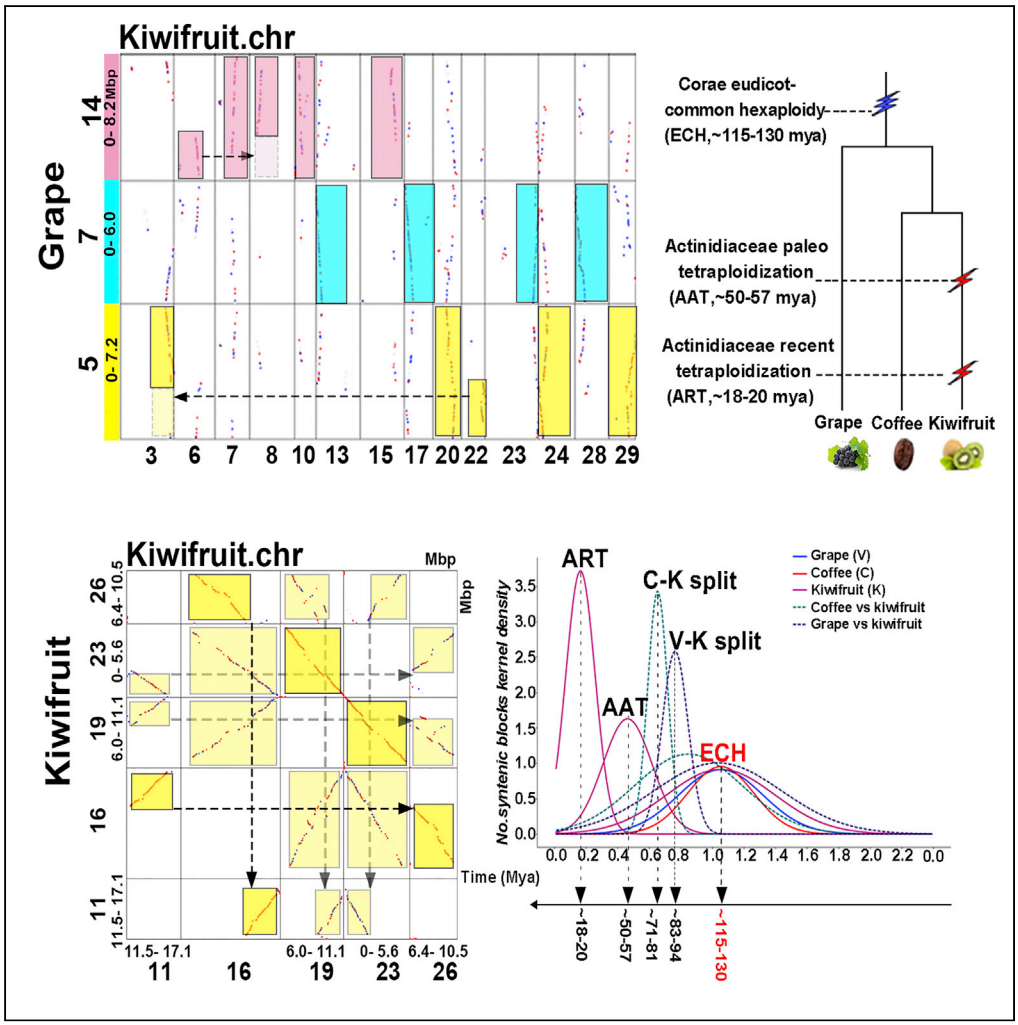
iScience

CellPress

Article

Two Likely Auto-Tetraploidization Events Shaped Kiwifruit Genome and Contributed to Establishment of the Actinidiaceae Family



Jin-Peng Wang, Ji-Gao Yu, Jing Li, ..., Jin Zhang, Jian-Yu Wang, Xiyin Wang

wangxiyin@vip.sina.com

HIGHLIGHTS

Two independent paleo-tetraploidization events may have occurred in Actinidiaceae

The tetraploidization events are likely autotetraploid in nature

These events contribute to the expansion of key trait genes

Hierarchical deconvolution allowed analysis of the kiwifruit genome interweaving homology

Wang et al., iScience 7, 230–240
September 28, 2018
https://doi.org/10.1016/j.isci.2018.08.003

iScience

CellPress

Article

Two Likely Auto-Tetraploidization Events Shaped Kiwifruit Genome and Contributed to Establishment of the Actinidiaceae Family

Jin-Peng Wang,[1,2,3] Ji-Gao Yu,[1,2,3] Jing Li,[1,2,3] Peng-Chuan Sun,[2,3] Li Wang,[1,2,3] Jia-Qing Yuan,[1,2] Fan-Bo Meng,[1,2] Sang-Rong Sun,[1,2] Yu-Xian Li,[1,2] Tian-Yu Lei,[1,2] Yu-Xin Pan,[1,2] Wei-Na Ge,[1,2] Zhen-Yi Wang,[1,2] Lan Zhang,[1,2] Xiao-Ming Song,[1,2] Chao Liu,[1,2] Xue-Qian Duan,[1] Shao-Qi Shen,[1] Yang-qin Xie,[1] Yue Hou,[1] Jin Zhang,[1] Jian-Yu Wang,[1] and Xiyin Wang[1,2,4,*]

SUMMARY

The genome of kiwifruit (*Actinidia chinensis*) was sequenced previously, the first in the Actinidiaceae family. It was shown to have been affected by polyploidization events, the nature of which has been elusive. Here, we performed a reanalysis of the genome and found clear evidence of 2 tetraploidization events, with one occurring ~50–57 million years ago (Mya) and the other ~18–20 Mya. Two subgenomes produced by each event have been under balanced fractionation. Moreover, genes were revealed to express in a balanced way between duplicated copies of chromosomes. Besides, lowered evolutionary rates of kiwifruit genes were observed. These findings could be explained by the likely auto-tetraploidization nature of the polyploidization events. Besides, we found that polyploidy contributed to the expansion of key functional genes, e.g., vitamin C biosynthesis genes. The present work also provided an important comparative genomics resource in the Actinidiaceae and related families.

INTRODUCTION

Kiwifruit (*Actinidia chinensis*) is known as "the king of fruits," with a remarkably high vitamin C content and a balanced nutritional composition of minerals, fiber, and some health-beneficial metabolites, which are essential for human health (Seddon et al., 1994; Fraser and Bramley, 2004). Kiwifruit is a member of the family Actinidiaceae, consisting of 3 genera and around 360 species (Anderberg et al., 2002), most of which are economically and nutritionally important crops. Actinidiaceae is commonly known as the Chinese gooseberry family and is the basal family in the order Ericales, which is one of the earliest orders to diverge from the asterids and shares the core-eudicot common hexaploidization (ECH).

Polyploidy plays an important role in the evolution of land plants, contributing to their origin and diversification (Paterson et al., 2004; Soltis et al., 2008; Jiao et al., 2011). Because of recursive polyploidization and genome repatterning, plant genomes are often considerably complex, hindering the efforts to understand their formation and evolution, and explore functional innovations of genes. The kiwifruit genome has a large number of collinear regions, which often are shared with other plant genomes. It was proposed that 2 whole-genome duplication events occurred during kiwifruit genome evolution (Tao et al., 2010; Huang et al., 2013).

However, the ancestral ploidy of kiwifruit has not been very clear because of the lack of deep exploration. Here, by using our newly proposed pipeline to unravel hierarchical genomic homology reconstruction, we performed a comprehensive analysis of the kiwifruit genome with grape and coffee genomes as outgroup references, produced a list of orthologous and paralogous genes, and explored the outcomes of genomic fractionation. Besides, the present effort provided an important comparative genomics resource for further biological exploration in the Actinidiaceae and related families.

RESULTS

Intra- and Intergenomic Gene Collinearity

Gene collinearity, showing the preservation of ancestral genome structure in the modern genome, is an important means of unveiling the cryptic nature of genomic evolution. We inferred collinear genes within the kiwifruit, coffee, and grape genomes, and between them, via ColinearScan (Wang et al., 2006), which is

[1]School of Life Sciences, North China University of Science and Technology, No.21 Bohai Road, Caofeidian, Tangshan, Hebei 063210, China

[2]Center for Genomics and Computational Biology, North China University of Science and Technology, No.21 Bohai Road, Caofeidian, Tangshan, Hebei 063210, China

[3]These authors contributed equally

[4]Lead Contact

*Correspondence: wangxiyin@vip.sina.com

https://doi.org/10.1016/j.isci.2018.08.003

an effective way to evaluate genomic blocks of collinear genes. The grape and coffee genomes were used as outgroup references because these genomes are relatively simple and have been affected only by ECH. Thus, we also inferred collinear genes within each of these genomes and all 3 genomes (Tables S1 and S2).

Kiwifruit has a much more highly preserved intragenomic homology than either grape or coffee genomes (Tables S1 and S2). We revealed 964 homologous blocks with ≥4 collinear genes, containing 9,998 collinear gene pairs in total. At the same parameter setting, we found 214 and 279 homologous blocks in grape and coffee genomes, containing 2,099 and 2,502 collinear gene pairs, respectively. The number of homologous genes in kiwifruit was about 4–5 times as many as those in grape and coffee, supporting the fact that the kiwifruit genome has undergone additional polyploidy events.

As to intergenomic gene collinearity, 1,631 blocks containing 18,626 collinear gene pairs were revealed between kiwifruit and grape and 1,869 blocks containing 18,869 collinear gene pairs were revealed between kiwifruit and coffee, showing similar collinearity between kiwifruit and the 2 referenced plants (Tables S1 and S2).

Evidence for Two Paleo-Tetraploidization Events

Our inference of ancient polyploidization features integration of sequence divergence and homologous gene dotploting. We first characterized the synonymous nucleotide substitutions on synonymous substitution sites (Ks) between the collinear genes inferred above. If all the Ks of collinear genes are pooled together, clustering analysis may result in absorbing effects. That is, a smaller cluster produced by one event (often an ancient one) would be covered by a bigger cluster produced by another event (often a recent one). Therefore, we calculated the median (relatively more stable than the mean) value of each block and used them to perform clustering analysis. The kiwifruit collinear blocks clearly produced 3 peaks (Figure 1A), and a curve-fitting process located them at 0.164 (+/−0.093), 0.462 (+/−0.156), and 1.55 (+/−0.297), showing 3 events. Then by mapping gene sequence onto whole-genome-based similarity inferred by BLASTN, we constructed the homologous gene dotplot within kiwifruit genome (Figure S1). We also mapped block median Ks values onto the dotplot. As to Ks and checking the dotplot, we could divide most of the blocks produced by 3 events, especially the longer blocks. In fact, if allowing the longer blocks to absorb their nearby smaller blocks, which could have been produced by genomic fractionation, we would distinguish all blocks by 3 events. Blocks produced by the most recent event covered 33.34% of the genomic DNA and 17.36% of total genes, and a covered region has a 1:1 correspondence, showing that it is tetraploidization. We called it as an Actinidiaceae recent tetraploidization (ART). Blocks produced by the mid-aged event covered 47.92% of the genomic DNA and 10.34% of total genes, and a covered region has a 1:1 correspondence, showing that it also is a tetraploidization event. We called it as an Actinidiaceae ancient tetraploidization (AAT). In contrast, we found that the blocks surely produced by the ECH covered 40.76% of the genomic DNA and only 6.96% of total genes.

To verify our inference of the 3 events ART, AAT, and ECH, we used grape as a reference to decipher kiwifruit genome, a rosid eudicot preserving the ECH-triplicated genome structure of a core-eudicot common ancestor. Besides, we used an asteroid plant, coffee, as another reference, which is more closely related to kiwifruit than grape and has not been affected by other polyploidizations after the ECH. Homologous gene dotplots generated from the BLAST results between genomes allowed us to locate the corresponding homologous regions and were used to distinguish orthologous regions, established due to the grape-kiwifruit split, and outparalogous regions, established due to different duplications (Figures 2A, S2, and S3). We mapped the inferred median Ks values of each collinear block onto the dotplots (Figure S1). Integrating the aforementioned 2 lines of information helps distinguish orthologous and outparalogous blocks. The grape chromosomes were denoted with blocks in 7 colors, corresponding to the 7 ancestral eudicot chromosomes before the ECH, therefore a grape genomic region mostly has 2 paralogous regions in the same color (Jiao et al., 2012).

As to the Ks of grape-kiwifruit collinear gene pairs, there is a bimodal distribution, and therefore can be decomposed into 2 distinct distributions that have means (0.764 ± 0.0802) and (1.034 ± 0.405) (Figure 1A), corresponding to the orthologous genes (originated through their split) and outparalogous genes (through the common ECH), respectively. According to the median Ks value of each block, it was not difficult to distinguish the orthologous and the outparalogous blocks, especially for the long ones with ≥10 collinear genes. Allowing the long blocks to absorb the nearby short ones, whose Ks often change much, we inferred orthology and outparalogy for all the predicted collinear blocks (Figures S2 and S3). With the orthologous

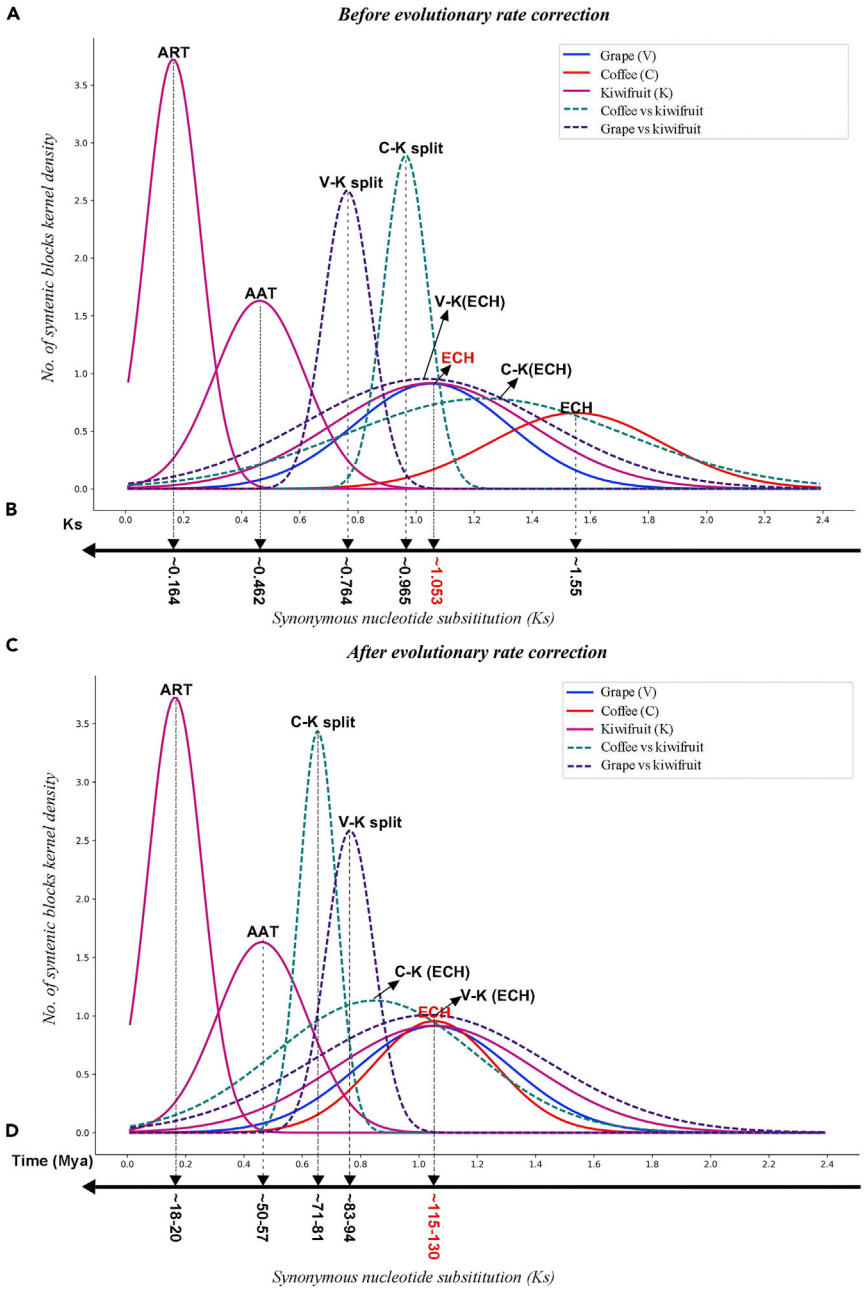


Figure 1. Original and Corrected Synonymous Nucleotide Substitutions between Collinear Genes (Ks)

