DISCUSSION FORUM

On clonal expansion and its effects on mutant frequencies, mutation spectra and statistics for somatic mutations in vivo

John A. Heddle
Department of Biology, York University, Toronto, Ontario M3J 1P3, Canada

Jackpots of mutations are the bane of transgenic mutation assays, for they can produce large mutant frequencies even in the absence of a mutagen. A jackpot is a large number of mutants that arose from a single mutation. They arise when there is a clonal expansion of a mutated cell. Jackpots as large as 1000 mutants/100,000 plaques have been observed, a mutant frequency of 1%. Such a jackpot can swamp not only the spontaneous mutant frequency but even the mutant frequency induced by potent mutagens, which are rarely >0.2% with even the most potent mutagens and are usually much lower (Hoorn et al., 1993; Tao et al., 1993; Morrison and Ashby, 1994; Douglas et al., 1995; Mursalis et al., 1995; Skopek et al., 1995; Cosentino and Heddle, 1999). Very large jackpots are not normally a problem, for they are statistical outliers and can be discarded legitimately by standard statistical tests. These large jackpots arise when a mutation occurs early in development in one of the progenitor cells for one or more tissues. The normal growth of the organism is by clonal expansion of the cells present during the preceding stage. Accordingly, mutations arising at an early stage of development produce a large number of mutant cells, i.e. a jackpot, when these are recovered together. Smaller jackpots, from mutations arising later in development, are harder to identify and are, thus, more problematic. Numerous proposals have been made to deal with this problem including complex statistical analysis and the sequencing of each and every mutant (de Boer et al., 1996; Knoll et al., 1996). There is a simple method of dealing with the data without such extraordinary efforts. In fact, current methods over-compensate for jackpots and thus distort both the mutant frequency and the mutation spectrum.

The transgenes present in the Muta™ Mouse and the Big Blue™ Mouse are bacterial genes embedded in a recoverable vector (Gossen et al., 1989; Kohler et al., 1991). They lack mammalian promoters or polyadenylation sites and are, presumably, unexpressed and unimportant to the mouse. Theoretically and practically they are genetically neutral, providing neither benefit nor harm to the cell containing them (Tao et al., 1993; Tao and Heddle, 1994; Cosentino and Heddle, 1996). This being the case, a cell containing a mutated transgene is neither more nor less likely to expand than any cell without such a mutation. In a large number of animals, therefore, clonal expansion will have no effect on the average mutant frequency: expansions of the numerous non-mutants will each tend to decrease the mutant frequency somewhat while expansions of the rare mutants will each tend to increase the mutant frequency to a larger extent (Figure 1). Nevertheless, in any one animal, expansion of a mutant cell into a large clone can have an enormous effect on the mutant frequency, for mutations are normally rare. Thus jackpots can and do occur and, when they do, both the mutant frequency in the animal and in the animal’s treatment group can be greatly increased. Obviously, smaller jackpots produce smaller effects and there is no precise dividing line between the quantitative contribution of the other sources of variation and that from jackpots. This means that mutant frequencies as measured are all being affected to a greater or lesser degree by jackpots. Most spontaneous mutations arise late in development, because more cells are at risk, so most will affect a single tissue and most will be small.

Because of the time required for cell proliferation to create a clone, most jackpots arise from spontaneous mutations during normal growth and development and not from expansion of induced mutations. Induced jackpots will be unimportant in adult animals unless there is a long interval between treatment and sampling. Most spontaneous mutations arise during embryogenesis and early growth, corresponding to the times of most rapid cell division (Zhang et al., 1995; Paasuis-Lew et al., 1997; Paasuis-Lew and Heddle, 1998). The earlier a mutation arises, the larger the jackpot it can produce, as there is more time and opportunity for the mutated cell to produce a large clone and thus to become a larger fraction of the cells present. Accordingly, the larger the jackpot, the earlier in development the mutation arose, other things being equal. This process is illustrated in Figure 2, where the alternating process of expansion of a small number of cells by rapid, possibly exponential, growth alternates with the bottlenecking (severe reductions in the number of cells in the population that will contribute to the next generation) produced when a few of these cells become the stem cells for the next stage of tissue differentiation. Once adulthood has been reached, most tissues of the mouse are of more or less constant size. In some, like the bone marrow and the epithelia, there is a rapid turnover of cells. In most tissues, there is a slow turnover which may be somewhat increased after toxic insult, like mutagenesis, but this is transient and does not lead to large clones in tissues where these can be visualized, like the skin and the small intestine. It is clear that if samples are taken soon after mutagenesis, as is usual, there is little time for a clone to become very large. Hence, large jackpots must arise during development. Since only large jackpots are recognizable from the mutant frequency alone, all of these must be spontaneous in origin. Accordingly, they should be no more frequent in the treated groups than in the controls. Small jackpots are another matter. Although these are expected to be much more frequent, each has a smaller effect on the mutant frequency and is, accordingly, harder to recognize from the mutant frequency.
Fig. 1. Illustration of the dual effects of clonal expansion. With no increase in the size of an organ, proliferation of one cell must be accompanied by the loss of another. To the left of the double line, normal cellular turnover leads to the loss of three mutants and their replacement by clonal expansion of non-mutants. To the right of the double line, the same process of normal cellular turnover leads to the loss of some non-mutant cells and their replacement by the clonal expansion of one of the mutant cells. On the left, the mutant frequency is reduced; on the right it is increased; overall, it is unchanged. These two effects balance out if the mutations are neutral because the probability of clonal expansion is the same for mutants and non-mutants. (Here 14% of the mutants expand but only 1.6% of the non-mutants, because of space limitations and the need to have enough mutants to see what is happening.)

