MODULATING GENE EXPRESSION IN HUMAN T CELLS WITH PETRA

M. Armas Reyes[1], L. Cubitt[1], M. Herger[1], and G. M. Findlay[1]
[1]The Genome Function Laboratory, The Francis Crick Institute, 1 Midland Rd, NW1 1AT, London, United Kingdom

Modulating human gene expression can provide a powerful means of engineering cellular function in the development of new therapies. For example, while chimeric antigen receptor (CAR) T cell therapies have shown promise for treating cancer, their ability to target solid tumours can be enhanced by altering the expression of key genes. One way to stably alter gene expression is to edit known regulatory regions of the genome. Yet, edits that induce desired effects on gene expression remain largely unknown. This is, in part, due to a paucity of experimental data linking genomic edits to gene expression changes.

Here, we present PETRA (Prime Editing of Transcribed Regulatory elements to Alter expression), a method to determine the effect of genomic edits on gene expression in a multiplexed manner. PETRA leverages Prime Editing to install precise edits in regulatory regions, coupled with targeted Next-Generation Sequencing to quantify gene expression levels. By targeting edits to transcribed genomic regions present in mRNA, we can measure the effect of each edit on target gene expression by quantifying its abundance in RNA. In this manner, PETRA allows us to determine the effects of thousands of edits on gene expression in a single experiment.

We first demonstrated PETRA in Jurkat cells (a human T cell line) by targeting a genomic site in the 5'UTR (untranslated region) of 4 genes with key roles in T cell function: *IL2RA*, *VAV1*, *OTUD7B*, and *CD28*. We tested a total of over 10,000 edits comprised of 6-nucleotide insertions at each of the 4 sites. We discovered edits causing both large increases and decreases in gene expression. For example, specific edits increased the expression of *VAV1* by over 400% while others decreased the expression of *OTUD7B* to less than 1%. Variants with strong effects on expression were highly correlated across replicate experiments and successfully validated when introduced independently. To gain further insight into the mechanisms underlying the observed changes in gene expression, we developed a computational tool to detect new Transcription Factor Binding Sites (TFBSs) generated by the insertions. TFBSs are genomic sequences bound by proteins that regulate gene expression, called Transcription Factors (TFs). This analysis indicated that the TFBS for MYBL2 increased gene expression of *IL2RA*. To validate this, we inserted MYBL2 binding sites in 3 different positions of the 5' UTR of *IL2RA*. This experiment revealed positional and combinatorial effects of MYBL2 binding sites, with specific combinations leading to a 10-fold increase in *IL2RA* expression. Next, we applied PETRA to human primary CD3+ T cells to test the effects of 6-nucleotide insertions in the 5'UTR of *IL2RA*. We found insertions that modify *IL2RA* expression to 450% and 1% of original levels, establishing the method's utility in primary cells. Lastly, we are now adapting PETRA to evaluate the effects of edits on protein levels and cellular function.

By developing and applying PETRA, we have established a new means of identifying genomic edits that can be stably engineered into human T cells to produce desirable effects on gene expression. We anticipate PETRA will be highly valuable for understanding the interplay between DNA sequence and gene expression across a variety of biological contexts by enabling the generation of large datasets that link edits to functional effects. Such datasets can be used to train machine learning models to predict gene expression changes from DNA sequence genome-wide. Furthermore, as PETRA can be applied to human cells of therapeutic relevance, top-performing edits can potentially be used to enhance the efficacy of cell therapies.