

Towards a medical grade human genome sequence

The substantial \$10 million purse of the Archon Genomics X PRIZE (AGXP) is being offered for the generation of rapid, accurate and complete human DNA sequences. Because so many genomics researchers have a stake, we offer to help with a process of community consultation to help evolve fair and efficient methods to validate contestant data for the competition.

We discovered that community comment using the public posting site in the archive *Nature Precedings* is a popular and efficient way to conduct peer review of certain papers. This process is important when we are proposing consensus standards for data release, data analysis and research practice (Editorial, "Discussing Standards," *Nat. Genet.* **42**, 915, 2010). We are now pleased to collaborate with the AGXP and the X PRIZE Foundation (p 175) to collect your ideas on ways in which to find the best practice in genome sequencing and assembly to be built into their validation protocol.

In this community review, we are asking those offering comments to trust that the AGXP authors will take all scientific comments into account. It is the journal's view that if the arguments are cogent and logical, that trust will be honored. Dissenting scientific arguments are welcomed by the corresponding authors and by us and may be posted to *Nature Precedings* (<http://precedings.nature.com/collections/human-genome-sequencing>) as comments or independent preprints. We will also post comments (you may choose to remain anonymous provided you identify yourself to the moderator) to *Free Association* (<http://blogs.nature.com/ng/freeassociation/>).

The AGXP is a contest as well as a standard-setting effort, so the advice from the research community will be interpreted and applied by the AGXP Foundation as they see fit. They have declared that the judging process will be as transparent as possible and we must accept that in the interests of fairness, we must allow the judges considerable discretion. As an example, the judges may need to keep secret some of the regions that will be sampled in the validation protocol, at least until the award is made.

How complete is a 'finished' genome? For discovery of human genetic variation and for association of gene variants to traits,

sequencing to increasing depth yields useful and asymptotically better results. Practically, sequencing efforts must stop somewhere and for prediction of traits and outcomes for an individual, the sequence generated must show not only all the alleles carried, but their genotypic context, including phase.

We anticipate that consensus reached for the purpose of judging the prize should also guide future best practice for individual genome sequencing. The fact that many will need to view and compare the same set of sequences in a short span of time will exert pressure to produce more efficient browsers and other software useful not only in meeting prize criteria for assembly and base calls, but for annotation and interpretation of the resulting high quality sequences.

Setting the choice of samples to be used in the contest through community consultation has potential benefit as well. We suggest that diverse groups not participating in the technological contest can gain from donating appropriately consented DNA samples for the competition because these will be sequenced many times over and thoroughly scrutinized during the competition. Partial sharing of samples provides a common standard and leverages the research in an obvious way. If ten groups were to donate just ten samples each, then each group would have access to a set of reference sequences ten times larger than they could have achieved by their own efforts. Add the prize contest and the reference set becomes many times deeper too.

Obviously, it is possible to make criteria more stringent, especially by suggesting validation tests to make the judging more onerous yet resulting in increased costs to run the prize. We very much hope that with many eyes on the prize, the reverse will be true and that stakeholders will streamline the judging process to make it more efficient and economical to achieve. ■