ART, Actinidiaceae recent tetraploidization; AAT, Actinidiaceae ancient tetraploidization; ECH, core-eudicot-common hexaploidization; Mya, million years ago. Continuous curves are used to show Ks distribution in a genome, and broken ones are between genomes. (A) Density fitted by using original Ks values; (B) Inferred means; (C) Density fitted by using corrected Ks values; (D) Inferred evolutionary dates.

blocks, by relating them to the AAT and ART blocks in kiwifruit dotplot, as to their locations on kiwifruit chromosomes, we separated the orthologous regions produced by different tetraploidization. For certain outparalogous regions lacking collinear genes, due to widespread and complementary gene losses (Maere et al., 2005), we transitively used the paralogy between grape chromosomes and orthology between grape and kiwifruit chromosomes to locate the regions where outparalogy should exist. The coffee-kiwifruit gene dotplot and Ks analysis revealed orthology and outparalogy between the 2 genomes (Figures 1, 2B, and S3). As expected, each grape or coffee genomic region often has 4 orthologous regions, 2 corresponding

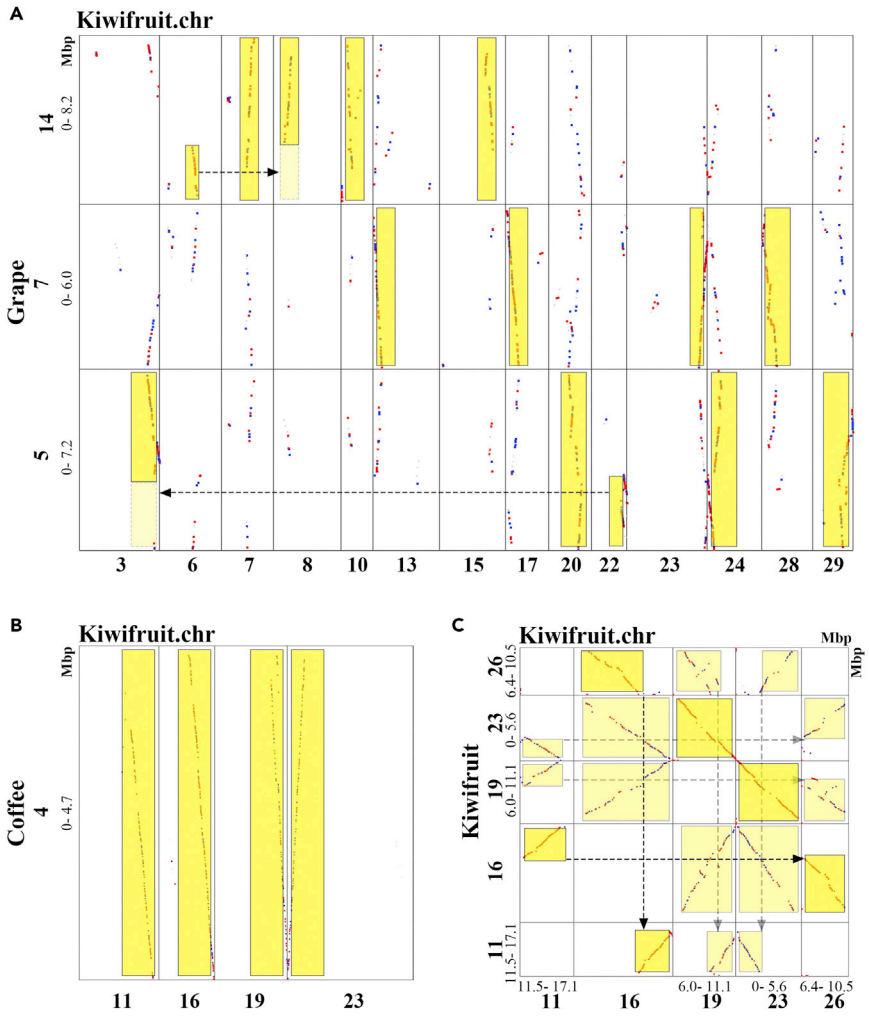


Figure 2. Examples of Homologous Gene Dotplots between Kiwifruit, Coffee, and Grape

Chromosome numbers and regions (in Mbp) were shown. Best-hit genes make red dots, secondary hits make blue dots, and the others are shown in gray. Highlights show the best matched chromosomal regions. Arrows show complement correspondence produced by chromosome breakages during evolution.

(A) Homologous dotplot between selected grape and kiwifruit chromosomes.

(B) Homologous dotplot between selected coffee and kiwifruit chromosomes.

(C) Homologous dotplot within selected kiwifruit chromosomes.

to the AAT and the other 2 to ART (Figure 3). These intergenomic analyses agree with the above inference of 2 sequential tetraploidizations in kiwifruit lineage.

Reconstructed phylogeny of homologous genes provided further support to 2 tetraploidizations in kiwifruit. We used 426 homologous groups, including a grape gene, its ortholog coffee gene, and at least 3 orthologous kiwifruit genes, to reconstruct gene trees. Actually, we successfully constructed gene trees for 311 (73.00%) homologous gene groups (Figure S4), and the other groups were discarded due to insufficient sequence similarity. Checking the topology of these 311 trees, we found that 33 (10.61%) of them each involved 4 kiwifruit genes and were separated into 2 subgroups with each subgroup having 2 genes. This is just what is expected if 2sequential tetraploidizations occurred in the kiwifruit lineage.

Multiple Genome Alignment

The above-mentioned homologous gene dotplotting and Ks analysis distinguished orthology from outparalogy between genomes and the ART-, AAT-, and ECH-produced paralogous regions in kiwifruit. We

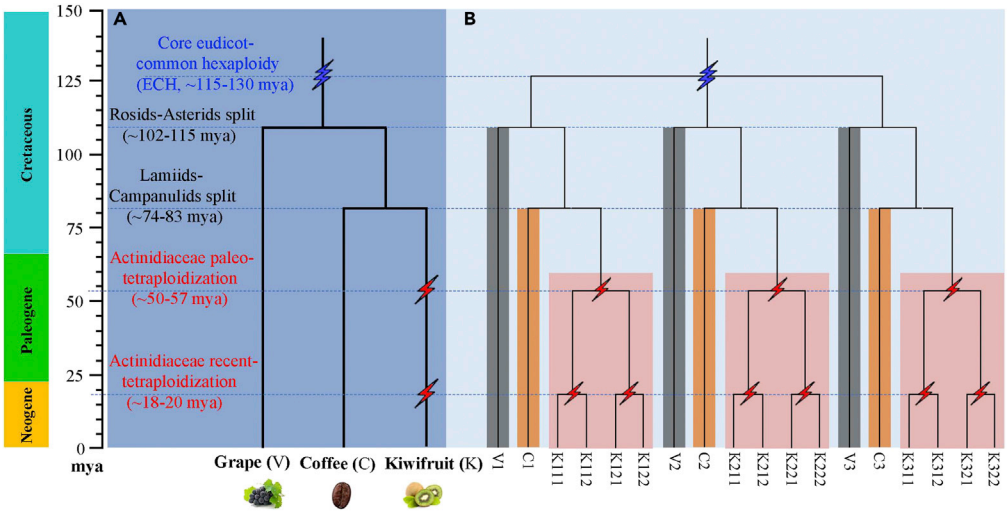


Figure 3. Species and Gene Phylogenetic Trees

(A) Phylogenetic tree of kiwifruit (K), coffee (C), and grape (V). The core-eudicot common hexaploidy (ECH) is denoted by a blue flash, and the two kiwifruit paleo-tetraploidizations are denoted by red flashes.

(B) Gene phylogeny: three paralogous genes in the grape and coffee genomes are denoted by V1, V2, and V3, and C1, C2, and C3, respectively, produced by the ECH, and each has 4 orthologs and 8 outparalogs in the kiwifruit genome (e.g., V1 has 4 orthologs, K111, K112, K121, and K122 and 8 outparalogs, K211, K212, K221, K222, K311, K312, K321, and K322 in kiwifruit). The species tree is produced based on our present analysis of homologous genes.

retrieved detailed information of orthologous and (out)paralogous regions from the dotplots (Tables S3–S5), and counted duplicated genes produced by different events. We found that the ECH event produced 1,505 collinear paralogous pairs, containing 2,150 genes in 188 regions; the AAT event produced 1,983 collinear paralogous pairs, containing 3,196 genes in 153 regions; and the ART event produced 2,753 collinear paralogous pairs, containing 5,366 genes in 95 regions. The number of ECH-related kiwifruit paralog genes was much smaller than the 3,866 genes found in grape. Although the ECH-related paralogs were much less in number, there were more than twice as many, but much shorter, ECH-related regions in kiwifruit as in grape, showing the occurrence of severe kiwifruit genomic fractionation (Table 1).

Intergenomic homology is much better than intragenomic homology. For example, there were 9,322 (30.16%) kiwifruit genes (9,508 pairs in 309 blocks) having coffee orthologs, and 2,885 (9.33%) genes (3,259 pairs in 261 blocks) having coffee outparalogs; meanwhile, 8,923 (28.87%) kiwifruit genes (9,125 pairs in 402 blocks) had grape orthologs, and 4,324 (14.07%) genes (3,681 pairs in 367 blocks) had grape outparalogs. Similar findings appeared in the grape and coffee alignment. More information can be found in Tables S3–S5.

With the grape genome as reference and by entering collinear gene IDs into a table, we constructed a hierarchical and event-related multiple-genome alignment (grape, coffee, kiwifruit), producing a table of homologous genes (Figures 4 and S5; Table S6). The table related paralogous genes in each genome and outparalogs between genomes to each polyploidization, and orthologs between genomes to their ancestral speciation (Figure 3). The table was translated to a whole alignment of the 3 genomes graphically (Figure 4), and a slice of the genome-level alignment can be illustrated in a linear format to give a close view of genomic fractionation (Figure 5). Moreover, to accommodate kiwifruit-specific genes with no grape collinearity, we constructed another homology table with coffee as reference (Figure S6; Table S7), which further proved the paleo-tetraploidizations in kiwifruit (for details see Transparent Methods).

Genomic Fractionation

As described in the previous sections, we mapped the kiwifruit genes, e.g., the AAT- and ART-produced paralogs, onto the grape genome, making it possible to check possible gene retention or loss (Figure S7A). Notably, we found that after 2 rounds of tetraploidizations, kiwifruit preserved a high fraction of genes collinear to grape orthologs. Actually, different grape chromosomes had collinear gene retention rate of 74%–95% in each of their 4 sets of orthologous regions (Table 2). With regard to 2 sets of AAT paralogous

Species	[a]ECH Related	[b]AAT Related	[c]ART Related
Grape	[d]87/2,432/3,866	–	–
Coffee	54/1,189/2,095	–	–
Kiwifruit	188/1,505/2,150	153/1,983/3,196	95/2,753/5,366

Table 1. Number of Duplicated Genes within the Kiwifruit Genome Related to Recursive Polyploidization Events

[a]Core-eudicot common hexaploidization.

[b]Actinidiaceae ancient tetraploidization.

[c]Actinidiaceae recent tetraploidization.

[d]Numbers of blocks, gene pairs, and gene numbers are separated by a solidus.

regions, grape gene retention rates are very similar. For 12 of 19 grape chromosomes, the difference of their gene retention rates in 2 AAT paralogous regions is $\leq$0.05. This means that there exist 2 AAT subgenomes retaining similar numbers of ancestral genes. Actually, the retention rate difference between the subgenomes, noted as AAT-1 and AAT-2, with arbitrarily assigned paralogous regions is 4.3% (Table 2). What is more interesting is that the gene retention rates between 2 sets of ART paralogous regions are just like the AAT regions. For 12 of 19 grape chromosomes, the ART paralogous regions have retention rate difference $\leq$0.05 (Table 2), suggesting 2 subgenomes with balanced gene retention. The ART paralogous regions are therefore arbitrarily assigned to ART subgenome, noted as ART-1 and ART-2, displaying a whole-genome difference ~2.2%.

We performed a scrutiny analysis of gene retention/loss using a sliding window along chromosomes. We found that, in nearly all local regions, with the exception of large patches of DNA losses in one copy of the duplicated chromosomes, genomic retention rates were often $\leq$0.05 (Figures S8 and S9). Actually, the difference observed at the chromosome level should have been caused by large patches of alternative segmental DNA losses due to genomic instability (Figures S8 and S9). With the coffee genome as the reference, we obtained similar observation (Table S8).

Modeling Genomic Fractionation

To explore the mechanism underlying genomic fractionation, we characterized the runs of continual gene deletions in kiwifruit when compared with the reference genomes (Wang et al., 2015). Although there were patches of chromosomal segmental losses (Figure S7), most of the runs of gene deletions were of 15 continual genes or fewer. A statistical fitness regression showed that deletion patterns followed a near geometric distribution (Figure S7). With grape and coffee genomes as reference, kiwifruit had a gene deletion pattern following similar distributions (geometric parameter p = 0.252–0.299, the probability of deleting one gene at a time, and goodness of fit F-test p value 0.89 to accept the fitness). This showed that 38%–42% of genes were deleted containing one or two genes, suggesting a mechanism of fractionation removing short DNA segment about 5–10 kb DNA in length. It seems that short deletion runs accounted for the majority initially, and then recursive deletion runs overlapping previous deletions elongated the observed length of runs, as revealed by a much-diverged reference genome.

Evolutionary Rate Divergence and Dating

By checking Ks, we managed to estimate the occurrence times of the AAT, ART, and other evolutionary events, e.g., speciation. As distribution of Ks between paralogs or orthologs related to a specific event, fitting by using normal distribution function, and the locations of the means (or peaks) and their variances were determined statistically (Figure 1A; Table S9). As to the divergent locations of the common ECH in different distributions, we found that the evolutionary rates of kiwifruit and grape were similar and the slowest among the 3 genomes, with evolution of the coffee genome faster by 47.20%.

