Tenaillon et al. present results from a transcriptomic analysis of an artificial selection experiment for flowering time in maize. The authors perform RNA-seq of the shoot apical meristem of artificially selected progenies of two selection lines. The authors then identify differentially expressed genes by using variance partitioning and contrasts, and validate their RNA-seq results using RT-qPCR. Overall, the experiment provides quite a unique material for asking interesting evolutionary questions regarding divergence in gene expression as a result of artificial selection, but I found that the authors could take more advantage of this experimental setting in the discussion of their results. Overall, I enjoyed reading the manuscript and have only a single rather technical major comment. I found that the manuscript is overall well written, but that sometimes the main message is lost in the numerous acronyms and technical terms used in describing the results (e.g. "The Status category combined DE genes whose expression varied between two SAM status either in the F252 ([StatusF]) or MBS ([StatusM]) genetic backgrounds." , or in general terms like "[SelM][StatusProg] category"). I would encourage the authors to try minimizing technical expressions and bring out more of the interesting biological and evolutionary results.

Major comments:

p. 7: After trimming, filtering and mapping steps, we recovered between 29.87% and 47.22% of the reads (Table S1) that were used to estimate gene expression.

Question: I'm surprised about the low percentage of retained reads after mapping and trimming. I would like to see a summary of the different criteria that account for reads to be not mapped/filtered out (e.g. are most of the reads of bad quality or otherwise unmappable?).

I would in addition like to see the authors address the potential problems that such low percentage of usable reads pose: Are transcript models well-covered by sequencing reads, or is mapping coverage "patchy" over gene models? The concern here is that if coverage is unequal over gene models it's unclear how biologically relevant the observed "expression differences" actually are.

Minor comments:

page 3: One important observation from these experiments is that the response to selection is generally steady over generations …

Question: It's unclear what's meant by "steady".

p. 4: This is particularly intriguing for the Saclay's DSEs that started from inbred lines with limited standing variation (<1.9%) and evolved under very small population size.

Question: It's unclear what's referred to by "<1.9%".

p. 4: The combination of both new mutations and standing variation is consistent with the known complexity of flowering time determinism in maize, and a high mutation target, i.e. >100 loci (Buckler et al., 2009).

Question: I understand from this sentence that the assumption is that changes in the expression of most of the DE genes are due to new mutations. In this context what's the authors take in the proportion of DE

genes showing selection effect - is this more than expected by chance, suggesting convergence? Or is expression divergence mutation-limited in this system?

p. 5: Second paragraph about candidate genes.

Question: I wonder could the authors come back to these genes and mechanisms in the discussion - I have trouble to follow how these genes were relevant in the light of the results.

p. 7: FT occurred at the same plant developmental stage in Early (FE) and Late (FL) genotypes in F252 (8 visible leaves), but occurred at an earlier plant developmental stage (9 visible leaves) in Early (ME) than in Late (ML) MBS progenitors (10 visible leaves).

Question: This is an interesting observation and I wonder if this reflects different mechanisms for response to selection? How this difference could be reflected in the different DE genes under selection in the two lines?

p. 8: Altogether, our methodology revealed repeatable patterns and a visible signal of differential gene expression.

Question: What is meant by "visible" signal?

p. 8: We performed 27 contrasts to detect DE genes (Table 2).

Question: It would be helpful for the reader if these contrasts would be summarized in some sort of schematic form. In addition, what is the contrast in SAM status (floral vs vegetative?)?

p. 9: For MBS, we found 446 DE genes within the Selection category. For F252, we found 2,120 in the Selection category, that comprised 748 DE genes between Early and Late or Very Late F252 progenitors ([SelF]). Considering both F252 and MBS, there were 2,451 DE genes falling into the Selection category (Table 2 & Table S4).

Question: I would like to see the authors discuss why there is so big difference in the number of Selection-category gene between the two lines (e.g. because expression divergence is mutation-limited? Because of single big trans-acting mutations in one line but not the other?).

p. 9: We performed a principal component analysis on the set of 7,370 DE genes, and attributed each one of them to a principal component axis based on correlation coefficient values (Table S4).

Question: What is the logic/advantages to use correlation instead of simply the PC loadings of the individual genes for this analysis?

p. 10: Out of 2,451 DE genes of the Selection category, the PC1 exhibited the greatest proportion of all PCs (46.3%, Table 3) followed by PC2 (30.1%, Table 3).

Question: I have troubles following the logic of this analysis attributing DE genes of certain category to PC's. I'm assuming I'm missing some key point here.

p. 11: Out of 70 and 984 of the FT_candidates and GWA_candidates, 54 and 294 respectively displayed differential gene expression (Table 3).

Question: Is this difference in proportions statistically significant?

p. 12: We focused on four organs: the immature part, mature part and the sheath of the last visible leaf, and the SAM.

Question: What is meant by "part"?

p. 18: We found an overall significant overdispersion (dispersion parameter=1.17, P-value=1.04 10^{-7}), albeit with noticeable differences among chromosomes (P-value=0.0006). Chromosomes 1, 6 and 10 were significantly enriched for DE genes of the Selection category, while chromosomes 4, 7, 9 were significantly depleted.

Question: I think this is an interesting analysis, could be given more weight and would better fit the results section on its own.

p. 18: …the majority of these genes (91.1%) belonged to the 4 first PCs…

Question: What does "belong" here mean?

p. 21: Genetic convergence between inbred lines is detected.

Question: I like the discussion of possible cis vs trans sources of transcriptomic convergence - but I can't follow which types of changes the authors think are more prevalent.