Review: Hwang, S., Kim, E., Lee, I., & Marcotte, E. M. (2016). Systematic comparison of variant calling pipelines using gold standard personal exome variants. *Scientific Reports*, *5*(1), 17875. https://doi.org/10.1038/srep17875

Foremost, I thought this paper did a very thorough job of testing and exposing the strengths and weaknesses of each of the variant calling pipelines including both the read aligners and variant callers. It was interesting that the two previous studies in this area used only a single data set for their testing, this later was shown to have an effect on the resulting recommendations[1,2]. In trying to understand how the variant calling pipelines were evaluated using precision-recall curves I found a very helpful forum post online that made the idea much clearer, helping me to understand the precision and recall "tradeoff"[3]. It seems from the paper that lower sequencing coverage is a big issue in variant calling, as it is in other areas of genomics, low coverage results in calls that are more likely to be false positives (e.g. sequencing errors). One of the most interesting results from the paper was the much higher amount of concordance across variant callers when looking across the different illumina datasets (~82-97%), this was in stark contrast to the levels reported from the previous studies (~57-70%) whereas the ion proton dataset shows lower concordance across all programs. I think this illustrates a valuable point when considering the results or recommendations of a given comparison paper, you should scrutinize the methods to see how it fits what you are trying to accomplish and the sequencing technology you have used.

References

1. Cornish, A. & Guda, C. A Comparison of Variant Calling Pipelines Using Genome in a Bottle as a Reference. BioMed Research International 2015, 11 (2015).
2. Highnam, G. et al. An analytical framework for optimizing variant discovery from personal genomes. Nat Commun 6, 6275 (2015).
3. https://www.quora.com/What-is-Precision-Recall-PR-curve