gel resulted in the artifactual increase in the size of PCR products owing to a concatemer formation. PCR products were separated in polyacrylamide gel and stained with ethidium bromide. Size markers are indicated in base pairs.

synthesize at least 20 μg of the product for cloning. Avoid increasing the number of PCR cycles to increase the yield of the product. Instead, use more ssDNA template to obtain the desired amount by scaling up the number of PCR reactions or the reaction volume.

3.2.4. Cloning

In order to construct a plasmid library with a complexity $\geq 10^7$ clones, it is necessary to prepare large amounts of vector and insert DNAs, to optimize carefully the conditions of ligation and transformation, and to have competent cells with very high transformation efficiency. We have obtained adequate results using the InVitrogen transformation kit, which provides competent cells for electroporation with an efficiency of 10^9 transformants/μg of plasmid DNA. This kit also includes competent cells for chemical transformation that allow one to select optimal ligation conditions using minimal amounts of material in pilot transformations. It is usually enough to have 5–10 μg of restriction enzyme-digested inserts and 20–30 μg of digested and alkaline phosphatase-treated vector.

1. Estimate optimal vector-to-insert ratio in the ligation mixture according to the manufacturer's (InVitrogen) recommendations.
2. Determine the proportion of recombinant clones after transformation (this should be no less than 60%) by PCR analysis of 20–40 individual colonies, using the sense strand of the adapter as the PCR primer.
3. Do large-scale ligation under predetermined conditions, and purify DNA from the ligation mixture by phenol–chloroform extraction and ethanol precipitation Dissolve the DNA in TE buffer at a concentration of 0.3–1 μg/μL. Divide ligated DNA into several batches (approx 1 μg of DNA in each), and electroporate competent cells using one batch at a time.
4. Plate bacterial suspension containing $2–5 \times 10^6$ transformed cells on L-agar with the appropriate antibiotic, at no more than 10^3 colonies/cm^2.
5. To maintain the representativity of the library, isolate and purify plasmid DNA directly from bacteria grown on agar plates, rather than from bacteria grown in suspension. Estimate the quantity of DNA obtained after each transformation, and keep each batch of plasmid DNA separately.

3.3. GSE Library Screening

Successful isolation of biologically active GSEs from random fragment cDNA libraries requires careful and accurate tuning of the selection conditions that depend on the nature of the selectable phenotype, cell type, and so on. The examples of several selection schemes are given in our papers *(5–10)*. Here we describe some general considerations and essential experimental steps involved in any GSE selection protocol.

It is important to keep in mind that since most GSEs are relatively weak and because the levels of transgene expression vary widely in a retrovirally transduced cell population, one cannot expect to see the selectable phenotypic alterations in every cell that carries a GSE. Therefore, in order to be selected, every element in the GSE library should be delivered to numerous (at least 30–100) cells. This is best achieved by increasing the number of infected cells and by reaching a very high efficiency of viral transduction.

The relatively weak effect of most GSEs also means that the phenotypic alterations caused by the GSEs may not be strong enough to allow the application of selection conditions that would be so stringent as to eliminate all spontaneous background (as determined with the same recipient cells infected with an insert-free retroviral vector). Although in some cases we were able to select GSEs under zero-background conditions, it is prudent to assume before the start of a new selection that GSEs would only work under the conditions that give a selection background of 10^{-5} to 10^{-4} or even higher.

It is often difficult to combine both requirements (a large number of library-infected cells and relatively mild selection conditions) in the design of the selection. This problem is overcome by the possibility of performing multiple rounds of selection. After each round, all the GSEs and "passenger" cDNA inserts integrated in the selected cells are recovered and reintroduced into fresh recipient cells. The cycle of selection and recovery is repeated until the number of clones becomes small enough to provide a clear increase in selectability of library-transduced cells over the control. At this time, the selected elements can be individually tested for the GSE activity. An example of a two-round selection is the isolation of GSEs conferring resistance to etoposide from a normalized GSE library of NIH 3T3 cDNA *(7)*.

