Opaque-2 Regulates a Complex Gene Network Associated with Cell Differentiation and Storage Functions of Maize Endosperm

Junpeng Zhan, Guosheng Li, Choong-Hwan Ryu, Chuang Ma, Shanshan Zhang, Alan Lloyd, Brenda G. Hunter, Brian A. Larkins, Gary N. Drews, Xiangfeng Wang, and Ramin Yadegari

Plant Cell. Advance Publication Sept. 27, 2018; doi: 10.1105/tpc.18.00392

Corresponding author: Ramin Yadegari, yadegari@email.arizona.edu.

Review timeline:

TPC2018-00392-RA	Submission received:	May 18, 2018
	1st Decision:	July 13, 2018 *revision requested*
TPC2018-00392-RAR1	1st Revision received:	Aug. 20, 2018
	2nd Decision:	Sept. 7, 2018 *accept with minor revision*
TPC2018-00392-RAR2	2nd Revision received:	Sept. 11, 2018
	3rd Decision:	Sept. 13, 2018 *acceptance pending, sent to science editor*
	Final acceptance:	Sept. 27, 2018
	Advance publication:	Sept. 27, 2018

REPORT: (The report shows the major requests for revision and author responses. Minor comments for revision and miscellaneous correspondence are not included. The original format may not be reflected in this compilation, but the reviewer comments and author responses are not edited, except to correct minor typographical or spelling errors that could be a source of ambiguity.)

TPC2018-00392-RA 1st Editorial decision – *revision requested* July 13, 2018

We have received reviews of your manuscript entitled "Opaque-2 regulates a complex gene network associated with cell differentiation and storage function of maize endosperm." Thank you for submitting your best work to *The Plant Cell*. The editorial board agrees that the work you describe is substantive, falls within the scope of the journal, and may become acceptable for publication pending revision, and potential re-review.

We ask you to pay attention to the following points in preparing your revision. Reviewer #2 and #3 raised concerns on the data for Figure 6 as to why was W22 used rather than B73 as well as the developmental stage presented as to the limitations in comparing kernels from different plants. Reviewer #3 raised multiple concerns throughout your manuscript that would greatly enhance the clarity of your manuscript.

Please note the following:

-Supplemental materials should be restricted to large datasets and tables, presentation of replicates, and validation of reagents, methods, or genotypes. Any data that are used to support claims must be in the main manuscript. Supplemental figure legends must indicate what figure in the main manuscript is supported by the supplemental data presented. With the large number of supplemental files in your manuscript, I ask that you seriously consider if all of them are required. If they do not meet the criteria, please move them to the main paper, and if they are not required, delete them entirely.

-Note that the fonts are too small in many of the figures, especially in subscripts and various lettering. Please consider the reader.

-Think about each figure as a single illustration and match individual parts of the same figure for size of individual elements.

-Please note that in some tables you have too many significant figures, only the first uncertain digit should be shown. When you take ratios, as in fold change, there is a rule for how to carry significant figures and the number shown

depends on the significant figures in the data that went into the numerator and denominator. Please think carefully about this.

-Sampling methods and nature of "biological replicates" should be described precisely (i.e. different plants, parts of plants, pooled tissue, independent pools of tissue, sampled at different times, etc.), along with a clear description of and rationale for any statistical analyses conducted. The reader should know exactly what was sampled; what forms the basis of the calculation of any means and statistical parameters reported. This is also necessary to ensure that proper statistical analysis was conducted.

Please contact us if there are ambiguous comments or if you wish to discuss the revision.

---------- Reviewer comments:

[Reviewer comments shown below along with author responses]

TPC2018-00392-RAR1 1st Revision received Aug. 20, 2018

Reviewer comments and **author responses**:

Reviewer #1:

In the history of maize genetics research, the *o2* mutant has been known for 100 years! The valuable trait of high lysine in this mutant was recognized in 1964 and since then *o2* was extensively studied in many aspects, such as nutrition evaluation, cell biology, biochemistry and molecular biology. Before cloning of this gene, o2 has been applied for breeding Quality Protein Maize (QPM), which resulted in the award of 2000 Food Prize. Understanding of the molecular basis of O2 took a long way and is still ongoing. *O2* was cloned in 1987 and then was determined to encode a bZIP TF in 1990 and regulate 22-kD alpha-zein genes in 1992. Later, O2 was found to regulate a few non-storage protein genes as well, like b32 and cyPPDK1.

