[56] Lineage Analysis Using Retrovirus Vectors

By Constance L. Cepko, Elizabeth F. Ryder,
Christopher P. Austin, Christopher Walsh,
and Donna M. Fekete

Introduction

The complexity and inaccessibility of the murine embryo has made lineage analysis through direct approaches, such as time-lapse microscopy and injection of tracers, almost impossible. A genetic and clonal solution to lineage mapping in mice is through the use of retrovirus vectors. The basis for this technique is summarized in this chapter, and the strategies and current methods in use in our laboratory are detailed.

Transduction of Genes via Retrovirus Vectors

A retrovirus vector is an infectious virus that transduces a nonviral gene into mitotic cells *in vivo* or *in vitro*.[1] These vectors utilize the same efficient and precise integration machinery of naturally occurring retroviruses to produce a single copy of the viral genome stably integrated into the host chromosome. Vectors that are useful for lineage analysis have been modified so that they are replication-incompetent and thus cannot spread from one infected cell to another. They are, however, faithfully passed on to all daughter cells of the originally infected progenitor cell, making them ideal for lineage analysis.

Retroviruses use RNA as their genome, which is packaged into a membrane-bound protein capsid. They produce a DNA copy of the genome immediately after infection via reverse transcriptase, a product of the viral *pol* gene included in the viral particle. The DNA copy is integrated into the host cell genome and is thereafter referred to as a provirus. Complete synthesis of an integration-competent viral genome requires an S phase, and thus only mitotic cells will serve successfully as hosts for retroviral integration.

Most vectors began as proviruses that were cloned from cells infected with a naturally occurring retrovirus. Although extensive deletions of proviruses were made, vectors retain the cis-acting viral sequences necessary for the viral life cycle. These include the ψ packaging sequence

---

[1] R. Weiss, N. Teich, H. Varmus, and J. Coffin, "RNA Tumor Viruses," Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1984 and 1985.

(necessary for recognition of the viral RNA for encapsidation into the viral particle), reverse transcription signals, integration signals, and viral promoter, enhancer, and polyadenylation sequences. A cDNA can thus be expressed in the vector using the transcription regulatory sequences provided by the virus. Because replication-incompetent retrovirus vectors usually do not encode the structural genes whose products make up the viral particle, these proteins must be supplied through complementation. The structural proteins gag, pol, and env are typically supplied by "packaging" cell lines. These lines are stable mouse fibroblast lines that contain the *gag, pol,* and *env* genes as a result of the introduction of these genes by transfection. However, these lines do not contain the packaging sequence, ψ, on the viral RNA that encodes the structural proteins. Thus, the packaging lines make viral particles that do not contain the genes *gag, pol,* or *env.*

Retrovirus vector particles are essentially identical to naturally occurring retrovirus particles. They enter the host cell via interaction of a viral envelope glycoprotein (a product of the viral *env* gene) with a host cell receptor. The murine viruses have several classes of env glycoprotein which interact with different host cell receptors. The most useful class for lineage analysis of rodents is the ecotropic class. The ecotropic env glycoprotein allows entry only into rat and mouse cells via the ecotropic receptor. It does not allow infection of humans and thus is considered relatively safe for gene transfer experiments. Until 1988, the packaging line most commonly in use was the $\psi2$ line.[2] It encodes the ecotropic *env* gene and, in our experience, makes the highest titers of vectors, relative to other packaging lines (for unknown reasons). However, it can also lead to the production of helper virus (discussed below). Two newer ecotropic packaging lines, ΨCRE[3] and GP + E-86,[4] have not been reported to lead to production of helper virus to date. A third such "helper-free" packaging line, ΩE,[5] with some improvement in design over the others, has been produced and will probably prove to be very useful. Any of the ecotropic packaging lines can be used to produce vector stocks for lineage analysis. Regardless of the packaging line used, however, all stocks should be assayed for the presence of helper virus.

Production of Virus Stocks for Lineage Analysis

Replication-incompetent vectors that encode a histochemical reporter gene, such as *Escherichia coli lacZ,* are the most useful for lineage studies

[2] R. Mann, R. C. Mulligan, and D. Baltimore, *Cell (Cambridge, Mass.)* **33,** 153 (1983).
[3] O. Danos and R. C. Mulligan, *Proc. Natl. Acad. Sci. U.S.A.* **85,** 6460 (1988).
[4] D. Markowitz, S. Goff, and A. Bank, *J. Virol.* **62,** 1120 (1988).
[5] J. P. Morgenstern and H. Land, *Nucleic Acids Res.* **18,** 3587 (1990).

as they allow analysis of individual cells in tissue sections or whole mounts. Stocks of such vectors are typically produced by packaging lines stably transduced with the vector genome. (A detailed description of protocols for making such "producer" lines, titering and concentrating virus stocks, and checking for helper virus contamination has been published[6] and is not given here.)

It is best to obtain lines that make high-titered stocks of lineage vectors from the laboratories that have created them. We have placed $\Psi 2$ and ΨCRE producers of BAG,[7] a *lacZ* virus that we have used for lineage analysis, on deposit at the ATCC (Rockville, MD). They can be obtained by anyone and are listed as ATCC CRL Nos. 1858 (ψCRE BAG) and 9560 ($\psi 2$ BAG). Similarly, $\psi 2$ producers of DAP,[8] a vector encoding human placental alkaline phosphatase (PLAP, described further below), is available as CRL No. 1949. Both of these vectors transcribe the reporter gene from the viral long terminal repeat (LTR) promoter and are generally useful for expression of the reporter gene in most tissues. We compared the expression of *lacZ* driven by several different promoters[9,10] and found that the LTR was generally the most reliable and non-cell type-specific. This is an important consideration as it is desirable to identify all of the cells descended from an infected progenitor, and thus a constitutive promoter is the most useful for lineage studies. However, even with constitutive promoters, it has been noted that some infected cells do not express detectable β-galactoside (β-Gal) protein, even among clones of fibroblasts infected *in vitro*. Thus, it is important to restrict conclusions about lineage relationships to cells that are marked and not to make assumptions about their relationships to cells that are unmarked.

For lineage applications it is usually necessary to concentrate virus in order to achieve sufficient titer. This is typically due to a limitation in the volume that can be injected at any one site. As viruses are macromolecular structures, they can be concentrated fairly easily by a relatively short centrifugation step. Virions also can be precipitated using polyethylene glycol or ammonium sulfate, and the resulting precipitate collected by centrifugation. Finally, the viral supernatant can be concentrated by centrifugation through a filter that allows only small molecules to pass [e.g.,

[6] C. L. Cepko, *in* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl, "Current Protocols in Molecular Biology," Supplement 17 (1992). Greene, Publishing New York, 1989.

[7] J. Price, D. Turner, and C. L. Cepko, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 154 (1987).

[8] S. C. Fields-Berry, A. L. Halliday, and C. L. Cepko, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 693 (1992).

[9] D. L. Turner, E. Y. Snyder, and C. L. Cepko, *Neuron* **4**, 833 (1990).

[10] C. L. Cepko, C. P. Austin, C. Walsh, E. F. Ryder, A. Halliday, and S. Fields-Berry, *Cold Spring Harbor Symp. Quant. Biol.* **55**, 265 (1990).

Centricon (Amicon, Danvers, MA) filters]. Regardless of the protocols that are used, one must keep in mind that retroviral particles are fragile, with short half-lives even under optimum conditions. To prepare the highest titered stock for multiple experiments, we usually concentrate several hundred milliliters of producer cell supernatant. Concentrated stocks are titered and tested for helper virus contamination. They can be stored indefinitely at $-80°$ in small (10–50 μl) aliquots, although we have noticed reduction in titer upon freeze–thaw for some stocks.

Replication-Competent Helper Virus

Replication-competent virus is sometimes referred to as helper virus as it can complement ("help") a replication-incompetent virus and thus allow it to spread from cell to cell. It can be present in an animal through exogenous infection (e.g., from a viremic animal in the mouse colony), expression of an endogenous retroviral genome (e.g., the *akv* loci in AKR mice), or recombination events between two viral RNAs encapsidated in retroviral virions produced by packaging lines. The presence of helper virus is an issue of concern when using replication-incompetent viruses for lineage analysis as it can lead to horizontal spread of the marker virus, creating false lineage relationships.