3.3.1. GSE Library Delivery

1. Plate the ecotropic (for mouse or rat recipient cells) or amphotropic (for recipient cells from other mammals) packaging cells (Note 9) at 6×10^6 cells/100-mm plate. (The number of plates with packaging cells should match the number of plates with the target cells that will be used for infection and selection. This number is calculated from the library size, the requirements of the selection procedure, and the predetermined efficacy of retroviral infection)
2. The next day, transfect each plate with 10 μg of plasmid DNA using standard calcium-phosphate procedure with modifications suggested by W Pear for BOSC23 and BING cells (*see* ref. *11* and Note 10).
3. Prior to transfection, change media on packaging cells by adding 9 mL of fresh DMEM with 10% serum. Dissolve plasmid DNA in 375 μL of sterile water, add 125 μL of 1*M* $CaCl_2$, and add 0.5 mL 2X HBS by bubbling. Immediately (within 1–2 min) add this solution to the cells.

4. Plate 2×10^6 target cells/150-mm plates.
5. Collect media from packaging cells 24 and 48 h after transfection, filter it through 0.45-mm filters, dilute twofold with fresh media, add polybrene to 4 μg/mL, and put on target cells for at least 2 h. Rock plates gently several times during this period. Initiate selection as soon as possible, but not earlier than 48 h after the last infection In most cases, the target cells do not require G418 selection (or other selection depending on the nature of the selectable marker in the vector) (*see* Note 10). A small subpopulation of the infected cells, however, may be subjected to G418 treatment to determine the efficiency of infection.

3.3.2. Virus Rescue from Selected Cells

Although we have used several procedures to rescue the integrated virus from infected cells (*see* Note 11), only one method so far has proven to be reliable for all the different types of target cells. This method involves PCR amplification of the cDNA inserts, followed by recloning into the retroviral vector. The same procedure is used both to generate a GSE-enriched library at intermediate rounds of selection and to derive retroviral clones carrying potential GSEs for individual testing. Primers for this PCR correspond to vector sequences flanking the inserts. It is important to have sense primer sequence partially overlapping with the beginning of the adapter. This allows the avoidance of PCR artifacts caused by annealing of adapter-derived inverted repeats flanking the insert. One of several possible primer designs that allows oriented cloning of cDNA fragments back into the same vector is given in our work on topo II GSE isolation (ref. *3*; Fig. 3).

3.3.3. Control of Library Representativity

This is important at different stages of library delivery and screening. It is based on the analysis of PCR-generated population of cDNA inserts by gel electrophoresis and, in the case of the total cDNA library, by Southern hybridization with probes for several cellular genes. As long as good representativity is maintained in the library, both the electrophoretic pattern and the hybridization profiles should show a smear within the GSE size range. Loss of representativity (which occurs during the selection as a normal and desired event, but may also arise as an unfortunate consequence of uneven infection or unequal growth of bacteria during library amplification) is accompanied by the appearance of discrete bands on the gel or in Southern blots. An alternative method is based on PCR using several pairs of primers, one of which is uniform and corresponds to vector sequence, and another representing internal sequences of certain cDNA sequences that are present in the GSE library. Until representativity is preserved, each of these pairs generates multiple bands in PCR.