But these findings only tackled a tip of the iceberg of the O2 regulatory network. In 2002, Hunter et al. used an Affymetrix GeneChip containing more than 1400 selected maize gene sequences to investigate the differentially expressed genes between WT and a number of opaque mutants including *o2*. This work resulted in the identification of a large number of genes that were directly and indirectly affected by *o2*. In 2015, Li et al used the RNA-Seq and Chip-Seq to investigate the transcriptional regulatory network of O2 and only identified 35 O2-modulated target genes. This number is much under the expectation. The patterns by which O2 regulates its modulated target genes and the association with cell differentiation and storage function remains unknown.

This work done by Zhan et al using a similar approach (RNA-Seq coupled with Chip-Seq) to analyze a different *o2* background (B73*o2*) performed an extensive and intensive analysis of the resulting data and provided a large amount of useful information. They identified 186 modulated and bound targets, of which 134 are additional O2-activated direct targets. These have unprecedentedly broadened our understanding of O2 regulatory network compared with any prior research. As we know more about O2 functions, the *o2* mutant appears as not an ideal high lysine donor for QPM breeding. The fine O2 network atlas constructed here is very helpful to guide engineering of maize nutritional quality in the future.

I would like to recommend two major merits in this work: 1. It determined at least two distinct modes of O2-mediated gene activation. Based on my own research, I indeed discovered a few genes regulated by O2 are activated before *O2* expression. This study has done a nice analysis and summary of this gene group. 2. They discovered that bZIP17 and NKD1 are activated by O2, which in turn co-activate other O2-network genes. This indicates that O2 is required for proper aleurone differentiation. Indeed, they found alterations in cell wall, PSV and oil bodies in aleurone. This has ranked O2 up to a higher hierarchy in regulation of maize endosperm development and storage reserve synthesis. They first linked the O2 regulatory network with endosperm development and cell differentiation. There is one conclusion in this work I don't agree or I am not sure based on currently presented data. The authors claimed that they identified 800 O2-repressed genes, of which 39 are direct targets. It is difficult for me to believe that. In the

o2 endosperm, zein proteins are dramatically reduced and non-zein proteins are compensatorily increased, which reasonably leads to a delayed turnover of some non-zein mRNA and as a consequence higher accumulation of their transcripts. A good example for this is *glb1* and *glb2*, which are abundantly expressed in embryo and but also have some in endosperm. The transcripts and proteins of GLBs are elevated in *o2*, but they are not regulated by this TF. Their increase is due to proteome rebalancing rather than derepression from the *o2* mutation. In Supplemental Figure 9, the data is not convincing for the "repressed" genes, because the fold changes are so small. I am not asking the authors to do more experiments, but they have to consider more possibilities to explain their data.

We thank the reviewer for all the comments. We agree with the reviewer that the genes up-regulated in *o2* are not necessarily transcriptionally repressed by O2 in the wild type, and the increase in their mRNA levels could be due to proteome rebalancing. We have added a sentence to the Discussion (Page 21) to indicate this possibility: "In addition, it is possible that at least a subset of the putative O2-repressed genes were detected as a result of proteome rebalancing processes in the *o2* mutant endosperm rather than as a result of loss of repression of O2—that is, their mRNA levels could be elevated in the *o2* endosperm as part of an overall increase in the levels of non-zein mRNAs to enhance synthesis of non-zein proteins {Wu, 2014 #179}{Larkins, 2017 #173}."

In Figure 2, some words overlap.

We have changed the terms "Molecular Function," "Biological Process," and "Cellular Component" to abbreviations "MF," "BP," and "CC" and re-positioned them in the figure to avoid overlap with the names of the GO terms.

In Figure 3, what do the red and blue lines represent?

The red and light blue lines in Figure 3 represent the RT-qPCR data (Ct values) of individual genes in wild-type and *o2* endosperm, respectively. We indicated this using a legend at the bottom right of the figure in the submitted version of the manuscript, and to make it clearer, we have increased the size of the legend in the revised version.

Figure 4 is too vague.

The previous Figure 4 is now Figure 6 in the revised version. We have enlarged the entire figure to make it clearer. Also, we will be sure to provide the journal with the figure of the highest resolution we can for publication.

In Supplemental Table 4, SSIII is missing, which was published in Zhang, et al. PNAS (2016), 113: 10842-10847.