The most likely source of helper virus is the viral stock used for lineage analysis. The genome(s) that supplies the *gag, pol,* and *env* genes in packaging lines does not encode the Ψ sequence, but it can still become packaged, although at a low frequency. If it is coencapsidated with a vector genome, recombination in the next cycle of reverse transcription can occur. If the recombination allows the Ψ^- genome to acquire the Ψ sequence from the vector genome, a recombinant that is capable of autonomous replication is the result. This recombinant can spread through the entire culture (although slowly, owing to envelope interference). Once this occurs, it is best to discard the producer clone as there is no convenient way to eliminate the helper virus. As would be expected, recombination giving rise to helper virus occurs with greater frequency in stocks with high titer, and with vectors that have retained more of the wild-type sequences (i.e., the more homology between the vector and packaging genomes, the more opportunity there is for recombination). Note that the helper genome itself will not encode a histochemical marker gene, as apparently there is no room, or flexibility, within murine viruses that allows them to be both replication-competent and capable of expressing another gene like *lacZ*. The way that spread would occur is by a cell being infected with both the *lacZ* virus and a helper virus. Such a doubly infected cell would then produce both viruses.

When performing lineage analysis, there are several signs that can indicate the presence of helper virus within an individual animal. If one allows an animal to survive for long periods of time after inoculation, particularly if embryos or neonates are infected, the animal is likely to acquire a tumor when helper virus is present. Most naturally occurring replication-competent viruses are leukemogenic, with the disease spectrum being at least in part a property of the viral LTR.

If one analyzes either shortly or long after inoculation, the clone size, clone number, and spectrum of labeled cells may be indicative of helper virus. For example, the eye of a newborn rat or mouse has mitotic progenitors for retinal neurons, as well as mitotic progenitors for astrocytes and endothelial cells. By targeting the infection to the area of progenitors for retinal neurons, we only rarely see infection of a few blood vessels or astrocytes, as their progenitors are outside of the immediate area that is inoculated and they become infected only by leakage of the viral inoculum from the targeted area. However, if helper virus were present, we would see infection of a high percentage of astrocytes, blood vessels, and, eventually, other eye tissues since virus spread would eventually lead to infection of cells outside of the targeted area. One would expect to see a correlation between the percentage of such nontargeted cells that are infected and the degree to which their progenitors are mitotically active after inoculation, owing to the fact that infection requires a mitotic target cell. If one were to examine tissues other than ocular tissues, one would similarly see evidence of virus spread to cells whose progenitors would be mitotically active during the period of virus spread. In addition, the size and number of "clones" may also appear to be too large for true "clonal" events if helper virus were present. This interpretation of course relies on some knowledge of the area under study. Finally, if one performs the infections with two different histochemical marker viruses, one can also assess the likelihood of helper virus problems, as discussed below.

Determination of Sibling Relationships

When performing lineage analysis, it is critical to unambiguously define cells as descendents of the same progenitor. This can be relatively straightforward when sibling cells remain rather tightly, and reproducibly, grouped. An example of such a straightforward case is the rodent retina, where the descendents of a single progenitor migrate to form a coherent radial array.[9,11] In such a system, lineage analysis can be performed using one or two distinct histochemical marker viruses, as described below.

[11] D. L. Turner and C. L. Cepko, Nature (London) 328, 131 (1987).

For more complex systems, particularly where cell migration is important, many more markers are needed, as described in a later section (see Clonal Analysis Using Polymerase Chain Reaction/Library Method).

If only a single histochemical marker virus is used, one can perform a standard virological titration, in which a particular viral inoculum is serially diluted and applied to tissue. In the retina, the number of radial arrays, their average size, and their cellular composition were analyzed in a series of animals infected with dilutions that covered a 4-log range. The number of arrays was found to have a first-order relationship to the dose of virus, whereas the size and composition of the radial arrays were independent of the viral dose. Such results indicate that the working definition of a clone, in this case a radial array, fulfilled the statistical criteria expected of a single hit event.

There are several difficulties involved with the use of a single marker virus. First, there must be a wide range of dilutions that can be injected to give countable numbers of events, which is required for determination of a first-order relationship between clone number and viral dose. This type of analysis also relies critically on controlling the exact volume of the injection. These problems may be avoided in cases where it is possible to perform the analysis solely on infections with very small amounts of virus, although generating large amounts of data under these conditions is tedious. Another potential problem with the dilution approach is that aggregates may form during virus concentration which are not separated by dilution. Thus, even at low dilution, a "clone" may be the result of infection of adjacent cells by two or more members of such an aggregate. In addition, it is difficult to calculate an error rate for the assignment of clonal boundaries, which may make interpretation difficult, particularly of events which occur rarely in the data set.

For these reasons, additional viral vectors with histochemically distinguishable marker genes have been developed for use in mixed infections. In addition to addressing the problems mentioned above, use of two marker viruses can often provide a much better initial idea of what clonal boundaries are likely to be, by simple visual inspection of the tissue. By using two markers, one can visualize errors arising from formation of viral aggregates, since some of these aggregates will include two distinguishable virions. Various criteria for clonal boundaries can be tested, and error rates for each can be computed. Extensive dilutions are not needed.

Two viruses that have been used for this approach encode cytoplasmically localized versus nuclear localized β-Gal. This can work when the cytoplasmically localized β-Gal is easily distinguished from the nuclear

localized β-Gal.[12,13] We have found that this is not the case in rodent nervous system cells, as the cytoplasmically localized β-Gal quite often is restricted to neuronal cell bodies and is therefore difficult to distinguish from nuclear-localized β-Gal. To overcome this problem, we created the above-mentioned DAP virus,[8] which is distinctive from the lacZ-encoding BAG virus.

To perform the two-marker analysis in the rodent retina, a stock containing BAG and DAP was made by growing a Ψ2 producer clone for BAG and a Ψ2 producer clone for DAP on the same dish. The resulting supernatant was concentrated and used to infect rodent retina. The tissue was then analyzed histochemically for the presence of blue (owing to BAG infection) and purple (owing to DAP infection) radial arrays. If radial arrays were truly clonal, then each one should be only one color. Analysis of 1100 arrays indicated that most were clonal. However, 5 comprised blue cells and purple cells, presumably from infection of adjacent progenitor cells with BAG and DAP. (Infection of one cell with both BAG and DAP would not be a problem, since the resulting cells would be clonally related.) The 5/1100 figure will be an underestimate of the true frequency of incorrect assignment of clonal boundaries, as sometimes two BAG or two DAP virions will infect adjacent progenitor cells; the resulting arrays will be a single color, but not clonal. A closer approximation of the true frequency can be obtained by using the following formula (for derivation see Ref. 8):

$$\frac{(\# \text{ bicolored arrays})[(a + b)^2/2ab]}{\# \text{ total arrays}} = \% \text{ errors}$$

where a and b are the relative titers in the virus stock. The relative titer of BAG and DAP used in the coinfection was 3 : 1, and thus the value for percent errors in clonal assignments was 1.2%.

Such a low error rate shows that the choice of clonal boundary was reasonable, and that viral aggregation during concentration (or at any other step) was not a problem for this viral stock. To take an extreme counterexample, if every array were the result of infections by two particles, 6/16 of the arrays would be composed of both blue cells and purple cells. (For a stock with relative viral titers of 1 : 1, half the arrays would have blue cells and purple cells). In analyzing the composition of arrays (e.g., by cell type), one need not worry, as one would if using one marker,

[12] S. M. Hughes and H. M. Blau, Nature (London) 345, 350 (1990).
[13] D. S. Galileo, G. E. Gray, G. C. Owens, J. Majors, and J. R. Sanes, Proc. Natl. Acad. Sci. U.S.A. 87, 458 (1990).

that rare cell type combinations may be due to occasional errors in assignment of clonal boundaries. If such rare arrays are always of one color (when the error rate is very low), then they are very likely clones.

