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

Hi,

Thank you very much for putting together this paper!

I think the main reason for posting a response relates to **how "low" the coverage sequencing is (such as 0.1x-0.5x, versus 4-6x)**, and I think your **Figure 4 and Table S4/S5** may be matching my own experiences.

So, I will **first** explain my experiences and reason for posting this, and I will **then** ask more specific questions in terms of this study:

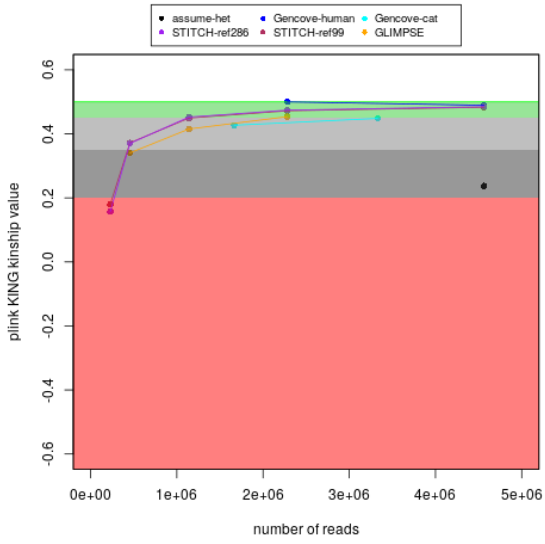
1) Problems with Nebula lcWGS results (before it was discontinued).

For example, this blog post indicates things like being predicted to *lack* an APOE E4 variant with "high confidence" the T/T genotype for rs429358 , when I in fact I have the C/T genotype.

The FDA MedWatch report that I submitted for my Nebula lcWGS results is registered in the MAUDE database under MW5093887.

To be fair, I thought some things were OK. For example, some broad ancestry applications described in this blog post, which uses some chromosome painting plots based upon code provided by the first author for this paper. However, I was concerned enough to say that I don't think the Nebula / Gencove genotypes that I was provided should be used for medical applications, especially if somewhat rare variants are involved.

2) Testing of down-sampling with my human and cat WGS data (with various imputation methods for human, and Gencove for cat):



The above plot and code is available to here.

That is more for self-identification (which I think is relevant for counting the genomic data as PHI), even if the utility for other applications is more limited.

Nevertheless, these are the possibly relevant points:

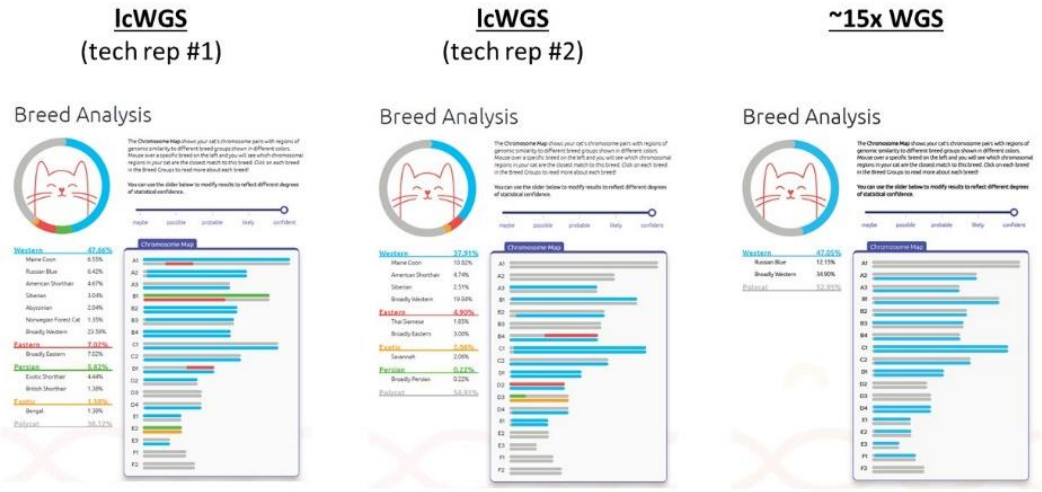
- The cat performance was lower than the human performance, which I guessed might relate to the cat variation not being understood as well as human variation (and/or I think the fragment size for that library was a bit on the small side).

- For some of the human imputations, I also tried to pick reference panels similar to my own ancestry (mostly European, with a little bit of African ancestry).

