Dear Dr. Perry,

Thank you for considering our manuscript as a potential paper on Nature Biotech. We especially appreciate the time you and other editors took to read and to assess our manuscript. While we agree that the manuscript described some of our findings after applying the newly developed method, to our understanding, the method we developed is, in fact, the main result being presented.

With TeXP, we are pledging for a dramatic change on how transposable elements and, in matter of fact, how any repetitive sequence activity could be transcriptionally gauged in regular RNA-seq datasets. The underlying message of this manuscript is that by leveraging how reads cross-map across families and repetitive instances, we can reliably quantify the activity of LINE-1 and other elusive elements of the human genome. This approach was never used before. It is also important to note that we validated our method with ddPCR and that, to date, not a single method to gauge LINE-1 activity has taken pervasive transcription into account. As it might be evident in Figure 2, we show that most of reads mapped to these repetitive elements are in fact derived from passive transcription. If methods do not take this into account, they will most likely overestimate the activity of repetitive elements in tumors and healthy tissues. We believe our method is not only a game changer in the Transposable Elements community, but it will also foment the development of similar strategies in other genomes and other elusive repetitive elements (which account, for example, to more than half of the human genome).

That is not to say that our findings after being able to reliably quantify LINE-1 expression are less important. The comprehensive landscape of LINE-1 activity in human tissues and tumors will be useful to inform many of the future analysis accounting from Somatic Structural Variation. Moreover, our theory that LINE-1 creates small insertions and deletions (INDELs) has been provoking a lot of enthusiasm. The proposed mechanism, as we describe in the manuscript, is reasonable: LINE-1 was previously shown to create double strand breaks in the genomic DNA and there is independent evidence that double strand breaks are fixed by NHEJ creating small insertions and deletions.  Our intention is to show that our method is powerful and can be used shed light on LINE-1, and other elements, biology. However, we could tone it down these findings to fit the manuscript into Nature Biotech.

Furthermore, it is important to stress that, as part of the ENCODE project, we believe that this paper could be a great addition to the whole package. Here, we are addressing the “dark matter” of the genome, which has been routinely ostracized. The ENCODE package might be the perfect venue to address the issue of how much pervasive transcription is impacting the estimation of the activity of these repetitive elements and how we can circumvent it. We will be looking forward to hearing from you soon.

Yours Sincerely,

Mark Gerstein