We have now added the information about *SSIII* to Supplemental Table 4.

Reviewer #2:

Zhan et al. report O2 transcription factor target genes based on RNA-seq and ChIP-seq from developing endosperm tissues. Overall, the data are of high quality, but this is essentially repeating an experiment published by the *Plant Cell* in 2015. The main conclusion from both studies is that O2 directly activates most of the zein storage proteins in maize endosperm. The authors found a similar number of differentially expressed genes and show direct regulation of target genes including several new targets of O2. The study also found two new O2-interacting transcription factors, bZIP17 and NKD2. The evidence for these interacting proteins to co-activate target genes is convincing. O2 is known to interact with several transcription factor partners, and the data shown in this manuscript is similar in that specific promoter elements identified by ChIP-seq show greater transactivation when both O2 and the interacting transcription factor are transiently expressed.

The NKD2 transcription factor is known to have a role in endosperm development, and this interaction with O2 is novel. In an attempt to show that O2 has a functional role overlapping with NKD2, Figure 6 quantifies ultrastructural differences between *o2* mutants and a reference inbred.

I am not convinced by Figure 6. The authors measured ultrastructural characteristics from nominally comparable kernels. The cellular structures measured are known to change with developmental time. Older aleurone has a thicker anticlinal cell wall with larger lipid bodies, consistent with the W22 sample being slightly older in developmental time. However, the developmental staging of the W22 and W22*o2* samples is highly dubious. The

comparison is between different W22 and W22*o2* plants that sampled based on the same DAP. Only subtle quantitative changes in cellular morphology are reported. Developmental stage is sensitive to abiotic stresses. Differences in weather, water availability, and nutrients are likely to make the same DAP ear be a different developmental stage in different plants. In addition, aleurone cells are known to have multiple developmental gradients. Kernels have a positional gradient of development along the long axis of the ear with kernels at the distal tip maturing more quickly. Aleurone also shows developmental gradients within the kernel with cells at the crown maturing more quickly. The authors quantified sections from 1-2 kernels from two plants without detailing how they controlled for all of these developmental gradients. This is insufficient sampling to attribute the subtle changes to *o2* instead of environmental noise.

Thus, the main novel points of the study are the new O2 target genes identified and the bZIP17 and NKD2 interacting transcription factors. The target genes were defined by different criteria than the previous Li et al (2015) study. It is not clear whether the same targets can be identified with the same criteria from both data sets. The lack of a convincing biological function for *o2* in the aleurone or for the new interacting proteins in requiring O2 to function reduces the scientific impact of their findings.

We thank the reviewer for the insightful comments. After a reevaluation of the available data, we recognize the issue raised by the reviewer regarding the need for a more comprehensive analysis of the phenotype using multiple plants with close attention paid to the sampling regimen, and we have determined that in order to adequately address this issue, we would have to grow the required genetic material and reanalyze the phenotype. This will take several months and cannot be done in time to resubmit the current manuscript. However, we feel that the manuscript is substantive without these data. Therefore, we have removed the aleurone phenotype data from the revised version.

With respect to the comment that "It is not clear whether the same targets can be identified with the same criteria from both data sets," we applied a similar cutoff criterion as used by Li et al. (2015) to determine O2-bound genes, which is that an O2-bound gene is defined as a gene with one or more O2 peak(s) detected in the genic regions or within 1 kb (instead of 5 kb which was eventually used in our manuscript) up- or downstream of the annotated gene models. Based on this criterion, we detected 1,057 O2-bound genes [versus 1,143 detected by Li et al (2015)], including 98 putative O2-activated [versus 35 detected by Li et al (2015)] and 17 putative O2-repressed [versus 4 detected by Li et al. (2015)] direct target genes. Comparison of the 98 O2-activated direct target genes that we detected versus the 35 O2-activated direct targets yielded only 13 genes commonly detected in both studies. These results suggest substantial differences in O2 target genes detected in the two studies. We have now included a description of these results in Supplemental Text 2, and a new Supplemental Figure 5 as supporting data.

Minor points: What does supplemental figure 1 provide? These are just introgressions of the same *o2* allele in different genetic backgrounds? The F1 hybrids were not analyzed in the study.