The error rate being computed here is the rate of "lumping" errors, that is, the frequency with which the criteria defining a clone lumps together daughters of more than one progenitor. However, this computation does not allow assessment of "splitting" errors, where clones that are more spread out (presumably owing to migration) are split by the definition of a clone into two or more subclones. If the criteria used to define clonal relationships are found to generate too many lumping errors, a more restrictive definition can be made, and then this definition can be tested for percent lumping errors. This process can be done iteratively until the error rate reaches a level that is acceptable relative to the point that is being tested. However, it should be kept in mind that the more restrictive a definition is for clonal relationships, the more it is prone to generate splitting errors. Splitting errors may become obvious in animals injected with very dilute viral stocks; for example, one blue array in a large, otherwise unlabeled area is probably a clone, but it may be divided into several by too restrictive clonal definitions. Thus, combining two-marker analysis with a few dilution experiments may be useful to help balance splitting and lumping errors. If it is necessary to avoid completely both splitting and lumping errors, a much greater number of vectors and more tedious detection method must be used, as detailed below (see Clonal Analysis Using Polymerase Chain Reaction/Library Method).

The value of 1.2% for lumping errors in assignment of clonal boundaries places an upper limit of 1.2% as the frequency of aggregation for this viral stock, since this figure will include errors due to both aggregation and independent virions infecting adjacent progenitors. The presence of helper virus would probably increase the error rate, as it would have the effect of enlarging apparent clone size and creating overlap between clones of different markers. The percentage of errors on injecting this stock in other areas of an animal will depend on the particular circumstances of the injection site and on the multiplicity of infection (MOI, the ratio of infectious virions to target cells). Most of the time the MOI will be quite low (e.g., in the retina it was approximately 0.01 at the highest concentration of virus injected). Concerning the injection site, injection into a lumen, such as the lateral ventricles, should not promote aggregation nor a high local MOI, but injection into solid tissue, in which the majority of the inoculum has access to a limited number of cells at the inoculation site, could present problems. By coinjecting BAG and DAP, one can monitor the frequency of these events and thus determine if clonal analysis is feasible.

The above analysis was performed using viruses that were produced on the same dish and concentrated together. This was done because we felt that the most likely way that two adjacent progenitors might become infected would be through small aggregates of virions. We grew both virions together on the same dish in order for the assay to be sensitive to any aggregation that might occur prior to concentration, but this is probably not necessary in most cases. Aggregation most likely occurs during the concentration step, as one often can see macroscopic aggregates after resuspending pellets of virions. Thus, when the two-marker approach is used to analyze clonal relationships, it is best to coconcentrate the two vectors together in order for the assay to be sensitive to aggregation arising from this aspect of the procedure. [Although aggregation of virions may frequently occur during concentration, it apparently does not frequently lead to problems in lineage analysis, presumably because of the high ratio of noninfectious particles to infectious particles found in most retrovirus stocks. It is estimated that only 0.1–1.0% of the particles will generate a successful infection. Moreover, most aggregates are probably not efficient as infectious units; it must be difficult for the rare infectious particle(s) within such a clump to gain access to the viral receptors on a target cell.]

To determine the ratio of two genomes present in a mixed virus stock (e.g., BAG plus DAP), there are several methods that can be used. The first two methods are performed *in vitro,* and are simply an extension of a titration assay. Any virus stock is normally titered on NIH 3T3 cells to determine the amount of virus to inject. The infected NIH 3T3 cells are then either selected for the expression of a selectable marker when the virus encodes such a gene (e.g., *neo* in BAG and DAP), or are stained directly, histochemically, for β-Gal or PLAP activity without prior selection with drugs. If no selection is used, the relative ratio of the two markers can be scored directly by evaluating the number of clones of each color on a dish. Alternatively, selected G418-resistant colonies can be stained histochemically for both enzyme activities and the relative ratio of blue versus purple G418-resistant colonies computed. A third method of evaluating the ratio of the two genomes is to use the values observed from *in vivo* infections. After animals are infected and processed for both histochemical stains, the ratio of the two genomes can be compared by counting the number of clones, or infected cells, of each color.

When all the above methods were applied to lineage analysis in mouse retina[8] and rat striatum,[14] the value obtained for the ratio of G418-resistant colonies scored histochemically was almost identical to the ratio observed *in vivo*. Directly scoring histochemically stained, non-G418-selected NIH

[14] A. Halliday and C. L. Cepko, *Neuron* **9**, 15 (1992).

3T3 cells led to an underestimate of the number of BAG-infected colonies, presumably because such cells often are only faintly blue, whereas DAP-infected cells are usually an intense purple. *In vivo,* this is not generally the case, as BAG-infected cells are usually deep blue.

Regardless of which method is used to score sibling relationships, one further recommendation to aid in the assignments is to choose an injection site that will allow the inoculum to spread. If one injects into a packed tissue, the viral inoculum will most likely infect cells within the injection tract, and it will be very difficult to sort out sibling relationships (i.e., too many lumping errors). In addition, one must inject such that the virus has clear access to the target population; the virus will blind to cells at the injection site and will not gain access to cells that are not directly adjoining that site.

The procedures described below are those that we have used for infection of rodents, histochemical processing of tissue for β-Gal and PLAP visualization, and preparation and use of a library for the polymerase chain reaction (PCR) method.[15]

Infection of Rodents

Injection of Virus in Utero

The following protocols may be used with rats or mice. Note that clean, but not aseptic, technique is used throughout. We routinely soak instruments in 70% (v/v) ethanol before operations, use the sterile materials noted, and include penicillin/streptomycin (final concentrations of 100 units/ml each) in the lavage solution. We have not had difficulty with infection using these precautions.

Materials

Ketamine hydrochloride injection (100 mg/ml ketamine)
Xylazine injection (20 mg/ml)
Animal support platform
Depilatory
Scalpel and disposable sterile blades
Cotton swabs and balls, sterile
Tissue retractors
Tissue scissors
Lactated Ringer's solution (LR) containing penicillin/streptomycin
Fiber optic light source
Virus stock

[15] C. Walsh and C. L. Cepko, *Science* **255**, 434 (1992).

Automated microinjector
Micropipettes, 1–5 μl
3-0 Dexon suture
Tissue stapler

1. Mix ketamine and xylazine 1 : 1 in a 1-ml syringe equipped with a 27-gauge needle; lift the tail and hindquarters of the animal with one hand and with the other inject 0.05 ml (mice) or 0.18 ml (rats) of anesthetic mixture intraperitoneally. One or more additional doses of ketamine alone (0.05 ml for mice and 0.10 ml for rats) is usually required to induce or maintain anesthesia, particularly if the procedure takes over 1 hr. Respiratory arrest and spontaneous abortion appear to occur more often if a larger dose of the mixture is given initially, or if any additional doses of xylazine are given.

2. Remove the hair over the entire abdomen using depilatory agent (any commercially available formulation, such as Nair hair remover, works well); shaving of remaining hair with a razor may be necessary. Wash the skin several times with water, then with 70% ethanol, and allow to dry.

3. Place the animal on its back in the support apparatus. For this purpose, we find that a slab of Styrofoam with two additional slabs glued on top to create a trough works well. With the trough appropriately narrow, no additional restraint is needed to hold the anesthetized animal.

4. Make a midline incision in the skin from xyphoid process to pubis using a scalpel, and retract; attaching retractors firmly to the Styrofoam support will create a stable working field. Stop any bleeding with cotton swabs before carefully retracting the fascia and peritoneum and incising them in the midline with scissors (care is required here not to incise the underlying bowel). Continue the incision cephalad along the midline of the fascia (where there are few blood vessels) to expose the entire abdominal contents. If necessary to expose the uterus, gently pack the abdomen with cotton balls or swabs to remove the intestines from the operative field, being careful not to lacerate or obstruct the bowel. Fill the peritoneal cavity with LR, and lavage until clear if the solution turns at all turbid.

Wide exposure is important to allow the later manipulations. During the remainder of the operation, keep the peritoneal cavity moist and free of blood; dehydration or blood around the uterus increases the rate of postoperative abortion.

