Article

Transcriptome Analyses throughout Chili Pepper Fruit Development Reveal Novel Insights into the Domestication Process

Octavio Martínez [1,*] , Magda L. Arce-Rodríguez [2] , Fernando Hernández-Godínez [1], Christian Escoto-Sandoval [1],
Felipe Cervantes-Hernández [1], Corina Hayano-Kanashiro [3] , José J. Ordaz-Ortiz [1] ,
M. Humberto Reyes-Valdés [4] , Fernando G. Razo-Mendivil [3], Ana Garcés-Claver [5]
and Neftalí Ochoa-Alejo [2,*]

1   Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (Cinvestav),
    Unidad de Genómica Avanzada (Langebio), Irapuato, Guanajuato 36824, Mexico;
    fernando.hernandezg@cinvestav.mx (F.H.-G.); christian.escoto@cinvestav.mx (C.E.-S.);
    felipe.cervantes@cinvestav.mx (F.C.-H.); jose.ordaz.ortiz@cinvestav.mx (J.J.O.-O.)
2   Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (Cinvestav),
    Departamento de Ingeniería Genética, Irapuato, Guanajuato 36824, Mexico; lisette02@hotmail.com
3   Departamento de Investigaciones Científicas y Tecnológicas de la Universidad de Sonora,
    Universidad de Sonora, Hermosillo, Sonora 83000, Mexico; angela.hayano@unison.mx (C.H.-K.);
    razoto@gmail.com (F.G.R.-M.)
4   Department of Plant Breeding, Universidad Autónoma Agraria Antonio Narro,
    Saltillo, Coahuila 25315, Mexico; mathgenome@gmail.com
5   Centro de Investigación y Tecnología Agroalimentaria de Aragón, Instituto Agroalimentario de Aragón-IA2
    (CITA-Universidad de Zaragoza), Unidad de Hortofruticultura, 50059 Zaragoza, Spain;
    agarces@cita-aragon.es
*   Correspondence: octavio.martinez@cinvestav.mx (O.M.); neftali.ochoa@cinvestav.mx (N.O.-A.)

Abstract: Chili pepper (*Capsicum* spp.) is an important crop, as well as a model for fruit development studies and domestication. Here, we performed a time-course experiment to estimate standardized gene expression profiles with respect to fruit development for six domesticated and four wild chili pepper ancestors. We sampled the transcriptomes every 10 days from flowering to fruit maturity, and found that the mean standardized expression profiles for domesticated and wild accessions significantly differed. The mean standardized expression was higher and peaked earlier for domes-ticated vs. wild genotypes, particularly for genes involved in the cell cycle that ultimately control fruit size. We postulate that these gene expression changes are driven by selection pressures during domestication and show a robust network of cell cycle genes with a time shift in expression, which explains some of the differences between domesticated and wild phenotypes.

Keywords: capsicum; fruit; gene expression; transcriptome; RNA-Seq; domestication

1. Introduction

Chili peppers of the genus *Capsicum* and the Solanaceae family are native to the American continent. Of the approximately 30 chili pepper species, five have been domesticated: *C. annuum* L., *C. frutescens* L., *C. baccatum* L., *C. chinense* Jacq. and *C. pubescens* Ruiz and Pav [1]. Among these species, *C. annuum* is the most important worldwide as a vegetable and spice crop, and production of this type of chili pepper has been steadily increasing in terms of both area harvested and yield [2]. In addition to their economic importance, chili peppers are a source of antioxidants, such as flavonoids, phenolic acids, carotenoids, and vitamins [3,4], as well as a plant model for studying the genetic and biochemical basis for the synthesis of these compounds [5–7]. Capsaicinoids, which are synthesized only in *Capsicum* species, impart pungency to chili peppers and are a focus of active research [8–10]. Furthermore, differential expression analysis during chili pepper fruit development al-

lowed the identification of regulatory and biosynthetic genes related to different metabolic processes during fruit development and ripening, including fruit softening, pigmentation, ripening time, and metabolite accumulation [11–13].

Chili peppers were domesticated from an ancestral variety, *Capsicum annuum* L. var. *glabriusculum*, locally known as "piquín" or "chiltepín" [14] in northeastern Mexico and/or central-east Mexico [15]. The oldest chili pepper macroremains date to around the time of their first cultivation or domestication in the mid-Holocene, 9000–7000 before present (BP) [15,16]. The exact time of chili pepper domestication is a subject of debate [17], as starch microfossils of domesticated *Capsicum* dating from 6000 BP have been found at seven sites [18], and there is evidence indicating that the fruit size of domesticated genotypes has increased considerably in the last 1500–1000 years BP [17]. The larger fruit size in domesticated compared with wild ancestors is part of the "domestication syndrome" [19].

Domestication, which involves breeding and selection of wild ancestral forms to modify phenotypes for human use, is not only a key achievement of modern civilization [20], but also provides a unique opportunity to identify the genetic basis of adaptation [21]. Examples of studies of plant domestication include maize [22–25], common bean [26,27], tomato [28–31], and *Capsicum* [32–35]. The *Capsicum* genome ($\approx$3.5 Gb) has been sequenced and annotated [36,37]; currently, there are nine genomic assemblies available from the NCBI (https://www.ncbi.nlm.nih.gov/genome/?term=Capsicum accessed on: 17 March 2021), and further sequencing of different genotypes has been reported [38,39]. In particular, the authors of [36] provided insights in order to evaluate the adaptive landscape of cultivated peppers [40] and reported a set of 511 genes that have a strong genomic domestication footprint. On the other hand, using allele-specific expression and network analyses from one domesticated and one wild chili pepper accession, as well as their F_1 cross, at 40 days after anthesis (DAA), the authors of [41] proposed that gene expression differences associated with the cultivated form could be explained by *cis*-regulatory hubs acting through *trans*-regulatory cascades.

To study the divergence caused by domestication in gene expression profiles during chili pepper fruit development, we examined fruit transcriptomes of six domesticated and four wild accessions using RNA sequencing (RNA-Seq) every 10 days from anthesis until fruiting at 60 days after anthesis (DAA). Our data show that there are significant differences in the mean expression profiles of domesticated and wild accessions that affect a set of interrelated biological processes, particularly the cell cycle. We postulate that such differences in expression profiles could partially explain the large differences in fruit sizes between domesticated and wild chili pepper varieties.

2. Results

We constructed and analyzed biological replicates of RNA-Seq libraries from 10 accessions (six domesticated (D) and four wild (W)), sampling every 10 DAA until maturity (see Materials and Methods and Supplementary S-1). Table 1 shows the accession, type, key, and time at which full maturity was reached in DAA.

A total of 22,427 genes, representing approximately 64% of the genes annotated in the *Capsicum* genome, were consistently expressed during fruit development in all of the accessions analyzed.

Gene expression varies by orders of magnitude [42], making it essential to perform data normalization [43], while differential gene expression allows one to contrast the relative expression of each gene under different conditions [44]. However, our aim was to evaluate changes throughout the time of fruit development, i.e., changes in gene expression profiles. With this purpose, we improved the method in [7] and defined the "standardized expression profile" (SEP), which is a seven-dimensional vector formed by the means of expression at each time point (0, 10, 20, 30, 40, 50, and 60 DAA), and it has a mean of 0 and standard deviation of 1 (see Materials and Methods and Supplementary S-2). The use of SEPs allows statistical comparisons between genes—or groups of genes—to be made independently of the relative gene expression of each gene.

Table 1. Chili pepper accessions used in this study.

Accession	Type	Key	Maturity (DAA)
Ancho San Luis	D	AS	80
Criollo de Morelos 334 (CM334)	D	CM	60
California Wonder	D	CW	70
Jalapeño Espinalteco	D	JE	70
Serrano Tampiqueño 74	D	ST	60
Zunla-1	D	ZU	60
Piquín Coahuila	W	CO	60
Piquín Queretaro	W	QU	60
Piquín Sonora Red	W	SR	60
Piquín Sonora Yellow	W	SY	60

For each gene within each accession, we estimated SEPs and analyzed differences between the D and W genotypes. Gene expression changes in times larger than 60 DAA in the late-maturing accessions (Ancho San Luis (AS), California Wonder (CW), and Jalapeño Espinalteco (JE)) were analyzed separately.

2.1. Analysis of Late Maturing Times

As seen in the "Maturity (DAA)" column in Table 1, three D accessions—AS, CW, and JE—take more than 60 DAA to reach the fully ripe fruit state; 70 DAA for CW and JE and 80 DAA for AS.