Divergent evolutionary rates among plants affect dating ancestral events. Here, based on a modified version of an approach that we previously developed (Wang et al., 2015, 2018), we performed evolutionary rate correction by aligning the peaks in different genomes corresponding to the ECH event to the same locations (see Transparent Methods for details) (Figure 1B; Table S10). The present correction aligned the ECH peaks to the same location, correcting rate differences that have accumulated after the ECH event between rosids and asterids. Supposing that the ECH occurred at ~115–130 million years ago (Mya) (Jiao et al., 2012; Vekemans et al., 2012), adopted by previous publications (Jaillon et al., 2007; Potato Genome

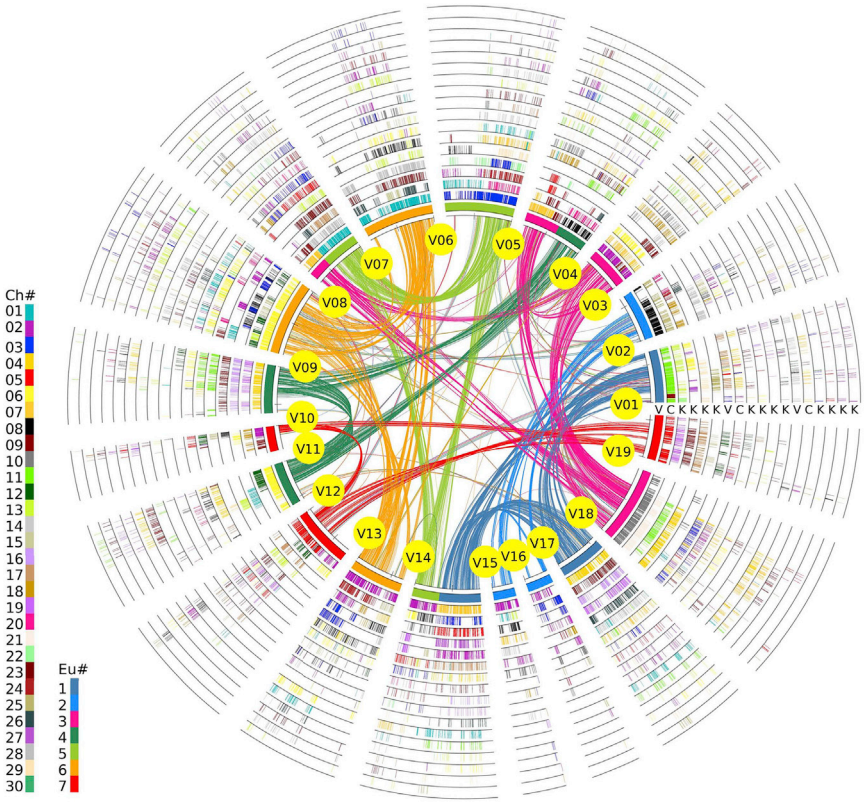


Figure 4. Homologous Alignments of Kiwifruit, Coffee, and Grape Genomes

With grape as a reference genome, genomic paralogy, orthology, and outparalogy information within grape and with coffee and kiwifruit are displayed in 18 circles. The curved lines within the inner circle, colored according to the 7 ancestral eudicot chromosomes (Jaillon et al., 2007), link paralogous pairs on the 19 grape chromosomes produced by the ECH. The short lines forming the innermost circle represent all predicted genes in grape, which have 2 sets of paralogous regions, forming another 2 circles. Each of the 3 sets of grape paralogous chromosomal regions has 1 orthologous copy in coffee and 4 orthologous copies in kiwifruit. Therefore, the 3 genomes result in 18 circles in the figure. Circles are respectively denoted by V, C, and K, corresponding to source plants in Figure 3. Homologous genes are denoted by short lines standing on a chromosome circle and are colored according to the chromosome number in the source plant shown in the inset legend.

Sequencing Consortium et al., 2011; Paterson et al., 2012), we inferred that the ART event occurred at ~18–20 Mya and the AAT event occurred at ~50–57 Mya. In addition, we inferred that kiwifruit (asterids) and grape (rosids) ancestors split at ~83–94 Mya, and then kiwifruit (campanulids) and coffee (lamiids) ancestors split at ~71–81 Mya. Therefore, the tetraploidization events might be shared by Actinidiaceae plants after the separation of Actinidiaceae and Roridulaceae sisters at 84.8 Mya (Magallón et al., 2015).

Balanced Gene Expression between Duplicated Chromosomes

Balanced gene expression was observed between duplicated copies of chromosomes produced in AAT and ART (Tables S11A–S11D). In AAT, between duplicated copies of chromosomes, 38.1% duplicated genes show no difference in gene expression, and the significantly diverged-expressed genes are distributed without significant difference. Between the ART duplicated copies of chromosomes, 33.7% and 40.1% of duplicated genes show no difference in expression and the significantly diverged-expressed genes are distributed without significant difference.

Ascorbic Acid (Vitamin C) Biosynthesis and Recycling Pathway

To check whether recursive polyploidization contributed to the expansion of key traits, such as the ascorbic acid (VC) biosynthesis and recycling pathway in kiwifruit, we used 18 VC-related previously reported gene families as seeds (Huang et al., 2013) and detected their homologs in grape, coffee, and kiwifruit at BLASTP $E < 1e\text{-}10$ and a score greater than 150 (Table S12). Kiwifruit had the most VC-related genes (123), 3.5–6.5 times more than in coffee

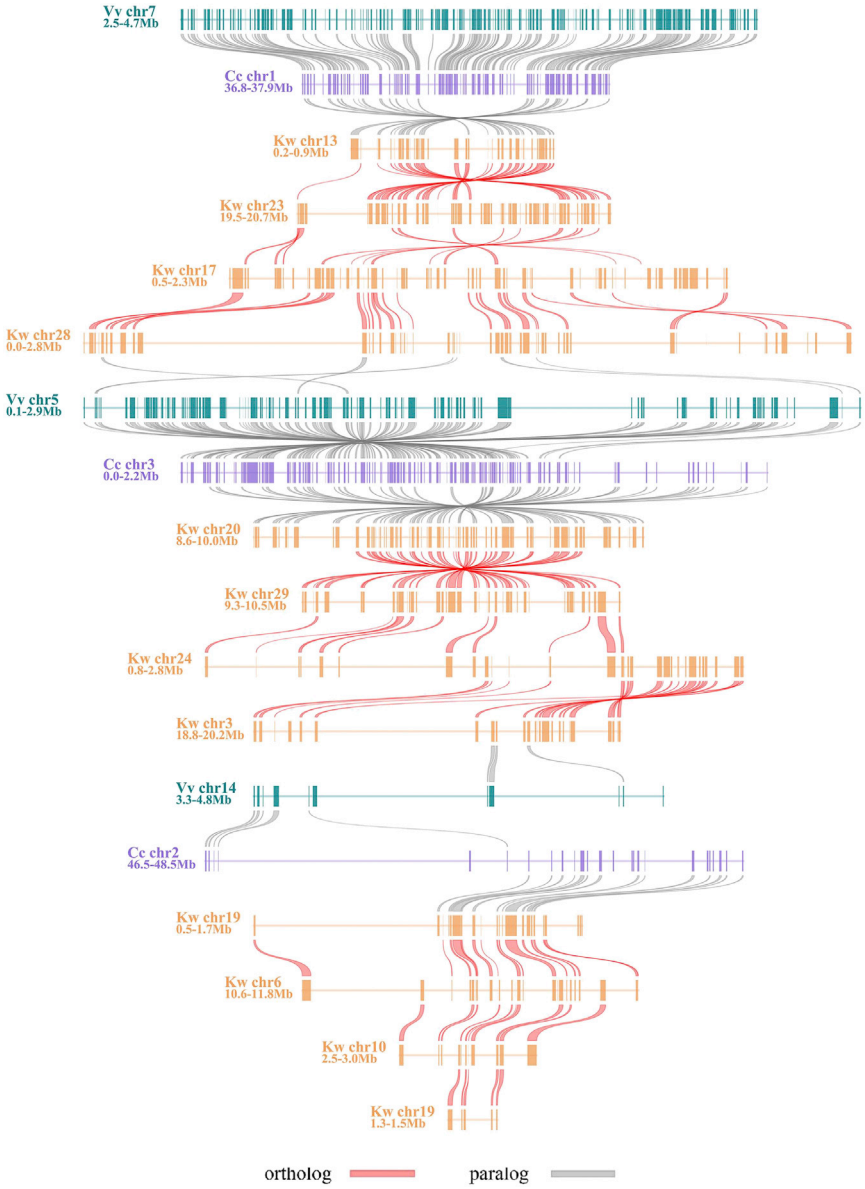


Figure 5. Local Alignment of the Kiwifruit Genome

Species are shown in different colors; homologous genes between adjacent chromosomes are ligated with Bezier curves.

and grape. Notably, 72.22%, 66.67%, and 61.11% of VC-related genes could be located at paralogous chromosomal locations related to the 3 polyploidization events (ART, AAT, and ECH), respectively. The gene copy numbers increased by 56 (44.53%) through the ECH, 35 (28.46%) through the AAT, and 52 (42.28%) through the ART. Moreover, among 18 VC-related families, 11, 12, and 13 of them exhibited an expansion after the ECH, AAT, and ART events, respectively. Specifically, for 22 kiwifruit GalUR-related genes, we found that 36.36%, 45.45%, and 27.27% of the copies were located in paralogous chromosomal regions related to the 3 polyploidization events (ART, AAT, and ECH), respectively. Comparatively, there are only 11 and 3 genes in grape and coffee, respectively.

DISCUSSION

About Deciphering Complex Genomes

Plants often have complex genomes, due to recursive polyploidization and genome repatterning (Soltis et al., 2014; Soltis et al., 2015). This makes it difficult to deconvolute their genome structures, understand their formation, and explore their gene functional evolution. Here, we used a recently proposed pipeline to reanalyze the kiwifruit

Grape		Kiwifruit Homoelogous Subgenomes					
Chromosome #	Gene #	AAT- 1	AAT-2	AAT 1–2 Difference	ART-1	ART-2	ART 1–2 Difference
1	1,327	0.86	0.88	0.02	0.80	0.74	0.06
2	1,237	0.88	0.87	0.01	0.80	0.78	0.02
3	1,000	0.79	0.86	0.07	0.80	0.90	0.10
4	1,638	0.85	0.81	0.04	0.86	0.81	0.05
5	1,748	0.86	0.89	0.03	0.84	0.75	0.09
6	1,779	0.95	0.89	0.06	0.76	0.79	0.03
7	1,409	0.88	0.82	0.06	0.83	0.83	0.00
8	1,867	0.80	0.79	0.01	0.83	0.95	0.12
9	1,221	0.85	0.87	0.02	0.92	0.90	0.02
10	632	0.81	0.87	0.06	0.77	0.75	0.02
11	1,107	0.88	0.74	0.14	0.79	0.83	0.04
12	1,481	0.84	0.84	0.00	0.84	0.90	0.06
13	1,329	0.87	0.79	0.08	0.89	0.80	0.09
14	1,729	0.84	0.86	0.02	0.85	0.81	0.04
15	561	0.88	0.78	0.10	0.89	0.90	0.01
16	647	0.92	0.95	0.03	0.78	0.87	0.09
17	1,168	0.82	0.81	0.01	0.85	0.86	0.01
18	1,886	0.74	0.73	0.01	0.87	0.92	0.05
19	1,135	0.95	0.80	0.05	0.85	0.86	0.01
Total/Average	24,901	0.86	0.83	0.04	0.83	0.84	0.05

Table 2. Gene Retention Rates in Subgenomes as to the Grape Genome

genome and showed that the kiwifruit genome was affected by 2 tetraploidization events. The pipeline features of integrative use of homologous gene dotplots and sequence divergence analysis make it possible to distinguish orthologous regions and (out)paraologous regions while comparing the genome under study with a well-characterized reference genome, or to distinguish paralogous regions in the studied genome resulting by different events. Therefore, this would help find the orthology (and outparalogy) ratio between genomes.

The produced event-related list of homologous genes among genomes is often very helpful in that it provided information about the origin and expansion of genes. Although it may not be complete as it does not include all genes produced by these events due to widespread fractionation over time, the list tells how and when a pair of homologs were produced and diverged, and whether there has been likely gene loss after certain events. Therefore, it can be useful information to reveal the evolutionary and function-innovation trajectories of genes, gene families, regulatory pathways, and economically and agriculturally important traits. As an example, we showed that genes related to VC biosynthesis and recycling pathway were expanded through each polyploidization event. By integrating more sophisticated phylogenetic and evolutionary analyses, many valuable genes could be re-checked at certain phylogenetic nodes to clarify specific evolutionary changes correlated with their functional innovation.

Two Likely Auto-Tetraploidization Events

A large number of genome sequencing analyses have shown that almost all plants on the earth have a polyploid ancestor. It has been proposed that polyploidization has contributed to the origin, divergence, and success of seed and flowering plants (Jiao et al., 2011) and their domestication (Kellogg, 2016).

The above-mentioned findings of balanced gene loss and balanced gene expression between the subgenomes in each tetraploidization event suggest little if any dominance between members of homologous chromosome pairs, which raises the likelihood of the 2 events, AAT and ART, being autotetroploid in nature.

Here we observed much lower evolutionary rates in kiwifruit genes when compared with the coffee homologs. This may be explained by the prolonged generation time of kiwifruit, which might further relate to auto-tetraploidization, as discussed below. Often, elevated evolutionary rates of genes are observed after polyploidization (Wang et al., 2016; Paterson et al., 2012). This may result from homoeologous recombination between duplicated chromosomal regions. Theoretically, the existence of a duplicated copy would buffer mutations in a gene. Therefore, allo-polyploidization would increase evolutionary rates due to homoeologous recombination and the existence of extra gene copies. However, (genetic) auto-tetraploidization results in multiple copies of homologous chromosomes, causing multisomic inheritance (Doyle and Egan, 2010). Multisomic inheritance, with multiple homologous chromosomes to interact and pair with one another in a reticular manner, may prolong the cell cycle time and therefore the generation time. In contrast, allopolyploids, although having doubled chromosomes, execute diploid inheritance and do not have prolonged cell cycle time or generation time. Actually, the effect in evolutionary rates may be a key difference between auto- and allopolyploids.

Autopolyploids may suffer from reduced fertility, whereas allopolyploids are thought to have advantages during the establishment phase owing to their potential for heterosis. These thoughts are consistent with the observation (Barker et al., 2016) that more crops are allopolyploid (e.g., wheat, cotton, tobacco, strawberry, and oilseed rape) than autopolyploid (e.g., potato, sugarcane, and banana). Based on information from sequenced genomes, maize, wheat, and the common ancestor of grasses were proposed to result from allopolyploidy, with only the most recent duplication in sugarcane proposed to be autopolyploidy (Schnable et al., 2011; Chalhoub et al., 2014; Mayer et al., 2014). Recently, we provided genomic evidence that soybean and Salicaceae plants may have autopolyploid origin (Wang et al., 2017; Liu et al., 2017). Here, the observation of balanced fractionation throughout the subgenomes after each of the AAT and ART provided evidence of their likely auto-tetraploidization nature. This is a singular observation so far in all sequenced genomes. However, we have to note that widespread genomic fractionation over millions of years does not allow the credible inference of auto- or allopolyploidy (Doyle and Egan, 2010) (Woodhouse et al., 2010).

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Transparent Methods, 9 figures, and 13 tables and can be found with this article online at https://doi.org/10.1016/j.isci.2018.08.003.

