**Press, Carlson, and Queitsch response to reviewers (TIGS-D-14-00076R1).**

**July 23rd, 2014**

Reviewer’s comments are in Verdana font.

**Our responses are in Arial bold.**

Reviewer 1

Great paper and very interesting ideas.

My only complaint is that it is excessively negative about the state of STR genotyping and makes some statements about the methods and somatic variability that are not well supported with citation.

**We thank the reviewer for this assessment. We address these concerns point-by-point below.**  
  
Here are the comments:  
  
Page 4: Line 14 and 15: Technical error problem does not come from somatic variation (unless you are doing tumor sequencing).

**We respectfully submit that somatic variation is a long-recognized feature of STRs. We have added citations and moved up a citation for the earliest work on this subject that we could locate from later parts of the manuscript (Slagboom *et al.* 1991, *Mutat. Res.*). This study estimates “the occurrence of DNA sequence changes in mini- and microsatellite regions of the rat genome at an average frequency of 2.7 × 10-3 per analyzed DNA fragment” for primary tissues isolated from largely juvenile rats. A PubMed search yielded several other examples, some studies even use somatic mutations across STRs to reconstruct cell lineages or assess age (16261192, 20817752). See also PMIDs 20302627, 19210789, 24860168, and 1570846. There is, however, insufficient space to cite all of these references.**

Technical error in germ line measurement comes from biological error (due to amplification steps), sequencing error, and mapping error (usually this is reference bias). This measure of technical error has been explicitly quantified by pubmed ID 20980555 and 23090981. I also think it is a bit excessive to say that utility of the methods is limited. While not perfect, the current methods have a FDR of 5% or less. You might also check out the latest validation of STR measurements: <http://biorxiv.org/content/early/2014/05/01/004671>

**We appreciate the reviewer’s comments, and have altered the language of this section to portray current methods in a more positive light. We believe that our portrayal of technical error is much the same as the reviewer’s, and have added further text and citations to reflect the reviewer’s points. Nonetheless, we might contend that in the presence of somatic variation, the technical error from the sources described is somewhat harder to estimate in an error model (see below for a continued discussion of this point).**

**We are not sure of the source of the 5% FDR statistic- it may originate from the bioRxiv preprint alluded to by the reviewer, which states that 95% of (necessarily homozygous) Y chromosome calls are accurate for the LobSTR method. We do not dispute this number, except to note that not all of the human genome is homozygous. The Highnam *et al.* RepeatSeq paper has some potentially relevant statistics for assessing mapping accuracy, but they appear to be based on simulated reads, which do not take into account all sources of error.**

**Our point is less that the methods are not useful than that their imperfections are best ameliorated by moderate-to-high coverage genomes with moderate-to-long read lengths, which don’t seem to exist in large numbers at the moment. In this sense, the methods are ‘limited’. See for instance the Willems *et al* preprint, where confident results are primarily population-level. We have modified the text to attempt to communicate this more effectively.**   
  
Page 5: Line 9 and 10: I believe that 9207800 also reported STR mutation in RunX2 that correlated to craniofacial phenotypes in humans. Also in dogs, runx2 repeats correlate with craniofacial phenotypes 15596718

**We appreciate the reviewer’s attention to detail- however, it appears that reference 9207800** **is talking about SNV missense variants in the gene. We have added a citation to the Fondon and Garner 2004 PNAS paper here.**

Page 10: Line 11 and 14: I think it is an overstatement to say that the requirement "severely" limits STR length. Likewise, evidence suggests most STRs can be captured with 75bp or larger reads. I would not exclude 75bp from the analysis. Both of these points were quantitatively measured in 22427938. See Figure 1 and 2.

**We have removed the word ‘severely’. 76 bp reads were included in the analysis – a close examination reveals that the figure 1 referred to in 22427938 leads to a similar interpretation to our figure 2 (though we are more generous, because we assume zero error from any source).**

**We do not dispute that a majority of STRs in a particular reference genome (for instance, in Drosophila, which has relatively short STRs, as indicated in 22427938) may be accessible using 76 bp reads and deep sequence coverage- our point is that population-scale sequencing data today, for humans at least, is decidedly not deep enough. Moreover, at least some of the biologically interesting repeats have a natural allele size range that cannot effectively be ascertained by 76bp reads (consider our favorite, the *ELF3* STR, which has a repeat that varies from 21bp to 87 bp). This of course entirely ignores repeat hyperexpansions (which we intentionally treat lightly in this review).**

Page 10 Line 20: If substantial somatic variation was present in the sequenced sample you would expect to capture it in the error models based on homozygous and haploid genomes as well, right? Why do author presume high somatic STR variation? I think this is in conflict with the literature. While some tissue types in the brain and liver exhibit substantial somatic variability, theses tissues are not typical utilized to for germline sequencing. STRs are fairly stable in blood and saliva, at least in non-cancer genomes.

