Response to reviewers: Rival *et al.* manuscript GENETICS/2014/167866.

We thank both reviewers for their thoughtful and knowledgeable evaluation of our manuscript. We have taken their comments into account to make various corrections, for an improved result. The referees have brought up some good points and we appreciate the opportunity to clarify our research objectives and results. As detailed below, we have responded to all comments provided by the referees and have made necessary changes.

Throughout, reviewer’s comments are in Helvetica, and **our responses are in Times New Roman Bold**.

**Reviewer 1:**

**We thank this reviewer for their attention to detail and comprehensive evaluation of our manuscript. We respond to specific points below.**

***Major Comments:***

**General**

The poly-Q tracts in PFT1 are translated from heterogenous codons on the cDNA sequence. Thus, whereas there is a clear homorepeat (“QQQ...”) in the protein sequence, there is no clear triplet-repeat on the coding sequence. Rather, as proposed by tandem repeat detection algorithms such as T-REKS (http://bioinfo.montp.cnrs.fr/?r=t-reks/) or TRF (http://tandem.bu.edu/trf/trf.html) run on the PFT1 cDNA sequence, there seem to be tandem repeats with repeat units of different sizes, e.g. ~18. This might have implications on the discussion of the results. First, it seems the tandem repeat might be rather a minisatellite (~ >= 10n) than a microsatellite (~ <10n), or a combination of both, so the naming STR could be misleading. Tracks with a single codon are much shorter than ~90aa.

The composition of the coding sequence in terms of codons is very important for several reasons: It will infuence the tandem repeat mutation rate, with more identical codons expected to undergo more tandem repeat duplication/deletion mutations. Along the same lines, an accumulation of diverse codons might be resulting from positive selection to inhibit further duplication/deletion mutations to stablize the length of the repeat tract. This might be in line with the authors's main conclusion “the PFT1 STR is constrained to its approximate wild-type length by its various functional requirements” (page 3, 18). Also, the codon composition is a means to understand the evolutionary history of the tandem repeat. It might be possible decipher the most recent repeat duplication/deletion mutations, and by this pinpoint where exactly they took place within the protein sequence.

The authors mention that the PFT1 gene encodes for two splice forms (page 16, 2). Thus, the exon structure might be important for gene function. At the same time, the exon structure is important to determine tandem repeat duplication/deletion mutability. Tandem repeat duplications/deletions using typical, fast mechanism (such as replication slippage, unequal recombination) are not expected to occur across exon boundaries. Thus, introns may subdivide the poly-Q tract into multiple poly-Q tracts, that mutate independently. This, together with the codon composition, mighy strongly infuence the poly-Q tracts mutational behaviour.

In sum, including the composition and structure of the codon sequence into results/discussion could help to strengthen the manuscript:

* Is the AA homorepeat truly a DNA STR (microsatellite) on the sequence level?

**We consider that the question of microsatellite vs. minisatellite for *PFT1* in the reference genome is to some extent a matter of taste- it could perhaps be fit into either category. One can find an impure large repeat structure, or a small impure repeat structure, or a series of smaller Q-encoding microsatellites interrupted by other amino acids. Given that the variation we do observe consists of single amino acid insertions and deletions, we came to the conclusion that it is behaving like an impure STR of 3bp units. While this discussion may be of interest to a small sector of readers, we considered that 1) it was not particularly of interest to the general audience of *Genetics,* 2) our conclusion is a heuristic conclusion at best (chiefly serving to call into question the usefulness of the various categories), and 3) a variety of databases (TAIR/UniProt/NCBI/etc.) exist which have complete information on *PFT1* sequence, so an interested reader dissatisfied with the notion of an “impure STR” is empowered to make their own judgment. We also include our own raw sequencing data, which allows a reader to assess the question across *A. thaliana* strains.**

* What mutation rates are expected considering the heterozygosity of the codon composition and the exon structure (Legendre et al (page 4, 22) may need revisiting)?