Point being, if **the reference variation was not understood and/or the reference panel was less of a good fit for the individual**, then it seems to me like the lcWGS results could be worse than the higher coverage results (between groups with different ancestries). For example, I think you could kind of see that with off-target reads from my Exome sample (with noticeably different ADMIXTURE assignments for me, depending upon the reference set).

3) **Noticeably discordant lcWGS reports** for my cat, at least some of which appear to be resolved with the use of higher coverage sequencing.

This is for cats (not humans), but I will try to briefly show the discordance below:



Since the details may be a bit off topic, I think it might be best if I refer to this blog post that relates to the 2nd and 3rd points.

Nevertheless, this is something that people (the pet's owners) can currently receive, and it relates to my concern in that sense.

So, for this paper, these are my questions:

a) Do you think the **~0.5x** sequencing is a fair representation of what you see currently available to consumers, across various companies? If so, do you think the "low coverage" being offered to consumers is possibly too low, and that relates to my own personal experiences with lcWGS data?

My concern is how this might be used to justify lcWGS results that are currently available, when you might be imposing stricter requirements?

b) For my human imputations, picking samples with similar ancestry seemed to make a difference. Are you saying that the imputation panels should be selected with **similar ancestry (different for different individuals)**? If so, does that mean individuals with mixed ancestry might want higher coverage sequencing results (and/or noticeable development is needed for different groups)?

Similarly, if ancestry was not known in advance, could that cause issues and/or increase costs? For example, would you then run multiple analyses with different reference panels and then select a "best" guess?

c) Is there a sense of what you think is most likely to be a **useful/robust application**, for what you are recommending as "low coverage"?

I am asking because I noticed the author list, and my understanding was that the samples were collected for the purposes of mental health?

In other words, I am concerned that misuse or misinterpretation of results that are not highly predictive can cause harm to individuals that may be unfairly receiving a stigma based upon a genomic lcWGS result (about a disease they may never have, or whose severity might vary over time with treatment, as with my personal experience). For example, I have been concerned about applications where I didn't see as much discordance as I have seen for my lcWGS (such as discussed in this blog post about the book *Blueprint*). Do you think that is a fair concern?

To be fair, I think medium-to-large copy number alterations could be reasonable with 4-6x WGS, but I think that would be a different set of benchmarks.

d) For your comparisons, do you always **include imputations** (both lcWGS and SNP chip)? For example, if you have previously characterized variants that were *directly measured* on the SNP chip that are approved by the FDA to return and take action upon, then it seems to me that the SNP chip might have higher specificity than the lcWGS (maybe even at higher coverage than my own lcWGS samples)?
In other words, for medically actionable results, I would probably group imputed SNP chip *or* lcWGS in one category, and directly measured SNP chip variants (perhaps for tested populations) in another category.

I am also assuming that the 1000 Genomes reference samples are used for the lcWGS imputations (as I believe is mentioned in "***Genotype refinement, phasing, and imputation***"), so that these really are independent lcWGS imputations. It is also possible that I might be slightly confused about the use of SNP chip data *without* imputation because of what I saw in **Table S7**, in which case I apologize.

That said, Exome genotypes are not being compared, and perhaps Exome and Amplicon-Seq genotypes might be a relevant comparison for directly measured SNP chip genotypes (and the paper already covers a lot)?

e) This is something that I have added to the post more recently (on 3/28/2021, instead of 3/26/2021). However, is the **variability in reads per sample** considered in the cost estimates?

For example, in my regular work, there is noticeable variability per sample. For example, if you want to say >95% of samples have >0.8M reads, then you might really have to aim for more like 2.0M reads on average.

So, are the calculations based assuming the observed and expected reads are a good match for a large number of samples, or are you calculating the cost with trying to guarantee that >95% or >99% of samples have above a certain threshold?

I don't think this matters so much with the higher coverage sequencing, but I think it might affect the 0.5x sequencing?

I think batch effects can have more of an impact on lower coverage WGS data, but that may be hard to simulate.

I apologize that this comment is quite long, but I very much appreciate any feedback that you can provide.

I also want to thank you again for posting this publication. For example, as I understand it, I think seeing the **4x lcWGS / 6x lcWGS / Omni 2.5 SNP** chip results in different populations was interesting.

Sincerely,

Charles