The precise time alignment of the last phenological states in all accessions is questionable; thus, we decided to perform a full analysis of seven time points that were uniformly sampled in all accessions $(0, 10, 20, \cdots, 60 \text{ DAA})$ with the SEP methodology and to independently evaluate gene expression changes in late stages in these three accessions. With this aim, we performed contrasts between the times 60 and 70 in CW and JE, as well as 60 and 70 plus 70 and 80 in AS. Only 633 (2.82%) genes in total were differentially expressed in one or more of the tests performed in these contrasts, with a false discovery rate (FDR) threshold of 0.01 [45].

Moreover, by performing tests for all SEPs between the normal (Criollo de Morelos 334 (CM334) (CM), Serrano Tampiqueño 74 (ST), and Zunla-1 (ZU)) and late-maturing (AS, CW, and JE) accessions, only 37 (0.16%) of these gene expression profiles were significantly different at 0.01 FDR. In addition, the analysis of the expression profiles for the gene coding for "capsanthin/capsorubin synthase" (protein XP_016577952.1), which is expressed only in chromoplasts in late maturing states in *Capsicum* [5,46,47], showed no significant differences ($p = 0.24$) between the normal and late-maturing accessions, implying that the analysis of SEPs from the 10 accessions using times between 0 and 60 DAA was reasonable, and would not induce a strong bias in the results. Details of these analyses are presented in Supplementary S-3.

2.2. Domesticated (D) and Wild (W) Accessions Have Different Mean SEPs

To study similarities in gene expression profiles between accessions, we calculated the mean Euclidean distances between the SEPs for all 22,427 genes expressed during fruit development to generate a dendrogram (Figure 1).

The D and W accessions form two distinct cluster groups, with a mean normalized distance of 2.85 on the Y-axis (Figure 1). The four W accessions (in blue) form a cluster at a mean distance of 2, whereas the six D accessions form a cluster at a mean normalized distance of approximately 2.4 (Figure 1 and Supplementary S-4).

To perform statistical analyses of genes or sets of genes, we considered contrasts between two groups of accessions: 6 D (AS, CW, JE, ST, and ZU in Table 1) and 4 W (Piquín Coahuila (CO), Piquín Queretaro (QU), Piquín Sonora Red (SR), and Piquín Sonora Yellow (SY) in Table 1). In all cases, the null hypothesis was that at each time point, the mean expressions of the D and W groups were equal, whereas the alternative was that these

parameters differed. A t-test was used to obtain confidence intervals (CIs) for the means and to evaluate the significance at each of the seven time points sampled. We determined the mean SEPs for different gene groups in the D and W accessions (Figure 2).
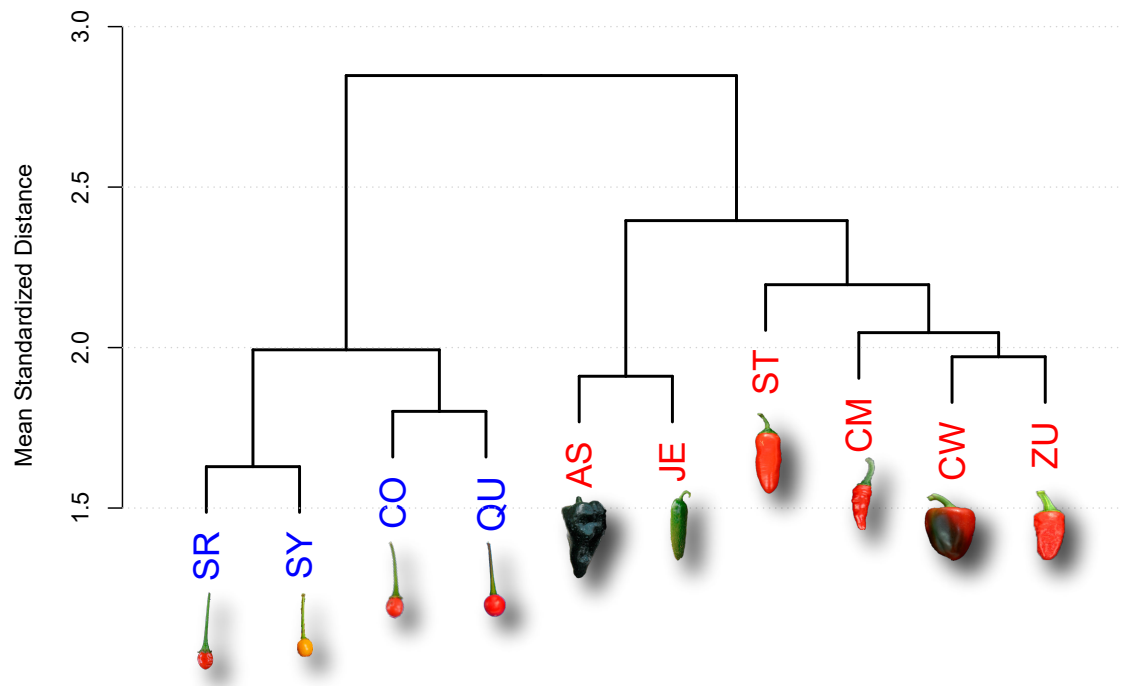

Figure 1. Dendrogram for the domesticated (D, red) and wild (W, blue) accessions. The dendrogram was obtained from the Euclidean distances between the full set of 22,427 standardized expression profiles (SEPs) of genes expressed during fruit development. Representative miniature photographs illustrate the fruits of the corresponding accessions. See Table 1 for accession names. Photographs of fruits are not at the same scale.

The means for the D and W groups differed significantly (continuous line in Figure 2). At the mature flower state (0 DAA), the standardized mean expression for D was much higher than for W, implying that the average transcription activity in this state is substantially larger for the D genotypes. In the interval between 0 and 10 DAA, the mean standardized expression increased for both groups, although the rate of increase was higher for D. At 10 DAA, the mean expression for D reached a peak value, but for W, the increase continued, although at a slower rate, to peak at 20 DAA. From the peak at 10 DAA, the mean expression for D decreased at different rates, and was lower at all subsequent time points. The lowest value was seen at 60 DAA. In contrast, decreases in the mean expression for W began later, occurring from 20 up to 50 DAA, and reached a minimum of -0.27, which is smaller than the minimum for the D group, -0.25, seen at 60 DAA. The more relevant differences in the mean expression profiles between D and W were seen during the intervals 10 to 20 and 50 to 60 DAA, when the trend was inverted such that D was decreasing while W was increasing. On the other hand, less marked differences between D and W were seen between 30 and 50 DAA, when the mean standardized expression decreased nearly in parallel for both groups. The average of the time at which the maximum expression was reached in each group was five days earlier for D than for W. All observed differences were significant (see Supplementary S-4).
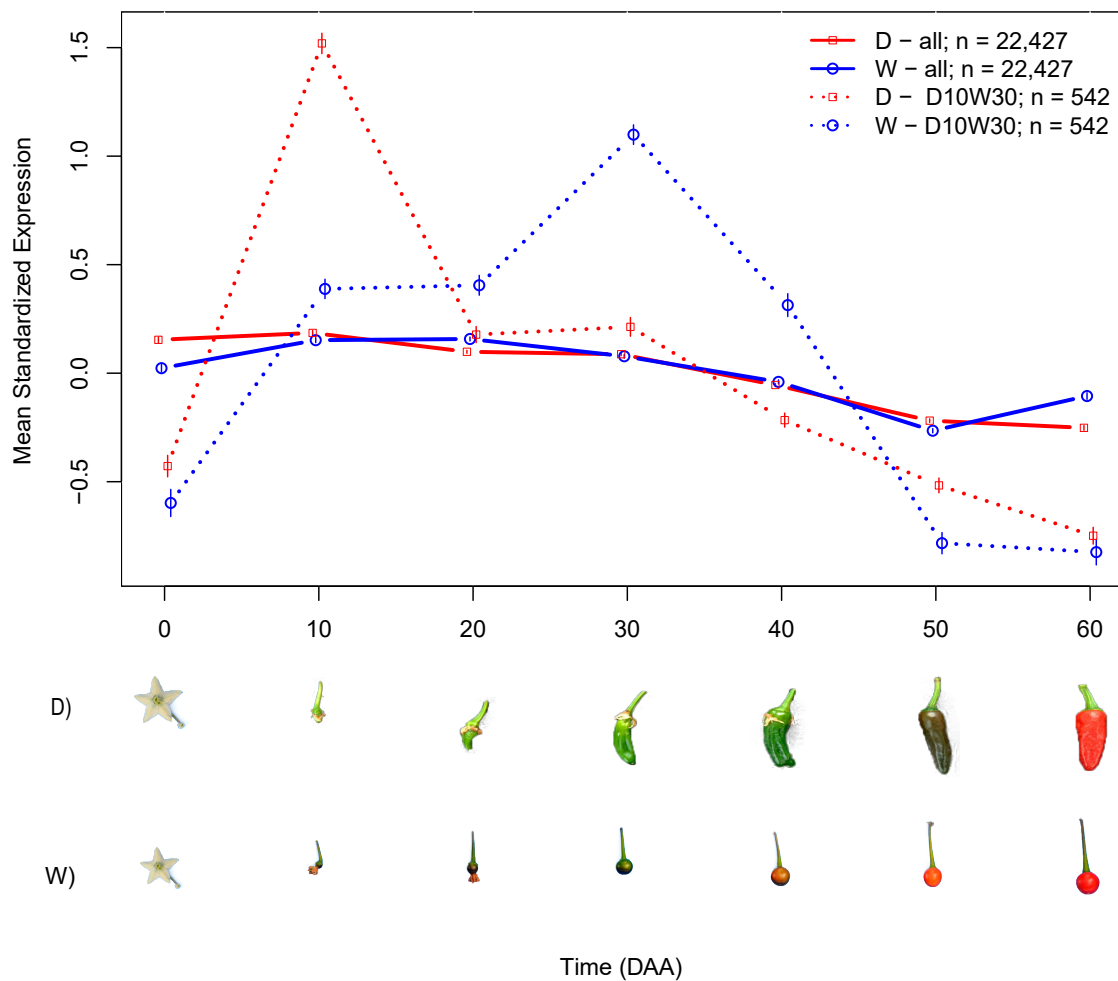
Figure 2. Average SEPs (Standardized Expression Profiles) for groups of genes in domesticated (D) and wild (W) accessions. Continuous colored lines (red and blue for domesticated (D) and wild (W), respectively) link the means of standardized gene expression at each time point for the complete set of expressed genes (n = 22,427). Dashed lines link the means of standardized gene expression at each time point for the n = 542 genes that presented the maximum expression at 10 DAA in D, while the maximum expression was reached at 30 DAA in W. These 542 genes form the group "D10W30" (see text). Representative miniature photographs illustrate the approximate fruit development of the D and W accessions at each time point. Photographs of fruits are not at the same scale.