Because the B73*o2* mutant has not been reported before, we confirmed the *o2* mutation by (a) performing a (non)complementation analysis with two of the better studied *o2* mutant alleles published previously, and (b) sequencing the genomic DNA and cDNA of the *O2* gene in the B73*o2* mutant in comparison to a W22*o2* mutant (and their respective wild type). We argue that because *o2* produces a recessive phenotype, analysis of the F1 kernels would be sufficient to confirm allelism. The rationale and relevant results are described in detail in Supplemental Text 1.

Fig 1A: RPKM is the wrong normalization for paired-end reads.

Since edgeR is one of the most widely used R packages for both single-end and paired-end read normalization (of which the output is CPM/RPKM) and analysis of differential gene expression, we utilized it in our analysis. In fact, RPKM has been used in many paired-end Illumina RNA-Seq studies not only for maize (*e.g.*, The Plant Cell, Vol. 30: 360–374), but also for other plant and animal systems (*e.g.*, The Plant Cell, Vol. 30: 620–637; Nature, Vol. 508: 345-350).

Fig 1B: The O2 ChIP-seq peaks map according to target bin size. Other intergenic is the largest target sequence followed by the 5 kb up- and down-stream of genes. How is this supposed to convince the reader that the signal is specific?

We used three approaches to make sure that we are working with the highest-confidence ChIP-Seq signals for the most closely associated genes. First, we used the B73*o2* mutant as a negative control in ChIP-Seq peak calling to make sure the ChIP-Seq signal is specific to O2-bound regions. Second, we focused on only the genes with peaks localized less than 5 kb away from the annotated gene models. Third, we used EMSAs and DLR assays to verify the ChIP-Seq peaks (signals) for some of the direct targets. Based on these criteria, we are confident that our ChIP-Seq signals are not due to artifacts. On the other hand, it is formally possible that there are functionally important distal O2-bound regions that are localized >5 kb away from gene models. However, the identification and validation of the genes they regulate and the mode of regulation vis-à-vis other transcription factors will require further experiments that are beyond the scope of this manuscript.

Line 180-181: What is the definition of a promoter region for the analysis?

We agree with the reviewer that perhaps the term "promoter" should not have been used in this context as we meant to indicate regulatory sequences encompassing "1 kb upstream and 0.5 kb downstream of the annotated transcription start sites" as indicated in the Methods section (Page 27). Accordingly, we have modified the revised version to replace all the relevant term "promoter" with "5'-regulatory region/sequence" where appropriate.

Why does Supplemental Figure 3 not include the most directly comparable Li et al, (2015) study?

We indeed considered including the DEGs from the Li et al (2015) study. However, unfortunately, a list of all DEGs identified in this previous study is not available either via the Plant Cell website or GEO. In fact, we think the DEGs available from the Zhang et al. (2016) study is as comparable as the Li et al. (2015) study, since the biological replicates of the endosperm material analyzed in Zhang et al. (2016) were handled separately (similar to ours), whereas in the Li et al. (2015) study the three biological replicates of each RNA sample were pooled before RNA-Seq.

Discussion of O2 "indirectly-repressed" genes needs to retain perspective. These are genes that have an apparent increase in transcription in *o2* mutants. The zein transcripts represent approximately 40% of the total transcriptome and are severely reduced in the *o2* mutant. The entire remaining transcriptome has an apparent activation that EdgeR is supposed to take care of with TMM modeling factors that optimize on global fold change. TMM gives a linear transformation of the library size, which handles typical comparisons and typical genes well. However, gene expression follows a Power Law distribution with the zeins being at the top of the distribution in endosperm. The authors need to provide some evidence that indirectly-repressed genes are not artifacts due to the massive differences in library complexity in normal and *o2* samples.

We recognize the potential limitations of the existing tools for analysis of differential gene expression (*i.e.*, mRNA levels) where high-prevalent RNAs may interfere with the analysis. For this reason, we used RT-qPCR to verify the mRNA levels for 22 (out of 186) direct O2 target genes that we detected (Figure 4 and Supplemental Data Set 8). Using this analysis, we show in the revised version that the differential expression patterns (at 15 DAP) obtained through our computational analysis are generally consistent with the experimental determination of mRNA levels using RT-qPCR (Supplemental Data Set 8, columns AR and AS). The gene set we analyzed included 4 putative O2-repressed direct targets. Therefore, we did not (and do not) think that differences in library complexity caused any significant artifacts in the differential gene expression analysis.

Reviewer #3:

The manuscript by Zhan et al provides a comprehensive characterization of targets and genes regulated by the Opaque-2 (O2) bZIP protein. The authors identified a number of regulatory motifs and new players in endosperm and aleurone development that significantly help clarify how O2 functions.

