A Y-encoded suppressor of feminization arose via lineage-specific duplication of a cytokinin response regulator in kiwifruit

Takashi Akagi, Isabelle Marie Henry, Haruka Ohtani, Takuya Morimoto, Kenji Beppu, Ikuo Kataoka, Ryutaro Tao

Plant Cell. Advance Publication April 6, 2018; doi: 10.1105/tpc.17.00787

Corresponding author: Takashi Akagi, takashia@kais.kyoto-u.ac.jp

Review timeline:

TPC2017-00787-RA	Submission received:	Oct. 9, 2017
	1st Decision:	Dec. 3, 2017 *revision requested*
TPC2017-00787-RAR1	1st Revision received:	Jan. 31, 2018
	2nd Decision:	Mar. 7, 2018 *accept with minor revision*
TPC2017-00787-RAR2	2nd Revision received:	Mar. 9, 2018
	3rd Decision:	Mar. 13, 2018 *acceptance pending, sent to science editor*
	Final acceptance:	April 5, 2018
	Advance publication:	April 6, 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.)

TPC2017-00787-RA 1st Editorial decision – *revision requested* **Dec. 3, 2017**

We have received reviews of your manuscript entitled "A Y-encoded sex determinant in kiwifruit arose via lineage-specific duplication of a cytokinin response regulator." 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 that you address each of the reviewers' comments outlined below. The reviewers are impressed with your work and most of their concerns can be addressed with straightforward analyses or more detailed explanations. Reviewer #2 recommends a couple of simple experiments in order to clarify the nature of fruit development in CPPU-treated kiwifruit plants, and the function of SyGI in sex determination. I agree that the work would be improved with better illustration of pollen germination and embryo development in experimental plants.

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. You have a large number of supplemental figures. If they do not meet the criteria stated above and in the IFA, please move them to the main paper.

-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.

-The attached file has some suggestions for improving figure aesthetics. Please apply the comments to the supplemental figures as well. The tables can be difficult to read because of the small font size for some of them, so you may consider supplying them in excel format. If not, please re-print the headers on the 2nd page if a table scrolls on to a second page.

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]

TPC2017-00787-RAR1 1st Revision received Jan. 31, 2018

Reviewer comments and **author responses**:

Thank you for your help with our manuscript, it was deeply appreciated by all authors. As per your suggestions and the reviewers' comments, we have revised our original manuscript extensively. We believe that we have addressed all comments and that the resulting manuscript is much improved.

Specifically, we have added results from RNA *in situ* hybridization using antisense-*SyGI* probes as well as report from a more in-depth characterization of the *N. tabacum* transgenic lines, including pollen growth test in pistils, to further improve our understanding of the function of *SyGI*. Finally, the phenotypic changes associated with the synthetic cytokinin treatments to both male and female flowers are now described in more details. Throughout the manuscript, as suggested by the reviewers, we have restricted our interpretation and discussion of these scenarios throughout the manuscript and focused our discussion on the regulatory mechanisms of this epigenetic sex determination in *D. kaki* instead.

All main modifications from the original manuscript are highlighted in yellow for ease of visualization.

We hope that you and the reviewers will find these modifications satisfactory and this manuscript now suitable for publication in *The Plant Cell*,

Reviewer 1:

The insight about gene duplication initiating plant Y chromosome evolution is interesting. It parallels the evolution of the W chromosome in strawberries, which is tied to polyploidization and is characterized by a duplicated female-specific retrogene (https://www.biorxiv.org/content/early/2017/07/15/163808). This study thus adds to a small but growing number of known sex-determining genes in plants, and helps to illuminate general genetic patterns in the evolution of dieocy.

We appreciate this suggestion. The concept of the suggested paper on turnover of sex chromosomes in *Fragaria* (Tennessen *et al.* 2017) is consistent with our results, where lineage-specific duplications or repeated translocation could drive diversification of sex determination systems. We have added this paper to our discussion section.

Most readers do not know which species are diploid versus polyploid. Please indicate this somewhere, at least for the key species used in the study.

We have added information about the ploidy levels of each species in the legend of Figure 3, in which a variety of the species firstly appeared.

Figure 3E: Only one ploidy level (6x) has both males and females shown. Why? Please explain.

Thank you for this question. This is simply due the lack of plant germplasms. Actinidia is a woody (perennial) crop, maintained by vegetative propagation.

Figure 6B: Where is the black circle labeled 'd'?

We have clarified this, it represents one of the duplication events.