2.3. Differences in SEPs of Individual Genes Varied between D and W

A total of 463 genes, representing approximately 2.06% of the total, showed significant differences between D and W, with an FDR threshold of 0.05, which, for the individual tests, corresponds to a p value < 0.000002. The differences in expression profiles between D and W were well defined and significant; the peak of the mean expression for D occurred at 10 DAA, while the peak for W occurred later at 30 DAA. The average times of maximum expression were 11.06 DAA for D and 28.33 DAA for W, or a difference of −17.27 DAA. Of the 463 selected genes, 36 ($36/463 \approx 0.08$ or 8%) are transcription factors (TFs). This percentage is higher than that for TFs annotated in the *Capsicum* genome ($1859/34,986 \approx 0.05$ or 5%). A list and description of the 463 selected genes and details of statistical analyses are presented in the Supplementary files SG and S-5, respectively.

We established that the main differences in SEPs between D and W were due to a set of 542 genes that presented an expression peak at 10 DAA in D, while the expression peak was at 30 DAA in the W accessions; we named these groups of genes "D10W30" (dashed line in Figure 2).

The results of this experiment showed differences in expression profiles between D and W at the levels of entire gene sets, groups of particular genes, and individual genes (see Supplementary S-5). Taking these findings together, we can thus conclude that there are relevant differences in expression profiles between domesticated and wild varieties of chili peppers during fruit development. We also found that gene expression diversity, expressed as the coefficient of variation of gene expression, is significantly ($p = 0.002$) smaller in the D than in the W accessions, corroborating the findings presented by [48] for different species of plants and animals.

2.4. Differences in Expression of Genes Related to Cell Reproduction Appear Earlier and Are Larger in Domesticated than in Wild Genotypes

Based on the evidence that mean SEPs differ between the D and W accessions, we investigated differences in the expression profiles in groups of genes related to particular biological processes. We first examined the mean SEPs of a group of 1125 genes associated with cell reproduction (Supplementary S-5.1). The mean expression value for 235 genes that are directly annotated in the cell cycle, but not in other cell reproduction processes, was significantly higher and occurred earlier for D compared to W, as evidenced by the peak of 0.3 standardized units at 10 DAA for D and 0.2 standardized units 30 DAA for W (Supplementary S-5.1). Similarly, the mean expression for 69 kinesins or kinesin-related proteins among the 1125 genes associated with cell reproduction exhibited a differential expression peak at 10 DAA for the D accessions, but for the W accessions, the peak was later at 30 DAA (Supplementary S-5.1). Thus, changes in the expression of genes associated with cell reproduction were significantly larger and occurred earlier for the D relative to the W accessions, not only for the full set of genes, but also for particular bioprocesses and gene families (Supplementary S-5–S-7).

2.5. Biological Processes Were Enriched in Genes That Are Expressed Earlier in Domesticated Genotypes

The results presented before indicate that the SEPs in the D and W accessions undoubtedly differ (Figure 2), and the genes for which expression peaks at 10 DAA for D but at 30 DAA for W (denoted here as "D10W30") play an important role in cell reproduction. To validate and expand our study, we considered the D10W30 expression pattern in a Gene Ontology enrichment analysis (for details, see Supplementary S-5.2, S-6, S-7, and SG).

A total of 86 biological processes (BPs) were significantly enriched (FDR = 0.05; $p < 0.0015$) in the D10W30 set, with a median odds ratio of 9.5. As such, these genes were much more abundant in these BPs than would be expected by chance. Apart from the aforementioned BPs related to cell reproduction, 43 of the enriched BPs, or 50% of the total, are involved in either positive or negative regulation of various biological processes. Of these, four (5%) are related to cellular component organization or biogenesis, three are associated with cellular component assembly, and another three play roles in organelle organization or fission. The general bioprocess, "cellular process" (GO:0009987), is also highly enriched in the D10W30 gene set, with an odds estimate of 2.25 and a highly significant p-value of 2.76×10^{-8}.

These results show that genes with the pattern D10W30 are over-represented in important BPs, which, in turn, implies that the expression of such BPs occurs earlier and at higher levels in the D compared to the W genotypes.

Interesting examples of D10W30 genes involved in cell reproduction are the high-mobility group B protein 6 (XP_016555757.1), the MYB-related protein 3R-1 (XP_016537977.1), and the kinetochore protein NDC80 (XP_016539151.1); see Supplementary S-5.2 for the SEP plots. The gene encoding the "high-mobility group B protein 6" is a *WRKY* transcription factor involved in the nucleosome/chromatin assembly that was annotated in 12 of the 86 aforementioned BPs—particularly the cell reproduction BP. The gene encoding the transcription factor "MYB-related protein 3R-1" was included in 6 of the 86 enriched BPs and is mainly related to cellular, chromosome, and organelle organization. The "kinetochore protein NDC80" is part of the multiprotein kinetochore complexes that couple eukaryotic

chromosomes to the mitotic spindle to ensure proper chromosome segregation. NDC80 is part of the outer kinetochore and forms a heterotetramer with the proteins NUF2, SPC25, and SPC24 [49,50]. Interestingly, the genes encoding NUF2 and SPC25 also exhibit the D10W30 expression pattern. NDC80 is conspicuously present in 74 of the 86 enriched BPs.

2.6. A Network of Cell Cycle Genes with the D10W30 Expression Pattern

The availability of genome-wide gene expression technologies, as RNA-Seq, make it possible to identify gene interactions and represent them as gene networks [51]. In functional genomics, it is axiomatic that genes with highly similar expression profiles are likely to be regulated via the same mechanisms, and this hypothesis is the basis for the discovery of regulatory networks [52]. To show the concerted co-variation through time of cell cycle genes' expression during fruit development, we estimated a gene network comprising six structural genes and eight TF candidates for regulating three of the six structural genes. Figure 3 presents the estimated gene network, while Table 2 gives the descriptions of the genes involved (Supplementary S-8–S-10).