While this study might be perceived as not novel enough, given that Li et al published in Plant Cell in 2015 a similar analysis of O2 targets, the new manuscript adds several new aspects, and is probably done with significantly more rigor that Li et al. In fact, many of us have significantly questioned ChIP-Seq aspects of the Li et al study. In this new manuscript, Zhan et al make a careful comparison of their results and those published by Li et al, and I commend them for not being overly critical of the Li et al study.

Comments

1. Supplemental Fig. 1C is lacking loading controls

We thank the reviewer for the insightful comments. In the revised version we have now included a stained gel in Supplemental Figure 1E as loading controls for Supplemental Figures 1C and 1D.

2. Lane 745: Unclear what the RNA-Seq replicates are - different maize ears, same material divided into three?

The biological replicates were derived from separate ears from individual plants. This has been described in the Methods: "…For these experiments, each biological replicate of a given sample included multiple endosperms from a single ear (plant)" (Page 23).

3. Lane 197 - 198: I don't agree that the data supports a primary role of O2 as an activator. The number of activated and repressed genes is comparable (877 vs 800), and it is often the case that repression has a significantly lower fold change than activation, for example because to see repression, short mRNA half-life is also required. This is also in contradiction with the discussion (Lanes 613 - 614).

Based on the reviewer's comment, we have reevaluated the relevant data and changed the concluding sentence to: "These data suggest that the O2-activated, direct target gene set is most crucially dependent on regulation by O2 compared to the other O2-modulated and/or bound gene sets," and modified the preceding sentence accordingly (Page 7).

4. The authors may want to take advantage of existing maize transcriptional initiation data derived from CAGE to map binding sites in the targets with respect to transcription start sites when possible, rather than to translation start sites. If available, it can be viewed in the maize genome Browser at MaizeGDB in v3 of the maize genome.

In our analysis, the O2-binding sites (peaks) were indeed mapped with respect to the annotated transcription start sites (based on the annotation available via Ensembl Plants/Gramene). We only referred to the positioning with respect to translation start sites where we cited previously published data. With respect to CAGE data as suggested by the reviewer, we were able to find one report of CAGE analysis of maize transcription start sites (The Plant Cell, Vol. 27: 3309–3320), but we were not able to use the data because the data were derived using root and shoot tissues, which likely has different patterns of transcription initiation than endosperm; and the data are not in the form that could be readily used to compare to the O2-binding sites (*i.e.*, the O2 peaks) identified here.

5. In Supp Fig. 8, it would be nice if the shifted bands (and free probes) could be labeled with arrows. If not explained in the text, the legend should explain the presence of multiple shifted bands. The sequences (or at least a diagram indicating the coordinates and putative O2 binding sites) of the probes should be shown in each case to be able to interpret the results. Perhaps it would have been good to include one or more sequences not recovered from the ChIP-Seq as a negative control?

As suggested by the reviewer, we have now added arrows and asterisks to Supplemental Figure 8, which is now Supplemental Figure 9 in the revised version, to label the shifted bands and free probes, respectively. The coordinates of the EMSA probes and the putative O2-binding sites are shown in Supplemental Table 5, and we have now indicated this in the figure legend of Supplemental Figure 9.

6. In Supp Fig. 9, it would have been nice to include just O2 (no VP64 activation domain) to see if the activation/repression can be recapitulated.

Supplemental Figure 9 is now Supplemental Figure 10 in the revised version. For five of the O2 peaks (the ones associated with the O2-activated genes *FL2*, 18-kD δ-zein, *bZIP17*, *NKD2*, 15-kD β-zein, and *TAR3*), we indeed tested them using intact O2 protein (without fusion to VP64 domain) as the effector, and all of them turned out to be activated by O2 (when O2 is the only effector expressed in addition to YFP; Supplemental Tables 8 and 10).

7. Lanes 258-261: It is rather obvious that other regulatory proteins will participate in the control of O2 targets - the sentence really adds little and makes it sound that this is unique to O2 (it would be unique if it would work without other TFs!). The last part of this section is rather speculative and I would suggest to limit it unless the authors are ready to validate some of the predictions, e.g., showing co-localization of PBF or any of the other TFs mentioned.

In response to this comment, we have edited down the sections in Results (Page 9) and Discussion (Page 21) about the involvement of other TFs inferred from the motif analysis.