ACKNOWLEDGMENTS

We appreciate financial support from the Ministry of Science and Technology of the People's Republic of China (2016YFD0101001), China National Science Foundation (31510333 to J.W. and 3117022 to X.W.), and Natural Science Foundation of Hebei Province (C2015209069 to J.W.). We thank the helpful discussion with researchers at the iGeno Co. Ltd., China.

AUTHOR CONTRIBUTIONS

Conceptualization: X.W. and J.-P.W.; Formal Analysis: J.-P.W., J.-G.Y., J.L., P.-C.S., L.W., F.-B.M., J.-Q.Y., S.-R.S., Y.-X.L., Y.-X.P., W.-N.G., Z.-Y.W., S.-Q.S., L.Z., X.-M.S., C.L., X.-Q.D., Y.-Q.X., Y.H., J.Z., and J.-Y.W.; Methodology: J.-P.W., and L.W.; Software: P.-C.S., T.-Y.L. and L.Z.; Data Curation: J.-G.Y. and J.L.; Writing – Original Draft: J.-P.W., J.-G.Y., and J.L.; Writing –Review & Editing: X.W., J.-P.W.; Supervision: X.W.

DECLARATION OF INTERESTS

The authors declare no conflict of interest.

Received: May 5, 2018
Revised: July 31, 2018
Accepted: August 2, 2018
Published: September 20, 2018

REFERENCES

Anderberg, A.A., Rydin, C., and Källersjö, M. (2002). Phylogenetic relationships in the order Ericales SL: analyses of molecular data from five genes from the plastid and mitochondrial genomes. Am. J. Bot. *89*, 677–687.

Barker, M.S., Husband, B.C., and Pires, J.C. (2016). Spreading wings and flying high: the evolutionary importance of polyploidy after a century of study. Am. J. Bot. *103*, 1139.

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I.A., Tang, H., Wang, X., Chiquet, J., Belcram, H., Tong, C., and Samans, B. (2014). Plant genetics. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. Science *345*, 950–953.

Doyle, J.J., and Egan, A.N. (2010). Dating the origins of polyploidy events. New Phytol. *186*, 73–85.

Fraser, P.D., and Bramley, P.M. (2004). The biosynthesis and nutritional uses of carotenoids. Prog. Lipid Res. *43*, 228–265.

Huang, S., Ding, J., Deng, D., Tang, W., Sun, H., Liu, D., Zhang, L., Niu, X., Zhang, X., and Meng, M. (2013). Draft genome of the kiwifruit Actinidia chinensis. Nat. Commun. *4*, 2640.

Jaillon, O., Aury, J.M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., Choisne, N., Aubourg, S., Vitulo, N., Jubin, C., et al. (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature *449*, 463–467.

Jiao, Y., Wickett, N.J., Ayyampalayam, S., Chanderbali, A.S., Landherr, L., Ralph, P.E., Tomsho, L.P., Hu, Y., Liang, H., and Soltis, P.S. (2011). Ancestral polyploidy in seed plants and angiosperms. Nature *473*, 97–100.

Jiao, Y., Leebens-Mack, J., Ayyampalayam, S., Bowers, J.E., McKain, M.R., McNeal, J., Rolf, M., Ruzicka, D.R., Wafula, E., Wickett, N.J., et al. (2012). A genome triplication associated with early diversification of the core eudicots. Genome Biol. *13*, R3.

Kellogg, E.A. (2016). Has the connection between polyploidy and diversification actually been tested? Curr. Opin. Plant Biol. *30*, 25–32.

Liu, Y., Wang, J., Ge, W., Wang, Z., Li, Y., Yang, N., Sun, S., Zhang, L., and Wang, X. (2017). Two highly similar poplar paleo-subgenomes suggest an autotetraploid ancestor of salicaceae plants. Front. Plant Sci. *8*, 571.

Maere, S., De Bodt, S., Raes, J., Casneuf, T., Van Montagu, M., Kuiper, M., and Van de Peer, Y. (2005). Modeling gene and genome duplications in eukaryotes. Proc. Natl. Acad. Sci. USA *102*, 5454.

Magallón, S., Gómez-Acevedo, S., Sánchez-Reyes, L.L., and Hernández-Hernández, T. (2015). A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. New Phytol. *207*, 437.

Mayer, K.F.X., Rogers, J., Doležel, J., Pozniak, C., Eversole, K., Feuillet, C., Gill, B., Friebe, B., Lukaszewski, A.J., and Sourdille, P. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science *345*, 1251788.

Paterson, A.H., Bowers, J.E., and Chapman, B.A. (2004). Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc. Natl. Acad. Sci. USA *101*, 9903.

Paterson, A.H., Wendel, J.F., Gundlach, H., Guo, H., Jenkins, J., Jin, D., Llewellyn, D., Showmaker, K.C., Shu, S., and Udall, J. (2012). Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. Nature *492*, 423.

Potato Genome Sequencing Consortium, Xu, X., Pan, S., Cheng, S., Zhang, B., Mu, D., Ni, P., Zhang, G., Yang, S., Li, R., et al. (2011). Genome sequence and analysis of the tuber crop potato. Nature *475*, 189.

Schnable, J.C., Springer, N.M., and Freeling, M. (2011). Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. Proc. Natl. Acad. Sci. USA *108*, 4069–4074.

Seddon, J.M., Ajani, U.A., Sperduto, R.D., Hiller, R., Blair, N., Burton, T.C., Farber, M.D., Gragoudas, E.S., Haller, J., and Miller, D.T. (1994). Dietary carotenoids, vitamins A, C, and E, and advanced age-related macular degeneration. Eye Disease Case-Control Study Group. JAMA *272*, 1413–1420.

Soltis, D.E., Bell, C.D., Kim, S., and Soltis, P.S. (2008). Origin and early evolution of angiosperms. Ann. New York Acad. Sci. *1133*, 3.

Soltis, D.E., Visger, C.J., and Soltis, P.S. (2014). The polyploidy revolution then…and now: stebbins revisited. Am. J. Bot. *101*, 1057–1078.

Soltis, P.S., Marchant, D.B., Van de Peer, Y., and Soltis, D.E. (2015). Polyploidy and genome evolution in plants. Curr. Opin. Genet. Dev. *35*, 119–125.

Tao, S., HongWen, H., Barker, M.S., and HeslopHarrison, J.S.P. (2010). Ancient genome duplications during the evolution of kiwifruit (Actinidia) and related Ericales. Ann. Bot. *106*, 497–504.

Vekemans, D., Proost, S., Vanneste, K., Coenen, H., Viaene, T., Ruelens, P., Maere, S., Van de Peer, Y., and Geuten, K. (2012). Gamma paleohexaploidy in the stem lineage of core eudicots: significance for MADS-box gene and species diversification. Mol. Biol. Evol. *29*, 3793.

Wang, X., Shi, X., Li, Z., Zhu, Q., Kong, L., Tang, W., Ge, S., and Luo, J. (2006). Statistical inference of chromosomal homology based on gene colinearity and applications to Arabidopsis and rice. BMC Bioinformatics *7*, 447.

Wang, X., Wang, J., Jin, D., Guo, H., Lee, T.H., Liu, T., and Paterson, A.H. (2015). Genome alignment spanning major Poaceae lineages reveals heterogeneous evolutionary rates and alters inferred dates for key evolutionary events. Mol. Plant *8*, 885–898.

Wang, X., Guo, H., Wang, J., Lei, T., Liu, T., Wang, Z., Li, Y., Lee, T.H., Li, J., and Tang, H. (2016). Comparative genomic de-convolution of the cotton genome revealed a decaploid ancestor and widespread chromosomal fractionation. New Phytol. *209*, 1252.

Wang, J., Sun, P., Li, Y., Liu, Y., Yu, J., Ma, X., Sun, S., Yang, N., Xia, R., Lei, T., et al. (2017). Hierarchically aligning 10 legume genomes establishes a family-level genomics platform. Plant Physiol. *174*, 284–300.

Wang, J., Sun, P., Li, Y., Liu, Y., Yang, N., Yu, J., Ma, X., Sun, S., Xia, R., Liu, X., et al. (2018). An overlooked paleo-tetraploidization in Cucurbitaceae. Mol. Biol. Evol. *35*, 16–26.

Woodhouse, M.R., Schnable, J.C., Pedersen, B.S., Lyons, E., Lisch, D., Subramaniam, S., and Freeling, M. (2010). Following tetraploidy in maize, a short deletion mechanism removed genes preferentially from one of the two homologs. PLoS Biol. *8*, e1000409.

ISCI, Volume *7*

Supplemental Information

Two Likely Auto-Tetraploidization Events Shaped Kiwifruit Genome and Contributed to Establishment of the Actinidiaceae Family

Jin-Peng Wang, Ji-Gao Yu, Jing Li, Peng-Chuan Sun, Li Wang, Jia-Qing Yuan, Fan-Bo Meng, Sang-Rong Sun, Yu-Xian Li, Tian-Yu Lei, Yu-Xin Pan, Wei-Na Ge, Zhen-Yi Wang, Lan Zhang, Xiao-Ming Song, Chao Liu, Xue-Qian Duan, Shao-Qi Shen, Yang-qin Xie, Yue Hou, Jin Zhang, Jian-Yu Wang, and Xiyin Wang

SUPPLEMENTAL INFORMATION

Two likely auto-tetraploidization events shaped kiwifruit genome and contributed to establishment of the Actinidiaceae family

Jin-Peng Wang[1,2,3], Ji-Gao Yu[1,2,3], Jing Li[1,2,3], Peng-Chuan Sun[2,3], Li Wang[1,2,3], Jia-Qing Yuan[1,2], Fan-Bo Meng[1,2], Sang-Rong Sun[1,2], Yu-Xian Li[1,2], Tian-Yu Lei[1,2], Yu-Xin Pan[1,2], Wei-Na Ge[1,2], Zhen-Yi Wang[1,2], Lan Zhang[1,2], Xiao-Ming Song[1,2], Chao Liu[1,2], Xue-Qian Duan[1], Shao-Qi Shen[1], Yang-qin Xie[1], Yue Hou[1], Jin Zhang[1], Jian-Yu Wang[1], Xiyin Wang[1,2,*]

[1] School of Life Sciences, North China University of Science and Technology, No.21 Bohai Road, Caofeidian, Tangshan, Hebei 063210, China

[2] Center for Genomics and Computational Biology, North China University of Science and Technology, No.21 Bohai Road, Caofeidian, Tangshan, Hebei 063210, China

[3] These authors contributed equally.

[*] Lead contact/Corresponding Author: Xiyin Wang

Email: wangxiyin@vip.sina.com

TRANSPARENT METHODS

Materials

Genomic sequences and annotations were downloaded from the respective websites for each genome project, for which complete information can be found in Table S13. Transcriptome data of 4 kiwifruit tissues (Mature leaves, immature fruits (20 days after pollination (DAP)), mature green fruits (120 DAP) and ripe fruits (127 DAP)) from a 5-year-old 'Hongyang' plant generated by Huang *et al.* was obtained from the NCBI Short Read Archive (https://www.ncbi.nlm.nih.gov/sra/?term=SRA065642) (Huang et al., 2013).

Gene colinearity

Colinear genes were inferred by using ColinearScan (Wang et al., 2006), a statistically well-supported algorithm and software. Maximal gap length between genes in colinearity along a chromosome sequence was set to be 50 genes apart, as adopted in many previous studies (Wang et al., 2005; Wang et al., 2015; Wang et al., 2016a; Wang et al., 2016b; Wang et al., 2017). Homologous gene dotplots within a genome or between different genomes were produced by using MCSCANX toolkits (Wang et al., 2012), of which the corresponding author contributed to direct development. The gold-standard pipeline to decipher complex genomes was carefully followed as described previously (Wang et al., 2017).

Construction of an event-related, colinear gene table

To construct the table with the grape genome as a reference (Table S6), all grape genes were listed in the first column. Each grape gene may have two extra colinear genes due to the hexaploidy of the genome, and we assigned two other columns in the table to contain this information. For each grape gene, when there was a corresponding colinear gene in an expected location, a gene ID was entered in a cell of the corresponding column in the table. When a colinear gene was missing, often due to gene loss or translocation in the genome, we entered a dot in the cell. For the

coffee genome, without extra duplications, we assigned one column. Meanwhile, for the kiwifruit genome, with two paleo-terapoidization events, we assigned four columns. Therefore, the table had 18 columns, reflecting layers and layers of threefold and then fourfold homology due to recursive polyploidies across the genomes. The coffee-referenced table was constructed similarly (Table S7).

Nucleotide substitution

Evolutionary divergence between homologous genes and kernel function analysis of Ks was estimated as described previously (Wang et al., 2018; Wang et al., 2015).

Calculation of Gene Expression Levels

For all expression sets, reads were aligned to the kiwifruit genome using hierarchical indexing for spliced alignment of transcripts (HISAT2) (Kim et al., 2015). The align results were sorted and transfer 'bam' to 'sam' format by SAMTOOLS (Li et al., 2009). Following the alignment, estimates the expression levels for each kiwifruit gene model using StringTie (Pertea et al., 2015).

Evolutionary dating correction

By aligning the peaks of the kiwifruit, coffee, and grape ECH events from different Ks distributions with the corresponding location in the grape Ks distribution, we performed evolutionary rate correction. Supposing that the grape and kiwifruit peak appears at k_{VK} and that the peak for coffee appears a k_C, the relative evolutionary rate of coffee can be described with

$r = (k_C - k_{VK})/k_{VK}$.

We then performed rate correction to find the corrected rate $k_{C\text{-}correction}$ of coffee relative to k_{VK}:

(1) For the Ks between duplicates in coffee, we can define the correction coefficient W_C as

$$\frac{k_{C\text{-}correction}}{k_C} = \frac{k_{VK}}{k_C} = W_C;$$

therefore, we get

$$k_{C\text{-}correction} = \frac{k_{VK}}{k_C} \times k_C = \frac{1}{1+r} \times k_C$$

and $W_C = \frac{1}{1+r}$.

(2) For the Ks between homologous genes from grape (kiwifruit) and coffee, if the peak was located at $k_{C\text{-}VK}$, supposing the correction coefficient W_C in coffee, we then calculated a corrected evolutionary rate $k_{C\text{-}VK\text{-}correction} = W_C \times k_{C-VK}$.

SUPPLEMENTAL TABLES

Table S1. Number of homologous blocks and gene pairs within each genome or between genomes, Related to Figure 1.

Homologous blocks within and among genomes	BL[a] > 4	BL > 10	BL > 20	BL > 50	ACGP[b]	LDB[c]	LDB chromosomal locations
Kiwifruit	9,998/964[d]	6,023/278	3,622/91	1,440/16	10.37, 21.67, 39.80, 90.00	195	Kw19-Kw23
Grape	2,099/214	1,279/58	825/26	111/2	9.81, 22.05, 31.73, 55.50	61	Vv5-Vv7
Coffee	2,502/279	1,330/47	1,049/26	408/6	8.97, 28.30, 40.35, 68.00	95	Cc01-Cc06
Kiwifruit* vs *grape	18,626/1,631	12,492/541	8,142/211	2,862/40	11.42, 23.09, 38.59, 71.55	145	Vv18-Kw04
Kiwifruit* vs *coffee	21,448/1,869	14,108/520	10,220/232	5,180/57	11.48, 27.13, 44.05, 90.88	191	Cc04-Kw16
Grape* vs *coffee	15,605/1,262	10,256/233	9,044/143	6,310/56	12.37, 44.02, 63.24, 112.68	328	Vv5-Cc03

[a] BL: block length; [b] ACGP: average colinear gene pairs, respectively, per block; [c] LDB: number of colinear gene pairs residing in the longest duplicated block; [d] number of gene pairs/number of homologous blocks.