(1) – MAP/ASE1 (1).
(2) – DUF566.
(3) – Kinesin 3, isoform X3.
(4) – MAP/ASE1 (2).
(5) – POLLENLESS 3.
(6) – Shugoshin.

(7) – AP2/B3–TF (1).
(8) – HMG B, 6.
(9) – Cyclin S13–7.
(10) – SET domain protein 35.
(11) – AP2/B3–TF (2).
(12) – FHA, PS1.
(13) – Myb 3R–1.
(14) – Tesmin / TSO1 CXC 3.
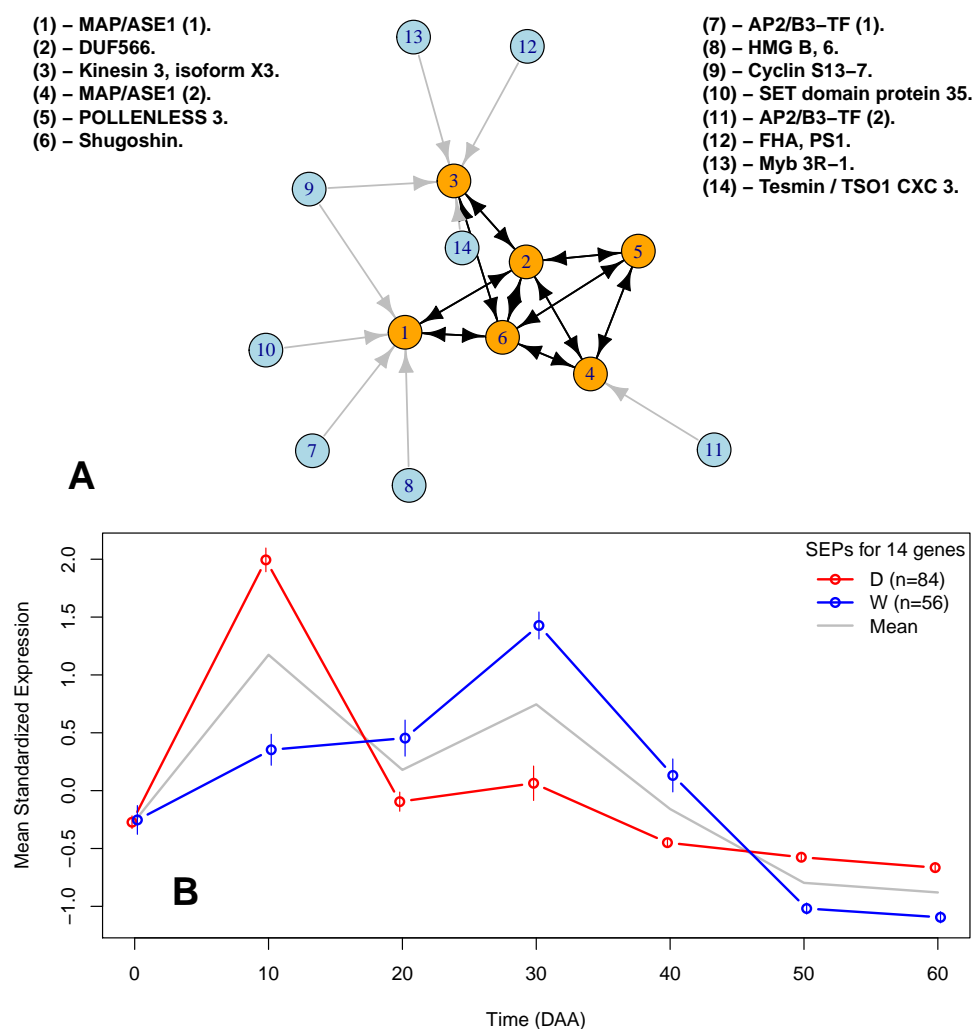
Figure 3. Gene network of the D10W30 cell cycle genes in the D and W accessions. (**A**) Graphic representation of the network. Orange circles represent structural genes, while blue ones represent transcription factor (TF) candidates for regulating the structural genes represented by orange circles (see Table 2 and Supplementary S-9). (**B**) Mean standardized expression of the genes in the network in the D and W accessions.

**Table 2.** Identifiers and descriptions for genes in the network of Figure 3A.

Figure 3A Legend	Protein ID	Description	Ortholog
(1)-MAP/ASE1 (1).	XP_016564755.1	65-kDa microtubule-associated protein 3	AT5G51600
(2)-DUF566.	XP_016538322.1	ENDOSPERM DEFECTIVE 1 (DUF566)	AT2G44190
(3)-Kinesin 3, isoform X3.	XP_016541615.1	Kinesin 3 isoform X3	AT4G21270
(4)-MAP/ASE 1 (2).	XP_016575449.1	65-kDa microtubule-associated protein 3 isoform X1	AT5G51600
(5)-POLENLESS 3.	XP_016577799.1	Protein POLLENLESS 3	AT4G20900
(6)-Shugoshin.	XP_016548908.1	Shugoshin-1; chromosome segregation	AT3G44960
(7)-AP2/B3-TF (1).	XP_016568750.1	AP2/B3-like TF family protein	AT5G42700
(8)-HGM B, 6.	XP_016555757.1	High-mobility group B protein 6	AT4G11080
(9)-Cyclin S13-7.	XP_016543946.1	G2/mitotic-specific cyclin S13-7	AT3G11520
(10)-SET domain protein 35.	XP_016547461.1	SET domain; methyltransferase activity.	AT1G26760
(11)-AP2/B3-TF (2).	XP_016575946.1	AP2/B3-like TF family protein	AT5G58280
(12)-FHA, PS1.	XP_016574880.1	FHA domain-containing protein PS1	AT1G34355
(13)-Myb 3R-1.6	XP_016537977.1	Myb-related protein 3R-1	AT4G32730
(14)-Tesmin / TSO1 CX3 3.	XP_016565918.1	Protein tesmin/TSO1 CXC 3	AT3G22780

Figure 3A presents the network formed by 14 D10W30 genes, in which the orange circles represent six structural genes involved in the cell cycle, and blue circles represent the TF gene candidates for regulating three of the structural genes (see the legend of Figure 3A and Table 2 for gene descriptions). Arrows between structural genes are drawn for genes with highly significant ($p < 0.0001$) positive Pearson's correlation coefficients ($r > 0.96$) between mean SEPs within the D and W accessions, while Pearson's correlation coefficients of the same gene pairs between the D and W accessions were small ($r < 0.4$) and not significant ($p > 0.5$). The origin and robustness of the network presented in Figure 3A can be appreciated by observing Figure 3B, which presents the mean SEPs for the 14 genes involved, which represent $n = 84$ independent SEPs estimated from the six D accessions (red line; $84/6 = 14$ genes) and $n = 56$ independent SEPs estimated from the four W accessions (blue line; $56/4 = 14$ genes). In Figure 3B, the 95% CI for the mean standardized expression of the genes—vertical lines at each time point—shows that these 14 genes are highly correlated within the D and W accessions, but have a very low correlation between the D and W groups. For details, see Supplementary S-8 and S-9.

For a better appreciation of the changes that occur in the individual standardized expression of the genes included in the network of Figure 3, Figure 4 presents a grid of the time and the accession groups, but in this case, the sizes of the circles representing genes vary in proportion to their individual standardized expression over time (rows; from 0 to 60 DAA) and sets of accessions (columns; D on the left-hand side and W on the right -hand side). Representative miniature photographs show the approximate fruit development stage for the D and W accessions.

