Chapter 8

Discussion

8.1 Enrichment for homozygous mutants in *Blm*-deficient ES cells

In Chapters 5 and 6 I described the development of a method to enrich for homozygous mutants by selection for the copy number increase that occurs during loss-of-heterozygosity (LOH). I developed a transposon carrying a double selection construct that can be used to isolate cells with two (or more) copies of the construct. I first used clonal cultures, in which all cells had the same heterozygous transposon insertion site to begin with. After expansion of the culture, I was able to isolate double resistant cells that had increased the copy number of the construct.

From these experiments it was clear that LOH leading to segregation of a homozygous daughter cell is not the only pathway for copy number gain in *Blm*-deficient ES cells. The wild type copy of the mutated locus could still be detected by PCR or Southern blot in some subclones isolated from the double resistant population. This was in addition to the two distinct forms of the selection construct, indicating that three alleles were present in some cells, including a wild-type. I refer to these subclones as 'wild-type retaining' clones to distinguish them from the genuine homozygous mutants. I interpreted these results as arising from chromosomal instability, and indeed I found that some of the wild-type retaining subclones had a near-tetraploid karyotype.

The average proportion of genuine homozygous subclones isolated from a given clone was 34%. While this level of enrichment is sufficient to easily obtain homozygous mutants by subcloning, this may not be enough to use the unsubcloned double resistant population directly in screening assays. Therefore I sought to adapt the method to produce clonally pure populations that would be suitable for screening directly. This required several technical improvements to strictly limit the initial copy number of the transposon to one by chromosomal mobilisation of the transposon specifically in G1 phase of the cell cycle (Chapter 6). Doing this removes the requirement to subclone cultures immediately after mutagenesis, and enables a mixed pool of mutants to be grown together. Analysis of 45 double resistant subclones revealed 19 mutants with two allelic insertions, representing 16 different insertion sites. Thus, this procedure can produce clonally pure mutants with a single subcloning step, without severe redundancy with respect to the number of different insertion sites.

8.1.1 Future improvements to library generation

The most obvious improvement required is the generation of large libraries with tens of thousands of mutants. The limiting factor in the experiments reported here was the low mobilisation efficiency using mRNA. As thousands of new transposon insertions can be obtained from transfection of 10^7 cells with PB-CDT1 plasmid, it should be possible to improve the efficiency of mobilisation using mRNA.

In the library analysed, the remaining 26 of the 45 subclones had two non-allelic insertions. However, in 25 of these cases, one of the insertions remained at the donor locus. Finding a way to eliminate these would help to increase the proportion of useful mutants in the enriched library. I plan to investigate these to see if these cells arise from aneuploidy present prior to mobilisation, and thus whether sorting cells by DNA content could reduce the problem.

8.2 Using enriched libraries for screens

Although not complete, the level of enrichment for homozygous mutants achieved here is high enough to consider using these libraries to investigate phenotypes that are not strongly positively selectable.

8.2.1 New technologies applicable to genetic screens

The traditional way to screen collections of homozygous mutants would be to pick and assay each individually. In cell culture, this means using multiwell plates (96- or 384-well for high throughput). However, some new technologies incorporate elegant

solutions to this requirement, particularly new sequencing technologies. With high throughput sequencing of transposon insertion sites, for example using the Illumina method described in Chapter 3, the number of cells belonging to each clone can be determined by counting the number of reads from their associated insertion site. This makes it a promising method for investigating phenotypes linked to survival or fitness in prolonged culture. Some phenotypes that could be interesting to investigate are differentiation into different lineages. Using a suitable differentiation protocol that is efficient in bulk culture, such as those for neural or mesodermal lineages, a differentiated library could be isolated. Sequencing all insertion sites in the differentiated population, and comparing to the starting population and an expanded, undifferentiated population, could identify mutants unable to progress to the differentiated stage. Assays for sensitivity to drugs should also be possible—the experiment described in Chapter 3 is a proof of principle of this type of screen.

Another class of phenotype that can be screened by this system is weak positive selection. Mutant clones with fitness advantages under a selective condition will expand and increase their representation in the pool. One potential area of application is screens for infection by viruses and other pathogens, or resistance to toxins. ES cells are not the natural hosts for pathogens, and may not be killed effectively enough to conduct a traditional resistance screen using a non-enriched library. Using an enriched library with a chronic treatment may produce better results than relying on complete acute killing.