5. Elevate the embryos one at a time out of the peritoneal cavity, and transilluminate with a fiber optic light source to visualize the structure to be injected. For lateral cerebral ventricular injections, for example, the cerebral venous sinuses serve as landmarks. When deciding on a structure to inject, keep in mind that free diffusion of virus solution through a fluid-

filled structure lined with mitotic cells is best for ensuring even distribution of viral infection events throughout the tissue being labeled. The neural tube is an example of such a structure; when virus is injected into one lateral ventricle, it is observed to quickly diffuse throughout the entire ventricular system.

6. Using a heat-drawn glass micropipette attached to an automatic microinjector, penetrate the uterine wall, extraembryonic membranes, and the structure to be infected in one rapid thrust; this minimizes trauma and improves survival. Once the pipette is in place, inject the desired volume of virus solution, usually $0.1-1.0$ μl. Coinjection of a dye such as 0.005% (w/v) trypan blue or 0.025% fast green aids determination of the accuracy of injection and does not appear to impair viral infectivity; coinjection of the polycation Polybrene (80 μg/ml) aids in viral attachment to the cells to be infected.

The type of instrument used to deliver the virus depends on the age of the animal and the tissue to be injected. At early embryonic stages, the small size and easy penetrability of the tissue makes a pneumatic microinjector (such as the Eppendorf 5242) best for delivering a constant amount of virus at a controlled rate with a minimum of trauma. Glass micropipettes should be made empirically to produce a bore size that will allow penetration of the uterine wall and the tissue to be infected. At later ages (late embryonic and postnatal), a Hamilton syringe with a 33-gauge needle works best.

When injecting through the uterine wall, all embryos may potentially be injected except those most proximal to the cervix on each side (injection of these greatly increases the rate of postoperative abortion). In practice, it is often not advisable to inject all possible embryos, if excessive uterine manipulation would be required. At the earliest stages at which this technique is feasible [embryonic day (E) 12 in the mouse or E13 in the rat], virtually any uterine manipulation may cause abortion, so any embryo that cannot be reached easily should not be injected.

7. Once all animals have been injected, lavage the peritoneal cavity until it is clear of all blood and clots, ensure that all cotton balls and swabs have been removed, and move retractors from the abdominal wall/fascia to the skin. Filling the peritoneal cavity with LR with penicillin/streptomycin before closing increases survival significantly, probably by preventing maternal dehydration during recovery from anesthesia as well as preventing infection.

8. Using 3-0 Dexon or silk suture material on a curved needle, suture the peritoneum, abdominal musculature, and fascia from each side together, using a continuous locking stitch. After closing the fascia, again lavage using LR with penicillin/streptomycin.

9. Close the skin using surgical staples (such as the Clay-Adams Autoclip) placed 0.5 cm apart. Sutures may also be used, but these require much more time (often necessitating further anesthesia, which increases abortion risk) and are frequently chewed off by the animal, resulting in evisceration.

10. Place the animal on its back in the cage and allow the anesthesia to wear off. Ideally, the animal will wake up within 1 hr of the end of the operation. Increasing time to awakening results in increasing abortion frequency. Food and water on the floor of the cage should be provided for the immediate postoperative period.

11. Mothers may be allowed to deliver progeny vaginally, or the offspring may be harvested by cesarean section. Maternal and fetal survival are approximately 60% at early embryonic ages of injection and increase with gestational age to virtually 100% after postnatal injections.

Injection of Virus Using exo Utero Surgery

Injections into small or delicate structures (such as the eye) require micropipettes that are too fine to penetrate the uterine wall. In addition, it is impossible to target precisely many structures through the rather opaque uterine wall. These problems can be circumvented, though with a considerable increase in technical difficulty and decrease in survival, by use of the *exo utero* technique.[16] The procedure is similar to that detailed above, with the following modifications to free the embryos from the uterine cavity.

1. The technique works well in our hands only with outbred, virus antigen-free CD-1 and Swiss-Webster mice, but even these strains may have different embryo survival rates when obtained from different suppliers or different colonies of the same supplier. This variability presumably results from subclinical infections, which may render some animals unable to survive the stress of the operation. We have had no success with this technique in rats.

2. After the uterus is exposed and *before* filling the peritoneum with LR, incise the uterus longitudinally along its ventral aspect with sharp microscissors. The uterine muscle will contract away from the embryos, causing them to be fully exposed, surrounded by the extraembryonic membranes.

3. Only two embryos in each uterine horn can be safely injected, apparently because of trauma induced by neighboring embryos touching

---

[16] K. Muneoka, N. Wanek, and S. V. Bryant, *J. Exp. Zool.* **239**, 289 (1986).

each other. Thus, all other embryos must be removed. Using a dry sterile cotton swab, scoop out each embryo to be removed, with its placenta and extraembryonic membranes, and press firmly against the uterine wall where the placenta had been attached for 30–40 sec to achieve hemostasis. It is important to stop all bleeding before proceeding. From this point on, the embryo must be handled extremely gently, as only the placenta is tethering the embryo to the uterus, and it detaches easily.

4. Fill the peritoneal cavity with LR, and cushion each embryo to be injected with sterile cotton swabs soaked in LR. Keeping the embryos submerged throughout the remainder of the procedure is essential for survival.

5. The injection should then be done with a pneumatic microinjector and heat-pulled glass micropipette. This may usually be performed by puncturing the extraembryonic membranes first and then the structure to be injected; for some very delicate injections if may be necessary to make an incision in the extraembryonic membranes, which is then closed with 10-0 nylon suture after the injection.

6. At the time of desired fetal harvest, or at the latest early on the last day of gestation, sacrifice the mother by cervical dislocation, rapidly incise the abdomen, and deliver the fetuses. If survival to a postnatal time point is desired, it is necessary to foster the pups with another lactating female. This is best done with a mouse that has delivered at the same time as the experimental animal, but we have successfully fostered pups with mothers that delivered several days to a week previously. Attempts to reanesthetize the experimental mother for delivery, thus allowing her to survive and obviating the need for fostering, have been unsuccessful owing to poor survival of the pups. This is probably due to the deleterious effect of the anesthesia on the pups, as well as poor milk production by the mother after multiple operations.

Human Placental Alkaline Phosphatase as Histochemical Marker Gene

Human placental alkaline phosphatase (PLAP) has only recently been adapted for use as a histochemical marker for lineage studies. As such, neither its benefits nor its potential drawbacks have been exhaustively analyzed, and this should be kept in mind as the enzyme enjoys wider use. Our laboratory has undertaken a systematic lineage analysis in the postnatal mouse retina[8] and rat striatum[14] comparing β-Gal and PLAP, and little difference between the two has been observed. Thus, for these particular combinations of tissue and vectors (BAG and DAP), there appears to be no effect of ectopic PLAP expression on the choice of cell fate during development. We do not yet know whether this conclusion

will hold for other tissues. In the chick retina and cerebellum, for example, clonal analysis does show differences between β-Gal and PLAP in the ease of detection of different cell types (C. L. Cepko, D. Fekete, and E. Ryder, unpublished observations). It is not yet clear whether these differences reflect cell-specific inactivation of the promotor or the enzymes, differential distribution of the two enzymes within intracellular compartments, or (most worrisome) the perturbation of cell fate.

Human PLAP was initially chosen as a potential histochemical marker for several reasons. Among the variety of isoenzymes of alkaline phosphatase that have been studied, human PLAP is by far the most heat stable,[17] by a factor of about 100. Thus, although many tissues express endogenous alkaline phosphatase(s), it is possible in most cases to greatly reduce this background reactivity by preincubating the tissue at 65° for 30 min. In addition, PLAP is resistant to a variety of substances that act to inhibit other isoenzymes of alkaline phosphatase.[18] We have tested the following inhibitors on mouse and chick neural tissue as recommended by Zoellner and Hunter[18]: 0.5 mM levamisole [L(−)-2,3,5,6-tetrahydro-6-phenylimidazo[2,1-b]thiazole], 2 mM mercuric chloride, 5 mM L-leucylglycylglycine, 1 mM EDTA, 1 mM L-phenylalanineglycylglycine, 0.2 M lysine hydrochloride, and 0.3 mM sodium arsenate. Levamisole was the most useful in reducing the background staining in brains, although it also reduced PLAP staining slightly in some cases; it was less effective in retinas. None of the inhibitors was nearly as effective as heat treatment in reducing background in the central nervous system. Nonetheless, it is certainly possible that their use may facilitate staining of other tissues.