The authors repeatedly refer to male and female "pools". This wording is confusing, since "pooled" sequencing usually means all samples are mixed together without barcodes to separate individuals. Instead, what I think the authors mean here is that individual reads from each species were combined bioinformatically after being sequenced. I suggest re-wording.

We appreciate this suggestion. According to the suggestion, we have renamed these "male reads and "female reads", with simple explanations in the methods section.

Line 465: Unclear what "residues" means in the context of nucleotide sequence

According to this suggestion, we have substituted "residues" with "nucleotide sequences"

Line 160: "fully-sex linked" should be "fully sex-linked"

We have revised this description.

Reviewer 2:

SyGI expression pattern could be better described by *in situ* hybridisations. Such technique will allow to visualize where and when SyGL is expressed in developing kiwi fruit flowers.

We agree with the suggestion for these additional experiments, and conducted *in situ* RNA hybridization in developing gynoecia by using anti-sense *SyGI* probes. These experiments confirmed that *SyGI* was substantially expressed in the early developmental stages of the gynoecia, especially at the surface (or possibly meristematic region). These results were added to Supplementary Figure S8.

The authors use *Arabidopsis thaliana* and *Nicotiana tabacum* to try to shed light into the role of SyGl. They have introduced *SyGI* under the control of its native promoter in two quite distant plant species. In which cells and when pSyGl is active in *Arabidopsis* and *Nicotiana*? This information is missing, the promoter activity could be explored analyzing *pSYGL-GUS* (or GFP) plants.

Thank you for proposing another aspect of the functional analysis on SyGl. We have used RT-PCR analysis to investigate the activation pattern of the *SyGI* native promoter (*pSyGI*) in *Arabidopsis/Nicotiana* and, indeed, found that it was similar in the early developmental stages of stigmsa/styles but was different from what we observed in Actinidia in other organs. For instance, in contrast to the situation in Actinidia, *pSyGI* is not active in young ovary including the ovules in Arabidopsis. This is very relevant to the answers below. In *N. tabacum*, the expression patterns in flower organs is more similar to the situation in Actinidia, but not in vegetative organs (Supplementary Figure S12). Nevertheless, the activation of pSyGl in stigma/styles or carpels was conserved in all three taxa (Actinidia, Arabidopsis, and *N. tabacum*), and could induce gynoecium sterility in all three cases as well, which is most relevant to the present study. We have added information about the expression pattern of the SyGl under the control of pSyGl in *N. tabacum* (Supplementary Figure S12).

Unfortunately the Arabidopsis double mutant *arr22 arr24* does not show any obvious phenotype to be rescued with *SyGI* (Gattolin *et al.*, 2006). It has been previously reported that plants over-expressing *ARR22* (*ARR22*-ox) are dwarf (Kiba *et al.*, 2004), but this is not reported for *35S-SyGI* and this aspect should be discussed.

We appreciate this suggestion on the functions of the *ARR22/24* genes. From our results on *p35S-SyGI* in *N. tabacum* transgenic lines, we could very occasionally detect a dwarf phenotype (N=1/24), which could be consistent with at least partial functional conservation with *ARR22*-ox in Arabidopsis, although the fact that we have only observed this in one transgenic line does not preclude the possibility that the observed phenotype is due to the location of the transgene in that particular line. We have added this information in the text and Supplementary Figure S10.

The description of *Arabidopsis thaliana pSyGI-SyGI* transgenic plants needs some integrations. In the legend of figure 4, the authors declare that *pSyGI-SyGI* transgenic plants develop rudimentary pistils that appear definitely shorter than the normal ones, although the reason for this phenotype is not indicated: for instance, are the pistil cells smaller? It has been demonstrated that in Arabidopsis cytokinin content impacts ovule primordium number, is this the case? (Bartrina *et al.*, 2011). Clearing analyses (Schneitz *et al.,* 1995) might be added to exclude defects of the

developing ovules and of the embryos sacs. Aniline blue staining of pollen tubes *in vivo* will allow to clarify whether Col0 pollen grains can germinate on the stigmatic papillae of *pSyGI-SyGL* pistils. The aniline blue staining can eventually nicely monitor the pollen tube growth inside the transmitting tract.