In addition, depending upon the turnover of the cells, there is a greatly reduced chance of obtaining jackpots when tissues are sampled relatively soon after treatment. Although ‘relatively’ is not very precise, some simple calculations give some dimensions to this term.

Consider one of the most actively dividing tissues of the body, the small intestine. In this tissue, the turnover of the reproductive cells is approximately daily, so a clone can double in size each day. After 10 days, a clone could be 1024 cells if all of the daughter cells continued to reproduce at this rate. Since packaging typically captures <0.1% of the copies in the DNA and not all the DNA is captured or used, even this expansion is insufficient to generate a jackpot. (For example, in a packaging reaction using 8 µl of 2 mg DNA/ml, there are \(-16 \times 10^6\) mouse genomes each containing 40 copies of \(\lambda\) DNA for a total of \(640 \times 10^6 \lambda\), yet this rarely produces as many as \(640 \times 10^3\) plaques, which would be 0.1%.) In tissues with slower turnover times, the generation of jackpots will be less likely.

Fig. 2. Illustration of the growth of a mouse from the zygote through the blastula, most of which forms placenta, to the gastrula and the beginnings of tissue differentiation, to an embryo in which each tissue and organ is formed from a small number of precursor stem cells. Growth and development leading to any tissue consists of alternating phases of rapid (possibly exponential) growth and bottlenecks in which a few cells become the founder stem cells for the next phase of development leading to a particular tissue. If one of these stem cells is mutant at any stage, a large fraction of the tissue formed from it will also be mutant, as shown in the heart.

Fig. 3. A photograph of typical mutant clones in the small intestine as revealed by the method of Winton et al. (1988). Almost all of the mutant clones are of this size and shape in the first few weeks after mutagenesis (cf. Cosentino et al., 1996), showing the limitations on clonal expansion that exist in adult animals. Normal patterns of cell turnover determine the size and morphology of mutant clones in all tissues.

The intestine also provides a visual demonstration of clonality when the Dlb-1 locus is used (Winton et al., 1988; Cosentino et al., 1996). Mutations of this locus are revealed by a histological staining procedure (Figure 3) which demonstrates that each mutant occupies a more or less fixed proportion of the tissue. In other words, a mutation occurring in an adult animal can expand only so far. Occasionally, after treatments that are very severe, two or three Dlb-1+ ribbons can be
Fig. 4. Relative frequency of clones of different sizes in a cell population that arose by equal proliferation by all cells in the population. Those that arose in the most recent cell division will be the most common, but each mutation will be present in a single cell. In the previous generation, there were half as many cells, so half as many mutations arose, but each mutation is now present in two cells. Two generations previously, there were 25% of the number of cells and of new mutations, but each mutation is now present in four cells. And so on.

seen emerging from a single crypt, indicating a small clonal expansion beyond the normal limits (Cosentino et al., 1996). While mutations in transgenes cannot be visualized in this way, the expansion of clones must be similarly limited. These limitations will be true of most tissues of the body where either very slow rate of turnover or the presence of multiple stem cells will limit the size of the jackpot that can be obtained. An exception may be the bone marrow where there is an extensive hierarchy of stem cells resting on a final base of a small number of pluripotential cells. Nevertheless, even in this tissue a large jackpot is unlikely to be induced because of the small number of pluripotent cells involved and the long time required for clonal expansion to occur from these cells. For an induced jackpot to become important in a short time would require a tissue with a small number of very rapidly dividing cells and a rapid turnover. Such a tissue does not exist. Hence, the efforts to correct for induced jackpots is unnecessary under normal conditions, with the possible exception of bone marrow and tissues containing cells that arise from it, such as the thymus, blood and spleen. The only abnormal condition that would change this would be an event that led to extensive replacement of a tissue, for example the liver after partial hepatectomy.