Another benefit of PLAP as a marker is the fact that its activity is probably minimal at normal intracellular pH, since this is considerably below the optimal pH for enzyme function. This may suggest that ectopic expression of PLAP during development is not likely to perturb normal physiological processes, but this remains to be tested systematically for different tissues.

Double Staining of Infected Tissues for β-Galactosidase and Alkaline Phosphatase Activities

The following protocol was adopted for the double staining of β-Gal and PLAP in nervous system tissue. Cells expressing β-Gal will be rendered a bright blue, whereas cells expressing PLAP will be rendered purple owing to the presence of nitro blue tetrazolium (NBT) in the reaction. The order

[17] R. B. McComb, G. N. Bowers, Jr., and S. Posen, "Alkaline Phosphatase." Plenum, New York, 1979.
[18] H. F. A. Zoellner and N. Hunter, *J. Histochem. Cytochem.* **37**, 1893 (1989).

in which the staining is done is critical, since β-Gal is inactivated by the heat treatment that is required to inhibit endogenous alkaline phosphatases. Obviously, if only a single marker is needed, the protocol can be minimally adapted for use with either enzyme individually.

Detecting both enzymes in the same cell is not always possible using the following protocols, because the PLAP reaction product is so intense that it usually obscures the β-Gal reaction product. For example, it is difficult to detect cells cotransfected with both genes *in vitro*. This is not a problem when the two enzymatic markers are used in retroviral vectors for lineage mapping studies, because individual cells are unlikely to be infected with both vectors and express both enzymes. (Even if they were, the interpretation of the results would not be affected, since one is still observing a clone.) If detection of cells containing both enzymes is necessary, use of different chromogenic substrates might circumvent this difficulty. PLAP can be reacted with naphthol-AS-BI-phosphate/New Fuchsin (Sigma, St. Louis, MO), which produces a red stain,[19] and NBT can be added to the β-Gal reaction to produce a purple precipitate. Red/blue or red/purple might be distinguishable from the corresponding single stains; however, we have not tested these possible combinations.

We have found that performing immunohistochemistry in conjunction with either β-Gal or PLAP histochemical staining is very difficult. The colored precipitates block fluorescence, thus preventing the use of fluorescently conjugated antibodies. Using horseradish peroxidase (HRP)-conjugated antibodies, it is difficult to distinguish blue/brown double-labeled cells from brown labeled ones, whereas purple and brown/purple cells would probably look identical. Situations in which the subcellular localization of β-Gal and an antigen are significantly nonoverlapping (e.g., a nuclear antigen) may allow for detection of both.[20] In addition, one can try to limit the intensity of the β-Gal or PLAP reaction product by shortening the time of the reactions to allow for more sensitivity in detecting an antigen within the same cell by HRP-conjugated antibodies.[21] Alternatively, antibodies exist for both β-Gal (rabbit polyclonal, Cappel, Malvern, PA; mouse monoclonal, Boehringer-Mannheim, Indianapolis, IN) and PLAP (rabbit polyclonal, Zymed, San Francisco, CA; mouse monoclonal, Medix), although using such reagents is obviously more time consuming than relying on histochemical reactions.

Different species, tissues, or parts of tissues (i.e., brain regions) can

[19] E. Harlow and D. Lane (eds.), "Antibodies: A Laboratory Manual." Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1988.
[20] E. Y. Snyder, D. L. Deitcher, C. Walsh, S. Arnold-Aldea, E. A. Hartwieg, and C. L. Cepko, *Cell (Cambridge, Mass.)* **68**, 33.
[21] P. J.-J. Vaysse and J. E. Goldman, *Neuron* **5**, 227 (1990).

have varying amounts of background labeling.[22–24] This is especially noticeable with alkaline phosphatase reactions, but is also true for β-Gal. We recommend always including negative controls in order to assess the extent of background for a particular tissue or region of interest. Important variables that affect the signal-to-noise ratio include the following: type of fixative, length of fixation, length of washes, length of heat treatment, prolonged exposure to light, and prolonged storage after staining. For a particularly problematic tissue, it may prove advantageous to try different inhibitors of endogenous alkaline phosphatases (see above), as well as different substrates.

Solutions

Phosphate-buffered saline (PBS) (10×): 80 g NaCl, 2 g KCl, 11.5 g Na_2HPO_4, and 2 g KH_2PO_4 in 1 liter water; adjust to pH 7.2–7.4 and dilute 1 : 10 before using

0.5% Glutaraldehyde: 25% stock (Sigma, St. Louis, MO) can be stored at $-20°$ and frozen/thawed many times; make dilution immediately before use

4% Paraformaldehyde: 4 g solid paraformaldehyde, 2 mM $MgCl_2$, and 1.25 mM EGTA (0.25 ml of a 0.5 M EGTA stock, pH 8.0) in 100 ml PBS, pH 7.2–7.4

Heat ~80 ml water to 60° and add paraformaldehyde; add NaOH to get paraformaldehyde in solution. Cool to room temperature, add 10 ml of 10× PBS, adjust the pH with HCl, add $MgCl_2$ and EGTA, and make up to 100 ml with H_2O. The solution can be stored at 4° for several weeks

X-Gal detection buffer: 35 mM potassium ferrocyanide (can vary from 5 to 35 mM), 35 mM potassium ferricyanide (can vary from 5 to 35 mM), 2 mM $MgCl_2$, 0.02% Nonidet P-40 (NP-40) (diluted from 10% stock solution), and 0.01% sodium deoxycholate (diluted from 10% solution) in PBS; can be stored for at least 1 year at room temperature in a foil-covered container

X-Gal Stock (40×): 40 mg/ml X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) in dimethylformamide; store at $-20°$ in a glass container covered with foil

X-Gal reaction mix: Make a 1 : 40 dilution of X-Gal stock into X-Gal detection buffer immediately before using; the final concentration of X-Gal is 1 mg/ml

[22] B. Pearson, P. L. Wolf, and J. Vazquez, *Lab. Invest.* **12,** 1249 (1963).

[23] E. Robins, H. E. Hirsch, and S. S. Emmons, *J. Biol. Chem.* **213,** 4246 (1968).

[24] A. M. Rutenburg, S. H. Rutenburg, B. Monis, R. Teague, and A. M. Seligman, *J. Histochem. Cytochem.* **6,** 122 (1958).

X-P detection buffer (Buffer 3, Genius Kit, Boehringer-Mannheim): 100 mM Tris-HCl, pH 9.5, 100 mM NaCl, and 50 mM MgCl$_2$. Store at room temperature. Salts tend to precipitate over several weeks, but this does not seem to markedly affect the staining

X-P stock (100×): 10 mg/ml 5-bromo-4-chloro-3-indolylphosphate (X-P; also referred to as BCIP) in water. Store in the dark as aliquots at −20°; can also be frozen and thawed several times

NBT stock (50×): 50 mg/ml nitro blue tetrazolium in 70% dimethyl-formamide, 30% water. Store at −20° in a glass container covered with foil (the stock does not freeze at this temperature)

X-P reaction mix: 50 μl X-P stock (100×) 100 μl NBT stock (50×), and 50 μl of 50 mM levamisole (100×) (if desired) in 5 ml X-P detection buffer. Make fresh immediately before using. Final concentrations are as follows: X-P, 0.1 mg/ml; NBT, 1 mg/ml; and levamisole, 0.5 mM. Levamisole was ineffective at reducing background in chick cerebellum, while slightly inhibiting PLAP, and was therefore not used in this tissue. Any other desired inhibitor (see above) would also be added to the X-P reaction mix.