8. Figure 2 is confusing, adds little to what is discussed in the text, and probably belong better in the supplementary- where is "bound repressed"?

Figure 2 shows the functionalities that are enriched among genes that are putatively directly or indirectly regulated by O2 based on the GO term enrichment analysis, and there is an entire section in the manuscript that is devoted to description and discussion of the data. In fact, there is a Supplemental Data Set 4 that supports this figure, so we would prefer to keep this figure as a main figure. The putative O2-repressed direct target ("bound repressed") genes did not show any significant enrichment of GO terms at FDR < 0.05, and we have described this in the Results section: "The subset of genes identified as directly repressed by O2 did not show significant enrichment of any GO terms... (Page 12)."

9. It is difficult to understand what the O2-activated indirect targets mean - are they controlled by a TF controlled by O2, or are they a secondary consequence of changes that happen in endosperm as a consequence of O2-induced zein accumulation? Too much space in the paper is being used in discussing indirectly affected genes.

Based on our definition, the O2-activated indirect targets may include genes that are regulated by other TFs which in turn are regulated directly by O2. We have added a sentence to the Results section (Page 7) to clarify this: "…which are likely to include genes whose transcription is regulated by TFs downstream to O2." We consider the regulation of downstream genes by O2 through direct regulation of other TF genes (*e.g.*, *bZIP17* and *NKD2*) important and novel components of the O2-regulated network and, thus, we would rather leave unchanged the relevant discussion of this topic.

10. The authors use steady-state mRNA accumulation levels as a proxy to conclude that there two modes of O2-mediated gene activation. This is risky as post-transcriptional changes can significantly impact the conclusions. For example, does the absence of *O2* mRNA at early stages correlate with no O2 protein? If the authors wanted to make conclusions about gene activation, then at the very least transcriptional run-off experiments should be included. The section in the results that discusses these modes should be revised to accommodate for the possibility that mRNA and protein levels are not linear.

In response to this comment, we have added a sentence to the Discussion (Page 20) to indicate the possibility that mRNA and protein levels may not be directly related: "However, it is also possible that post-transcriptional regulatory mechanisms may also contribute to the modulation of the encoded gene products as mRNA accumulation patterns do not correlate strictly with the respective protein levels for many endosperm genes {Walley, 2013 #180}". We also have accordingly added "…using both transcriptomic and proteomic approaches" to the next sentence in the text.

11. More technical, the error bars in Fig. 3B-M are surprisingly small to correspond to true biological replicates (i.e., coming from separate maize ears, in my opinion the appropriate way to do this) - how were the replicates obtained?

The replicates in fact were derived from separate maize ears. This has been indicated in the Methods section as in our response to this reviewer's comment #2.

12. Why were the aleurones investigated in W22 rather than in B73?

We agree with the reviewer in the sense that ideally the phenotypic analysis should be carried out using B73*o2* and B73 kernels, given that nearly all other analyses in this paper were performed using the B73 background. Considering this issue and a comment by Reviewer #2 (comment #1), we have decided to remove the aleurone phenotype data from the revision. The rationale is further described in our response to Reviewer #2's comment #1.

13. Any mutants available for bZIP17?

To our knowledge, there is no mutant available for bZIP17.

14. There is no bar in Supp Figure 18. I can't see the cell walls in the figures shown.

We have removed the previous Supplemental Figure 18 in the revised version due to the reasons described in response to Reviewer #2's comment #1 and Reviewer #3's comment #12.

15. Not sure that Fig. 7A adds much.... too many "?"

Figure 7 is now Figure 8 in the revised version. In response to this comment, we have removed panel A from the figure.

16. The regulation of VP1 by O2 is unexpected. VP1 controls the expression of anthocyanins (which usually occur during late aleurone development), but to my knowledge O2 has no effect on anthocyanin accumulation. How is this explained?

We think one explanation is that although *VP1* is significantly downregulated in the *o2* mutant, redundant/compensatory factors of *VP1* may function to maintain anthocyanin accumulation in the *o2* mutant.

Minor comments 1. I don't understand in Supp Fig 1B the structure of O2 gDNA in B73, based on what is shown for *O2* cDNA in B73*o2* and *O2* gDNA in B73*o2*. Is the annotation of *O2* gDNA in B73 wrong in v3? Was this corrected in v4? this is a classic maize gene, structure known for decades.