Table S2. Number of homologous genes residing in blocks within each genome or between genomes, Related to Figure 1.

Homologous blocks within and among genomes	BL[b] > 4	BL > 10	BL > 20	BL > 50	LDB [c]	LDB chromosomal locations
***Kiwifruit* (30906)[a]**	11,535	9,047	6,301	2,795	195	Kw19-Kw23
***Grape* (24901)**	3,404	2,284	1,552	222	61	Vv5-Vv7
***Coffee* (21971)**	3,710	2,236	1,812	800	95	Cc01-Cc06
Kiwifruit* vs *Grape	1,3121 vs 9,214	10,460 vs 7,417	7,555 vs 5,571	2,858 vs 2,371	145	Vv18-Kw04
Kiwifruit* vs *Coffee	14,314 vs 10,076	11,748 vs 8,040	9,306 vs 6,511	5,168 vs 3,916	191	Cc04-Kw16
Grape* vs *Coffee	9,876 vs 9,946	8,340 vs 8,278	7,612 vs 7,585	5,822 vs 5,796	328	Vv5-Cc03

[a] Total gene numbers anchored on chromosomes in each genome; [b] BL: block_length; [c] LDB: number of colinear gene pairs residing in the longest duplicated block.

Table S3. Paralogous, orthologous, and outparalogous gene pairs within a genome or between genomes, Related to Figure 1.

Genome	Grape	Coffee	Kiwifruit
Grape	2432	7494	9125
Coffee	2965	1172	9508
Kiwifruit	3681	3259	5905

Numbers on the main diagonal denote paralogous gene pairs within a genome, numbers above the diagonal denote orthologous gene pairs between two genomes, and numbers below the diagonal denote outparalogous gene pairs between two genomes.

Table S4. Paralogous, orthologous, and outparalogous genes within a genome or between genomes, Related to Figure 1.

Genome	Grape	Coffee	Kiwifruit
Grape	3866	7366/7395	6476/8923
Coffee	2587/2666	2103	6702/9322
Kiwifruit	3235/2412	2885/2193	8057

See the legends of Table S3.

In non-diagonal cells, gene numbers in two corresponding species are shown, from vertical and horizontal lists respectively.

Table S5. Paralogous, orthologous, and outparalogous blocks within a genome or between genomes, Related to Figure 1.

Genome	Grape	Coffee	Kiwifruit
Grape	87	153	402
Coffee	152	54	309
Kiwifruit	367	261	422

See the legends of Table S3.

In non-diagonal cells, gene numbers in two corresponding species are shown, from vertical and horizontal lists respectively.

Table S6. Homologous alignments of kiwifruit and coffee genomes with grape as reference. (Excel table), Related to Figure 4.

Table S7. Homologous alignments of kiwifruit genome with coffee as reference. (Excel table), Related to Figure S5.

Table S8. Kiwifruit gene retention rates in subgenomes with coffee as reference genome, Related to Figure S9.

Coffee		Kiwifruit homoelogous subgenomes					
Chr #	**Gene #**	**T1- A**	**T1-B**	**T1 A-B difference**	**T2-A**	**T2-B**	**T2 A-B difference**
1	2198	0.84	0.84	0.00	0.74	0.78	0.04
2	4000	0.81	0.79	0.02	0.74	0.82	0.08
3	1632	0.87	0.75	0.12	0.79	0.79	0.00
4	1727	0.78	0.77	0.01	0.86	0.85	0.01
5	1661	0.90	0.72	0.18	0.82	0.86	0.04
6	2389	0.73	0.86	0.13	0.79	0.80	0.01
7	2146	0.77	0.64	0.13	0.92	0.74	0.18
8	1718	0.86	0.73	0.13	0.81	0.85	0.04
9	1094	0.76	0.76	0.00	0.78	0.90	0.12
10	1653	0.75	0.71	0.04	0.83	0.84	0.01
11	1753	0.66	0.88	0.22	0.78	0.76	0.02
Total/Aver.	21971	0.79	0.77	0.09	0.81	0.82	0.05

Table S9. Kernel function analysis of Ks distribution related to duplication events within each genome and between genomes (before evolutionary rate correction), Related to Figure 1.

Intragenomic/intergenomic colinear gene pairs	Weight coefficient related to duplication event or speciation	Peak of Ks distribution (μ)	Deviation (σ)
Grape ECH-related	0.610	1.053	0.266
Coffee ECH-related	0.490	1.55	0.297
Kiwifruit ECH-related	0.787	1.053	0.342
Kiwifruit AAT-related	0.637	0.462	0.156
Kiwifruit ART-related	0.865	0.164	0.093
Grape- Kiwifruit	0.531	0.764	0.082
Grape- Kiwifruit (ECH)	1.023	1.034	0.405
Coffee- Kiwifruit	0.552	0.965	0.076
Coffee- Kiwifruit (ECH)	0.944	1.238	0.481

Table S10. Kernel function analysis of Ks distribution related to duplication events within each genome and between genomes (after evolutionary rate correction), Related to Figure 1.

Intragenomic/intergenomic colinear gene pairs	Weight coefficient related to duplication event or speciation	Peak of Ks distribution (μ)	Deviation (σ)
Grape ECH-related	0.610	1.053	0.266
Coffee ECH-related	0.502	1.053	0.210
Kiwifruit ECH-related	0.787	1.053	0.342
Kiwifruit AAT-related	0.637	0.462	0.156
Kiwifruit ART-related	0.139	0.164	0.093
Grape- Kiwifruit	0.531	0.764	0.082
Grape- Kiwifruit (ECH)	1.023	1.034	0.405
Coffee- Kiwifruit	0.544	0.653	0.063
Coffee- Kiwifruit (ECH)	0.966	0.848	0.341

Table S11. Gene retention and expression of different tissues in homoeologous Actinidia chinensis regions (Excel table), Related to Figure 2.

Table S12. Ascorbic acid (vitamin C) biosynthesis and recycling pathway genes related to duplication events in each genome, Related to Figure 2.

Gene family	Description	Grape		Coffee		Kiwifruit			
		All	ECH-related	All	ECH-related	All	ECH-related	AAT-related	ART-related
AIase	Aldonolactonase	2	2 (100.00%)	0	0 (0.00%)	8	0 (0.00%)	3 (37.50%)	0 (0.00%)
AO	L-Ascorbate oxidase	1	0 (0.00%)	0	0 (0.00%)	7	0 (0.00%)	2 (28.57%)	1 (14.29%)
APX	L-Ascorbate peroxidase	4	2 (50.00%)	3	2 (66.67%)	21	7 (33.33%)	6 (28.57%)	10 (47.62%)
DHAR	Dehydroascorbate reductase	0	0 (0.00%)	1	0 (0.00%)	2	0 (0.00%)	0 (0.00%)	0 (0.00%)
GalLDH	L-Galactono-1,4-lactone dehydrogenase	1	0 (0.00%)	1	0 (0.00%)	1	0 (0.00%)	0 (0.00%)	1 (100.00%)
GalUR	D-Galacturonic acid reductase	11	10 (90.91%)	2	1 (50.00%)	22	6 (27.27%)	10 (45.45%)	8 (36.36%)
GDH	L-Galactose dehydrogenase	0	0 (0.00%)	0	0 (0.00%)	1	0 (0.00%)	0 (0.00%)	0 (0.00%)
GGP	GDP-L-galactose phosphorylase	1	1 (100.00%)	1	1 (100.00%)	4	1 (25.00%)	2 (50.00%)	1 (25.00%)
GLOase	L-Galactono-1,4-lactone dehydrogenase	1	0 (0.00%)	1	0 (0.00%)	1	0 (0.00%)	0 (0.00%)	1 (100.00%)
GME	GDP-D-mannose-3,5-epimerase	3	3 (100.00%)	1	1 (100.00%)	2	2 (100.00%)	2 (100.00%)	2 (100.00%)
GMP	GDP-D-mannose pyrophosphorylase	1	1 (100.00%)	1	1 (100.00%)	3	2 (0.00%)	2 (0.00%)	2 (66.67%)
GPP	L-Galactose-1-phosphate phosphatase	0	0 (0.00%)	0	0 (0.00%)	3	0 (0.00%)	0 (0.00%)	0 (0.00%)
IPS	Inositol-3-phosphate synthase	1	0 (0.00%)	1	0 (0.00%)	4	1 (25.00%)	2 (50.00%)	0 (0.00%)
MDHAR	Monodehydroascorbate reductase	4	2 (50.00%)	2	2 (100.00%)	10	3 (30.00%)	9 (90.00%)	8 (80.00%)

MIOX	*myo*-Inositol oxygenase	2	2 (100.00%)	2	1 (50.00%)	14	6 (42.86%)	6 (42.86%)	5 (35.71%)
PGI	Glucose-6-phosphate isomerase	1	0 (0.00%)	1	0 (0.00%)	6	1 (16.67%)	3 (50.00%)	4 (66.67%)
PMI	Mannose-6-phosphate isomerase	1	1 (100.00%)	2	2 (100.00%)	12	5 (41.67%)	8 (66.67%)	8 (66.67%)
PMM	Phosphomannomutase	1	0 (0.00%)	1	1 (100.00%)	2	1 (50.00%)	0 (0.00%)	1 (50.00%)
TOTAL		35	22（62.86%）	20	12（60.00%）	123	35 (28.46%)	56 (44.53%)	52 (42.28%)

Table S13. Genomic data information, Related to Figure 3.

Order	Species name	Common name	Version	Data source	Genes	Anchored Genes[a]	Reference
1	*Actinidia chinensis*	Kiwifruit	v1.0	JGI (http://bioinfo.bti.cornell.edu/kiwi)	39,040	30,906	(Huang et al., 2013)
2	*Coffea canephora*	Coffee	v1.0	(http://coffee-genome.org/)	25,574	21,971	(Denoeud et al., 2014)
3	*Vitis vinifera L*	Grape	Genoscope.12X	JGI (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Vvinifera)	37,829	24,901	(Jaillon et al., 2007)

[a] Total gene numbers anchored on chromosomes.

SUPPLEMENTAL FIGURE LEGENDS

Figure S1. Homologous dotplot within kiwifruit genomes, Related to Figure 2.
The best, secondarily, and other matched homologous gene pairs were shown by red, blue, and gray dots, respectively. Mean Ks of each inferred collinear block is shown besides. The Ks values <=0.30 are in red, >0.70 in black, and others in blue, often showing paralogy blocks related the three duplication events, respectively.

Figure S2. Homologous dotplot between grape and kiwifruit genomes, **Related to Figure 2.**
The best, secondary, and other matched homologous gene pairs output by Blast were dotplotted by using red, blue, and gray colors, respectively. Mean Ks of each inferred collinear block is shown besides. The Ks values <=0.85 are in red, and others in blue, often showing orthology and outparalogy, respectively. The 19 grape chromosomes are colored corresponding to the seven ancestral eudicot chromosomes, as described in the main text. Chromosome regions exhibiting genomic orthology were identified and are distinguished by solid and dashed rectangles corresponding to the ART and AAT events, respectively. Arrows of the same color indicate chromosomal segments on different kiwifruit chromosomes that correspond to the same grape chromosome, likely produced by chromosome breakages during evolution.

Figure S3. Homologous dotplot between coffee and kiwifruit genomes, **Related to Figure 2.**
The best, secondary, and other matched homologous gene pairs output by Blast were dotplotted by using red, blue, and gray colors, respectively. Mean Ks of each inferred collinear block is shown besides. The Ks values <=1.1 are in red, and others in blue, often showing orthology and outparalogy, respectively. The 11 coffee chromosomes are colored corresponding to the seven ancestral eudicot chromosomes, as described in the main text. Chromosome regions exhibiting genomic orthology were identified and are distinguished by solid and dashed rectangles corresponding to the ART and AAT events, respectively. Arrows of the same color indicate chromosomal segments on different kiwifruit chromosomes that correspond to the same coffee chromosome, likely produced by chromosome breakages during evolution.

Figure S4. Trees with topology supporting the AAT and ART, Related to Figure 2.
Homoeologous groups related to the eudicot-common hexaploidy are shown with hexagons of the same color, and a grape homoeolog and corresponding kiwifruit orthologs are shown with red rectangles and green rectangles, between them related AAT.

Figure S5. Homologous alignments of the kiwifruit genome with grape as reference, Related to Figure 4.

With grape as a reference genome, genomic paralogy, orthology, and outparalogy information within kiwifruit is displayed in 15 circles: The curved lines within the inner circle, colored according to the 7 ancestral eudicot chromosomes (Jaillon et al. 2007), link paralogous pairs on the 19 grape chromosomes produced by the ECH. The short lines forming the innermost circle represent all predicted genes in grape, which have two sets of paralogous regions, forming another two circles. Each of the three sets of grape paralogous chromosomal regions has four orthologous copies in kiwifruit. Therefore, the two genomes will result in 15 circles in the figure. Each circle is colored according to its source plant corresponding to the color of the 19 grape chromosomes. Homologous genes are denoted by short lines standing on a chromosome circle, and colored according to the chromosome number in the source plant shown in the inset legend.

Figure S6. Homologous alignments of the kiwifruit genome with coffee as reference, **Related to Figure 4.**

With coffee as a reference genome, genomic paralogy, orthology, and outparalogy information within kiwifruit is displayed in four circles. Detailed notation and explanation can be found in the legend of **Figure S5**.

Figure S7. Fitting a geometric distribution and gene loss rates in kiwifruit as compared with grape and coffee, Related to Figure 5.

(A) Kiwifruit with grape as reference genome. (B) Kiwifruit with coffee as reference genome.

Figure S8. Kiwifruit gene retention along corresponding orthologous grape chromosomes. Related to Figure 5,

Rates of retained genes in sliding windows of grape homoeologous region group 1 (red), homoeologous region group 2 (black), homoeologous region group 3 (green), homoeologous region group 4 (blue), and large patches of chromosomal segmental losses (yellow) are displayed; grape chromosomes 1-19.

Figure S9. Kiwifruit retention along corresponding orthologous coffee chromosomes. Related to Figure 5,

Rates of retained genes in sliding windows of grape homoelogous region group 1 (red), homoelogous region group 2 (black), homoelogous region group 3 (green), homoelogous region group 4 (blue), and large patches of chromosomal segmental losses (yellow) are displayed.