Spontaneous jackpots are another matter. These are inevitable. In an exponentially growing population, their distribution would be as shown in Figure 4. In essence, given that the mutation rate (mutations/cell division) is a constant, the chance of a mutation is proportional to the number of cells at risk of mutation. The size of the jackpot arising from a mutation is inversely proportional to the number of cells at risk of a mutation. Thus each generation the number of mutations doubles and the size of the jackpot halves. This generates a distribution in which the probability of a jackpot of size $N$ is exactly half of the probability of a jackpot of size $2N$, as shown in Figure 4. In reality, there are other factors. First, as already noted, only a small fraction of the cells in a clone contribute a copy to the measured sample, so this contributes a substantial sampling variation. Second, the organism does not grow by doubling with each cell division, so not all cells in the tissue have incurred the same number of divisions since fertilization, and bottlenecks occur (cf. Figure 2). Third, when sampling a tissue, there may be non-random recovery of clones. This would occur, for example, if only one part of the liver were used for DNA extraction. Nonetheless, these factors are relatively minor. Use of a single lobe representing one-eighth of the liver, for example, would give a relative expansion of only ~8-fold to a liver clone in that lobe. In most tissues, all or most of the tissue is used and this is an even smaller factor. Obviously, one way to reduce the influence of jackpots is to use as much as possible of the tissue.

Statistically, the problem with jackpots is the influence that they have on the mean. A single jackpot can easily double the mean, which is important in a field in which a doubling of frequency has become the touchstone of a real effect. Furthermore, in many experiments in which known carcinogens have been used, the mutant frequency in the target tissue has been increased by only a factor of two or so. In this situation, the means of dealing with jackpots becomes of great importance. Large jackpots can be eliminated by standard statistical methods, but these are very rare. As noted above, the smaller the jackpot, the more likely its occurrence. A simple and effective approach to this problem is to use the median value rather than the mean to represent the group. Although this is uncommon in the mutation field, this is a robust statistical measure which is often used when the distribution of observed values is not normal, as is the case here. Given the general acceptance among those who use transgenic animals that the statistical analyses of Piegorsch et al. (1994) provide the basis for the numbers of animals and plaques to be used to define a negative result with acceptable power, it is common now to use five animals in a group. With five animals the median value becomes a very reliable representation of the mutant frequency of the group. The advantage is that there is no need to define in advance the size of a jackpot to be discarded nor to adjust the data retrospectively in any way.

The disadvantages of the corrections being used to deal with jackpots have not been much discussed. For spontaneous mutation, where jackpots are most influential, removal of jackpots biases the mutant frequency downward. Assuming, as there is every reason to believe, that these mutations are neutral, removal of jackpots is a genuine bias: for each and every clonal expansion producing a jackpot of mutants is matched by a huge number of similar clonal expansions of non-mutants in the tissue. While each of these clonal expansions of non-mutants has only a minor effect on the mutant frequency, overall they are important. Jackpots of mutants are not experimental artifacts, they are real mutations and represent real biological events in the tissue. To the extent that an unbiased estimate of the mutant frequency is important, their inclusion in the database is also important, although they cannot be allowed to dominate the estimate in a small group of animals, such as a treatment group. In mutation spectra, as well, presumptive jackpots must be treated with care. While the elimination of genuine jackpots from the spectrum is reasonable, the importance of true hotspots is reduced and the spectrum distorted every time two mutations are attributed to a single event but are actually two independent events. This improperly reduces both the class of mutation being considered
(G→A transitions, for example) and the importance of the particular site. Unfortunately, the distortion is not even linear, as the hotter the site, the more likely it is that two independent events will occur and be falsely attributed to clonality. Given our knowledge of tissue structure, stem cells and turnover rates, it is incumbent upon those who make these corrections to take care not to over-correct and thus distort the mutation spectrum, especially for induced mutation. For spontaneous mutations arising during development, there can be no doubt that jackpots occur and must be considered. For spontaneous mutations arising during adulthood, the importance of clonality will vary from tissue to tissue and with the time involved. Recognition of the biological basis of jackpots, rather than consideration of them as statistical anomalies, leads to a far simpler means for dealing with their occurrence in mutational data than has been practiced to date. Specifically, the use of the median rather than the mean and the recognition that only large jackpots need to be excluded from mutational spectra should improve the database and make the presentation and analysis of data simpler.

Acknowledgements

I thank Jason Bielas, Beichen Sun and Roy R.Swiger for their helpful comments on this manuscript. The photograph was taken by Dara Dickstein and Lidia Cosentino. The research in my laboratory is funded by the Cancer Research Society Inc., the National Cancer Institute of Canada and the Natural Sciences and Engineering Research Council.

References


Received on July 17, 1998; accepted on January 7, 1999