In Figure 4, we can appreciate that individual gene expression is highly coordinated through time, but largely differs between the D and W groups of accessions. At 10 DAA in the D accessions (second row, left-hand side), all 14 genes have reached their maximum expression (largest circle sizes), while such expression levels are only reached 20 days later—at 30 DAA—for the W accessions (fourth row, right-hand side). Thus, the main cell cycle genes are well behind in maximum expression time in the W compared with the D accessions. In addition, expression changes occur faster in the D compared with the W accessions, a fact that can be verified by observing Figure 3B, where the slope of the line linking 0 and 10 DAA is more pronounced than the corresponding change to reach the maximum from 20 to 30 DAA in the W accessions. Departing from comparable standardized expression at the mature flower stage (0 DAA), individual expression rapidly diverges between the D and W groups to finally converge again to basal expression levels at 50 and 60 DAA. The concerted but highly divergent expression of cell cycle genes demonstrates that domestication has tailored this process to differ between the D and W genotypes, explaining, in part, the large differences in fruit size between these groups.
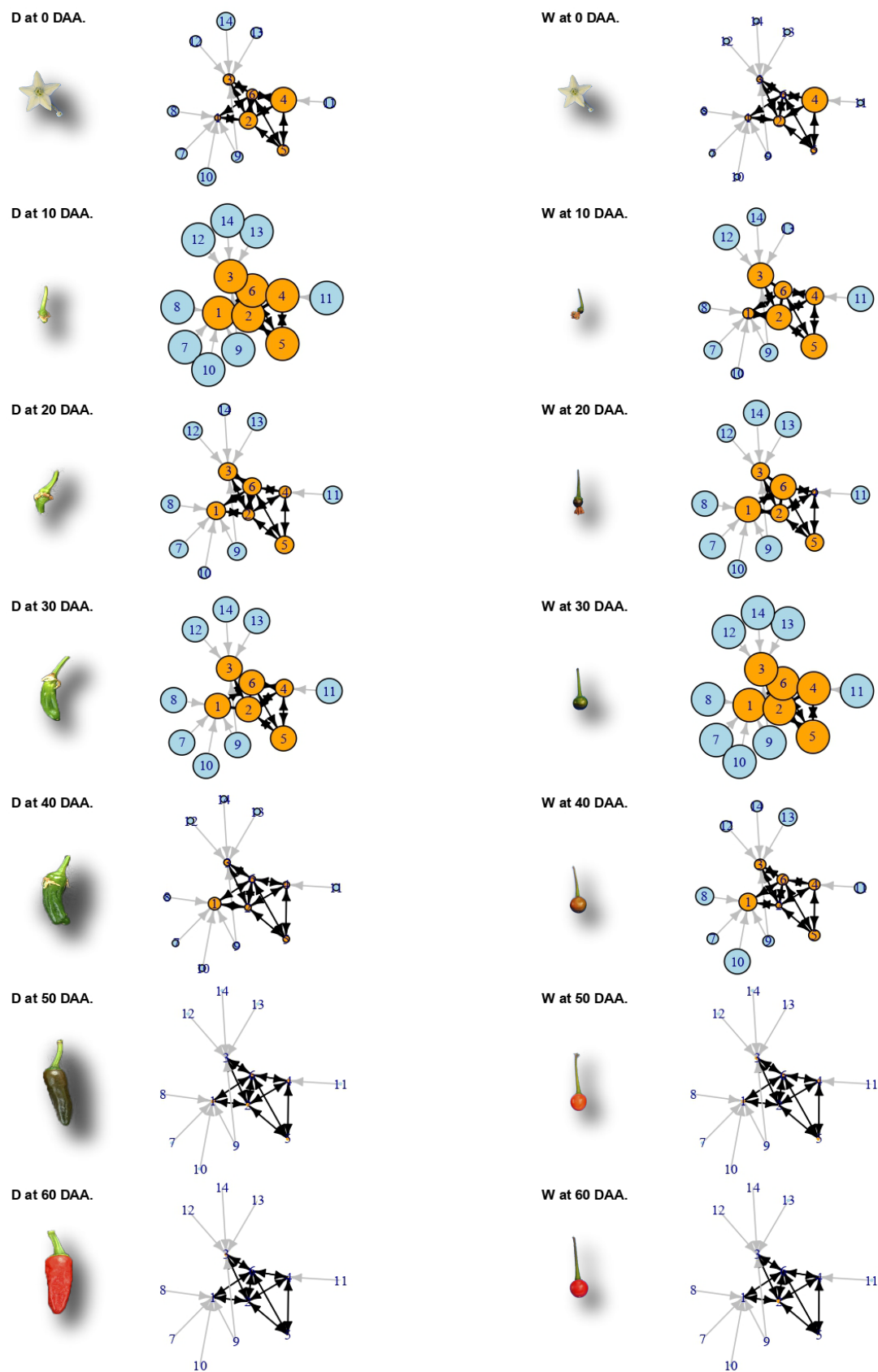
Figure 4. Change over time (rows) and groups of accessions (columns) of the network in Figure 3A. Gene expression is proportional to circle size. Photographs of fruits are not at the same scale.

3. Discussion

Building on our previously described method for estimating the dynamics of the chili pepper transcriptome [7], as well as our time-course experiment [53] and statistical analysis methods, in this study, we generated entire gene expression profiles (SEPs) across the full course of fruit development to examine differential gene expression patterns between domesticated (D) and wild (W) varieties of chili peppers (Figure 1).

Domestication, as initially studied by Darwin [54], is currently defined as a distinctive coevolutionary and mutualistic relationship between domesticator and domesticate, and marked a key transition in human history [20]. Here, we propose that the differences observed between gene expression profiles in sets of domesticated (D) and wild (W) accessions (Figure 2) can be attributed to domestication.

Gene expression is an intrinsically noisy process [55]; however, our experimental design systematically took into account variations between sets of fruits by examining two replicates of each accession at each time point, as well as variation within the target groups (D and W), by examining six and four genotypes for each group, respectively. Thus, the differences observed between the D and W gene expression profiles can be attributed to differences in the selection histories of the two groups, i.e., to domestication. The effect of domestication on gene expression patterns is also corroborated by the fact that the D and W accessions belong to well-segregated clusters in the dendrogram presented in Figure 1. Although the distance between the D and W groups is approximately 2.85, the maximum distance within the groups is only 2.4, again demonstrating that gene expression patterns during fruit development were modified by domestication.

Selection during domestication can alter molecular footprints at the genomic level. For example, in *Capsicum*, Qin et al. [36] identified 115 genomic regions containing 511 genes that show strong selective sweep signals due to domestication. The same method of searching for selective sweep signals identified candidate genes not only in plants, such as maize [23], sunflower [56], soybean [57], and Asian rice [58], but also in domesticated animals, such as dogs [59] and cattle [60]. Ross-Ibarra et al. [21] reviewed methods of identifying the genes responsible for adaptation to domestication and classified them in terms of the phenotype–genotype hierarchy as "top-down", in which the method starts with a phenotype to identify candidate genes, and as "bottom-up", in which genetic analyses are used to identify adaptive genes. Bioinformatics tools are then used to connect the selected genes to a phenotype. Our approach in this study can be considered as a hybrid between "top-down" and "bottom-up" methods in that we began with a molecular phenotype, i.e., a standardized gene expression profile (SEP), which showed that there are significant differences between the D and W groups of accessions, and then examined the biological relevance of these findings.

Expression levels of genes that are important during chili pepper fruit development have been reported in [11–13]. Differences in the accessions employed, times of development, genes analyzed, and methodologies preclude full and equitable comparisons between the results presented in those references with the ones presented here. Examination of 74 SEP plots for genes studied in the mentioned works resulted in coincidences of 62% with our results. This percentage can be considered to be high, given the dissimilarities of the above-mentioned factors (plots and results available upon request). On the other hand, Díaz-Valenzuela et al. [41] reported results from the analyses of one cultivated accession, one wild accession, and their F_1 only at 40 DAA. Given that this study does not include gene expression measurements at different times of fruit development, it is impossible to examine their findings in the framework of our work.

Differences in gene expression between crops and their wild ancestors were reported in tomato [29,61], which, like *Capsicum*, belongs to the Solanaceae family, as well as common bean [26,27], carrot [62], ramie [63], and cotton [64]. For animals, similar comparisons were done for turkey [65], trout [66], and other species [67]. All of these studies reported sets of differentially expressed genes between the domesticated and wild forms, but only for tomato did the study authors use a time-course experiment to evaluate the deceleration

of the circadian clock [29]. On the other hand, a decrease in gene expression diversity in domesticated forms of a set of animals and plants was reported by Liu et al. [48]. Analysis of our data confirmed that gene expression is significantly less diverse in domesticated chili peppers compared to that seen for wild accessions.

Normalization of the gene expression profiles of genes or sets of genes—discussed here as mean SEPs (see Results) —implies that the mean over time of the profile is equal to zero, and the standard deviation is equal to one. This transformation allows direct and unbiased comparisons of expression profiles in the D and W groups of accessions, with a statistical evaluation performed at each time point (Figure 2 and Supplementary S-4 and S-5).