Supplemental Figure 1B does not suggest that the annotation of *O2* gDNA is wrong in B73 v3, and we think that the annotation is correct. In this figure, all other sequences were aligned to *O2* gDNA in B73*o2*, and the gene model shown in the figure was indeed based on the annotation of *O2* in the B73 genome (v3). In fact, the overall gene model is consistent between *O2* gDNA in B73 and *O2* gDNA in B73*o2* except that there are many mismatches and indels. We have indicated these in the legend of Supplemental Figure 1B: "The alignment was performed using SnapGene 4.0.5 with the *O2* gDNA in B73*o2* as the reference sequence. The *O2* gDNA sequences of B73 and W22 were based on gene model annotation of the reference genomes B73 RefGen_v3 and Zm-W22-REFERENCE-NRGENE-2.0, respectively;" and in Supplemental Text 1 that "Alignment of the 2,711-bp sequence (*O2* gDNA in B73*o2*) with *O2* genomic sequence in B73 using SnapGene v4.0.5 (from GSL Biotech; available at snapgene.com) detected 39 mismatches and 23 gaps, whereas alignment of the same sequence with the *O2* genomic sequence in W22 detected 32 mismatches and 21 gaps."

2. Spell out aleurone whenever possible instead of using the abbreviation

As suggested, we now have replaced all the "AL" abbreviations with "aleurone" in the Results and Discussion sections.

How many times was o2 from W22 backcrossed into B73? Are the seeds available at the Maize COOP for others to be able to verify results using the same lines described here?

Frankly, we don't know how many times the *o2* mutant was backcrossed into B73. This information was not available and likely has been lost by the original breeder who backcrossed the line. However, we believe it has a clean B73 background based on the mapping of endosperm RNA-Seq data. The percentage of mapped reads was very similar to the wild-type B73 (around 93%; Supplemental Table 2), whereas for wild-type W22, even when using similar Tophat parameters, the percentage of mapped reads was only about 83% (data not shown in the current manuscript). Regarding the availability of the B73*o2* seeds, we have now deposited the seeds to the Maize Genetics Cooperation Stock Center, and have included a stock number in the Methods section (Page 23).

TPC2018-00392-RAR1 2nd Editorial decision – *accept with minor revision* Sept. 7, 2018

We have reviewed your manuscript entitled "Opaque-2 regulates a complex gene network associated with cell differentiation and storage function of maize endosperm." On the basis of the advice received, the board of reviewing editors would like to accept your manuscript for publication in *The Plant Cell*. This acceptance is contingent on revision based on the comments of our reviewers.

TPC2018-00392-RAR2 2nd Revision received Sept. 11, 2018

Reviewer comments and **author responses**:

In response to the editor's comments, we have made the following changes to the manuscript (and the data deposited on GEO):

1. We deleted the previous Supplemental Tables 1-3 and Supplemental Figure 2, and changed the numbering of all subsequent Supplemental Tables and Figures accordingly.

2. We uploaded the RNA-Seq and ChIP-Seq mapping statistics (presented in the previous Supplemental Tables 2 and 3) to GEO.

3. We incorporated the SCC values presented in the previous Supplemental Figure 2A into the relevant text (Supplemental Text 2, lines 1004-1007): "This analysis revealed SCCs ranging from 0.92 (biological replicates 2 versus 3) to 0.95 (biological replicates 1 versus 2, and 1 versus 3) for the WT samples and 0.96 (biological replicates 1 versus 3, and 2 versus 3) to 0.97 (biological replicates 1 versus 2) for the B73*o2* samples that indicated high reproducibility of the sampling."

4. We now cite Supplemental Data Set 1 in the sentence where we cited Supplemental Figure 2B (lines 1008-1011): "An examination of expression levels of several canonical direct O2 targets showed that almost all the known targets, as exemplified by *O2* itself, a 22-kD α-zein gene (*azs22-4*) and the *LKR/SDH* gene, were dramatically downregulated in our mutant data (Supplemental Data Set 1)."

TPC2018-00392-RAR2 3rd Editorial decision – *acceptance pending* **Sept. 13, 2018**

We are pleased to inform you that your paper entitled "Opaque-2 regulates a complex gene network associated with cell differentiation and storage function of maize endosperm" has been accepted for publication in *The Plant Cell*, pending a final minor editorial review by journal staff.

Final acceptance from Science Editor **Sept. 27, 2018**