Screening of Recombinant DNA Libraries

Modified after J.G. Seidman (2000) Current Protocols in Molecular Biology.

INTRODUCTION

The usual approach to isolating a recombinant DNA clone containing a particular gene is to screen a recombinant DNA library. A recombinant DNA library consists of a large number of recombinant DNA clones, each one contains a different segment of foreign DNA. Since only a few of the thousands of clones in the library encode the desired nucleic acid sequence, the investigator must devise a procedure for identifying the desired clone. The optimal procedure for isolating the desired clone involves a positive selection for a particular nucleic acid sequence. If the desired gene confers a phenotype that can be selected, then the desired clone can be isolated under selective conditions. However, most genes do not encode a gene with a selectable function. Clones encoding non-selectable functions are identified by: (1) hybridizing to a nucleic acid probe, (2) expressing a segment of protein that can be recognized by an antibody, or (3) amplifying the defined sequence by a particular set of primers.

Screening libraries involves a rapid assay to identify the recombinant DNA clone in the library and then to purify the clone. To test a large number of clones at one time, the library is spread out on agarose plates, then the clones are transferred to filter membranes. The clones can be hybridized to a particular probe. When the desired clone is first identified, it is usually found among many undesirable clones. Thus, to reach homogeneity (purity) in screening involves multiple rounds of screen. Next, a recombinant DNA library must be selected. When choosing which library to screen the investigator should consider whether he or she wants to isolate the cDNA clones encoding the mRNA sequences or genomic clones containing genomic sequences.

Screening a cDNA library. The optimal cDNA library is one made from a particular tissue or cell that expresses the desired mRNA sequence at high levels. In highly differentiated cells, some mRNAs may comprise as many as 1 of 20 of the poly$(A)^+$ mRNA molecules, while other mRNAs are either not present at all or comprise as low as 1 molecule in 100,000 poly$(A)^+$ mRNA molecules. When choosing a cDNA library, the investigator must make every effort to obtain a library from a cell where the mRNA is being expressed in large amounts. Of course, the number of clones that must be screened is determined by the abundance of the mRNA in the cell. The amount of protein that is found in the cell is frequently a good indicator of the abundance of the mRNA. Thus, proteins that comprise 1% of the total cell protein are made by mRNAs that usually comprise 1% of the total poly$(A)^+$ mRNA, and the desired cDNA clones should comprise about 1% of the clones in the cDNA library.

Screening a genomic library. In general, genomic libraries can be made from DNA derived from any tissue, because two copies of the gene are present in all cells. The predicted frequency of any particular sequence would be identical to the number of clones that must be screened to have a good probability of success. This number is a function of the complexity of the genome and the average size of the inserts in the library clones. Usually about 1 million bacteriophage clones must be screened to identify a genomic clone from a mammalian DNA library. Many of the clones that are screened from an amplified library will be screened more than once. In the typical genomic libraries maintained in *E. coli*, the size of the insert is limited to 20 to 25 kb for lambda vectors.

General considerations. If the desired clone is not found in a particular library, another independent library should be screened.

Having selected the library, the investigator is ready to begin screening for the desired clone. Libraries are plated out, transferred to nitrocellulose filters, and hybridized to ^{32}P-labeled or DIG probes. The major problem associated with these techniques is that "false" positives can be identified; the probe may hybridize to clones that do not encode the desired sequence. The investigator will be screening as many as one million clones. If the library contains any contaminating recombinant DNA clones that have been previously grown

in the laboratory, it will be identified in the screening procedure. Thus, extreme care must be exercised to prevent contamination of the library with previously isolated recombinant clones. Despite these problems the ability to screen large DNA libraries to isolate the desired clone provides a powerful tool for molecular biologists.