SUPPLEMANTAL REFERENCES

Denoeud F, Carreteropaulet L, Dereeper A, Droc G, Guyot R, Pietrella M, Zheng C, Alberti A, Anthony F, Aprea G (2014) The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. Science 345 (6201):1181

Huang S, Ding J, Deng D, Tang W, Sun H, Liu D, Zhang L, Niu X, Zhang X, Meng M, Yu J, Liu J, Han Y, Shi W, Zhang D, Cao S, Wei Z, Cui Y, Xia Y, Zeng H, Bao K, Lin L, Min Y, Zhang H, Miao M, Tang X, Zhu Y, Sui Y, Li G, Sun H, Yue J, Sun J, Liu F, Zhou L, Lei L, Zheng X, Liu M, Huang L, Song J, Xu C, Li J, Ye K, Zhong S, Lu BR, He G, Xiao F, Wang HL, Zheng H, Fei Z, Liu Y (2013) Draft genome of the kiwifruit Actinidia chinensis. Nature communications 4:2640. doi:10.1038/ncomms3640

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, Vezzi A, Legeai F, Hugueney P, Dasilva C, Horner D, Mica E, Jublot D, Poulain J, Bruyere C, Billault A, Segurens B, Gouyvenoux M, Ugarte E, Cattonaro F, Anthouard V, Vico V, Del Fabbro C, Alaux M, Di Gaspero G, Dumas V, Felice N, Paillard S, Juman I, Moroldo M, Scalabrin S, Canaguier A, Le Clainche I, Malacrida G, Durand E, Pesole G, Laucou V, Chatelet P, Merdinoglu D, Delledonne M, Pezzotti M, Lecharny A, Scarpelli C, Artiguenave F, Pe ME, Valle G, Morgante M, Caboche M, Adam-Blondon AF, Weissenbach J, Quetier F, Wincker P, French-Italian Public Consortium for Grapevine Genome C (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449 (7161):463-467. doi:10.1038/nature06148

Kim D, Langmead B, Salzberg SL (2015) HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12: 357-360.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078-2079.

Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL (2015) StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol 33: 290-295.

Wang X, Shi X, Li Z, Zhu Q, Kong L, Tang W, Ge S, Luo J (2006) Statistical inference of chromosomal homology based on gene colinearity and applications to Arabidopsis and rice. Bmc Bioinformatics 7 (1):447

Wang X, Shi X, Hao B, Ge S, Luo J. (2005) Duplication and DNA segmental loss in the rice genome: implications for diploidization. The New phytologist 165 (3):937

Wang J, Yu J, Sun P, Li Y, Xia R, Liu Y, Ma X, Yu J, Yang N, Lei T (2016a) Comparative Genomics Analysis of Rice and Pineapple Contributes to Understand the Chromosome Number Reduction and Genomic Changes in Grasses. Frontiers in Genetics 7

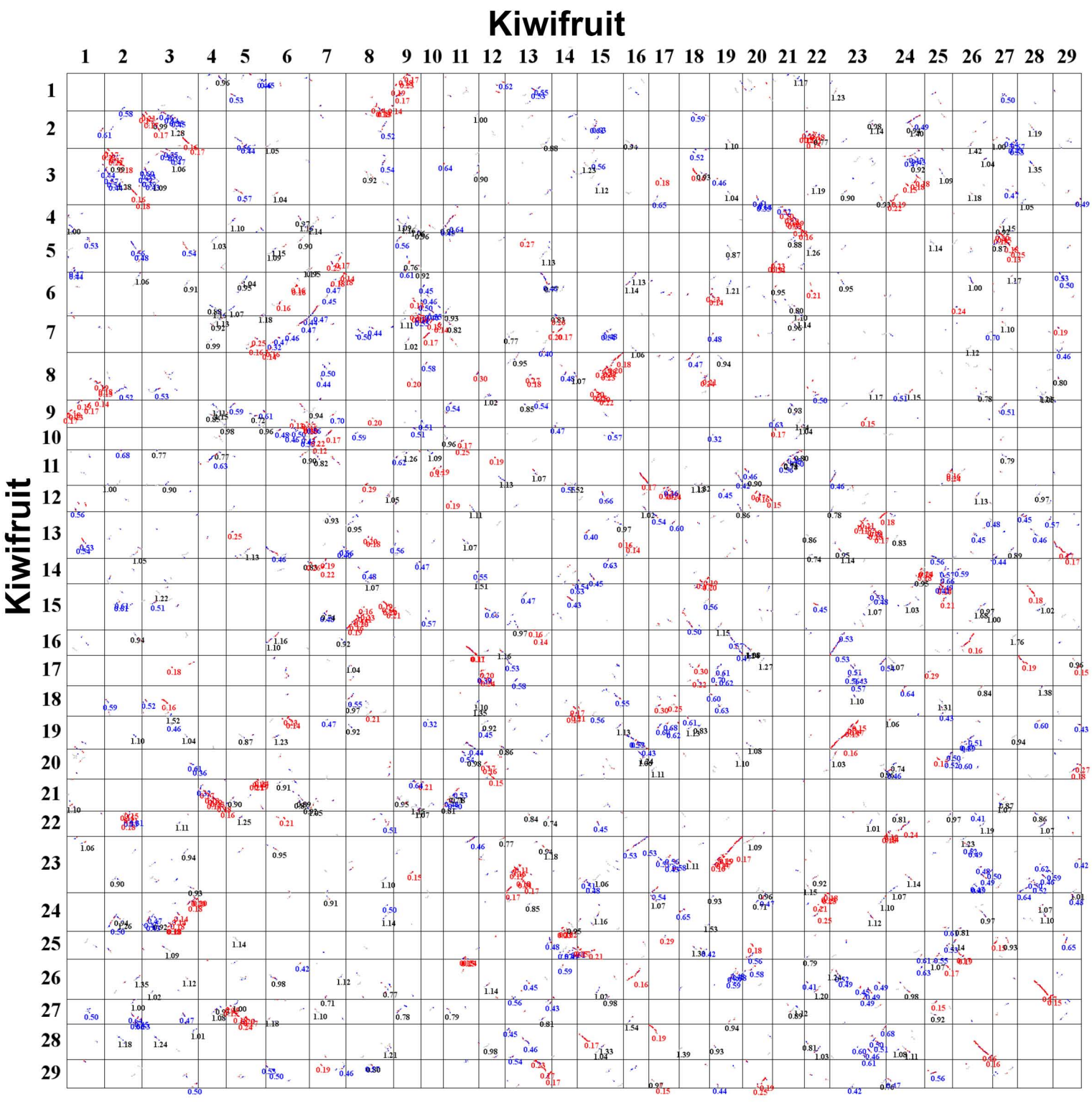
Wang X, Guo H, Wang J, Lei T, Liu T, Wang Z, Li Y, Lee TH, Li J, Tang H (2016b) Comparative genomic de‐convolution of the cotton genome revealed a decaploid ancestor and widespread chromosomal fractionation. New Phytologist 209 (3):1252

Wang J, Sun P, Li Y, Liu Y, Yu J, Ma X, Sun S, Yang N, Xia R, Lei T, Liu X, Jiao B, Xing Y, Ge W, Wang L, Wang Z, Song X, Yuan M, Guo D, Zhang L, Zhang J, Jin D, Chen W, Pan Y, Liu T, Jin L, Sun J, Yu J, Cheng R, Duan X, Shen S, Qin J, Zhang MC, Paterson AH, Wang X (2017) Hierarchically Aligning 10 Legume Genomes Establishes a Family-Level Genomics Platform. Plant physiology 174 (1):284-300. doi:10.1104/pp.16.01981

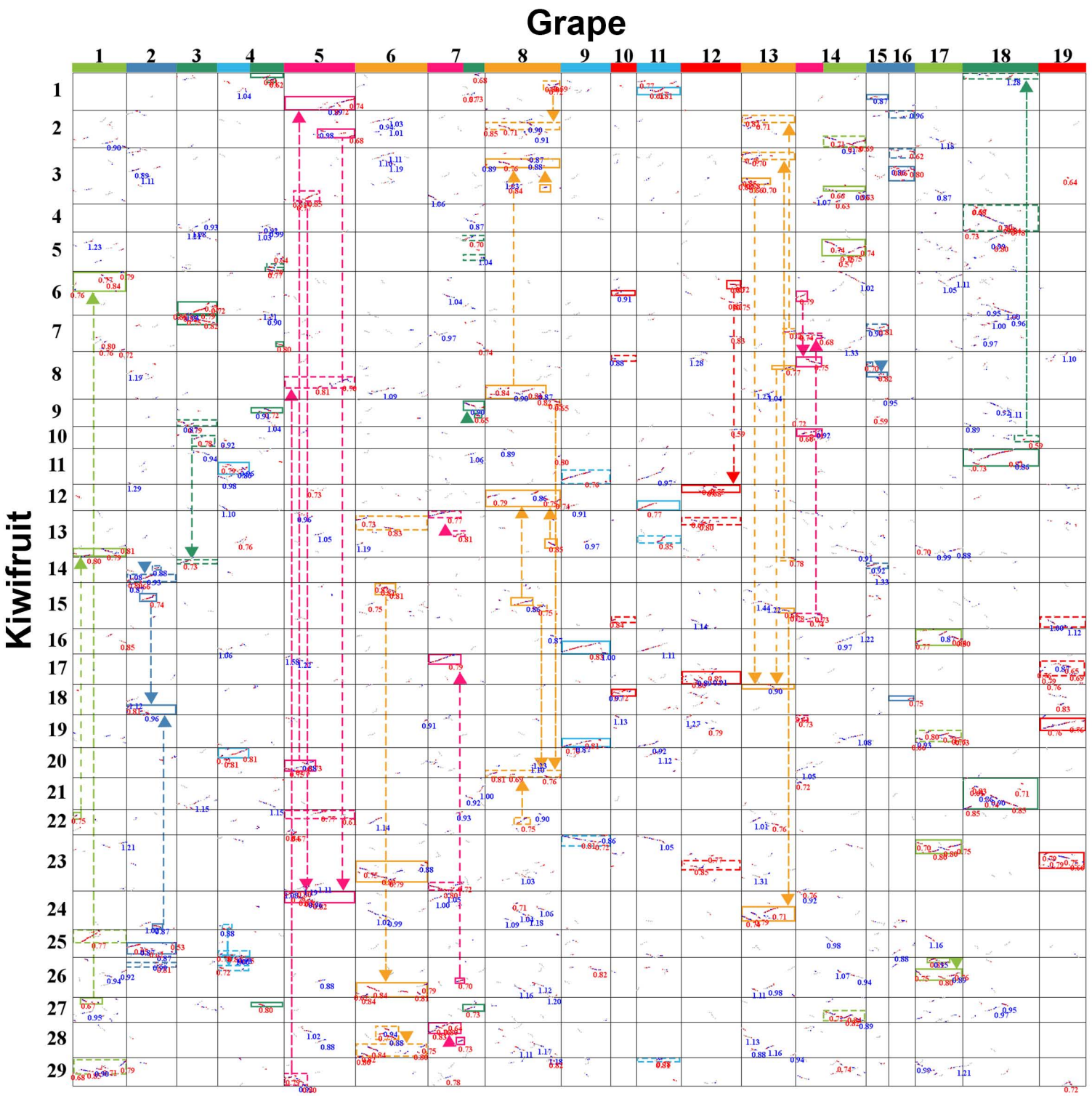
Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee T, Jin H, Marler B, Guo H (2012) MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Research 40 (7):e49-e49

Wang J, Sun P, Li Y, Liu Y, Yang N, Yu J, Ma X, Sun S, Xia R, Liu X, Ge D, Luo S, Liu Y, Kong Y, Cui X, Lei T, Wang L, Wang Z, Ge W, Zhang L, Song X, Yuan M, Guo D, Jin D, Chen W, Pan Y, Liu T, Yang G, Xiao Y, Sun j, Zhang C, Li Z, Xu H, Duan X, Shen S, Zhang Z, Huang S, Wang X (2018) An overlooked paleo-tetraploidization in Cucurbitaceae Molecular and Biological Evolution 35 (1):16-26

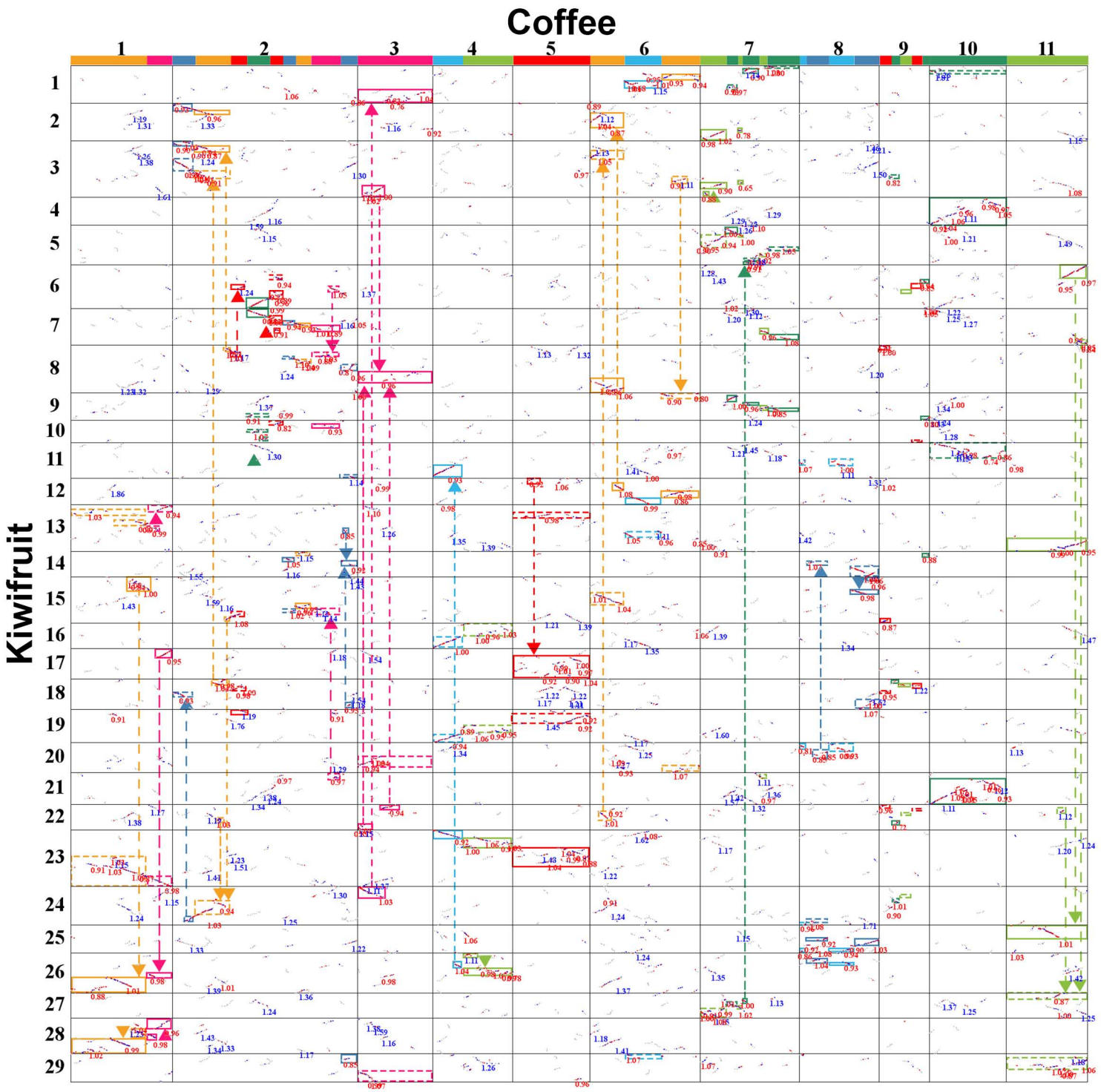
Wang X, Wang J, Jin D, Guo H, Lee TH, Liu T, Paterson AH (2015) Genome Alignment Spanning Major Poaceae Lineages Reveals Heterogeneous Evolutionary Rates and Alters Inferred Dates for Key Evolutionary Events. Molecular plant 8 (1):885-898



