Genotyping BayGenomics Mice
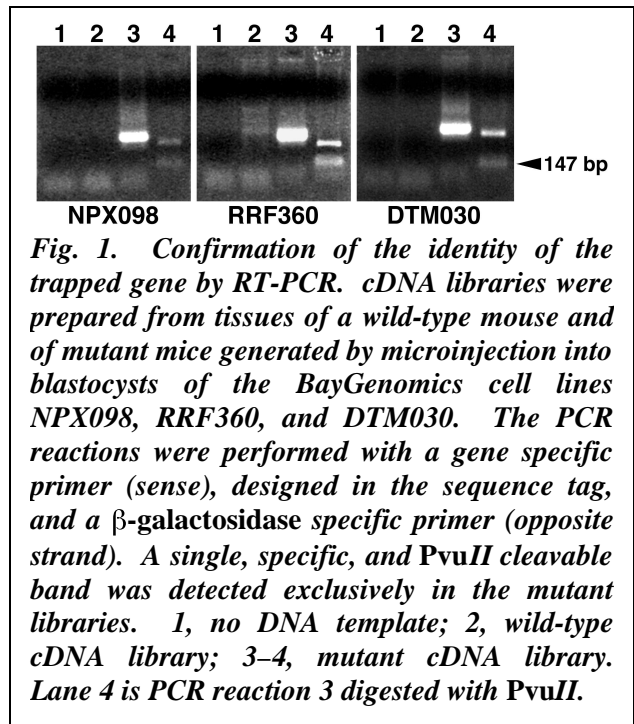
1. Introduction

The gene-trap vectors used within BayGenomics contain a splice-acceptor sequence upstream of a reporter gene, β-*geo* (a fusion of -galactosidase and neomycin phosphotransferase II)[1] . These vectors insert randomly into introns. The resulting insertional mutation creates a fusion transcript containing sequences from exons upstream to the insertion joined to the β-*geo* marker. As a result, the identity of the trapped gene can easily be determined by 5 rapid amplification of cDNA ends (5 RACE) followed by automated DNA sequencing. Of note, the sequence tags posted on the BayGenomics website have been trimmed of any "contaminating" vector sequences.

2. Confirming the identity of the trapped gene

The identity of the trapped gene should be systematically confirmed. This can easily be done by RT-PCR. First, prepare a cDNA library from your BayGenomics clone or from mutant mouse tissue, using a commercially available reverse transcriptase. Then perform a PCR reaction by designing a forward primer in the BayGenomics sequence tag, and by using the following reverse primer (in the β-*galactosidase* reporter):

5 -GAC AGT ATC GGC CTC AGG AAG ATC G-3

Fig. 1. *Confirmation of the identity of the trapped gene by RT-PCR. cDNA libraries were prepared from tissues of a wild-type mouse and of mutant mice generated by microinjection into blastocysts of the BayGenomics cell lines NPX098, RRF360, and DTM030. The PCR reactions were performed with a gene specific primer (sense), designed in the sequence tag, and a* β-**galactosidase** *specific primer (opposite strand). A single, specific, and* **Pvu***II cleavable band was detected exclusively in the mutant libraries. 1, no DNA template; 2, wild-type cDNA library; 3–4, mutant cDNA library. Lane 4 is PCR reaction 3 digested with* **Pvu***II.*

[1] Stryke, D., Kawamoto, M., Huang, C. C., Johns, S. J., King, L. A., Harper, C. A., Meng, E. C., Lee, R. E., Yee, A., L'Italien, L., Chuang, P. T., Young, S. G., Skarnes, W. C., Babbitt, P. C., and Ferrin, T. E. (2003) *Nucleic Acids Res.* **31**, 278–281

Such PCR reaction should yield a single band, exclusively in samples from mutant cells or mouse tissues.   Finally, the specificity of the amplified band should be further confirmed by restriction digest with *Pvu*II, which will clip a 147bp fragment off β-*galactosidase* (Fig. 1).

3.   *Identifying where the insertional mutation occurred*

The BayGenomics cell line's web page links directly the sequence tag to the coding sequences of the trapped gene that are available at the GenBank (NCBI) or the Mouse Genome Informatics (The Jackson Laboratory) websites.   From either of these sites it is straightforward to retrieve genomic sequences as well, since they both link out to LocusLink (http://www.ncbi.nlm.nih.gov/LocusLink). (Of note, for some BayGenomics cell lines, it is actually possible to link out directly to LocusLink.) Then,   align   the   coding   and   the   genomic   sequences   at http://www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html by entering the corresponding accession numbers in the designated fields.   Doing so will allow the determination of the structure of your gene of interest (namely, the number and size of exons and introns there are in that specific gene).   Then, aligning the sequence tag (copy and paste the tag sequence in the designated window) and the cDNA makes it possible to determine which exon(s) are upstream to the insertion, and thereby in which intron the insertion occurred.

