Assignment 1

Simon Lee, Daniel Cavaleri, Daniel Molinuevo, Lucille Niederhauser

1. What type of DNA vector (e.g. plasmid, cosmid, BAC) do you think is likely this DNA fragment is cloned in? Explain why.

Since the DNA clone is only 9.5kb, we can assume that a plasmid was probably used. This size could also fit in bigger vectors (e.g cosmids) but since it is much simpler to insert the sequence in a plasmid, we assume that the prior researcher chose this option. Other vectors are more appropriate for bigger fragments.

2. Having made a good guess at the vector, you can begin sequencing the piece of cloned DNA. What type of sequencing technology will you use? How many sequencing reactions are likely needed? Explain why.

Since we only have a fragment and it is not very long, using NGS seems simpler but would not make sense due to its size. Typically NGS is used for much larger sequences and is expensive. This leads us to believe that Sanger sequencing should provide us with satisfactory results if used properly.

Knowing the type of vector, we could start by using a set of primers known to work for the most commonly used plasmids to start the sequencing and hope for one of them to work. If the size of the cloned DNA was smaller than 1000bp this would be sufficient, but since we have a larger sequence (around 9.5Kb), we have to perform multiple Sanger sequencing cycles by breaking up our sequence into smaller fragments.

We estimate that at least 10 cycles of full Sanger sequencing will be required, each of them with a different subsequence. The initial reaction would be with the first 1000bps and would be initiated from the site where our starting primer binds. To proceed to the next cycle, we would use the already sequenced DNA to design a new primer that corresponds to its last base pairs. This would successfully initiate the reaction of the next cycle, thus yielding the next base pairs of our fragment (with some overlap with the previous fragment).

3. You successfully sequence the inserted DNA fragment (see attached file). From this sequence, can you tell

3.1 Which species this DNA fragment is from?

The DNA fragment is from the Homo-sapiens species. Source.

3.2 The name of the gene this fragment is from?

The DNA fragment is part of the gene coding for the transmembrane protein 41B. More precisely, the gene name is : TMEM41B. Source.

Explain how you determined this.

We ran the online tool, BLAST, against a relatively large set of different species and we obtained a 100% sequence match with the TMEM41B gene of the Home-sapiens.

4. This fragment of DNA has some unusual features – what are they? Explain how you determined these features and why they are not common.

First, we observed that the DNA fragment is an intron (the fragment is the end of exon 3 + the intron + the start of exon 4) of the gene. In addition this fragment has a unique property of containing many repeating A and T nucleotide bases [1]. These two features suggest that the sequence could be a scaffold/matrix attachment region (S/MAR), which is estimated to only constitute around 0.3% of the human genome. The main function of these sequences is in the formation of chromatin domains that facilitate enzyme interactions with DNA [2].

Moreover, we found out that such regions are used to form a special type of plasmid called an episome that could be used for gene therapy [3]. Episomes, unlike most types of plasmids, have the distinctive feature of being stable enough to not be degraded by the cell during mitosis and therefore be transferred to subsequent generations. This alludes to its application to gene therapy, since episomes allow the introduction of DNA sequences into human cells without the risk of integrating them into random places of the genome, which for example could lead to the activation of oncogenes.

Therefore, our guess is that the previous researcher was testing if this particular S/MAR sequence conferred stability to the resulting plasmid and could be later combined with a DNA coding sequence to be used for gene therapy, by having more efficient episomes.

References:

[1] Giannakopoulos, A., Stavrou, E. F., Zarkadis, I., Zoumbos, N., Thrasher, A. J., & Athanassiadou, A. (2009). The functional role of S/MARs in episomal vectors as defined by the stress-induced destabilization profile of the vector sequences. Journal of molecular biology, 387(5), 1239-1249.

[2] The functional role of S/MARs in episomal vectors as defined by the stress-induced destabilization profile of the vector sequences. Giannakopoulos A, Stavrou EF, Zarkadis I, Zoumbos N, Thrasher AJ, Athanassiadou A. J Mol Biol. 2009 Apr 17;387(5):1239-49. doi: 10.1016/j.jmb.2009.02.043. Epub 2009 Feb 25. PMID: 19248788.

[3] Advances in the Development and the Applications of Nonviral, Episomal Vectors for Gene Therapy. Grace Elizabeth Mulia, Virginia Picanço-Castro, Eleana F. Stavrou, Aglaia Athanassiadou, and Marxa Leão Figueiredo, Human Gene Therapy 2021 32:19-20, 1076-1095