**It is specifically using the statistical framework (SERV) described in the paper that the reviewer mentions that we came to the conclusion that the *PFT1* STR should be highly variable. We note that this framework does not explicitly estimate a mutation rate, though by eye on Figure 1D of the cited paper, we might expect a mutation rate of 1E-5+ / generation. However, there is not an obvious conversion between the SERV VARScore and mutation rate; all indications are that variability of specific STRs is highly context-dependent (see Eckert and Hile 2009, *Molecular Carcinogenesis*). The SERV output is intended more as a score for how likely an STR is to be variable, rather than the specific value of a parameter such as mutation rate. We are unclear what the reviewer means by ‘heterozygosity of the codon composition’ and the ‘exon structure’; the various strains of *A. thaliana* are inbred and thus expected to be homozygous. It is true that the Qs in PFT1 are encoded by both CAA and CAG codons. However, these themselves tend to be clustered with like codons. Again, SERV and similar algorithms tend to take into account repeat purity in their calculations, but the VARScore for *PFT1* is still very high relative to most repeats (particularly shorter ones like *ELF3*). Unfortunately, our unpublished database of these estimates, and the SERV algorithm itself, are not available at the moment due to security problems on our collaborator Matthieu Legendre’s servers (see below).**

* Can we pinpoint recent mutations within the poly-Q tract thanks to the codon sequence? E.g. have there been deletions/duplications of different lengths? In particular, can we pinpoint the mutations that occurred between the different A. *thaliana* strains (Suppl. Table S1)?

**The observed variation appears to be from the addition and deletion of single codons. However, given the relatively small sample size, we hesitate to give this as an unqualified answer, and the ancestral sequence is not known.**

- A plot showing the position, composition, exon structure of PFT1 may improve clarity.

**We have included a supplementary figure that shows certain notable characteristics of *PFT1*.**

The authors compare the length of the poly-Q tract across A. Thaliana strains for PFT1 (and ELF3) using Sanger (dideoxy) sequencing.

This approach is very useful to estimate the evolutionary conversation of the PFT1 STR:

Conservation across these strains would be a good indicator that little variation is expected on the population level. On the other hand variation across these strains would render variation on the population much more likely.

However, the authors do not discuss which part of the poly-Q tract is measured. As they mention before, the poly-Q tract is interrupted at multiple locations. Therefore, in the introduction, the authors refrain from presenting a precise range of poly-Q tract lengths (page 4, 21). To improve clarity, the authors may want to explain their result in more detail, and also discuss the used method and its reliability in more detail. Can the sequencing data be provided?

**The entire polyQ is measured by Sanger sequencing. The observed range of amino acids in the STR is 88-90, as displayed in Table S1. It is well-known that Sanger sequencing is the ‘gold standard’ for sequencing STRs. We have included the raw sequencing data as a supplementary file.**

Also, if the precise resulting numbers can be trusted, there seems to be some variation in the polyQ tract. Is it possible to tell where duplications/deletions occurred within the sequence? May duplications/deletions of single amino acid have phenotypic effects?

As mentioned above, a possible approach to compare evolutionary conservation of poly-Q tracts would be sequence comparison on the codon-level: Fast-mutating homorepeats are expected to be encoded by a single codon, whilst conserved homorepeats are expected to be encoded by different codons, to slow down the tandem repeat mutation rate.

**See general comments above. Variation of single amino acids may well have phenotypic effects, as has been shown in a variety of organisms (see for instance Fondon and Garner 2004, PNAS). However, variation on this scale is not the focus of our study. As noted above, the PFT1 polyQ is encoded by various Q codons. However, given the size overall of the polyQ domain, there is still dramatic variation expected by at least the SERV algorithm. We are not aware of a published approach such as the reviewer outlines. While this method is potentially interesting, developing and testing such an approach is not within the scope o this paper, given the existence of the well-described and accurate SERV algorithm.**

**Discussion**

*Page 14, 22: [...] while STR alleles perform worse with increasing distance from this length “optimum”.*

The authors name the wildtype TR length a length “optimum”. This interpretation seems far-fetched. First of all, only very severe changes in the TR length have been analysed in this study. The effect of slight changes (i.e. mutations/losses of single amino acids) has not been studied. Also, the authors might want to be careful to use judgemental descriptors such as “optimum”. The presentation and discussion of results should be adapted accordingly.