4.   *Genotyping of BayGenomics mice*

BayGenomics mice can be genotyped by PCR, Southern blot, or DNA Dot blot.

4.1.   *Genotyping by PCR*

The insertion site can be determined precisely with various commercial inverse PCR kits.   In the "lucky" event that the vector inserted in a short intron (~1 kb or less), it is possible to genotype BayGenomics mice by PCR, without knowing the exact site of insertion.   The wild-type allele will be detectable by designing a forward primer in the exon immediately upstream to the insertion, and a reverse primer in the exon immediately downstream to the insertion.   (Of note, under standard PCR conditions it will be impossible to amplified the mutated allele this way since the vectors used in

BayGenomics are at least 8kb long.) The mutated allele will be detected using the same forward primer than above, and by designing a reverse primer in the vector. We have used with success either of the following reverse primers:

5 -CAC TCC AAC CTC CGC AAA CTC-3

5 -CGT GTC CTA CAA CAC ACA CTC CAA CC-3

The BayGenomics vector sequences are readily accessible from the BayGenomics web site, should you need to design your own reverse primer.

4.2. Genotyping by Southern blot

There are two straightforward ways of genotyping BayGenomics mice by Southern blot. All is needed is the complete sequence of the trapped intron, and the sequence of the vector used to trap your gene-of-interest. One approach is to choose a restriction endonuclease that will cut in the vector but not in the trapped intron.
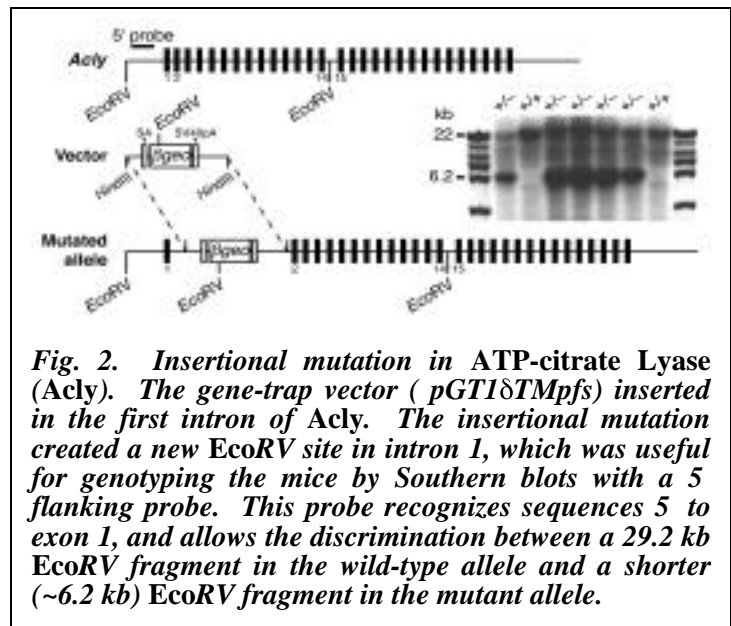

Fig. 2. Insertional mutation in ATP-citrate Lyase (Acly). *The gene-trap vector (pGT1δTMpfs) inserted in the first intron of* **Acly.** *The insertional mutation created a new* **EcoRV** *site in intron 1, which was useful for genotyping the mice by Southern blots with a 5 flanking probe. This probe recognizes sequences 5 to exon 1, and allows the discrimination between a 29.2 kb* **EcoRV** *fragment in the wild-type allele and a shorter (~6.2 kb)* **EcoRV** *fragment in the mutant allele.*

For example, there is a single *Eco*RV site in the vector backbone. In this case, a probe recognizing sequences 5 to the trapped intron will detect the wild-type allele and the shorter mutated allele (Fig. 2).

Another approach is to cut with an enzyme that cuts once (ideally) in the trapped intron but not in the vector.

For example, there are no *Spe*I sites in the vector backbone. It is necessary to test probes recognizing sequences 5 or 3 to the trapped intron, since it is not possible to know *a priori* whether

the vector inserted before or after the restriction site for *Spe*I. In this case, the appropriate probe will detect the wild-type allele and a longer mutant allele.

These strategies will prove difficult to use if:

—the complete sequence of the trapped intron is not available in the current databases (very unusual); or

—the vector inserted in a large intron. In this case, the chances of finding an enzyme that will not cut, or cut only once in such intron, are much lower.

4.3.   Genotyping by DNA Dot blot

Should genotyping by PCR or Southern not work out, it will *always* be possible to genotype BayGenomics mice with a DNA dot blot. The protocol for this technique is posted on the website at http://baygenomics.ucsf.edu/protocols/.