Gelatin subbing solution for slides: 2 g gelatin and 0.1 g chromium potassium sulfate (chrome alum) in 200 ml water. Heat the water to 60°, dissolve the chrome alum, and then gradually dissolve gelatin. Filter before use. The percentage of gelatin can be increased or decreased. Load slides in racks, dip quickly, and air-dry overnight

Gelatin/sucrose embedding medium: 7.5% gelatin (porcine skin, Sigma), 15% sucrose, and 0.05% sodium azide in 1 × PBS. Dissolve components gradually at 60°, with stirring. The medium solidifies at room temperature to a transparent gel. Store at room temperature. Liquify in a microwave oven with frequent swirling before embedding samples

Gelvatol mounting medium: Make up according to the protocol given by Rodriguez and Deinhardt.[25] Instead of Elvanol, we use Vinol grade 205 (polyvinyl alcohol) (Air Products and Chemicals, Inc., Allentown, PA).

Whole-Mount Staining Procedure

The primary protocol shown was worked out for staining of intact mouse retinas that had been coinfected with BAG and DAP.[8] The incubation times for X-Gal and X-P reactions are variables that may require adjustment for different tissues. Longer times may be necessary for large

---

[25] J. Rodriguez and F. Deinhardt, *Virology* **12**, 316 (1960).

or dense chunks of tissue, although there is a trade off as the background staining of either enzyme intensifies with increasing reaction time. Whole chick embryos (E7) and chick brains (E10) were found to be incompletely reacted in their centers after 4-hr incubation periods. For tissues that are difficult to stain completely as whole mounts, one can stain as sections (see below). Alternatively, one can do a 3- to 4-hr whole-mount stain to locate cells of interest, dissect out and section only those areas of interest, and restain the sections for 20–30 min to obtain optimal staining. This will only work for recovering more PLAP activity, as the β-Gal activity will be destroyed by the heat treatment.

1. Dissect the tissue into PBS containing 2 mM MgCl$_2$ (PBS + Mg^{2+}) on ice.

2. Fix in 0.5% glutaraldehyde in PBS + Mg^{2+} for not longer than 45 min on ice. The 0.5% glutaraldehyde decreases the alkaline phophatase activity in chick cells stained *in situ* (but not in chicken embryo fibroblasts stained *in vitro*). Therefore, fixation of chick whole mounts is typically done in 4% paraformaldehyde in PBS for 2–4 hr at 4°. In areas where background alkaline phosphatase activity is a problem, increasing the time in 4% paraformaldehyde, even up to several days, can decrease endogenous background without significantly decreasing PLAP. However, such long fixation times may decrease β-Gal activity.

3. Rinse in PBS + Mg^{2+}, 5 times, 5 min each time. Rinsing overnight is fine, but waiting for several days at this step may decrease β-Gal activity.

4. Stain in X-Gal reaction mix for 2–4 hr at 37°. (Tris buffer was also tried in place of PBS, with no success.)

5. Rinse many times in PBS until the solution no longer turns yellow. This usually takes about 5 changes. An overnight rinse is fine. It is important to remove the X-Gal, since residual β-Gal activity in the presence of X-Gal and NBT (added for the following reaction) may enable β-Gal$^+$, PLAP$^-$ cells to turn purple. Chick retinas and cerebella have been kept in PBS at 4° for at least 1 month at this point with no appreciable loss of signal in subsequent X-P staining. X-Gal staining can be easier to examine prior to carrying out the X-P reaction, as background alkaline phosphatase staining can obscure the X-Gal signal somewhat.

6. Heat the tissue in PBS at 65° for 30 min. This is usually done by floating the dish containing the tissue in a water bath preset to 65°. For staining of embryonic chick diencephalon (one of the areas of the brain with the highest background), this step has been increased to 1.5 hr.

7. Preincubate in X-P detection buffer for 15 min. Extending the time of this step results in diffusion of the alkaline phosphatase reaction product.

8. Incubate in X-P reaction mix for 3 hr at room temperature. Since background staining increases in light, cover with foil during and after staining.

9. Rinse in 20 mM EDTA in PBS for 2–4 hr. Background can be due to endogenous alkaline phosphatase or other reactions that generate hydride ions and thus reduce NBT to form a purple precipitate. We have noted that background staining appears more slowly in the presence of EDTA. Tissue can be stored in the dark at 4° in PBS plus EDTA or 30% sucrose in PBS plus EDTA for many months, although the background clearly increases over time.

10. Embed in paraffin wax using minimum necessary times for the tissue of interest. For mouse retina, which is approximately 250 μm thick, the following procedure is used. Dehydrate through graded ethanols: 50%, 70%, 95%, 100%, 100% for 20 min each. Clear in xylene, 2 times, 15 min each. Infiltrate with a 1:1 mix of xylene and paraffin at 65° for 30 min, then with paraffin, 2 times, 15 min each. Embed in paraffin.

Clearing and paraffin embedding are not recommended for tissues fixed with paraformaldehyde. In paraformaldehyde-fixed chick tissues, both the β-Gal and the PLAP reaction products were found to be very sensitive to xylene treatment; even relatively short exposures to xylene caused the reaction product to diffuse into the surrounding tissue. In some cases, this was true of ethanol as well. Frozen sections are a workable alternative (see below). Other embedding protocols, particularly aqueous-based procedures, may be worth testing. It is worth noting that strong staining in glutaraldehyde-fixed material can even withstand preparation for electron microscopy.[20]

11. Section onto slides coated with gelatin. Silane-treated slides are equally effective.[26]

12. Remove paraffin with xylene and mount with Permount. For frozen sections, fix sections to slides with 4% paraformaldehyde for 15 min. Rinse with PBS, then mount in Gelvatol (plus EDTA if desired). Storing slides at −80° helps prevent background staining from increasing.

Protocol for Staining Frozen Sections

The following protocol was worked out for embryonic rat striatum.[14] Fixation and staining times may need to be altered for other areas of interest.

1. Fix the tissue by perfusion followed by immersion in 4% paraformaldehyde at 4° for 8 hr. Rinse briefly in PBS, then sink in 30% sucrose in

[26] M. Rentrop, B. Knapp, H. Winter, and J. Schweizer, *Histochem. J.* **19**, 271 (1986).

PBS containing 2 mM MgCl$_2$ (PBS + Mg^{2+}) at 4°. Fixation times will vary with the size of the tissue. Perfusion may not be necessary for all tissues, especially in embryonic animals. Shorter fixation times may be preferable, as X-Gal staining may be decreased by lengthy fixation.

2. Embed brain in OCT or gelatin/sucrose mounting medium and freeze using liquid N$_2$. Gelatin/sucrose embedding gives better frozen sections for embryonic tissue than does OCT. Paraffin embedding destroys β-Gal activity. Cells in culture treated as if to embed in paraffin retain PLAP activity.

3. Cut cryostat sections and mount on gelatin-coated slides; air-dry overnight. Sections up to 90 μm thick (the thickest we have tried) have been successfully stained.

4. Fix sections to slides in 4% paraformaldehyde for 10–15 min at 4°.

5. Rinse slides in PBS + Mg^{2+} twice, for 10 min each, at 4°.

6. Stain slides in X-Gal reaction mix for 6 hr at 37°.

7. Rinse slides in PBS 3 times, for 10 min each, or until solution is no longer yellow. Slides can be left in PBS overnight. (See also comments for Whole-Mount Staining Procedure, Step 5.)

8. Transfer slides to preheated PBS at 65° and heat for 30 min.

9. Rinse slides in X-P detection buffer for 10 min.

10. Stain slides in X-P reaction mix for 12 hr. Because background staining increases in light, cover with foil during and after staining.

11. Rinse slides in PBS plus 20 mM EDTA 3 times, 10 min each. Mount in Gelvatol (plus EDTA if desired). (See also comments on background staining for Whole-Mount Staining Procedure, Step 9.) Storing slides at −80° helps prevent background staining from increasing.

Clonal Analysis Using Polymerase Chain Reaction/Library Method

Preparation of Retroviral Library for Polymerase Chain Reaction Analysis

In principle, any retroviral plasmid can be used to make a library. We start with the BAG plasmid,[7] which contains a unique cloning site (*Xho*I) downstream of the reporter gene.[15] The insertion of DNA into this site does not appear to interfere with expression of the upstream genes. The inserted DNA is genomic DNA from *Arabidopsis thaliana* digested with *Mbo*I. The digested DNA is run on an agarose gel, and DNA fragments less than 450 bp in size are used as inserts.[24]

Before making retroviruses with any of these constructs, approximately 100 constructs are identified whose inserts are conveniently distinguishable by size or by their pattern of digestion using restriction enzymes.

