

1. Many members of an isolated group of humans are subject to a fatal genetic disorder known as muscular maniosis (MM). A person with this disease is subject to periodic fits of uncontrolled and violent muscle contractions of the entire body accompanied by mental delirium. It is rare for people with this disease to live more than 35 years because they usually die from fatal injuries sustained during these fits of madness. It is not known what gene is defective in individuals with muscular maniosis. It is known, however, that the disease is inherited within families as an autosomal recessive disorder.

A world-renowned group of researchers headed by the illustrious biochemist Jules Mealcard and the celebrated biophysicist Sandy Nabob has just been given a huge research grant to identify the gene responsible for muscular maniosis.

a. Initially they will need to find out roughly where the MM gene is located in the human genome. How should they do this?

They should test the DNA of individuals who have MM with the goal of finding molecular markers (such as RAPDs or SNPs) that are linked to the MM trait. They would have to start out by testing markers from all of the autosomes. Once any significant linkage was established they would know what chromosome the MM was located on and could then try many other markers from that chromosome. Eventually they would want to find a marker that was very tightly linked with the MM trait. Ideally this would be a case where every individual with MM also carried the same allele at the molecular marker.

b. They have successfully narrowed down the location of the MM gene to within a 100 kbp region of chromosome 14, which they have isolated in a BAC clone. At this point they need to figure out what genes are actually present in this part of the genome and what type of proteins they encode. How could they do this?

They can obtain the sequence for the whole 100 kbp fragment from the public database containing the entire human genome. Even having the genomic sequence, however, might not really allow them to fully determine what actual mRNAs (and proteins) are encoded by the DNA. Since the genes would contain introns (and annotation of the genome might not be correct) it might be difficult to be sure what open reading frames are present.

They should subclone the 100 kb fragment into about 10 smaller pieces that would fit into plasmids for propagation in *E. coli*. This would enable them to prepare RNA blots from human muscle tissue and probe these blots with each of the 10 genomic subfragments. This would tell them which of these subfragments includes a gene that is expressed in muscle cells. To find out the sequence of the actual mRNA (and the protein) they would need to use the correct genomic clone to screen a human cDNA library and obtain the corresponding cDNA clone.

NOTE – they may be able to bypass some of this northern blotting by taking advantage of “electronic northern” that are available on the web.

c. After finishing the preceding analysis they have narrowed down the search to 2 genes. Both of these genes encode muscle proteins, but they can't figure out the exact function of either one. How can they figure out which gene is the MM gene?

They will have to obtain DNA clones of these two genes from a number of unaffected individuals and a number of individuals who suffer from MM. One of the two genes should have essentially the same sequence in all the unaffected individuals but have some type of significant mutation in all of the MM individuals. That is the MM gene.

4. Professor Ablo, whom we know from her previous work on Tasmanian devils, is now busy making a genomic library from these lovely beasts. The size of the TD genome is 1200 Mbp. In making her library, Dee has randomly broken up genomic DNA and cloned the pieces into a plasmid vector. Her average insert size is 10 kbp.

a. How many independent clones will she need to have in order to have a library with a total insert size that is equal to the size of the genome (1 genome equivalent)?

1,200 Mbp = 1,200,000 kbp

120,000 individual 10 kbp clones = 1 genome equivalent

b. If Dee wants to find a single-copy gene in the TD genome, can she be assured of finding it in a library with a size of 1 genome equivalent. Why or why not?

Since inserts are inserted randomly into the clones, it is NOT a certainty that any one particular gene will be present in a 1 genome equivalent library. It would be better to screen a library of 2 to 3 genome equivalents.

5. Professor Ablo has also made a cDNA library from Tasmanian devil brain tissue. The library consists of 120,000 independent cDNA clones.

a. She wants to obtain a cDNA clone for the fairly abundant *QED* mRNA. This mRNA is known to make up about 0.2% of all brain mRNA. In order to save on materials (grant money is running low for this project) how many cDNA clones should she screen in order to isolate a *QED* cDNA clone?

On average, 1 out of every 500 cDNA clones should contain the *QED* cDNA. If she screens 2000 clones she should be quite sure of finding a *QED* clone.

b. If she wanted to isolate a cDNA clone for the *BOB* mRNA, which represents approximately 0.0005% of all brain mRNAs, how many clones should she screen?

On average, 1 out of every 200,000 cDNA clones should contain the *BOB* cDNA. There is a very good chance that the *BOB* cDNA is NOT represented in her 120,000 clone library. She could give it a try, but she may need to get a bigger cDNA library before being able to successfully clone the *BOB* cDNA.