About Fig. S9 (panel B and C) the authors indicate with an arrow the developing embryos ("Representative phenotypes of ovaries from control and *pSyGI-SyGI* transgenic *Nicotiana tabacum* plants. Ovaries from *pSyGI-SyGI* plants were often smaller than control plants. On the other hand, dissected ovaries of both control and transgenic plants (B-C), carried normal embryos (Em)"). Without clearing the developing seeds is indeed not possible to visualise the embryos. Has pollination occurred? and therefore also in *Nicotiana* it could be very useful to monitor pollen tube growth *in vivo* by aniline staining

We deeply appreciate these suggestions on potential functional analyses of the transgenic lines. As described above, because the expression pattern of *pSyGI* in Arabidopsis was different, especially in the ovary, we focused on *pSyGI-SyGI* induced *N. tabacum* lines to further characterize SyGI function. We first made sections stained with toluidine blue to observe the structural changes in gynoecia. Our results suggest that SyGI mainly disrupts the structure of the stigma, and potentially reduce cell numbers in the pistils, while it does not affect cell size within the gynoecia. The size of the ovary was significantly reduced, but the size of the ovules was not affected. These results were summarized at Supplementary Figure S10.

Additionally, we conducted aniline blue staining to monitor pollen tube growth, in three crosses, *pSyGI-SyGI* x cont, cont x *pSyGI-SyGI* and cont x cont. The results clearly showed that fertile pollen tubes were arrested in the disrupted stigma of *pSyGI-SyGI* line, which is consistent with our observation of the altered stigma structure, and with the fact that SyGI does not affect male function. We have added these results in a main figure, Figure 4, and Supplementary Figure S11.

Regarding the cytokinin application, the authors say "Consistent with this possibility, exogenous synthetic cytokinin partially restored gynoecia in male kiwifruit". They have used N-(2-chloro-4-pyridyl)-N′-phenylurea (CPPU), in tomato application of CPPU to unfertilized pistils can trigger fruit formation even in absence of pollination (Matsuo *et al.,* 2012) thus I wonder whether the partial restoration here described is due to general effects of the CPPU molecule rather than a genuine compensation of the block exerted by SyGI in kiwi male flower pistils.
What happens if CPPU is applied to female kiwi flowers? CPPU should also be applied to Col0 and pSyGI- SyGI plants and the effects compared with those already reported in literature for BA application (Bartrina *et al.,* 2011).

We appreciate this expertise knowledge on the effect of cytokinin application. The effect of cytokinins on gynoecium development might be conserved, although application of cytokinin does not always induce feminization in dioecious plant species (Grant *et al.* 1994). We have added this point in the discussion section. Our model suggests that the Y-encoded SyGI specifically acquired strong expression in the gynoecium in order to counteract the cytokinin signal that induces gynoecium development, altogether resulting in repressed gynoecium development. Thus, it assumes that cytokinin promotes gynoecia development, and posits that the newly established Y chromosome-specific SyGI is an inhibitor of this general pathway, resulting in male flowers.

CPPU treatment on female plants did not further promote carpel development. We have added the results of CPPU treatment to male and female plants in Supplementary Figure S14.

Reviewer 3:

One definition that caused me a bit of confusion is "sex-linked pseudoautosomal", a region which is referred to multiple times throughout the manuscript. This also appeared as "Y-linked pseudoautosomal", or "fully sex-linked pseudoautosomal". As a reader the two words "sex-linked" and "pseudoautosomal" are at odds with each other, as the pseudoautosomal region is by definition recombinant and not perfectly linked to sex. Throughout the text there needs to be a much clearer distinction between the partially or fully recombining pseudoautosomal (PAR) region, and the non-recombining X and Y-linked sex determination regions. For instance, MSY is used on line 137 but the definitions get confused from thereon. A cartoon figure of the labeled X and Y, like in the persimmon *Science* paper, paired with improving the definitions in the text, would go a long way in helping to clarify the regions that the authors are describing in each test. As it stands, there is much confusion in the text with words like PAR, fully sex-linked

PAR, MSY, fully sex-linked region, Y-specific, etc. I believe it may have led to my misunderstanding a major result (see point 3).

We thank to this suggestion, and we apologize for the fact that our original version was confusing. We have redefined these terms and are now using only "Y-specific contigs" (for which there is X-homolog) and "Y-allelic sex-linked contigs". We are no longer using the term "PAR". The Y-allelic sex-linked contigs were further categorized into "fully sex-linked" and "partially sex-linked contigs. These definitions are included in the main text and summarized in Supplementary Figure S3.