**The reviewer rightly raises the point of “high somatic STR variation”- what is “high”? Based on the existing literature, we assume that A) STRs have a high expansion/contraction mutation rate relative to mutation mechanisms at other loci; B) somatic mutations occur in primary tissues; and C) somatic STR variation is to be expected at a high level relative to somatic SNV variation. The various articles referred to throughout support our expectation; STRs are expected to, on average, accumulate somatic mutations 104-105 times faster than SNVs (16261192).**

**To our knowledge, there is no study that addresses genome-wide somatic STR variation in different primary tissues, nor are we aware of a literature that prohibits STR variation in these tissues. However, there are many observations of locus-specific somatic STR variation that have been chronicled. For example, one study using a mouse CAG repeat instability model finds similar levels of STR instability across a variety of tissues (though, as the reviewer suggests, liver is elevated, PMID 20302627). Another study finds a somatic mutation rate in lymphoblast-derived cell culture of 2\*10-6 / cell division / dinucleotide STR, suggesting that somatic STR variation is to be expected among leukocytes (10749142). As detailed above, there are a variety of studies that report somatic STR variation in primary tissues from several organisms (PMIDs 16261192, 20817752, 20302627, 19210789, 24860168, 1570846).**

**We agree that a model may in principle capture the sum of technical and somatic variation, though we imagine that it would be difficult to generalize (unless we were sure that all STRs had rates of somatic variation on the same magnitude as technical error or less). Consider that somatic variation should contribute to overall sequencing variation by a different mechanism from technical error; it is basically ‘heterozygosity’ of unpredictable proportions. In sequencing, this scenario is similar to the difference between mutation mechanisms described in the classic Luria-Delbrück experiments: you are starting out with a heterogeneous population before your selection (*i.e.* for STR genotyping this would be tissue sampling, DNA extraction, library preps with amplification), which can lead to ‘jackpot’ mutations. We can of course attempt to model this, but we do not think the current error models consider this scenario. In any case, we do not believe that current data is sufficient to really address this point, because relatively few studies have addressed somatic variation, and those do not model technical error other than mapping (see 24209623). We look forward to a study that specifically addresses this question.**

**We also think that somatic variation, in addition to being a source of error in detecting germ-line alleles, is itself biologically interesting. Thus, even if the error model is working, it is throwing out interesting data (note that this is the thrust of our argument in the section “Somatic STR variation may be a sensitive marker for increased disease susceptibility”).**

**In the revised manuscript, we have modified the text to be clearer about these various points and included such citations as possible (given space limitations) to support our argument.**

Page 11: Line 1: Ref #35 is cited for calling repeats with as low as 1 read, however RepeatSeq requires 2 reads to call a genotype in diploid genome. It is true that Ref #36 permits 1 read (LobSTR).

**We have modified the text to clarify this point (note that Table 1 also describes this difference).**  
  
Reviewer 2  
This ms. clearly and forcefully presents the need for increased investment in accurately sequencing tandem repetitive DNA, to enable reliable research into the causal association between short tandem repeats (STRs) and phenotypic traits, including disease susceptibility. The authors first review compelling evidence that STRs can and do play a significant functional role (the "overdue promise" of the title). This ms. then explains in some detail the difficulties inherent in working with these "slippery" sequences. Researchers who would attempt genome wide association studies of STRs thus face a serious challenge. The authors finally suggest important questions to address while developing tools to enable future progress.   
  
This referee would be pleased to see this essay published directly in its present form.

**We thank the reviewer for their appreciation and approval.**

Nevertheless, I would also offer the following suggestions for consideration by the authors (and editors).  
  
The "Outstanding Questions" might be presented with a bit more force and clarity [e.g., at least in my opinion the words "How can we…" are unnecessary and vague, softening the effect]. I would also suggest distinguishing questions of method from questions of substance. Thus:  
\* Methodological questions:  
- In light of widespread epistasis, what statistical and experimental tools are needed to establish STR variation's effects on phenotype?  
- Can inexpensive, accurate tools be developed for germ-line and somatic STR genotyping?  
- [The meaning for the third question in the ms. is not clear to this referee.]  
\* Substantive questions  
- Will somatic STR variation prove useful as a readout for disease susceptibility?  
- [More than one substantive question is surely appropriate. I suggest another obvious possibility.] Will STR variation fulfill the promise of accounting for a significant proportion of heritable variation? 