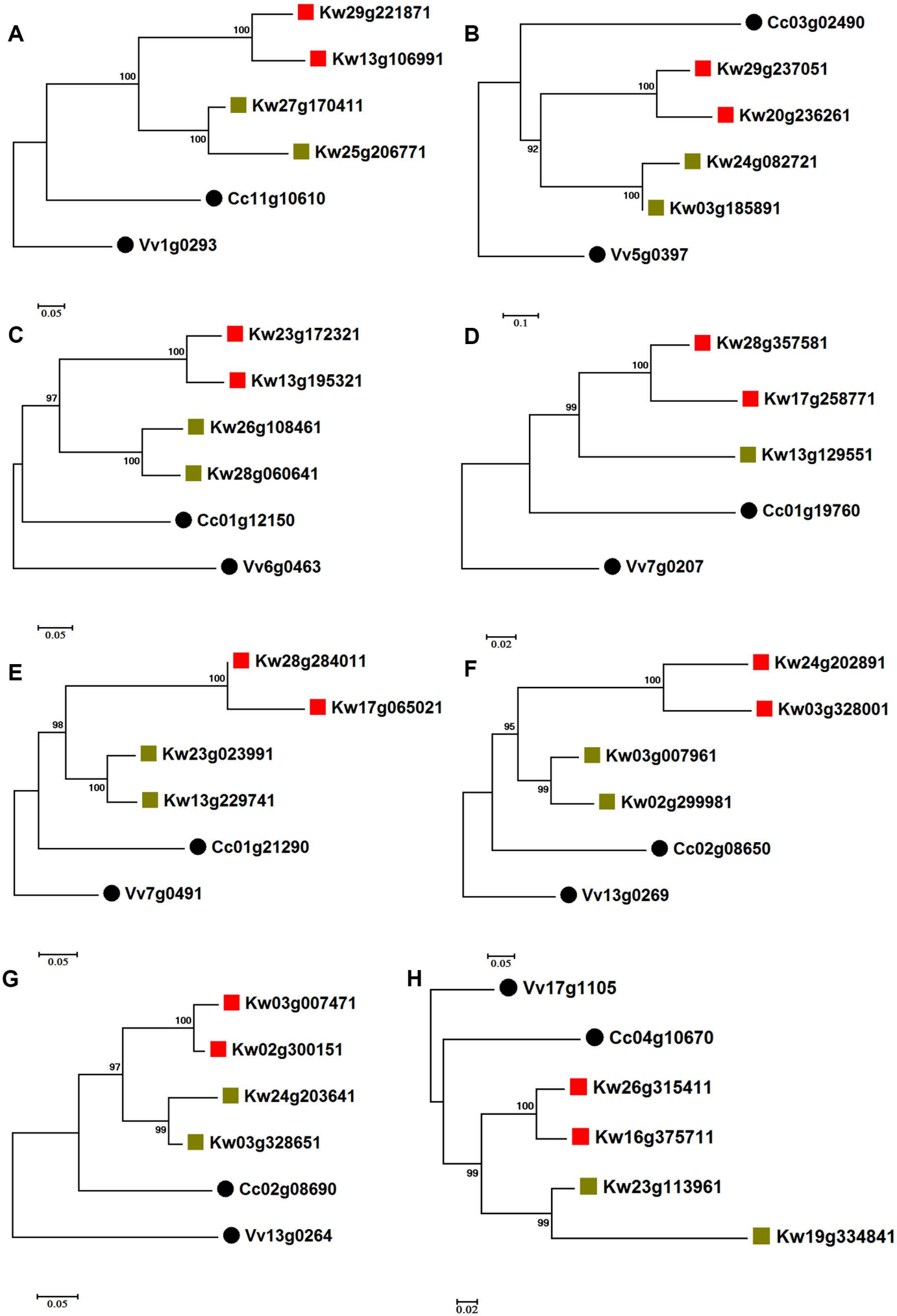
Supplemental Figure S1. Homologous dotplot within kiwifruit genomes. The best, secondarily, and other matched homologous gene pairs were shown by red, blue, and gray dots, respectively. Mean Ks of each inferred collinear block is shown besides. The Ks values <=0.30 are in red, >0.70 in black, and others in blue, often showing paralogy blocks related the three duplication events, respectively.



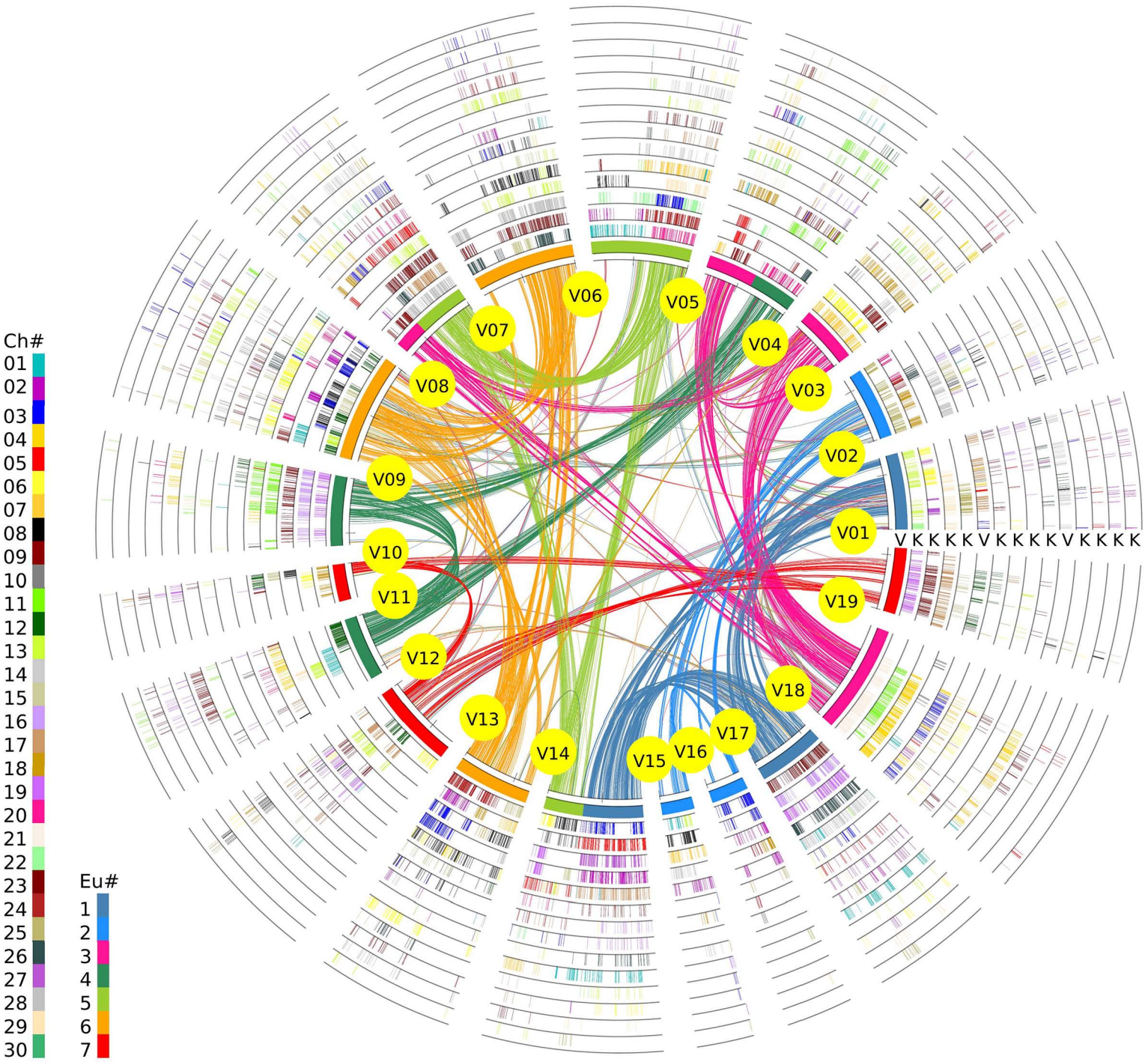
Supplemental Figure S2. Homologous dotplot between grape and kiwifruit genomes. The best, secondary, and other matched homologous gene pairs output by Blast were dotplotted by using red, blue, and gray colors, respectively. Mean Ks of each inferred collinear block is shown besides. The Ks values <=0.85 are in red, and others in blue, often showing orthology and outparalogy, respectively. The 19 grape chromosomes are olored corresponding to the seven ancestral eudicot chromosomes, as described in the main text. Chromosome regions exhibiting genomic orthology were identified and are distinguished by solid and dashed rectangles corresponding to the ART and AAT events, respectively. Arrows of the same color indicate chromosomal segments on different kiwifruit chromosomes that correspond to the same grape chromosome, likely produced by chromosome breakages during evolution.



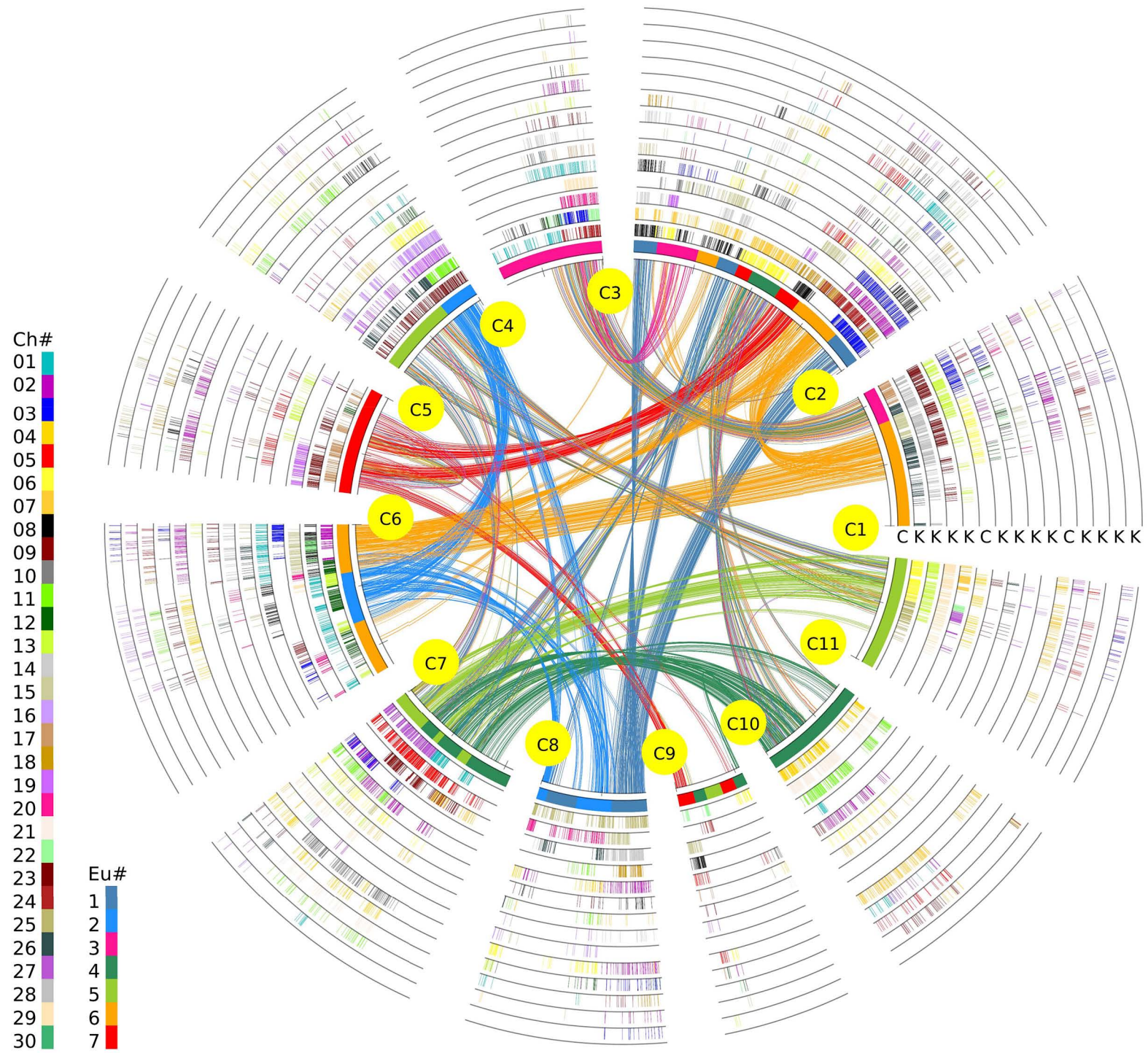
Supplemental Figure S3. Homologous dotplot between coffee and kiwifruit genomes. The best, secondary, and other matched homologous gene pairs output by Blast were dotplotted by using red, blue, and gray colors, respectively. Mean Ks of each inferred collinear block is shown besides. The Ks values <=1.1 are in red, and others in blue, often showing orthology and outparalogy, respectively. The 11 coffee chromosomes are colored corresponding to the seven ancestral eudicot chromosomes, as described in the main text. Chromosome regions exhibiting genomic orthology were identified and are distinguished by solid and dashed rectangles corresponding to the ART and AAT events, respectively. Arrows of the same color indicate chromosomal segments on different kiwifruit chromosomes that correspond to the same coffee chromosome, likely produced by chromosome breakages during evolution.



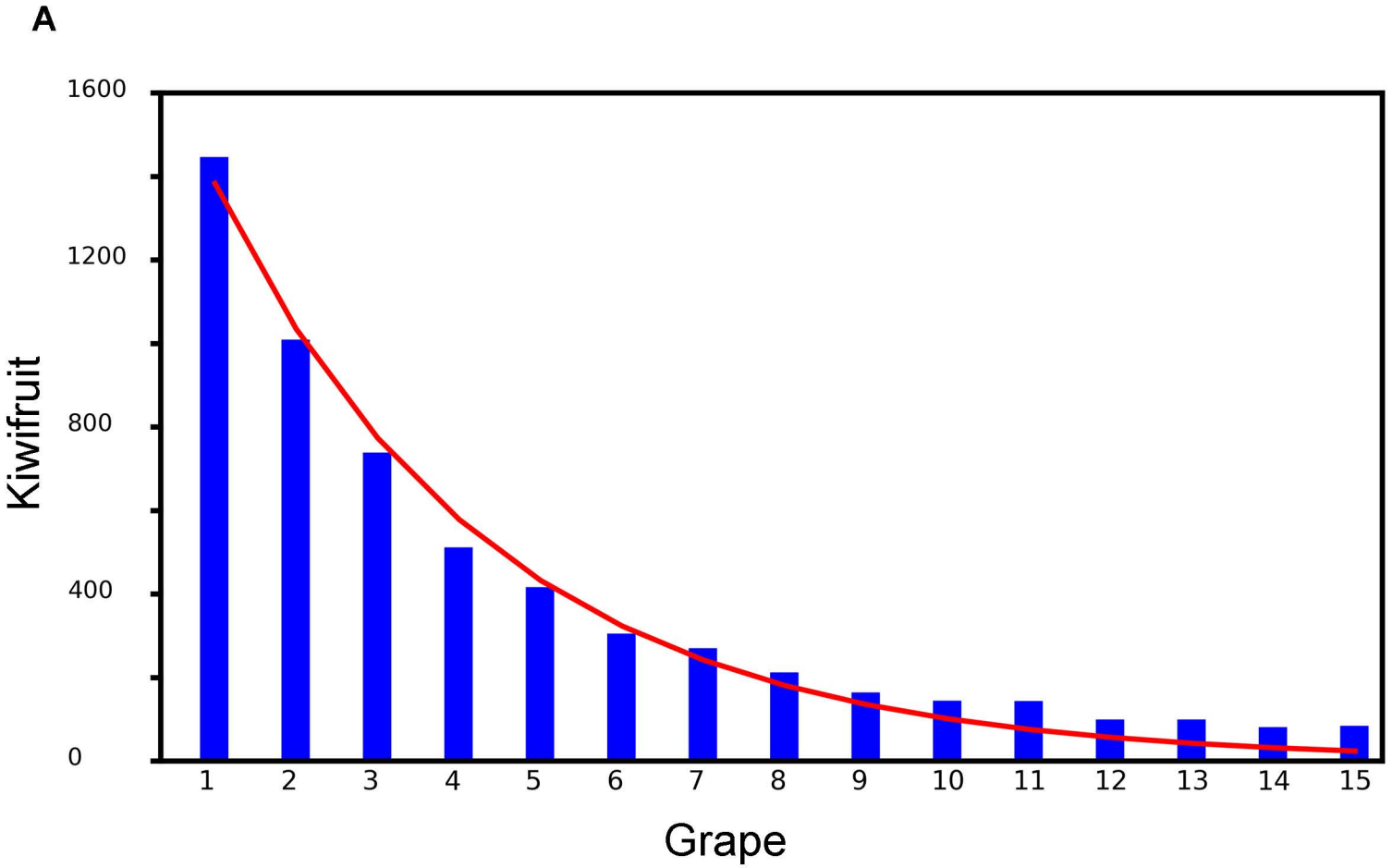
Supplemental Figure S4. Trees with topology supporting the AAT and ART. Homoeologous groups related to the eudicot-common hexaploidy are shown with hexagons of the same color, and a grape homoeolog and corresponding kiwifruit orthologs are shown with red rectangles and green rectangles, between them related AAT.