It is important to note that the differences in mean SEPs between D and W were produced by a set of genes that were likely affected by the domestication process (Figure 2). However, the large majority ($21,666/22,427$; 96.6%) of expressed genes did not exhibit significant differences in the mean expression profile (Supplementary S-5), implying that a large part of the transcriptome during fruit development was not affected by domestication.

In chili pepper, as in tomato, the first ten days after anthesis (DAA) are characterized by a period of very active cell division, in which the number of cell layers across the pericarp doubles compared with that at anthesis, and this number is maintained through the end of development [68]. Fruit growth then proceeds into the cell expansion phase and continues until ≈40 DAA, progressing to full ripening at ≈50 DAA before entering senescence at ≈60 DAA [7]. During the period of cell expansion—from approximately 10 to 40 DAA—the dominant process is cell endoreduplication [68–70]. In *Capsicum*, pericarp thickness has a high positive correlation with the degree of polysomaty [71], which implies that accessions that have a thick pericarp have higher endoreduplication compared to accessions with a thin pericarp. This relationship has been corroborated through the induction of different ploidy levels [72].

The time lag observed for the peak of mean standardized expression between the D and W accessions across the entire set of expressed genes (Figure 2) implies that domestication caused a shift in the time and intensity of gene expression, favoring an earlier and higher expression maximum in D. Considering 463 genes that had the largest differences in expression profiles between D and W confirmed that the peak of expression for D occurs at 10 DAA compared to 20 DAA for W, and the peak expression value is significantly higher for D than for W. A similar expression pattern is seen for a gene encoding a G2/mitotic-specific cyclin, which is essential for control of the cell cycle at the G2/mitosis transition (Supplementary S-5). The transcript of this gene accumulated steadily during G2 and abruptly declined at mitosis. In *Arabidopsis*, which exhibits a similar trend in expression, members of the cyclin family are thought to be part of a developmental mechanism that coordinates the switch between proliferation and endoreduplication [73].

Here, we found the peak gene expression for D at 10 DAA, a time when cell division is very active [68]. Furthermore, domesticated accessions bear substantially larger fruits than wild accessions [33], and this larger fruit size is primarily achieved through increases in cell numbers [74]. As such, we grouped a set of 1125 genes associated with cell reproduction by including genes annotated in nine bioprocesses and examined the gene expression profiles (Supplementary S-5). Genes annotated for the cell cycle presented an earlier increase and higher mean expression in D relative to W (Supplementary S-5). Given that both cell number and cell size, which are respectively determined by cell division and cell expansion [75], contribute to fruit size, it is compelling that, in the D accessions, the cell division expression profile peaked earlier and higher than in the W genotypes. This finding is consistent with that seen for tomato, in which cell division genes strongly influence fruit yield [76].

The mean expression of the genes related to the cell cycle presented a profile characterized by a large peak expression at 10 DAA for D, whereas the W accessions had a smaller peak that occurred later at 30 DAA. We isolated a set of 542 genes that presented this pattern, which we termed "D10W30". Gene Ontology enrichment analyses produced a

set of 86 biological processes (BPs) that were highly enriched among the D10W30 genes (Supplementary S-6 and SG). In addition to four cell reproduction BPs, this set included 43 BPs associated with regulation of different cell processes, including negative regulation of cellular process, as well as protein modification and transferase activity. Interestingly, such negative regulation has been associated with fruit development and ripening [77]. Other selected BPs include four that are related to cellular component organization, which has been linked to the accumulation of soluble sugars and organic acids in fruits [78], and three that are related to cellular component assembly and also showed differential expression in a proteomic study of *Capsicum* [79]. Another three in this set were identified with organelle fission or organization, and two were related to microtubule-based processes or movement, and are clearly associated with mitosis; they have been linked to floral development in the genus *Aquilegia* [80] and to the autophagy BP, which allows remodeling of intracellular structures during cell differentiation [81]. In *Arabidopsis*, this BP has been linked with the complete proteolysis of stromal proteins [82]; see Supplementary S-6, S-7, and SG for details.

We plotted the expression profiles of genes that follow the D10W30 pattern, i.e., expression peaks at 10 and 30 DAA for the D and W accessions, respectively (Figure 2). Interesting examples of this expression pattern are genes encoding: (i) the high-mobility group B protein 6 (HMG B, 6), which belongs to a group of chromosomal proteins that regulate DNA-dependent processes and display a highly dynamic nuclear localization [83]; (ii) the transcription factor "MYB-related protein 3R-1", which, in *Arabidopsis*, synergistically maintains G2/M-specific genes repressed in post-mitotic cells and restricts the time window of mitotic gene expression in proliferating cells, thus giving it a role in determining organ size [84]; and (iii) the kinetochore protein NDC80, which is an essential component of the kinetochore complex that mediates chromosome segregation and spindle checkpoint activity to ensure proper cell division. In *Arabidopsis*, the NDC80 mutant *mun-1* has a reduced cell division rate, aneuploidy, and defects in chromosome segregation [85]; see Figure 3 and Supplementary S-5.

The D10W30 set includes genes coding for "microtubule-associated protein TORTI-FOLIA1" and the "protein TPX2", which have genomic fingerprints for domestication in *Capsicum* [36]. TORTIFOLIA1 (TOR1) is a plant-specific microtubule-associated protein (MAP) that regulates cortical microtubule orientation and the direction of organ growth [86]. TOR1 also determines microtubule organization by modulating microtubule severing [87] and participates in organ elongation [88]. On the other hand, TPX2 performs multiple roles in microtubule organization [89], such as regulating prospindle assembly before nuclear envelope breakdown [90], and is linked to fruit development in European pear [91]. Of the 300 domestication genes reported by Qin et al. [36] and expressed during fruit development, 59 ($\approx$20%) also showed significant ($p < 0.05$) differences between the D and W accessions in this study.

An example of the coordinated time lag existent between the D and W accessions is presented in Figures 3 and 4. The 14 genes involved in this network are pivotal for the cell cycle process and are thus functionally related, but also present highly coordinated gene expression profiles within the D and W accessions, which markedly differ between these two groups by having the D10W30 expression pattern (Supplementary S-8, S-9, and S-10). It is important to assess the robustness of the links inferred between the genes in this network (Figure 3). With this aim, we must take into account the fact that such links were inferred from ten fully independent datasets, i.e., the ones corresponding to each one of the 10 accessions. Then, if we assume that the selection of a gene to be part of the network has an error probability "ε", then the probability of erroneously selecting a gene to be part of the network repeatedly in the ten accessions is ε^{-10}. Thus, even if the individual error probability is large—for example, $\varepsilon = 0.1$—the probability of having committed such errors repeatedly in the ten accessions is vanishingly small; for $\varepsilon = 0.1$, this equals 10^{-10}, or one in ten billion.

Of the genes involved in the network of Figure 3A, two of them—(1) and (4) in the figure legend—encode two versions of the microtubule-associated protein 3. The *Arabidopsis* orthologous of these genes, PLEIADE/AtMAP65-3, has been shown to have physical and genetic interactions with the Transport Protein Particle II (TRAPPII), which is required for coordinating cytokinesis with the nuclear division cycle [92]. Gene (2) in Figure 3A, labeled as *DUF566* in the legend, contains the InterPro domain IPR007573 and corresponds with the *Arabidopsis* orthologous *AT2G44190* (Table 2), which codes for the endosperm-defective 1 (*ede1*) gene; this encodes a microtubule-associated protein that is essential for microtubule function during the mitotic and cytokinetic stages, which generate the endosperm and embryo, and is thus essential for seed formation [93]. Gene (3) in Figure 3 encodes a kinesin 3, which is an *Arabidopsis* ortholog, *AT4G21270*, and encodes the *ATK1* gene, which has been demonstrated to be required for spindle morphogenesis [94]. Consistently, the large majority of chili pepper kinesins follow the D10W30 expression pattern (Supplementary S-5.1). The gene labeled as "(5)—POLLENLESS 3" in the legend of Figure 3A contains a tetratricopeptide repeat (TPR), and its closest *Arabidopsis* ortholog, *AT4G20900*, encodes the TDM1 gene, which has been previously shown to be essential for meiotic termination [95]. Structural gene (6) in Figure 3A encodes shugoshin, a conserved kinetochore protein that prevents dissociation of cohesin from centromeres during mitosis [96]. The closest *Arabidopsis* ortholog locus of this chili pepper sequence, *AT3G44960*, encodes five splicing variants of shugoshin, one of which has been reported to protect centromeric cohesion during meiosis [97]; see also Supplementary S-10. It is intriguing that three of the chili pepper genes, the ones labeled as numbers (3), (5), and (6) in the legend of Figure 3A, have as their closest orthologs in Arabidopsis genes that are reported primarily in meiosis [94,95,97], respectively, while the expression patterns of these *Capsicum* genes clearly correspond to genes involved in mitosis. This fact could be due to the large functional divergence between the *Capsicum* and *Arabidopsis* genomes, which diverged from a common ancestor more than 150 million years before present [36].