I conducted one pilot experiment for this approach to screening, which suggested several improvements (Chapter 3). First, all transposons in the library need to be stably integrated, such that no *de novo* events occur after library generation. For this reason, I generated all subsequent libraries using mRNA to express the transposase, to remove the possibility of stable expression of the transposase in some cells that integrate the expression plasmid that I used previously. Combined with further technical and biological replicates, and the addition of a sample prior to expansion, a high confidence set of transposon sites present at the start of the experiment could be formed to compare the treated population against. This should increase confidence in the identified insertion sites. Using larger libraries, with more than one insertion site per gene would also strengthen the evidence that loss-of-function mutations in that gene are causing the phenotype.

A drawback of this approach is that the mutant cell line cannot be directly obtained. In a traditional genetic screen, this would be a problem, but as single mutant ES cell lines can be obtained easily from the public resource for rapid confirmation, it is less important now. In my view, the emphasis should be on obtaining rapid leads to gene function, which is the role played by the screening systems described here. In any case, mutants need to be reconfirmed on a wild-type (*Blm*-proficient) background using the cell line described here. Transferring the system to the $Blm^{tet/tet}$ line, for example, would be a further improvement.

Another potential improvement for screening assays could be the use of micro-patterned agar to array single cells for screening (Wood *et al.*, 2010). This technique provides a simple method to seed single cells in a grid pattern. This then allows single cells to be screened, and reliably located by a computer-controlled microscope. Screening single cells, rather than a population, has many advantages. For example, in my clone-by-clone isolation experiments (Chapter 5) a mixture of homozygotes and aneuploid cells was obtained. Screening this population, for example for sensitivity to a drug conferred by the mutation, would show an intermediate survival phenotype depending on the relative amounts of homozygotes to aneuploid cells. At the single cell level, the structure of the population can be seen more accurately—in fact in the paper above, cells can also be stained for DNA content, raising the possibility that tetraploid cells could be detected directly during the analysis.

8.2.2 Comparison to other systems for recessive genetic screens

The system described here has several advantages compared to siRNA screens. By picking colonies from homozygote-enriched libraries, a clonally arrayed library of mutants could be constructed that would be usable in similar situations to siRNA screens. The main advantage here is robust mutagenesis. The transposon construct that I used effectively abolished transcription of the wild type allele when homozygous and inserted into an intron (Figure 5.14). When using siRNA the knockdown is often incomplete, and it is also possible that not all cells are transfected, or receive different amounts of siRNA (this can be improved to some extent by shRNA approaches with selection for transformation).

As enrichment of the library is incomplete, there will be some 'junk' clones in such a library, which will manifest as false negatives. This is also a prob-

lem for siRNA screens, however, as the effectiveness of a particular knockdown cannot be guaranteed. Furthermore, transposon mutagenesis deals effectively with false positives, as the insertion can easily be removed by remobilising the transposon (Li *et al.*, 2010). This provides a simple test for causality, which is not available with siRNA.

The recent discovery of a human haploid cell line may represent a powerful alternative system for screens, although it remains to be seen how these cells behave (Carette *et al.*, 2009). Transposon mutagenesis should be readily applicable in this cell line, and the generation of loss-of-function mutants is much more straightforward compared to the *Blm*-deficient ES cell system. The limitation of screens to a single cell type, derived from a tumour, appears to be the only major limitation of this system. Certainly screens for differentiation are not possible in these cells, and they may also have other mutations acquired during tumourigenesis that could make them unsuitable for screening other phenotypes. In these situations, using the ES cell system described here will be necessary.

8.3 Other uses of the copy number selection transposon

The ability to select for copy number increase using the construct that I developed could find wider applications in the field of chromosome instability and copy number variation. Such effects could be easily investigated at different loci in different cell lines through use of the transposon to make stable, single copy integrations. Some of the ES cell lines generated as part of this work could be used as reporters for induction of copy number instability by drugs or mutagens, or new ones could easily be generated in other mutant backgrounds. As described in Chapter 7, PB also induces double strand breaks that are repaired by the host machinery, so my construct can also be used to investigate repair of locus specific DNA damage in a similar way.

8.4 Conclusions

The experimental systems and protocols that I describe in this thesis further extend the genetic toolkit available for analysis of gene function in mice. The main technology, homozygote enrichment by copy number selection, will be useful for conducting recessive genetic screens, a powerful technique from other model organisms that has still not been completely translated to mammalian systems. Technical improvements that were necessary to solve problems associated with copy number instability in ES cells during this process could prove to be more generally useful for the study of genome instability and DNA repair. Application of the technologies that I have developed will assist in the ongoing task of functionally annotating the mammalian genome.