Hybridization

Modified after William M Strauss (2000) Current Protocols in Molecular Biology

Background Information

All hybridization methods depend upon the ability of denatured DNA to re-anneal when complementary strands are present in an environment near but below their T_m (melting temperature). In a hybridization reaction involving double-stranded DNA on a filter and a single-stranded DNA probe, there are three different annealing reactions occurring. First, there are the desired probe-DNA interactions, which result in signal. Second, there are mismatch interactions that occur between related but nonhomologous sequences; these mismatch hybrids are the ones that must be eliminated during the washing of the filters. Third, are the non-sequence-specific interactions; these result in noise. The ability to extract information from a particular hybridization experiment is a function of the signal-to-noise ratio. High background or poor specific signal can both result in uninterpretable results.

Washing filters is required to remove excess probe, as well as probe that has bound to the DNA on the filter as mismatch hybrids. Temperature and salt concentration dramatically affect the maintenance of specific hybrids. Detergents and other charged species can have a profound effect upon the nonspecific binding of species that contribute to background. Excess probe is rinsed away under low-stringency conditions. When washing is complete, the filters should produce very little “noise.”

Critical Parameters

Hybridization. Kinetically, the hybridization of DNA (or RNA) probes to filter-bound DNA is not significantly different from hybridization in solution. For single-stranded probes, the rate of hybridization follows first-order kinetics, since the probe is available in excess. Under conditions of excess probe, the time for hybridization is inversely proportional to the probe concentration. For double-stranded probes the rate of hybridization displays a more complex relationship to the initial probe concentration. However, to a first approximation, the initial probe concentration is inversely proportional to the rate of hybridization. The normal hybridization time is 16 hrs. In any hybridization, nonspecific interactions will occur and that sources of noise will be present. Thus, too much probe in a hybridization is as bad as too little.

One important source of background hybridization to filters is due to the hybridization of the probe to vector sequences or to *E. coli* DNA. Be certain that there is no vector or *E.coli* DNA sequences in the probe. This can be ensured by isolating the probe from one type of vector (e.g., plasmid) and screening a library made with a different type of vector (e.g., bacteriophage).

Washing temperature. Washing at low stringency is a straightforward proposition. Buffer is added at room temperature and washing proceeds at room temperature.

High-stringency wash is determined empirically. The relative homology between the probe and target sequence is a determining parameter. If the similarity is 100%, a high temperature (65 to 75 ºC) can be used. As the similarity drops, lower washing temperatures must be used. In general one starts at 37 to 40ºC, raising the temperature by 30 to 50 ºC intervals until background is low enough not to be a major factor.

The length of the probe is also important. Very short probes (<100 bp) must be washed at lower temperatures, even if the similarity is 100%. Washing strategy is the same as for probes of differing similarity.

Salt concentration. The lower the salt concentration, the higher the stringency. With this said, the protocols as outlined do not require readjustment of salt concentration for adjustment of stringency. Only the washing temperature is varied.

Probe. The nucleic acid probe must be of high specific activity and greater than 50 bp in length so that it can form stable hybrids.

Time Considerations.

Generally hybridizations are carried on overnight for 12 to 16 hr. This is sufficient for most probes and blots. However, with probes of increasing complexity longer hybridization times are required. This is preferable to increasing the probe concentration from experiment to experiment.

Principle of Nucleic Acid Labeling and Detection with the Genius System

Kessler, C. (1990) BioTechniques 9(6):762-768.

Genius System uses digoxigenin, a steroid hapten, to label RNA and DNA for subsequent dUTP *via* random primed labeling, nick translation, cDNA synthesis, or the Polymerase Chain Reaction (PCR). The DIG-labeled probes are hybridized to a membrane-bound nucleic acid through northern blotting, Southern blotting, dot/slot blotting, or colony plaque lift hybridization procedures. These hybridized probes are immuno-detected with an alkaline phosphatase-conjugated anti-digoxigenin antibody and then visualized with the chemiluminescent substrate Lumi-Phos 530 or with the colorimetric substrates NBT and X-Phosphate.