**We use the term “optimum” purely as a relative measure among studied STR alleles. We appreciate the reviewer’s point, which is the reason that we present the word in quotation marks (“”), indicating that it is to be taken figuratively.**

***Minor Comments***

**General**

Is it possible to deposit raw data used to create the results Figures openly accessible?

**We have included the raw phenotypic data in an additional .xls file.**

**Abstract**

*Page 3, 2: Despite their high mutation rate, some STRs show little to no variation in populations.*

Rather: Despite *these (generally)* high mutation rates,

**To the extent of our knowledge, all described STRs have mutation rates higher than (for instance) single nucleotide substitutions. This is a simple consequence of their architecture and the process of DNA replication. In light of this, we do not feel that it is necessary to hedge in this fashion.**

*Page 3, 15:*

Rather: establishing the function importance of the STR domain.

**We are unsure of the suggested change; this appears to be a quote from our text.**

*Page 3, 11: Transgenic plants carrying the endogenous PFT1 STR generally performed* *best across adult PFT1-dependent traits.* What is meant by *best* performance?

**As laid out in various places in the text, the ‘best’ is ‘most wild-type’, or ‘most effectively complementing *pft1-2*’. We have included another reference to this definition here.**

**Introduction**

*Page 4, 8: For example, in Saccharomyces cerevisiae natural polyQ variation in the FLO1* *protein underlies variation in focculation, which is important for stress resistance and bioflm formation in yeasts (Smukalla et al. 2008).*

The variable tandem repeat in FLO1 is of length 45aa, and therefore no microsatellite. There is no mentioning of polyQ variation in (Smukalla et al. 2008).

Please confer:

http://www.uniprot.org/uniprot/P32768

Verstrepen, K. J., Jansen, A., Lewitter, F. & Fink, G. R. Intragenic tandem repeats generate functional variability. Nat. Genet. 37, 986–990 (2005).

Also: This excellent review describes among others cases of micro- and minisatellite variation in coding sequence with a phenotypic effect:

Richard, G.-F., Kerrest, A. & Dujon, B. Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. Microbiol. Mol. Biol. Rev. 72, 686–727 (2008).

**We apologize for the oversight- the Verstrepen et al. 2005 reference was the intended paper, and is now cited instead.**

*Page 4, 12: elf3 loss-of-function phenotypes in a common reference background*

*(Undurraga et al. 2012).*

Please make sure to use upper/lower case consistently for protein names.

**This is standard *A. thaliana* mutant nomenclature.**

*Page 5, 1: http://www.igs.cnrs-mrs.fr/TandemRepeat/Plant/index.php*

The link seems to be broken (checked July 11th 2014)

**We have consulted with the maintainer of this unpublished resource (Matthieu Legendre, CNRS), who explained that the CNRS has been subject to electronic attacks and the servers have been down for an unexpectedly long period. A secondary host has volunteered to put up the resource, but has apparently not done so yet. The resource should be back online with time at the URL in question. Links were taken from this page: <http://www.kuleuven.be/verstrepen/en/tools>, which gives a rough description of the resource.**

*Page 6, 6: [...] the known propensity of polyQ tracts for protein-protein and protein-DNA interactions [...]*

What is meant by this subsentence? Could the authors please add relevant sources?

**We have included the relevant references.**

**Results**

*Pages 9-11: Figures 1-3:*

It is not clear how the results were split into Fig. 1-3. The authors may consider to merge all similar results (Fig. 1B, 1D, 2, 3A), to allow for easier comparison and decrease redundancy in results description.

**We debated this question, and decided that this format best served the presentation of our argument, by separating the phenotypes according to their expected physiological significance (i.e. flowering vs. shade avoidance vs. early seedling).**

E.g., one question would be why no data for petiole length under SD conditions is presented.

**These measurements were not performed.**

Also, the authors repeatedly use Tukey's HSD test. Can they explain why this is the appropriate test to analyse their results? In Figure 3B, would it be possible to indicate the NaCl concentration in the plot to improve clarity? Preferably, the same representation style as for (Fig. 1B, 1D, 2, 3A) may be used, and be merged with these.