PCR products are prepared from bacterial colonies, the products are separated on agarose gels, and the approximate size of each PCR fragment is recorded. PCR products are sorted by size, and each product is digested with a mixture of restriction enzymes with 4-base recognition sites that are chosen because they are inexpensive and compatible within the same digestion buffer (MspI, RsaI, AluI, CfoI, and MseI). The size and restriction pattern of each tag are recorded, and approximately 100 constructs are chosen that are indistinguishable in this standard assay.[15]

Retroviruses are next prepared from the 100 constructs by transfecting "mini-prep" DNA,[27] purified using crushed glass, into a packaging cell line. The 100 DNAs are transfected[6] pairwise into 50 dishes of the ΨCRIP amphotropic packaging line[3] so that each transfected plate contains a mixture of cells producing two viral constructs. The supernatant of each transfected plate is used to infect a dish of the ecotropic packaging cell line Ψ2. Infected Ψ2 producer cells are selected by growth in medium containing G418 for 7 to 10 days, and the population of resistant colonies is raised to confluence.

Viral supernatants are recovered and titered on NIH 3T3 cells in 6-well tissue culture dishes, according to previously described techniques.[6] DNA from the infected NIH 3T3 cells is amplified using the PCR to evaluate the passage of the genetic tags. After the titering reaction, 0.5 ml of a solution of 0.5% Tween 20 and 200 $\mu g/ml$ proteinase K in $1 \times$ PCR buffer is added to each plate of cells. After incubation for 60 min at 65°, the proteinase solution is transferred to a 0.6-ml microcentrifuge tube and incubated at 85° for 20 min, then 95° for 10 min. The undissolved X-Gal precipitate is pelleted by centrifugation at 10,000 rpm for 10 min, and 10 μl samples of the supernatant are used as templates for 50-μl polymerase chain reactions as described below.

The 50 viral supernatants are then mixed to generate approximately equal ratios of each tag, and the mixture is concentrated by centrifugation.[6] The concentrated stock is titered and tested for helper virus in vitro.[6] It is then used to perform experiments in the cerebral cortex.

The method described above can be easily modified in several ways. Any convenient restriction site in a replication-incompetent vector can be used for insertion of tags, so long as it would not interfere with expression of the reporter gene or transmission of the virus. We have successfully created libraries using BAG, DAP, as well as two avian replication-incompetent vectors. Any source of DNA that is not present normally in the host tissue or vector can be used. Any set of enzymes can be chosen for

[27] T. Maniatis, E. F. Fritsch, and J. Sambrook, "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1982.

the restriction enzyme characterization of the inserts, and any range of insert sizes can be used if the longer ones are not so long that their size would cut down on the efficiency of the PCR. Finally, it may be possible to transfect the entire library as one plasmid preparation into producer cells rather than going through the tedium of making individual stable producers or producer populations. We are currently checking this as a possible way to make more complex libraries.

Tissue Analysis for Polymerase Chain Reaction-Based Clonal Assignments

Animals are infected and processed histochemically for the appropriate reporter gene as described previously. Usually sections are made for this analysis, but in special circumstances whole mounts can be used. Once the histochemistry is complete, the labeled cells are analyzed and their position as well as the outline and landmarks of the tissue are recorded. The goal is to create a permanent record of the morphology and relative location of the labeled cells, since the PCR analysis will destroy the cells and some surrounding tissue. A standard microscope and camera lucida device allow morphological details of cells to be drawn a high magnification, while cell location can be plotted at low magnification. Alternatively, cells can be photographed or plotted on a computerized system (e.g., CARP[28]). The anatomical analysis is often the most time-consuming aspect of the entire analysis.

Polymerase Chain Reaction of Labeled Cells

After anatomical analysis, PCR amplification of the tags that distinguish the vectors allows clonal analysis. For these experiments, start with two nested pairs of oligonucleotides specific for the region of the vector in which the inserts were placed. PCR parameters for each oligonucleotide pair must be optimized, especially the Mg^{2+} concentration in the PCR buffer.[29,30]

Materials

Histological sections prepared and stained as described above, and coverslipped in Gelvatol
Sterile distilled water

[28] C. P. Austin and C. L. Cepko, *Development (Cambridge, UK)* **110,** 713 (1990).
[29] K. B. Mullis and F. A. Faloona, this series, Vol. 155, p. 335.
[30] F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl, "Current Protocols in Molecular Biology." Green, New York, 1989.

1 × PCR buffer (see below)

Two nested pairs of oligonucleotides (as 20 μM solutions)

Deoxyribonucleotide solution (20 mM each in dATP, dTTP, dCTP, and dGTP)

Proteinase K, 10 mg/ml solution, in sterile distilled water

Tween 20, 10% solution, in steriled distilled water

Mineral oil (light), from Sigma

Disposable, breakable razor blades and blade holder (Fine Scientific Tools)

Microfuge tubes, 600 μl (autoclaved and silanized), or 96-well microtiter dishes and lids (e.g., Falcon, Nos. 3911 and 3913)

Automated thermal cycler

Dissecting microscope

Centrifuge tubes, 50 ml

Note. All solutions and containers must be assembled and stored using the most stringent precautions to prevent contamination with DNA that could be amplified in the PCR. A dedicated set of solutions should be handled with dedicated, positive displacement (or other contamination-resistant) pipetting devices in a separate laboratory remote from the thermal cycler. Good discussions of minimizing PCR contamination are available elsewhere.[30]

Dissection of Cells and Digestion of Tissue

1. Prepare a lysis solution. For 50 samples, mix 50 μl, or optimum amount, of 10 × PCR buffer; 25 μl of 20 μM solution of each of the outermost oligonucleotides; 10 μl of 10 mg/ml solution of proteinase K (final concentration 0.2 mg/ml); 25 μl of 10% Tween-20 solution (final concentration 0.5%); and 375 μl distilled water.

2. Soak off coverslips in distilled water in a clean, sterile 50-ml centrifuge tube. After the slide has soaked about 30 min, the coverslip can be carefully pried off with a razor blade. After the coverslip comes off, soak the tissue about 5 min more to remove traces of Gelvatol.

3. Pipette 10 μl of the lysis solution into each 600-μl microcentrifuge tube (or each well of a 96-well microtiter dish).

4. Break off a fresh fragment (2–5 mm wide at the edge) of the breakable razor blade in the blade holder.

5. Under the dissecting microscope, locate a labeled cell. Well-stained cells can be seen with a low magnification objective (0.8 ×), but lighter cells may only be seen with a high magnification objective (5 ×).

6. Using the razor blade, cut a fragment of tissue that includes the nucleus of the labeled cell. If labeled cells are widely scattered, they can be dissected one at a time, in chunks that contain approximately 1000

unlabeled cells. Chunks are typically less than 500 μm in each dimension, but larger pieces can be used (PCR sensitivity may be less with larger pieces, since they do not dissolve as well). If labeled cells are immediately adjacent, dissect several cells in one chunk.

7. Transfer the tissue piece, on the razor blade, to the lysis solution. Confirm under the microscope that the labeled cell is in the lysis solution. It is important to keep careful notes of which cell goes where (drawings are helpful here).

8. Cover the lysis solution with 100 μl of mineral oil, and cap the tube (or cover the microtiter plate).

9. Replace the blade fragment with a fresh one, and dissect the next cell (return to Step 5). While dissecting cells, prepare negative controls. Intersperse samples that contain no tissue, or unlabeled tissue, among the positive samples.

10. After all cells have been dissected, and all lysis samples covered with mineral oil, cover the tubes (or plate).

11. Transfer samples to a thermal cycler. Digest for 2–3 hr at 65°. Inspect a few samples after this time to confirm that the tissue is totally dissolved. If not, digest longer (e.g., overnight at 37°). The X-Gal precipitate does not dissolve, but it does not interfere with the PCR.

12. Once the tissue is digested, heat it to 85° for 20 min, then heat to 95° for 5 min. This inactivates the proteinase K and denatures the genomic DNA. The samples are now ready for the PCR.