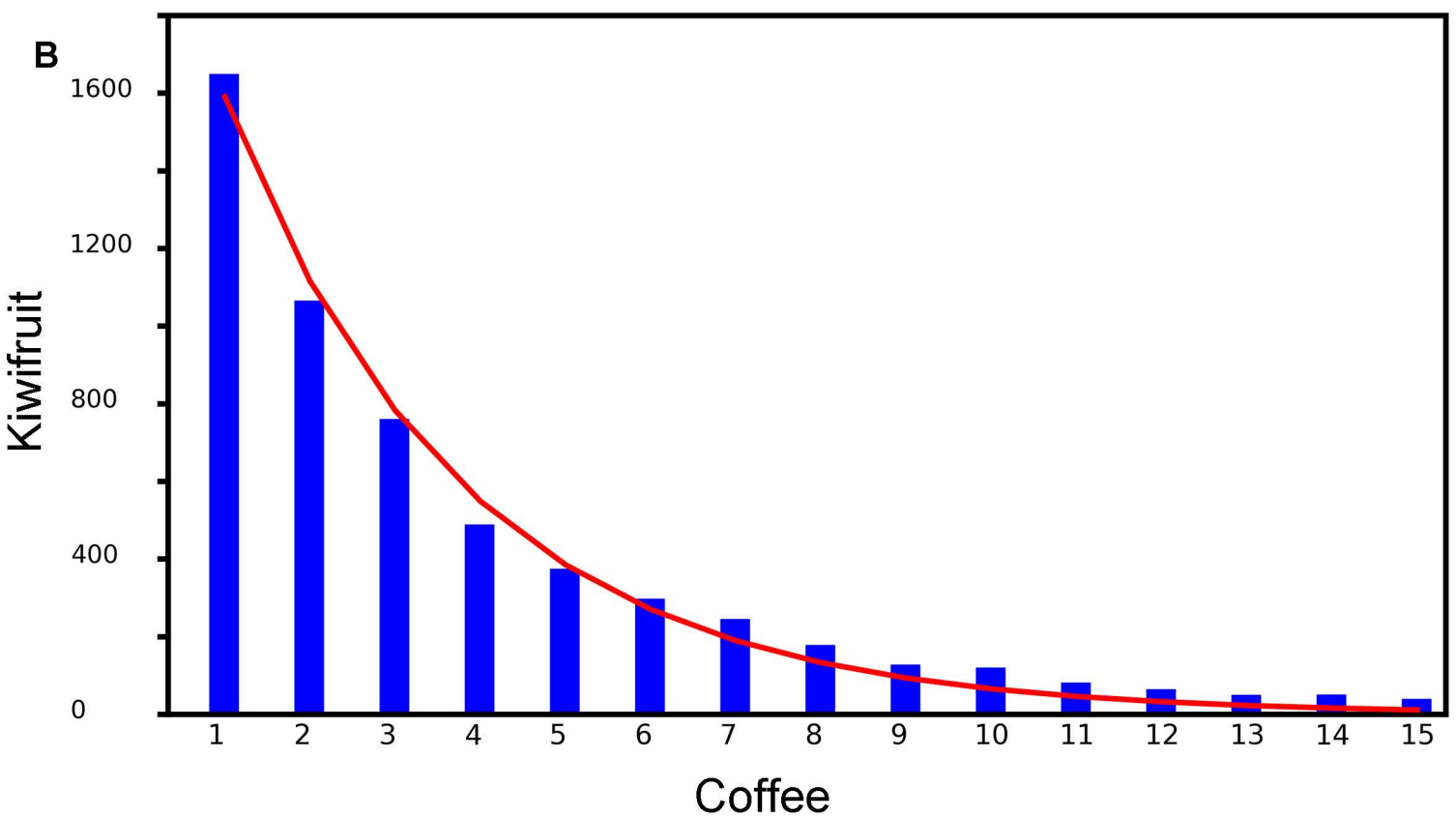


Supplemental Figure S5. Homologous alignments of the kiwifruit genome with grape as reference. With grape as a reference genome, genomic paralogy, orthology, and outparalogy information within kiwifruit is displayed in 15 circles: The curved lines within the inner circle, colored according to the 7 ancestral eudicot chromosomes (Jaillon et al., 2007), link paralogous pairs on the 19 grape chromosomes produced by the ECH. The short lines forming the innermost circle represent all predicted genes in grape, which have two sets of paralogous regions, forming another two circles. Each of the three sets of grape paralogous chromosomal regions has four orthologous copies in kiwifruit. Therefore, the two genomes will result in 15 circles in the figure. Each circle is colored according to its source plant corresponding to the color of the 19 grape chromosomes. Homologous genes are denoted by short lines standing on a chromosome circle, and colored according to the chromosome number in the source plant shown in the inset legend.

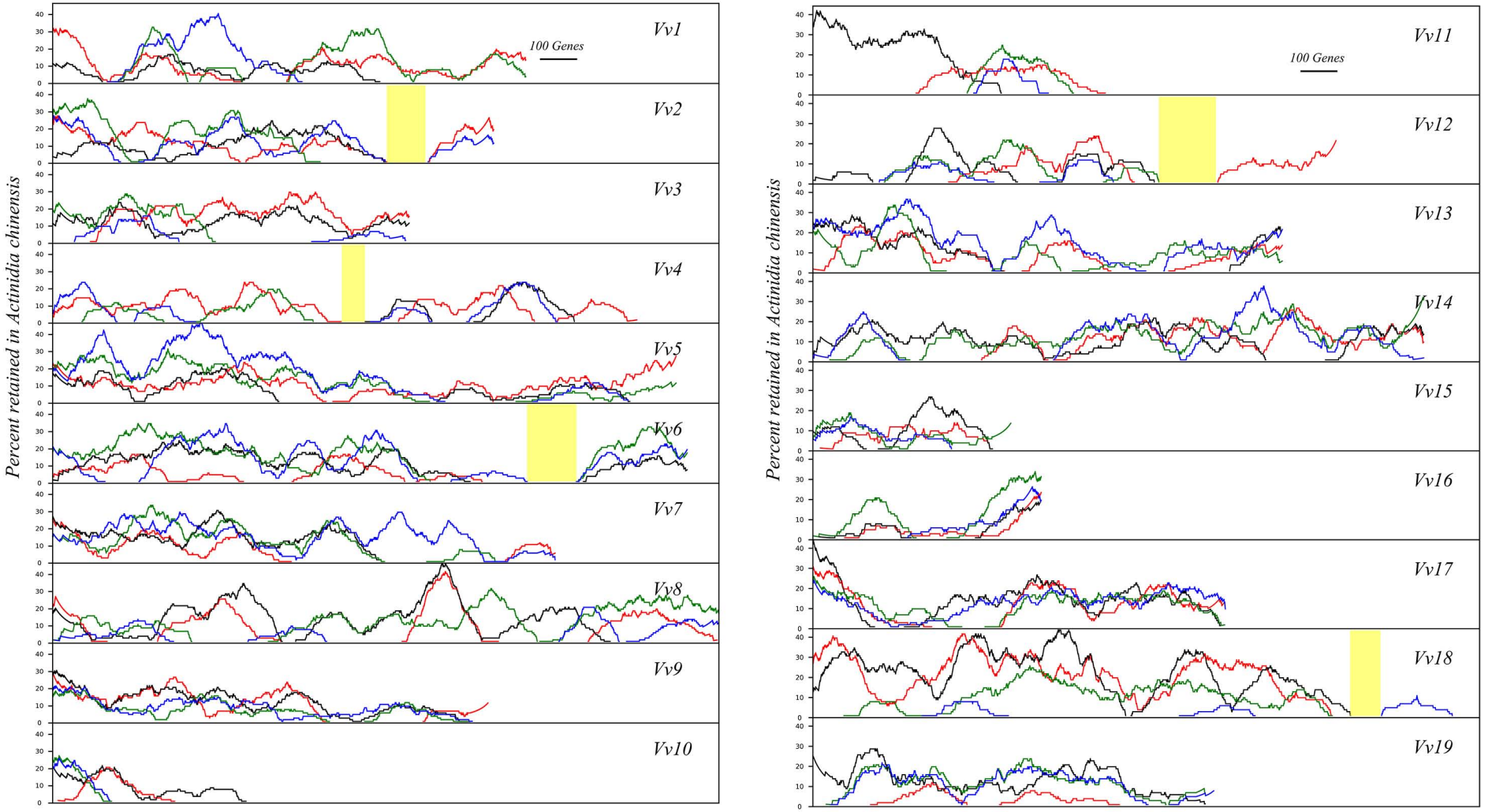


Supplemental Figure S6. Homologous alignments of the kiwifruit genome with coffee as reference. With coffee as a reference genome, genomic paralogy, orthology, and outparalogy information within kiwifruit is displayed in four circles. Detailed notation and explanation can be found in the legend of Supplemental Figure 5.

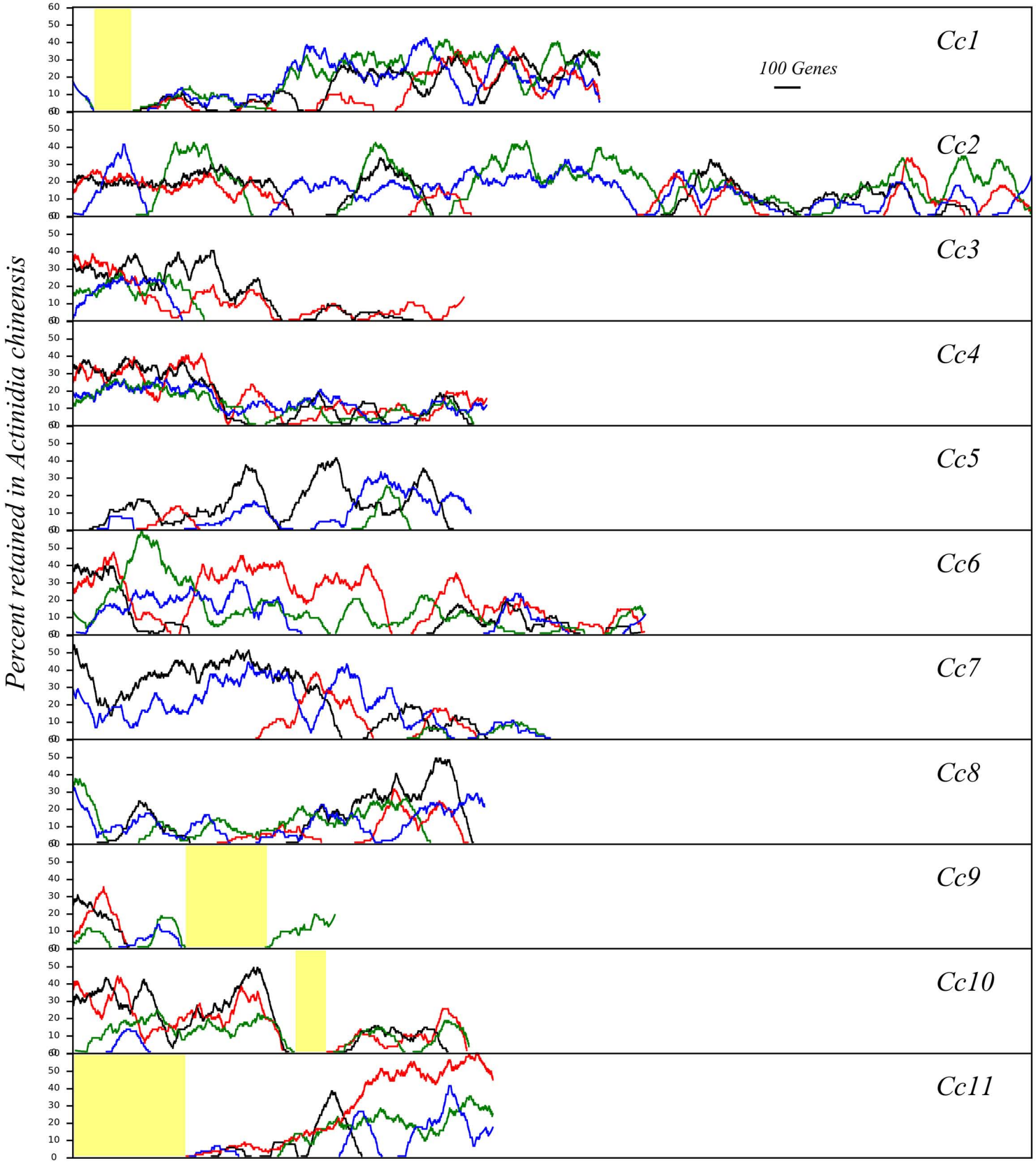




Supplemental Figure S7. Fitting a geometric distribution and gene loss rates in kiwifruit as compared with grape and coffee. **(A)** Kiwifruit with grape as a reference genome. **(B)** Kiwifruit with coffee as a reference genome.



Supplemental Figure S8. Kiwifruit gene retention along corresponding orthologous grape chromosomes. Rates of retained genes in sliding windows of grape homoeologous region group 1 (red), homoeologous region group 2 (black), homoeologous region group 3 (green), homoeologous region group 4 (blue), and large patches of chromosomal segmental losses (yellow) are displayed; grape chromosomes 1-19.



Supplemental Figure S9. Kiwifruit retention along corresponding orthologous coffee chromosomes. Rates of retained genes in sliding windows of grape homoelogous region group 1 (red), homoelogous region group 2 (black), homoelogous region group 3 (green), homoelogous region group 4 (blue), and large patches of chromosomal segmental losses (yellow) are displayed.