The measure of divergence between *A. chinensis* and *A. polygama/A. arugata* was based on Ks analysis of a single gene, linalool synthase. This gene is used as an autosomal reference to highlight the lower divergence of sex-linked contigs in Figure 2C and 2D. Is this gene representative of the autosomal divergence of a handful of tested genes, and how was it chosen? Surely there are autosomal genes which are also under significant constraint and could have PAR-like Ks values across species, which would make the figure look quite different. In other words, the comparative aspect of this figure is based on an arbitrary gene to highlight, or potentially exaggerate, the degree of divergence between autosomal and sex-linked genes in these species. This could be improved by taking a mean divergence from a handful of genes.

We agree with this suggestion. Our previous version used only one gene for calculation of silent divergence, because it was the only gene whose sequence was available from all three species (*A. chinensis, A. polygama*, and *A. arguta*). We now have used 20 putative single-copy genes from the Actinidia genome instead. These genes are conserved within the asterids (Wu et al. 2006.). We calculated average values of silent divergence and net divergence using these 20 genes. We have summarized this information in Supplementary Table S3. For Figure 2C and D, we have added the newly calculated values and associated standard error values.

I empathize that direct functional knockouts are not possible due to transformation and long generation time issues, and indeed the authors have done diligent experiments in Nicotiana and Arabidopsis to make a case for SyGI being a dominant female suppressor that has little/no effect on male fertility. I find this portion of the paper that focuses on testing functional aspects of this single gene to be incredibly elegant and clear. The complication that I run into is the logic cascade that led the authors to focus on this single gene in the first place.

For instance, line 172, when describing the 57 identified candidate transcriptome/genome contigs "All 57 PAR genes, including the five fully sex-linked transcripts [...] This results suggests that these sequences [...] make them unlikely candidate sex determinants". The conclusion of this paragraph seems to undermine the point of the analysis -- None of the candidate contigs are likely sex determinants? A more important point is that candidate sex determination genes were based off evidence of strong signal of differential expression paired with an additional/separate analysis of Y-linked cDNA assemblies. A main result of the paper is the assembly of Y-linked genomic kmer-containing reads to produce a Y pseudo-assembly comprising 249 Y-specific contigs, but this data did not seem to be used to its full potential in exploring all hypotheses. What other annotated, but perhaps lowly expressed/unexpressed genes remain in this ~500 kb assembled bin that could be alternate hypotheses for sex determination candidates? In other words, quickly narrowing the search area down to two expressed genes using the mRNA data makes a large assumption that the sex determining factor/gene is expressed at moderate levels in the sampled tissue and stage. The Y-contigs from the genomic analysis should be annotated using something like Augustus, trained on expressed full-length models.

We deeply appreciate the suggestion for a more in-depth exploration of the genes in the Y-specific regions. We conducted a more thorough search for genes located on the Y-specific contigs, using AUGUSTUS. This resulted in the identification of 61 putative genes, now listed in Supplementary Table S2. We mapped our Illumina mRNA-Seq reads from developing flowers and carpels in stage 1 to characterize their expression patterns. The results were consistent with our previous analyses, where only two genes, *SyGI* and *YFT*, substantially expressed in a male-specific manner. These results are reflected in the result section and in Supplementary Table S2. Some of the other putative genes were significantly expressed but not in a male-specific way. As pointed out, the other genes, which were not significantly expressed in early stages of developing flower and carpels, still constitute valid candidate genes especially of the male promoting factor. This point is raised in the discussion section.

How was SyGI orthology and paralogy determined across the angiosperms or within the family? Line 592 suggests that a single ortholog was manually searched for in every species, then line 614 details the use of a BLAST search. Figure 6A describes these as "*SyGI*-like genes".

Thank you for pointing out our ambiguous definition in the original version. We are now using "SyGI orthologs" throughout the manuscript, and described that we used BLAST search against the published draft genome sequences to identify the putative orthologous sequences, in the results section.

Throughout the manuscript, I appreciate that the *SyGI* gene is carefully discussed as a hypothesis, putative sex determinant, etc. The title of the manuscript does not reflect this, and a casual reader would assume that another master sex determination gene has been identified. This is not true, though, and instead a candidate gene has been identified.

We have revised the title to "A Y-encoded suppressor of feminization arose via lineage-specific duplication of a cytokinin response regulator in kiwifruit"

Line 646: Where does this mutation rate come from? This rate was used to estimate the 20 mya divergence, which leads to the result about a duplication event post-dating the divergence of Actinidiaceae. (upon further searching, I found this mutation rate in Shi *et al* 2010. which should be cited here).

