https://pubpeer.com/publications/D780DD5C4D665BBC415F45C707D0C4

Hi,

Thank you very much for communicating details about your research!

Because the journal doesn't have a comment system, I wanted to add a comment in this system.

First, here a comment on the earlier preprint:

https://www.biorxiv.org/content/10.1101/2020.04.29.068452v2#disqus_thread

So, I will try not to over-emphasize the question of *directly measured* genotypes versus Minimac4 imputed genotypes, since I asked that before.

I also have a comment on another recent paper (other authors describing lcWGS versus genotyping arrays):

https://www.cell.com/ajhg/fulltext/S0002-9297(21)00096-3#comments-heading

The questions posted in that comment are a little different (although some overlap the preprint version of this paper). If anybody cannot access the AJHG comment, I also have a draft posted here.

The Martin et al. 2021 paper mentions Gencove worked relatively better at lower coverage, and BEAGLE worked relatively better at the 4-6x WGS (with that being the highest imputed accuracy reported in Table S4 or Table S5).

So, I believe these would be the questions with updated information:

a) Since the preprint post, I have added at least 1 extra example of a lcWGS imputation having noticeable problems at 0.1x-0.5x. While hard to tell the exact genomic coverage (filtering adapters and non-genomic sequence) without the raw data, I am guessing the cat lcWGS may be roughly 0.2x-0.3x. I think might be roughly similar to the leftmost point in *Figure 1.*

   My impression from the Martin et al. 2021 paper was that some my ~0.5x lcWGS results (for myself and for my cat) might have been improved with **higher coverage lcWGS**. Do you think that is a fair assessment?

b) In the Martin et al. 2021 paper, the **GSA had the lowest performance among all of the arrays**. For example, in that paper, I would expect the Omni 2.5 array to outperform 0.5x to 2x lcWGS (with either Gencove or BEAGLE).

I believe this was also noted in the earlier preprint comment, although that was in reference to the EUR and ASW interactive plots from the GLIMPSE paper. However, to be fair, I am not sure where you would draw the line to be an unacceptably low correlation (for the less common variants), which I believe can affect the interpretation.

So, you do specifically mention the GSA array in the abstract, but I do not believe the title is universally true (even if you look at SNP chip genotypes that I believe include imputations).

Is there anything that I might have misunderstood?

Best Wishes,

Charles