**The Tukey HSD test is a standard post-hoc test for the differences of groups in the analysis of variance (ANOVA). We have included a short explanation in the Methods. The NaCl concentration is referred to in the figure legend.**

Can the authors explain the results for the empty vector control (VC) in Figure 3B?

**A previously described phenotype of the *pft1-2* mutant (i.e. the VC) is failure to germinate under high salt conditions. We have added a reference for this observation.**

*Page 13, 1; Figure 4*

To improve clarity, the plot and the description of the plot may be changed. E.g., better names could be chosen for the eight plotted phenotypical traits. Also, the exact position of the mean values for the eight A. *Thaliana* strains is not clear. It may be helpful to label all axes usefully. There is a sentence duplication typo within the fgure text (concerning red/blue arrows). It is not described what exact values were chosen for the PCA. The raw data is not available in supplementary material.

**The mean values are indicated by the positions of the text (0R, 1R, etc.) in the plot. The axes are described in the figure legend. The phenotype labels are unambiguous and informative abbreviations of the phenotypes that are also described in the legend- we are willing to change them if it will improve the figure, but there is not an obvious way to do so without detracting from the interpretability of the figure. The “duplication” we felt was necessary to fully and unambiguously describe the phenotype vectors and the meaning of their coloring. We have included the raw phenotype data (and the specific dataset of means used for PCA) in a supplemental file.**

**Discussion**

*Pages 17 2: Considering again the possibility that more conserved coding STRs have distinct functions from non-conserved STRs (Schaper et al. 2014).*

The mentioned study on protein function (in Schaper et al 2014) does not focus on proteins with STRs, but on proteins with tandem repeats with repeat unit length >= 15aa. To improve precision, the authors may want to replace STRs with tandem repeats in this sentence.

**We thank the reviewer for catching this error. We have corrected the text accordingly.**

**Methods**

*Page 18, 4: (also: Suppl. Mat.)*

Articles and spaces should be used consistently.

*Page 20, 8: [...](phenotypes are not measured on the same quantitative scale)*

Clarity and precision could be improved in this paragraph (See also comments on Fig. 4).

**We have altered the text to improve the clarity of this description.**

**Reviewer 2:**

**We thank this reviewer for their thoughtful evaluation of the experimental data, and in particular their background knowledge of PFT1 biology and literature.**

This manuscript by Rival et al nicely addresses a question of general   
importance in biology. How a protein subjected to molecular constrains evolves.  
  
However I have some concerns that may be addressed with relatively simple   
experiments.  
  
1) I’m concerned about the lack of complementation of seedling phenotypes. The   
authors have used a single allele of med25/pft1 for these experiments. There is   
still a chance that another not identified mutation in their lines, unrelated   
to pft1 could be responsible for these phenotypes. They obtained the alleles   
from ABRC insertion mutant lines and this is always possible. It should be easy   
to compare WT, pft1-2 and another allele of pft1 in the same conditions, to   
confirm that the observed phenotypes are due to a mutation in the PFT1 gene.

**The lines expressing the WT repeat do in fact complement well for the adult traits.**

**The *pft1-1*, *pft1-2*, *pft1-3* mutants have all been used in previous studies, and their phenotypes are very similar (Sundaravelpandian *et al*. 2012, Cerdan and Chory 2003). Furthermore, flowering time phenotypes for *pft1-2* and *pft1-1* are the same, so we are confident that the phenotype we observe is due to the *pft1* mutation alone.**

**Furthermore, it is known that different mutant alleles of the same gene can give rise to different phenotypes. It is not clear that observing different phenotypes for another *pft1* mutant would necessarily indicate that the phenotype we observe is due to another mutation, as opposed to *pft1* hypomorphism. To be entirely certain, we would probably need to reconstruct all transgenic lines in a different mutant background, which we consider to be too burdensome. As suggested in the text, we ascribe the differences to the fact that the *PFT1* gene has two different splice forms, of which we only complemented one.**

2) For hypocotyl length measurements, the authors used dark grown seedlings. I   
think the manuscript could be improved by performing simple light-response   
experiments, similar to those reported by Klose et al 2012 and Cerdan et al   
2003. The small differences they observe in dark grown seedlings could be   
indeed due to different seed batches, whereas a response to light (Red and far   
red light) would be more robust. The authors could also express their data   
relative to dark controls.