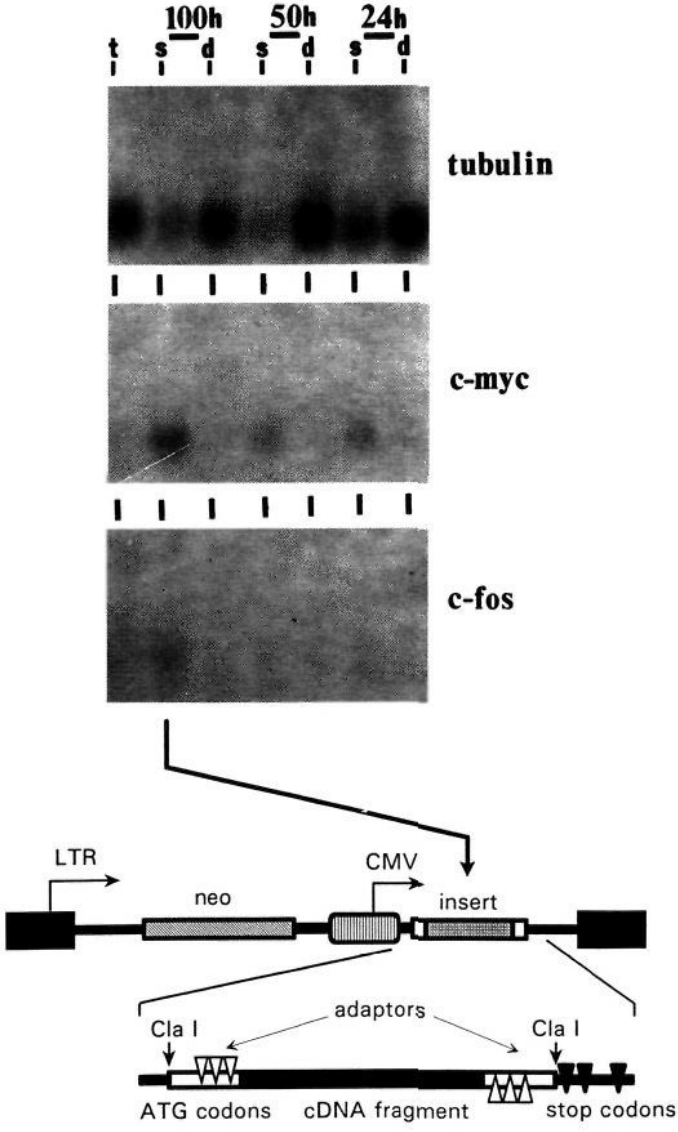

Fig. 3. Construction of normalized cDNA library. (Upper) Normalization of cDNA fragments from NIH 3T3 cells. Total cDNA (lane t) was reannealed for the indicated periods of time, and the single-stranded (lane s) and double-stranded (lane d) fractions were analyzed by Southern hybridization with the indicated probes. (Lower) Structure of an integrated provirus containing a cDNA fragment in LNCX vector.

3.3.4. Testing Individual Clones for GSE Activity

This is the final and most difficult stage of screening. Although the exact protocol for such testing depends on the specific features of the selection procedure used, there are some general observations and rules that are applicable to most of the GSEs.

GSEs are not conventional dominant selectable markers: in many cases, their biological activity is relatively weak and can only be demonstrated under mild selection conditions. This is particularly true for GSEs isolated from total cDNA libraries, rather than from individual genes. Do not necessarily expect that a GSE would have a stronger activity than that of the enriched virus population at the end of the selection. It is often easier to obtain a retroviral population that is capable of inducing the desired phenotype than to prove the activity of individual elements.

In most cases, GSEs are unable to induce the desired phenotype in every transduced cell. It is therefore important to deliver the putative GSEs into a large number of target cells (10^5 to 10^6) and to compare it with an equally large population of control cells. Avoid analyzing cell clones: it is often impossible to determine the relative impact of clonal variability and of the GSE effect on the phenotype.

Make sure that both experimental and control cells undergo exactly the same conditions of treatment. Even minor discrepancies in the treatment conditions (e.g., differences in the number of packaging cells taken for transfection with control and experimental plasmids, number of virus infections, etc.) may significantly jeopardize the result. To avoid difficulties in reproducing biological effects of GSEs during individual testing, it is necessary to maintain the same assay conditions throughout the whole process of selection (the stock of media, the lot of serum, the same generation of target cells, etc.).

Since the GSE enrichment during selection is often easier to demonstrate than direct GSE activity, it can be used as an assay for GSE confirmation: a mixture of control vector-transduced cells with the GSE-carrying cells (with the excess of control cells) may be subjected to the selection procedure in the expectation that cells with the GSE would be preferentially selected.

Some of the GSEs may have a detrimental effect on the cell growth. Even if such effect is very weak, cells with low GSE expression will gradually overgrow in the population, thus diluting the GSE-induced phenotype. Furthermore, spontaneous silencing of the GSE-expressing transcription unit may also occur over time. We therefore strongly recommend the investigator to analyze cell phenotypes immediately after GSE transduction (within several days). In most cases, G418 selection is not necessary for GSE testing and may even lead to a decrease in the GSE activity.