First Polymerase Chain Reaction

1. Prepare the PCR solution. For 50 samples, use 100 μl (or optimum amount) of 10× PCR buffer, 50 μl of 20 μM solution of each of the outermost oligonucleotide primers, 10 μl of the mixed deoxyribonucleotide solution (20 mM in each dNTP), 7.5 μl (37.5 units) of AmpliTaq DNA polymerase (Perkin-Elmer Cetus, Norwalk, CT), and 800 μl of distilled water. The reactions can be scaled up or down proportionately. Taq polymerase from other manufactures can be used but should be tested first.

2. Put the samples in the thermal cycler and start the PCR, which should begin with an initial denaturation at 92°–94° for 3 min. PCR conditions need to be optimized, but will comprise 45 repetitions of denaturation (92°–94° for 30–45 sec), annealing (55°–70° for 1–3 min), and extension (72° for 1 min).

3. Once the samples have reached at least 85°, uncover them and add 20 μl of the PCR solution to each sample. The "hot start" enhances the sensitivity and specificity of the PCR. The added PCR solution does not have to be mixed in. It will sink beneath the oil and join the aqueous phase.

4. Cover the samples and allow the PCR cycling to proceed.

Second Polymerase Chain Reaction. The samples now contain large amounts of amplified target DNA and should not be opened or even transported into the clean laboratory. Because contamination of them is not a concern, they may be handled in the main laboratory. The sensitivity of the second PCR is not critical, and *Taq* polymerase from any manufacturer can be used.

1. Prepare the following solution. For 50 samples, use 10 μl mixed deoxyribonucleotide solution (20 mM in each dNTP), 200 μl (or optimum amount) of 10× PCR buffer, 50 units *Taq* DNA polymerase, 1800 μl sterile distilled water, and 20 μl of each of the internal pair of oligonucleotide primers (final concentration 0.2 μM). The lower concentration of primers and deoxyribonucleotides used in the second reaction does not affect the PCR sensitivity or product yield.

2. Pipette 40 μl of this solution into 50 tubes (or microtiter wells).

3. Transfer 2–4 μl of the product of each first PCR into each tube (or well). Keep samples carefully labeled and take note of any pipetting errors.

4. Run the second PCR for 25–35 cycles. Conditions for the second PCR will also have to be optimized but will likely be similar to those used for the first reaction. If the same thermal cycler is used for both the first and second PCR amplifications, be sure not to open any of the samples of the second reaction near the machine. A "hot start" is not necessary for the second reaction.

5. Separate the PCR products on 3% NuSieve/1% SeaKem agarose gels (FMC, Rockland, ME). Use 1× Tris–borate– EDTA (TBE) as the running buffer,[28] and include pBR/*Msp*I or ΦX/*Hae*III DNA size standards. Anticipate that 40–70% of the samples will produce a band. Suspect a problem if yields are consistently below 40%; suspect contamination if many contiguous samples show the same product.

Direct Restriction Enzyme Digestion of Amplified Products

DNA inserts may be distinguished by size or sequence analysis. Restriction mapping offers the most convenient assay, and it is sufficiently specific for libraries with 100–300 different inserts. We chose five restriction enzymes with 4-base recognition sequences (*Cfo*I, *Rsa*I, *Alu*I, *Mse*I, and *Msp*I) that were active in similar buffer conditions (50–100 mM NaCl).[15] The mixture cuts small DNA fragments frequently, allowing them to be easily distinguished. Before loading the diagnostic gel, or prior to restriction digestion, sort the PCR products by size. When samples of similar initial size are run side by side on the gel, it allows the most direct comparison of the restriction fragments.

TABLE I
PROBABILITY OF TWO CLONES HAVING THE SAME TAG IN k
CLONES

Number of tags (n)	Probability			
	$k = 3$	$k = 5$	$k = 10$	$k = 20$
20	0.14	0.40	0.86	0.99
80	0.04	0.12	0.43	0.90
85	0.035	0.11	0.41	0.88
100	0.030	0.10	0.36	0.85
250	0.012	0.04	0.16	0.52

1. Prepare the following solution (for 50 samples): 250 units of each restriction enzyme; 150 μl of 10× restriction enzyme buffer, appropriate for the enzymes chosen; 15 μl of 10 mg/ml bovine serum albumin (BSA); and sterile distilled water to make a total of 500 μl.

2. Pipette 10 μl of this solution into individual wells of a microtiter dish (tubes can also be used).

3. Add 20 μl of one of the PCR products to each well.

4. Cover the microtiter plate and incubate at 37° for 3 hr.

5. Terminate the reactions by adding loading buffer, and separate the products of the reaction on a 3% NuSieve/1% SeaKem agarose gel run in 1× TBE.

6. Record which samples contain the same tags, along with the overall number of tags seen. Compare this information to the original plots of cell location to get clonal information.

Statistical Analysis

The tentative conclusion from the PCR analysis is that cells containing the same tag are members of the same clone. The confidence of this conclusion rests on (1) the number of clones in a given experiment, k, and (2) the number of tags in the library, n. There is a surprisingly large probability that the same tag will appear in two different clones by coincidence, and this probability needs to be considered in the clonal analysis.

The probability of "coincidental double infections" by one tag can be calculated using binomial theory, or by computer, using a Monte Carlo simulation. The simplest assumption for use in either method is that all tags are present in approximately equal ratios. This assumption must be verified experimentally. If it is not the case, the library may still be usable, but more complicated modeling must be performed. Detailed discussions of the statistical analysis of retroviral libraries are presented elsewhere.[31,32]

[31] T. B. L. Kirkwood, J. Price, and E. A. Grove, *Science* **258**, 317 (1992).
[32] C. Walsh, C. L. Cepko, E. F. Ryder, G. M. Church, and C. Tabin, *Science* **258**, 317 (1992).

A computer program (MONTAG) that aids in calculation of expected frequencies of coincidental infections given different distributions and numbers of tags has been written by George Church (Department of Genetics, Harvard Medical School, Boston, MA) and is available through anonymous internet ftp from rascal.med.harvard.edu. It will run on most VMS machines without recompiling. Type run MONTAG and answer the queries. If there are any problems, contact church@gnome.med.harvard.edu. Some typical results, assuming approximately equal ratios, are shown in Table I to provide guidelines. The probability that more than 1 clone show the same tag equals $1 - [n!/(n - k)! \, n^k]$.

To evaluate clonal data, it is important to evaluate the actual complexity of the viral library (i.e., how many tags are seen in total). The best results demand a highly complex library, with very few clones labeled in each experiment (3–4 clones). Under these conditions the probability is maximized that observed patterns faithfully reflect clonal patterns. However, be skeptical of patterns that are seen rarely, as they may only reflect rare coincidental double infections, rather than true clonal events.

Reagents and Solutions for Polymerase Chain Reaction

10× PCR buffer: 10 mM Tris buffer, pH 8.3 (purchase as premixed crystals from Sigma), 50 mM KCl, 0.01% gelatin, and 1.5–2.5 mM MgCl$_2$

Mixed deoxyribonucleotide solution: Deoxyribonucleotides may be purchased from Pharmacia (Piscataway, NJ) as separate 100 mM solutions of dATP, dCTP, dGTP, and dTTP. Mix them 1 : 1 : 1 : 1 : 1 (by volume) with distilled water to make a working mixture that is 20 mM in each. Store 10-μl aliquots at $-70°$

Proteinase K: Proteinase K can be purchased from many manufacturers, dissolved in sterile distilled water to make a 10 mg/ml solution, and stored as 20-μl aliquots at $-70°$

Oligonucleotide primer solutions: Deprotected oligonucleotides can be passed over a NAP-10 ion-exchange column (Pharmacia) and eluted with sterile distilled water. Adjust the concentration of the effluent to 20 μM by measuring the absorbance at 260 nm. We routinely use oligonucleotides without further purification. Oligonucleotides should be stored as 25- to 50-μl aliquots at $-20°$.

Note. Reagents, instruments, and glass microscope slides should be handled with scrupulous technique and UV-irradiated when needed to destroy contaminating DNA.