To select TF candidates for regulating the expression of the structural genes, represented by orange circles in the network of Figure 3, we employed an algorithm for TF imputation, which is detailed in Supplementary S-9. This algorithm relies on stringent thresholds for correlations in expression estimated for each accession, and has a very low probability of false-negative imputations, estimated in this case as 1.06×10^{-15}. The algorithm was successful in recovering a TF that was experimentally confirmed to be regulating the expression of the *AT3* gene in *Capsicum* [8,98,99] (see Methods and Supplementary S-9).

In summary, the functional network presented in Figures 3 and 4 demonstrates that domestication has produced a time lag in the expression of core cell cycle genes, thus anticipating by approximately 20 days the maximum expression of those genes and producing a higher standardized expression at the early fruit development stage (10 DAA) in D when compared with the W accessions.

Comparing gene expression profiles rather than focusing on the differential expression of single genes at a given time gives a better perspective on the complex interplay occurring in the transcriptome over time. Here, we demonstrated that a set of genes exhibits significant differences in expression profiles between the D and W accessions during the development of chili pepper fruit. Genes in this set are associated with processes that involve the cell regulation, cycle, localization, motility, and assembly, as well as with autophagy and organelle organization. In particular, differences in the time and intensity of the expression of genes are related to cell reproduction and provide an explanation at a molecular level of differences in fruit size between the D and W accessions, which is the main morphological difference between these two genotype groups [100].

4. Materials and Methods

4.1. Statistical Design

RNA-Seq was performed as a factorial experiment with time (seven levels: 0, 10, 20, 30, 40, 50, and 60 DAA) and accession (10 accessions: 6 domesticated and 4 wild;

see Table 1) as factors. Additional RNA-Seq libraries were constructed and analyzed for accessions with later maturation times: at 70 DAA for accessions CW and JE and at 70 and 80 DAA for AS. The RNA-Seq library was the experimental unit, and two libraries of every combination of time per accession were independently replicated two times; thus, we analyzed $7 \times 10 \times 2 = 140$ RNA-Seq libraries for time points between 0 and 60 DAA in all 10 accessions—plus four libraries at 70 DAA in accessions CW and JE and four libraries corresponding to 70 and 80 DAA in AS—for a total of $140 + 4 + 4 = 148$ RNA-Seq libraries. After quality control, the raw reads were mapped to the *Capsicum* genome (CM334 v1.6) to obtain reliable estimations of gene expression. The relative expression for each gene at each accession and time combination was considered here as the output variable.

4.2. Plant Materials and Cultivation

Seeds of each one of the D *Capsicum annuum* L. accessions (Table 1) were surface sterilized with a 70% ethanol solution for 10 s before treatment with a 10% hypochlorite solution for 10 s and six rinses with distilled water. Wild accession seeds were similarly treated after an initial treatment with 50% sulfuric acid solution to break seed dormancy. All accession seeds were germinated in plastic trays containing a mixture of three parts peat moss, one part perlite, one part vermiculite, one part sludge, and two parts forest soil in a growth chamber, with 16 h light (photon flux of 70 μmol m^{-2} s^{-1}) at 28 °C and 66% relative humidity. Three-week-old chili pepper plants were transplanted individually into 5 L plastic pots containing the same soil mixture as that described above. During transplantation, 15 g of a mixture of mycorrhizal fungi and beneficial bacteria was added to optimize root growth and development. Plants were fertilized with Long Ashton solution every two weeks. Flowers and fruits at each one of the times of development (0, 10, 20, 30, 40, 50, 60, 70, or 80 DAA, depending on the final maturation time; see Table 1) were collected, immediately frozen in liquid nitrogen, and stored at −80 °C.

4.3. RNA-Seq Library Construction and Processing

Total RNA was extracted from flowers and whole chili pepper fruits at different developmental stages using a NucleoSpin^{TM} RNA Plant kit (MACHEREY-NAGEL) according to the manufacturer's instructions. RNA was extracted from two biological samples comprising either flowers or fruits from 2–6 different plants. RNA quality was verified by determining the RNA integrity number (RIN) for each sample (Supplementary S-1). Samples of total RNA were shipped to Novogene (https://en.novogene.com/ accessed on: 17 March 2021) for library construction, sequencing, and mapping to a reference genome. At Novogene, libraries were prepared and sequenced using the Illumina NovaSeq platform to obtain at least 20 million raw paired-end reads of 150 bp per sample. These reads were subjected to quality control and then mapped to the *Capsicum* reference genome CM334 v1.6 (http://peppergenome.snu.ac.kr/ accessed on: 17 March 2021). Novogene provided the matrix of raw counts per library for each of the 35,883 *Capsicum* genes. These genes were identified by a protein product (when known) and annotated with Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) terms. For this study, a total 140 libraries were processed, yielding more than 2.29 billions of clean reads mapped to the genome. The number of reads mapped to the genome had a minimum of 10.33, a median of 16.26, a mean of 16.42, and a maximum of 23.86 million reads per library; see Supplementary S-1 and SG.

4.4. Estimations and Analyses of Standardized Expression Profiles (SEPs)

To avoid inclusion of genes that had very low or inconsistent expression patterns, only genes that had a raw count of > 0 in at least two of the replicates per accession were selected for analysis. After this filtering, 22,427 genes remained for analyses, and of these, $\approx$62.5% were identified in the *Capsicum* genome; see Supplementary SG.

All results were maintained in an in-site MySQL (https://www.mysql.com/ accessed on: 17 March 2021) relational database, and were analyzed with R version 3.4.4 [101]. For

each of the 10 accessions, we used the R package "edgeR" version 3.20.9 [44] to obtain *p* values for each gene in each accession in all contrasts between neighbor time intervals, i.e., 0–10, 10–20, 20–30, $\cdots$, 50–60 DAA. Following the method presented in [7], the gene expression tendency at each interval was classified as decreasing, steady, or increasing, with adjustment of the Type I error to 0.01. For each gene within each accession, the mean standardized expression was calculated, and the resulting seven-dimensional vector constituted the standardized expression profile (SEP) in downstream analyses. There were 10 SEPs for each of the 22,427 genes that were consistently expressed during fruit development, with one for each of the accessions, and these SEPs were classified into the groups of interest: domesticated (D) with six elements and wild (W) with four elements. To evaluate the difference between the D and W SEPs, we calculated the Euclidean distances between and within the groups and tested the hypothesis of equality of distances between and within SEPs with a t-test. Using an FDR of 1% (i.e., a *q* value $\leq$ 0.01), genes were classified as having an equal or distinct SEP in the D and W groups. For individual genes or groups of genes, we calculated the 95% CI for the means at each of the seven time points sampled, as well as the *P* value for the t-test of mean equality. A set of R functions was programmed to data-mine the results. The statistical procedures are described in detail in Supplementary S-2 to S-9, and a manuscript presenting the R package for data-mining the results is under evaluation [102].