In some cases, weak GSE effects may be enhanced by an increase in GSE expression. This can be achieved by introducing multiple GSE copies into the target cells through repeated infection *(7)*. Furthermore, when only a single GSE is available for the given gene (as is usually the case with the GSEs selected from total cDNA libraries), subsequent isolation of full-length cDNA sequence and screening of a new GSE library prepared from the cloned cDNA is likely to yield more active GSEs.

4. Notes

1. In the presence of Mn^{2+}, DNase I makes mostly double-strand breaks, whereas Mg^{2+} stimulates single-strand break formation *(11)*. It is important to prepare DNase I-buffer with Mn^{2+} just before mixing the reaction and to keep it in tightly closed tubes in an ice bath. The solution may have a light brown color
2. For GSE library preparation, we usually use random fragments whose size varies from 80–500 bp. No correlation between the GSE size and activity was found within this size range. PCR amplification of cDNA fragments after ligation with adapter (*see* Note 3) may lead to a substantial decrease in the average size of the fragments owing to more efficient amplification of shorter fragments This problem can be resolved by separate isolation and processing of cDNA fragments of two size ranges (e.g., 80–200 bp and 200–500 bp), which can be combined at the ligation stage.
3. Adapter design is determined by the vector and the restriction map of the cDNA target (if known). Generally, adapters should provide the insert with the cloning sites that are absent or rare in the insert, with translation initiator codon(s) at the beginning, and stop codons at the end of the cDNA fragment. The adapters also have a protruding end at one side to ensure directional ligation. Several suggested adapter designs are shown in Fig. 1. Adapters may have the initiator codon in one or in all three open reading frames. Initiator and stop codons (in different orientations) may be included into the same adapter sequence (Fig. 1A). Alternatively, initiator and stop codons may be separated between two different adapters (Fig. 1C). The latter design allows one to have shorter adaptor sequences and to use different cloning sites for directional recloning. Finally, adapters need not carry any of the initiation or termination codons if they are provided in the vector (Fig. 1B) The role of the adapter in such a case is reduced to supplying the cDNA fragments with uniform sequence tags for PCR amplification and cloning. Adapter sequence is a subject for future optimization and testing (Fig. 1D). For single-gene GSE libraries, we recommend the adaptor design shown in Fig. 1C, keeping in mind that the restriction sites for cloning should not be present within the target cDNA.
4. Ligation with adapters supplies cDNA fragments with inverted repeats. At the plateau of the PCR reaction, when the primer concentration is low and the concentration of the product is high, these repeats may link different cDNA fragments by crossannealing. This leads to formation of artifactual concatemeric PCR

products (*see* Fig. 3A). This problem does not occur if the proper number of cycles and primer concentrations are maintained.

5. We are primarily using two retroviral vectors constructed in A. D. Miller's laboratory (Fred Hutchinson Cancer Research Center, Seattle, WA): pLNCX and pLXSN *(12)*. Both vectors contain two promoters, one of which drives the expression of the *neo* (G418 resistance) gene. GSEs are expressed from the cytomegalovirus promoter in LNCX (*see* Fig. 2) and from the retroviral (Moloney murine leukemia virus) LTR in the LXSN vector. Both vectors have been successfully used for preparation and screening of single-cDNA GSE libraries. Note that GSEs in pLXSN are transcribed as a part of long proviral RNA whose secondary structure may in some cases interfere with the function of short antisense-oriented elements.
6. cDNA preparations initiated from oligo(dT) primer are known to underrepresent 5'-end mRNA sequences. Similarly, cDNAs synthesized on full-length mRNAs using random primers underrepresent 3'-end mRNA sequences. Fragmentation of mRNA allows us to reach more equal 5'- and 3'-end sequence representation if cDNA synthesis is done from a random primer. Random primers should be phosphorylated to ensure successful ligation with an adapter
7. The presence of cDNA fragments longer than the desired size range prior to PCR is no reason for concern: The average length of the final products will be significantly decreased as a result of PCR amplification that favors shorter fragments. To preserve longer fragments in the library (400–600 bp), it is recommended to isolate them from the gel after PCR amplification of normalized ssDNA and to amplify them separately. This fraction is added to the bulk of the DNA preparation before cloning.
8. Before HAP chromatography, phosphate buffer used for DNA elution must be calibrated. Prepare serial dilutions of a 1*M* stock of PBS ranging from 0.01–0.5*M* concentration, and determine PBS concentrations at which ssDNA and double-stranded DNA (dsDNA) are eluted from HAP. As a test DNA, use any ^{32}P-labeled PCR product generated under the conditions of asymmetrical PCR and containing a mixture of ssDNA and dsDNA. It is important to keep the temperature of the tubes with HAP at 68°C all the time, especially during ssDNA elution. At lower temperatures, ssDNA may form a net of fragments linked through annealed adapter regions that can behave as partially dsDNA and be retained in HAP. An example of cDNA normalization is shown in Fig. 3.
9. In our experience, ecotropic BOSC 23 *(9)* and amphotropic BING *(10)* packaging cell lines, developed by W. S. Pear in the laboratory of D. Baltimore (Rockefeller University) from human 293 cells, provide by far the highest retroviral titer (at least 10^6/mL) on transient transfection. The titer is maintained for 2 d, allowing two viral infections of target cells from one transfection. However, growth of some target cells (e.g., mouse embryo fibroblasts) is negatively affected by the media conditioned by BOSC23 or BING cells. Therefore, the following steps are recommended.