**We have edited the outstanding questions to reflect the reviewer’s concerns, and we thank them for their suggestions.**

In spite of the title, this ms. provides little sense of how long the "promise of STRs" has been "overdue." It might be appropriate to offer a bit more explicit history. It shouldn't take more than an introductory sentence or two, to mention that the potential promise of STRs to explain missing heritability was evident already by the early 1990s, as soon as the causal basis for genetic anticipation was identified in the so-called triplet-repeat diseases. By 1997, TiG could publish "Simple sequence repeats as a source of quantitative genetic variation," and that same year another Elsevier journal published "Evolutionary tuning knobs." Now, two decades later, we still have little more than a grab-bag collection of assorted individual cases rather than a well-supported research program to study the role of STRs.

**We have added text to specifically address this comment, and a reference to the 1997 Kashi paper, as well as a reference to genetic anticipation as the origin of the interest (which was also added to the Glossary at the end of the paper). We agree that this context more explicitly emphasizes the historical significance of recent developments in typing STRs.**   
  
Also unaddressed in this review is the extent to which the "promise of STRs" remains unfulfilled, at least in part because many researchers remain blind to this promise. This referee recently observed, at an international conference, how often the word "mutation" was used without qualification, as if it were simply a synonym for "single basepair substitution" (i.e., as if other styles of mutation, such as expanding and contracting STRs, either did not exist or were unworthy of notice). Similarly, the term "mutation rate" is often used as if one average rate were governed simply (simplistically) by a few genes that encode genome-wide processes of DNA replication. Both of these habits can foster blindness to patterns, such as STRs, which follow different rules from single-basepair mutation. 

**We appreciate the author’s insight, and hope that our manuscript may in small part educate the scientific community with regard to the potential importance of STR-phenotype associations. Specifically regarding the SNV/STR comparison, throughout our manuscript we have attempted to be explicit in distinguishing between STR mutation/variation and other varieties.**

Perhaps prudently, given limits of space, the authors of this ms. explicitly acknowledge that their review will avoid "evolutionary considerations." But evolutionary issues are deeply entangled with such blindness to the promise of STRs. For example, to quote a recent and quite relevant report [Elmore MH, Gibbons JG, Rokas A, 2012, Assessing the genome-wide effect of promoter region tandem repeat natural variation on gene expression, G3 2:1643], "in principle it is unlikely that a type of variation with high mutational instability, like [tandem repeats], would be a major contributor to phenotypic evolution." If this assertion were true, as Elmore et al. evidently believe, then the several well-known examples of phenotypically significant STR variation (which are duly acknowledged by Elmore et al.) must be unusual special cases and not indicative of a general pattern that has "overdue promise."   
  
However, this "principled" assertion by Elmore et al. depends on an old argument that applies only under highly restrictive conditions and is not irrelevant to understanding STRs. Quoting again from Elmore et al., "a well-established and supported tenet of evolutionary theory is that, because most new mutations are deleterious, selection in all organisms will act to reduce mutation rate toward the physiology- or selection-imposed minimum [Lynch 2010]." It is telling that the source cited by Elmore et al. in defense of this "well-established and supported tenet" [Lynch M, 2010, Evolution of the mutation rate, TiG 26:345] is quite explicit that "The focus here will be on base substitution mutations alone..." (i.e., that conclusions herein do NOT apply to STRs). Yet far too many biologists accept this "tenet" as a blanket generalization as if it were applicable to any mutational mechanism. It is also telling that this TiG review by Lynch, in spite of its explicitly limited scope, itself has a blanket title, "Evolution of THE mutation rate" (emphasis added). Hence the blindness is perpetuated.  
  
I leave it to the authors (and TiG editors) to decide whether addressing this concern would strengthen, weaken, or simply complicate this valuable review.

**As the reviewer has noted, we chose to avoid an extended evolutionary discussion specifically because these important ideas will require a substantial revisiting of such concepts as mutation rates and how they are related to the evolution of phenotypic novelty. However, in the revised manuscript, we focus on certain practical problems of the field, due to limitations of space.**

**We have ourselves experienced a ‘trivialization’ of STR variation in paper and grant reviews (“but that’s only a model organism phenomenon”, “can’t you just use linked SNVs?”), so we appreciate the reviewer’s ardor. It is our earnest hope that our contributions will stimulate further discussion of these issues, which biologists will have to tackle one day.**