It is essential to follow scrupulously the described protocol, as it has been optimized to ensure that the amount of DNA loaded on the membrane is saturating. This will allow the discrimination between heterozygous and homozygous knockout mice (Fig. 3).

We have used the following -galactosidase probe with success:

TTATCGATGAGCGTGGTGGTTATGCCGATCG

CGTCACACTACGTCTGAACGTCGAAAACCC

GAAACTGTGGAGCGCCGAAATCCCGAATCT
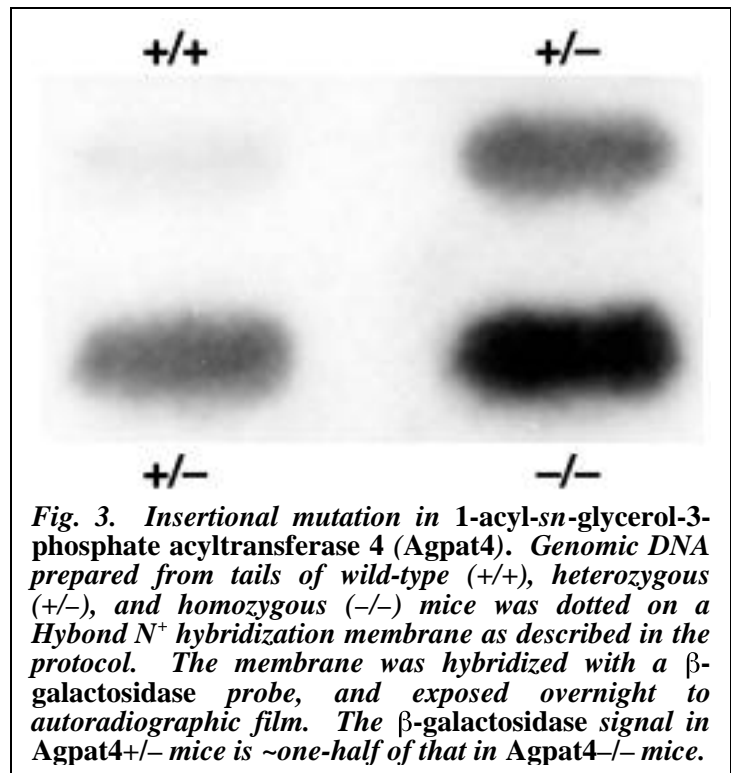
CTATCGTGCGGTGGTTGAACTGCACACCGCC

Fig. 3.   Insertional mutation in **1-acyl-*sn*-glycerol-3-phosphate acyltransferase 4** *(***Agpat4***). Genomic DNA prepared from tails of wild-type (+/+), heterozygous (+/–), and homozygous (–/–) mice was dotted on a Hybond N⁺ hybridization membrane as described in the protocol.   The membrane was hybridized with a* β-*galactosidase probe, and exposed overnight to autoradiographic film.   The* β-**galactosidase** *signal in* **Agpat4**+/– *mice is ~one-half of that in* **Agpat4**–/– *mice.*

GACGGCACGCTGATTGAAGCAGAAGCCTGCGATGTCGGTTTCCGCGAGGTGCGGATTGAAAATGGTCTGC

TGCTGCTGAACGGCAAGCCGTTGCTGATTCGAGGCGTTAACCGTCACGAGCATCATCCTCTGCATGGTCAG

GTCATGGATGAGCAGACGATGGTGCAGGATATCCTGCTGATGAAGCAGAACAACTTTAACGCCGTGCGCT

GTTCGCATTATCCGAACCATCCGCTGTGGTACACGCTGTGCGACCGCTACGGCCTGTATGTGGTGGATGAA

GCCAATATTGAAACCCACGGCATGGTGCCAATGAATCGTCTGACCGATGATCCGCGCTGGCTACCGGCGAT

GAGCGAACGCGTAACGCGAATGGTGCAGCGCGATCGTAATCACCCGAGTGTGATCATCTGGTCGCTGGGG

AATGAATCAGGCCACGGCGCTAATCACGACGCGCTGTATCGCTGGATCAAATCTGTCGATCCTTCCCGCCC

GGTGCAGTATGAAGGCGGCGGAGCCGACACCACGGCCACCGATATTATTTGCCCGATGTACGCGC

This probe was amplified by PCR, from genomic DNA extracted from BayGenomics mouse tail, with the following primers:

5 -TTA TCG ATG AGC GTG GTG GTT ATG C-3

5 -GCG CGT ACA TCG GGC AAA TAA TAT C-3

Dot blotting constitutes a fast, easy, and sure way of genotyping *any* BayGenomics mouse.

Remember, it is also possible to determine the exact site of insertion of the vector by inverse PCR with commercially available kits.