a. Dilute virus-containing media at least twofold with fresh media;
b. Expose target cells to the virus for 2–4 h with subsequent replacement with fresh media; and
c. Make sure that control cells are treated with an insert-free virus under exactly the same conditions.

In the cases where 293-derived cell supernatants have an adverse effect, one can also consider using NIH 3T3-derived packaging cells, such as GP+E86 (ecotropic) and GP+envAm12 (amphotropic) *(13)*, which are not efficient enough to ensure delivery of complex libraries, but can be used for smaller single-gene libraries *(3,4)*.

10 Transfection protocol for BOSC23 and BING cells suggested by their creators contains several differences from the conventional calcium-phosphate DNA transfection protocol *(11)*.
   a Cells should be about 80% confluent prior to transfection; otherwise significant cell death occurs after precipitate is added;
   b DNA precipitate is added to the cells immediately after mixing 2X HBS with DNA-$CaCl_2$ solution, and
   c. It is also recommended by W. Pear *(11)* to add 25 m*M* of chloroquine into media prior to transfection and keep cells with precipitate and chloroquine for 7–11 h, then media are replaced with a chloroquine-free medium, which is used as a source of the virus 48 h after transfection (We usually skip chloroquine treatment and media replacement and do infection twice: 24 and 48 h after transfection)

11 Fifty to 100% infection of most target cells can be reached after optimizing the conditions for infection This allows one to avoid the expensive and time-consuming G418 selection of target cells. In addition, it makes it possible to do GSE selection right after the delivery of the GSE library, before loss of complexity may occur owing to variations in cell growth rate or transcriptional inactivation of proviruses.

12. The alternative methods for provirus rescue from the selected cells are cell fusion with murine NIH 3T3-derived packaging cells *(13)* and long-range PCR of the entire integrated provirus, rather than just the cDNA insert. To rescue the virus by fusion, plate into a 100-mm plate a mixture of 4×10^6 ecotropic packaging cells GP+E86 *(13)* with 2×10^6 selected cells. Allow the cells to attach and spread (5–12 h), wash them three times with PBS to remove media with serum completely and load 1 mL of 10% polyethylene glycol (PEG, mol wt 6000, Sigma) for 1 min. Dilute and remove PEG, wash cells twice with PBS, and add fresh media. Collect media with rescued virus 24 and 48 h after PEG treatment, add polybrene, and infect amphotropic packaging cells, line GP + envAm12 *(13)*. Titer of rescued virus is usually quite low (up to 10^3). Select infected cells on G418 adding the drug 24 h after last infection. Virus produced by the resulting cell population usually has titer up to 10^5 and may be used for infection of fresh target cells and secondary screening Some GSE selections can be performed right on these packaging cell lines producing the library (for example, the first

round of selection of etoposide-resistant GSEs was done on the mixture of eco- and amphotropic packaging cells "ping-ponging" the NIH 3T3 GSE library *[7]*). In this case, virus rescue step may be omitted since selected cells produce virus by themselves Both virus rescue by fusion and long-range PCR are faster and less laborious than the main procedure, which involves recloning, but their efficacy has varied greatly among different target cell types and individual investigators. Thus, we would not yet recommend these procedures for rescuing complex mixtures of cDNA inserts until further optimization.