**Reviewer #2 is concerned by the lack of complementation of *pft1-2* by our transgenic lines and think that different batches of seeds could cause it. He suggests looking at light-grown seedlings to prevent the effect of different seeds batches.**

**We consider it unlikely that differences in seed batches are responsible for the effects observed. For example, the Col-0 and *pft1-2* line seeds were harvested at the same time, so the difference between them cannot be attributable to a batch effect. It is true that the seeds for the transgenic lines were harvested somewhat later, but it would be strange that the consensus of multiple transgenic lines is the same as one of the other batches. Furthermore, if the dark hypocotyl phenotype were so sensitive, one would expect substantial differences, or at least more variance, between the different transgenics, which we do not observe.**

**While light response is an interesting trait, it is not clear to us what it adds in addition to the various other phenotypes we have assayed. For germination, for instance, there is a dramatic difference between WT and *pft1* that we are able to mostly rescue with most of the lines (Figure 3B). While the hypocotyl data on its own may not be entirely convincing, we point to the large number of other phenotypes which do complement in our transgenic lines (even early seedling phenotypes). If we were specifically interested in the function of PFT1 in the biology of light response, photomorphogenesis, or stem elongation, the suggested experiments would be natural extensions. However, we are more interested in the performance of different STR variants in a diverse collection of PFT1-related phenotypes, for evaluating the importance of this conserved domain organismally.**

**First, we look at hypocotyl from dark grown seedlings because they are longer than the ones from light grown seedlings which enable to see small length variation. Light grown seedlings have short hypocotyls with very small length variation and thus the length differences will be even harder to observe. Since we don’t see differences in hypocotyl length of dark grown hypocotyls we don’t think that we will see differences by looking at hypocotyl from light grown seedlings. Furthermore, adding a 7th phenotype that we are unlikely being able to measure will complicate the manuscript instead of improving it.**

**Secondly, even though transgenic lines seeds are from the same batch but from a different batch than *pft1-2* and col-0, these seeds batches have been used for all the other experiments presented in the manuscript. In the others experiments we do see a complementation of *pft1-2* by the WT repeat, so it is unlikely that seed batches are why we can’t complement our mutant phenotype. In addition, we have three different transgenic lines per construct which show the same phenotype.**

3) For germination experiments: ¿Were seed batches obtained in the same   
conditions? ¿How many seed batches were used?

**As mentioned above, seeds of our transgenic lines are from the same batch (collected at the same time) but differ from *pft1-2* and Col-0 batches (which were in turn collected together). All batches are derived from plants grown in soil in a plant room under long days. All the lines have been grown in the same conditions. The batches used for the germination assay are the same as the ones we used for all the other assays. All the lines have a 100% germination rate in control conditions. Consequently, we observe that variation within like batches is in fact largely due to genotype (Figure 3B).**

3) They root length phenotypes could be discussed more in the light of a recent   
paper by Raya-Gonzalez et al (2014) which seems to differ.

**We do not think that our root length phenotypes are comparable to those of Raya-Gonzalez *et al* (2014) because our plants were dark-grown, whereas theirs were light-grown. This obviously has substantial consequences for root elongation.**

4) In page 5, last part of first paragraph is unclear (starts   
with “specifically, STRs that are conserved…”

**We wanted to point out that different class of STR exist in genomes. That is to say, STRs that have been conserved during evolution across speciation events *versus* STRs that have been gained/lost recently. Among *A. thaliana* STR-containing genes, we attribute PFT1 to the first class and ELF3 to second.**

4) minor points: genes are not always in italics, photoreceptors should be in   
lowercase letters (phyB vs PHYB, the apoprotein in capital letters). File S1   
could be converted to pdf to avoid changes in format. Could be also colored   
somehow.

**The text has been modified to make suggested changes this comment.**