We have added information about the estimated mutation rates, with the reference information (Shi et al. 2010) in the main text.

Figure 6A is difficult to reconcile and apply to our understanding of sex determination in this species, given that each of these disparate carpel and/or pistil datasets are not staged or comparable to each other. I do not find this figure convincing to the main point or objective of this manuscript, and certainly does not provide evidence that a lowly expressed transcript found in kiwifruit has neofunctionalized specifically to function in the carpel.

We agree with this point. It is not ideal to compare transcriptome information from these samples. We have deleted this figure from the manuscript.

Figure 6A: *Oryza*, *Solanum* incorrectly spelled

Thank you, we have corrected the mistake.

Several times throughout the manuscript, "custom R and python scripts" are cited and are significant to the results, but they do not exist to a reader. These should be deposited and viewable, in keeping with the deposition of the read trimming scripts for instance.

We have deposited the scripts for this study and added the information in the manuscript.

TPC2017-00787-RAR1 2nd Editorial decision – *accept with minor revision* Mar. 7, 2018

We have received reviews of your manuscript entitled "A Y-encoded suppressor of feminization arose via lineage-specific duplication of a cytokinin response regulator in kiwifruit." 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. In particular, please consider the following:

We appreciate the work you have put into this revision. As you will see in the reviewer comments, you have addressed most of their concerns thoroughly. Reviewers 1 and 2 do, however, have helpful suggestions to improve the precision of your figures. You should be able to make the minor requested adjustments quickly without too much effort. At the same time, please go over the manuscript one more time to be sure you are in compliance with the guidelines below for your supplementary material and statistical analyses.

TPC2017-00787-RAR2 2nd Revision received Mar. 9, 2018

Reviewer comments and **author responses**:

Thank you for accepting our manuscript, and for your suggestions to improve it. As per your suggestions and the reviewers' comments, we have revised our manuscript, as given below. We believe that we have addressed all comments and that the resulting manuscript is intensively improved.

Reviewer 1:

Please revisit my earlier comments about Figure 3 and about the phrase "fully-sex linked region". These are very minor points.

> **We have deleted Pi value from Figure 3. Regarding some polyploid *A. arguta* accessions in Figure 3E, because they are collections from the wild, we have only either of male or female individuals, except the hexaploid accessions. We have revised to "fully sex-linked" throughout the manuscript.**

Reviewer 2:

Fig 1; in D Em cannot indicate embryos but ovules! ovules contain the embryo sacs that will develop the two female gametes (the egg cell and the central cell)

Fig. 5 panel E is not clear. According to the authors, the arrows indicate embryo like structures, but to me they resemble seed like structures. Better pictures are needed

Supplemental Figure S1, in the female flower those labelled as Em=embryos are indeed ovule still not fertilised, once fertilised they will turn into seeds protecting the embryos, the new generation.

Supplemental Figure S10 again what indicated with Em are not embryos but ovules (if fertilisation did not occurred yet) or seeds if fertilisation has occurred. A closer view of the putative ovules/seeds might help to clarify whether the ovary contain ovules or seeds.

> **We have reflected these all suggestions. Especially, we have rephrased all "embryo" to "ovules" in our manuscript. For figure S10, the current pictures are the closest view of the ovules. Since the pollen tubes stopped on the stigma (see Figure 4), fertilization should not have occurred.**

Fig. 4 panel Q and R pollen grains have been acquired with different magnification; scale bars for them are needed

> **We have adjusted the size of pictures, and added scale bars on them.**

Supplemental Figure S8. Panel A the signal is quite faint; the photograph in A appears different from the other (B-G). Pictures from B to G suggest that *SyGI* is expressed also in specific cells of the petals, is this correct? Could be possible to add closer view of *SyGI* expression pattern? The authors must add the negative control, a section hybridised with *SyGI* sense probe.

> **We have adjusted the contrast of panel A to the others. The panels of E-H, which are closer views of *SyGI* expression patterns, have been revised to see clearer signals. According to the reviewer's suggestion, we have added negative controls using *SyGI* sense-probe in stage 1.**

TPC2017-00787-RAR2 3rd Editorial decision – *acceptance pending* **Mar. 13, 2018**

We are pleased to inform you that your paper entitled "A Y-encoded suppressor of feminization arose via lineage-specific duplication of a cytokinin response regulator in kiwifruit" has been accepted for publication in *The Plant Cell*, pending a final minor editorial review by journal staff.

Final acceptance from Science Editor **April 5, 2018**