Acknowledgments

The GSE methodology development was supported by National Cancer Institute grants R37CA40333, RO1CA39365, RO1CA56736, RO1CA62099 (I. B. Roninson) RO1CA60730, R21CA62045, RO3TW00475 (A. V. Gudkov), a grant from Ingenex, a US Army grant DAMD17-94-J-4038, a Faculty Research Award from the American Cancer Society (I. B. Roninson), and American Cancer Society (Illinois Division), grants 91-58 and 93-64 (A. V. Gudkov).

The authors thank all the members of their laboratories who contributed to the development of the GSE protocols and the investigators who provided vectors, cell lines, and other materials used in this work.

References

1 Roninson, I B , Gudkov, A. V , Holzmayer, T A., Kirschling, D J , Kazarov, A. R., Zelnick, C. R., Mazo, I A , Axenovich, S. A., and Thimmapaya, R (1995) Genetic suppressor elements: new tools for molecular oncology. *Cancer Res* **55,** 4023–4028

2 Holzmayer, T. A., Pestov, D. G., and Roninson, I. B. (1992) Isolation of dominant negative mutants and inhibitory antisense RNA sequences by expression selection of random DNA fragments. *Nucleic Acids Res.* **20,** 711–717.

3. Gudkov, A. V., Zelnick, C., Kazarov, A. R., Thimmapaya, R., Suttle, D. P., Beck, W. T., and Roninson, I. B. (1993) Isolation of genetic suppressor elements, inducing resistance to topoisomerase II-interactive cytotoxic drugs, from human topoisomerase II cDNA. *Proc. Natl. Acad Sci USA* **90,** 3231–3235.

4. Ossovskaya, V. S., Mazo, I. A., Strezoska, Z., Chernova, O. B., Chernov, M. V., Stark, G. R., Chumakov, P M , and Gudkov, A V (1996) Use of genetic suppressor elements to dissect distinct biological effects of separate p53 domains. *Proc. Natl Acad. Sci USA* **93,** in press.

5. Tarasewich, D. G., Schott, B., Xuan, Y., Chang, B., Salov, S., Holzmayer, T. A., and Roninson, I. B. (1996) Isolation of a genetic element inhibiting BCL-2 expression and sensitizing tumor cells to chemotherapeutic drugs. In preparation.

6. Pestov, D. G. and Lau, L F. (1994) Genetic selection of growth inhibitory sequences in mammalian cells. *Proc. Natl. Acad. Sci USA* **91,** 12,549–12,553

7 Gudkov, A. V., Kazarov A. R., Thimmapaya, R., Mazo, I. A., Axenovich, S., and Roninson, I. B. (1994) Isolation of genetic suppressor elements from a retroviral

normalized cDNA library: identification of a kinesin associated with drug sensitivity and senescence. *Proc Natl Acad Sci USA* **91,** 3744–3748.

8 Miller, A. D. and Rosman, G. J. (1989) Improved retroviral vectors for gene transfer and expression. *BioTechniques* **7,** 980–986.

9. Pear, W. S., Nolan, G. P., Scott, M. L., and Baltimore, D. (1993) Production of high titer helper-free retroviruses by transient transfection *Proc Natl Acad Sci USA* **90,** P8392–8396.

10 Pear, W S., Scott, M. L., Baltimore, D., and Nolan, G (1995) Production of high-titer helper-free retroviruses with an amphotropic host range by transient transfection In preparation

11. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

12. Patanjali, S. R., Parimoto, S., and Weissman, S M. (1991) Construction of a uniform abundance (normalized) cDNA library. *Proc Natl Acad Sci USA* **88,** 1943–1947

13. Markowitz, D., Goff, S., and Bank, A (1988) Construction and use of a safe and efficient amphotropic packaging cell line. *Virology* **167,** 400–406