4.5. Network Estimation

We began the network estimation with the 25 genes with expression profile D10W30, which was annotated with the cell cycle biological process in the 10 accessions. For all $25 \times (25-1)/2 = 300$ different pairs formed by these 25 genes, we calculated the Euclidean distance between their mean SEPs, selecting the pairs with a distance of less than 1 standardized units. This stringent criterion selected only 10 gene pairs (3.3% of the total, and shown as edges between orange circles in Figure 3A), which included the six structural genes (orange circles in Figure 3A). The SEPs of the 10 gene pairs linked with black double-headed arrows have a large ($r > 0.96$) and highly significant ($p < 0.0001$) mean Pearson's correlation within the D or W genotypes. In contrast, the corresponding mean correlation between the D and W accessions was small ($r < 0.4$) and not significant ($p > 0.5$), demonstrating that while these genes have a highly concordant expression within the D and W groups, the expression profiles are very different between those groups (Figure 3B and Supplementary S-8). Finally, the selection of the 8 TF candidates for regulating the expression of the structural genes in the network of Figure 3A was performed by employing an algorithm that selected TFs with a highly concordant expression profile between TFs and each target gene (see Supplementary S-9 for details).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Abbreviations

The following abbreviations are used in this manuscript:

D	Domesticated accession (Table 1)
W	Wild accession (Table 1)
DAA	Days after anthesis
SEP	Standardized expression profile
FDR	False discovery rate

References

1. Pickersgill, B. Relationships between weedy and cultivated forms in some species of chili peppers (genus *Capsicum*). *Evolution* **1971**, *25*, 683–691. [PubMed]
2. Jarret, R.L.; Barboza, G.E.; da Costa Batista, F.R.; Berke, T.; Chou, Y.Y.; Hulse-Kemp, A.; Ochoa-Alejo, N.; Tripodi, P.; Veres, A.; Garcia, C.C.; et al. Capsicum—An abbreviated compendium. *J. Am. Soc. Hortic. Sci.* **2019**, *144*, 3–22. [CrossRef]
3. Badia, A.D.; Spina, A.A.; Vassalotti, G. *Capsicum annuum* L.: An Overview of Biological Activities and Potential Nutraceutical Properties in Humans and Animals. *J. Nutr. Ecol. Food Res.* **2017**, *4*, 167–177. [CrossRef]
4. Cervantes-Hernández, F.; Alcalá-González, P.; Martínez, O.; Ordaz-Ortiz, J.J. Placenta, Pericarp, and Seeds of Tabasco Chili Pepper Fruits Show a Contrasting Diversity of Bioactive Metabolites. *Metabolites* **2019**, *9*, 206. [CrossRef]
5. Gómez-García, M.d.R.; Ochoa-Alejo, N. Biochemistry and molecular biology of carotenoid biosynthesis in chili peppers (*Capsicum* spp.). *Int. J. Mol. Sci.* **2013**, *14*, 19025–19053. [CrossRef]
6. Gómez-García, M.d.R.; Ochoa-Alejo, N. Predominant role of the l-galactose pathway in l-ascorbic acid biosynthesis in fruits and leaves of the Capsicum annuum L. chili pepper. *Braz. J. Bot.* **2016**, *39*, 157–168. [CrossRef]
7. Martínez-López, L.A.; Ochoa-Alejo, N.; Martínez, O. Dynamics of the chili pepper transcriptome during fruit development. *BMC Genom.* **2014**, *15*, 143. [CrossRef]
8. Arce-Rodríguez, M.L.; Ochoa-Alejo, N. An R2R3-MYB transcription factor regulates capsaicinoid biosynthesis. *Plant Physiol.* **2017**, *174*, 1359–1370. [CrossRef]
9. Tanaka, Y.; Nakashima, F.; Kirii, E.; Goto, T.; Yoshida, Y.; Yasuba, K.I. Difference in capsaicinoid biosynthesis gene expression in the pericarp reveals elevation of capsaicinoid contents in chili peppers (*Capsicum chinense*). *Plant Cell Rep.* **2017**, *36*, 267–279. [CrossRef]
10. Fayos, O.; Ochoa-Alejo, N.; Octavio, M.; Savirón, M.; Orduna, J.; Mallor, C.; Barbero, G.F.; Garces-Claver, A. Assessment of Capsaicinoid and Capsinoid Accumulation Patterns during Fruit Development in Three Chili Pepper Genotypes (*Capsicum* spp.) Carrying Pun1 and pAMT Alleles Related to Pungency. *J. Agric. Food Chem.* **2019**, *67*, 12219–12227. [CrossRef]
11. Osorio, S.; Alba, R.; Nikoloski, Z.; Kochevenko, A.; Fernie, A.R.; Giovannoni, J.J. Integrative comparative analyses of transcript and metabolite profiles from pepper and tomato ripening and development stages uncovers species-specific patterns of network regulatory behavior. *Plant Physiol.* **2012**, *159*, 1713–1729. [CrossRef] [PubMed]
12. Jang, Y.K.; Jung, E.S.; Lee, H.A.; Choi, D.; Lee, C.H. Metabolomic Characterization of Hot Pepper (Capsicum annuum 'CM334') during Fruit Development. *J. Agric. Food Chem.* **2015**, *63*, 9452–9460. [CrossRef] [PubMed]

13. Dubey, M.; Jaiswal, V.; Rawoof, A.; Kumar, A.; Nitin, M.; Chhapekar, S.S.; Kumar, N.; Ahmad, I.; Islam, K.; Brahma, V.; et al. Identification of genes involved in fruit development/ripening in *Capsicum* and development of functional markers. *Genomics* **2019**, *111*, 1913–1922. [CrossRef] [PubMed]

14. Hayano-Kanashiro, C.; Gámez-Meza, N.; Medina-Juárez, L.Á. Wild Pepper *Capsicum annuum* L. var. *glabriusculum*: Taxonomy, Plant Morphology, Distribution, Genetic Diversity, Genome Sequencing, and Phytochemical Compounds. *Crop Sci.* **2016**, *56*, 1–11.

15. Kraft, K.H.; Brown, C.H.; Nabhan, G.P.; Luedeling, E.; Ruiz, J.D.J.L.; d'Eeckenbrugge, G.C.; Hijmans, R.J.; Gepts, P. Multiple lines of evidence for the origin of domesticated chili pepper, *Capsicum annuum*, in Mexico. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 6165–6170. [CrossRef] [PubMed]

16. McClung de Tapia, E. The origins of agriculture in Mesoamerica and Central America. *Orig. Agric. Int. Perspect.* **1992**, *1*, 143–171.

17. Pickersgill, B. Chile peppers (*Capsicum* spp.). In *Ethnobotany of Mexico*; Springer: Berlin/Heidelberg, Germany, **2016**; pp. 417–437.

18. Perry, L.; Dickau, R.; Zarrillo, S.; Holst, I.; Pearsall, D.M.; Piperno, D.R.; Berman, M.J.; Cooke, R.G.; Rademaker, K.; Ranere, A.J.; et al. Starch fossils and the domestication and dispersal of chili peppers (*Capsicum* spp. L.) in the Americas. *Science* **2007**, *315*, 986–988. [CrossRef]

19. Doebley, J.F.; Gaut, B.S.; Smith, B.D. The molecular genetics of crop domestication. *Cell* **2006**, *127*, 1309–1321. [CrossRef] [PubMed]

20. Zeder, M.A. Core questions in domestication research. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 3191–3198. [CrossRef]

21. Ross-Ibarra, J.; Morrell, P.L.; Gaut, B.S. Plant domestication, a unique opportunity to identify the genetic basis of adaptation. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 8641–8648. [CrossRef]

22. Doebley, J.; Stec, A.; Wendel, J.; Edwards, M. Genetic and morphological analysis of a maize-teosinte F2 population: Implications for the origin of maize. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 9888–9892. [CrossRef] [PubMed]

23. Tian, F.; Stevens, N.M.; Buckler, E.S. Tracking footprints of maize domestication and evidence for a massive selective sweep on chromosome 10. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9979–9986. [CrossRef]

24. Studer, A.; Zhao, Q.; Ross-Ibarra, J.; Doebley, J. Identification of a functional transposon insertion in the maize domestication gene tb1. *Nat. Genet.* **2011**, *43*, 1160. [CrossRef] [PubMed]

25. Hufford, M.B.; Xu, X.; Van Heerwaarden, J.; Pyhäjärvi, T.; Chia, J.M.; Cartwright, R.A.; Elshire, R.J.; Glaubitz, J.C.; Guill, K.E.; Kaeppler, S.M.; et al. Comparative population genomics of maize domestication and improvement. *Nat. Genet.* **2012**, *44*, 808–811. [CrossRef]

26. Bellucci, E.; Bitocchi, E.; Ferrarini, A.; Benazzo, A.; Biagetti, E.; Klie, S.; Minio, A.; Rau, D.; Rodriguez, M.; Panziera, A.; et al. Decreased nucleotide and expression diversity and modified coexpression patterns characterize domestication in the common bean. *Plant Cell* **2014**, *26*, 1901–1912. [CrossRef] [PubMed]

27. Singh, J.; Zhao, J.; Vallejos, C.E. Differential transcriptome patterns associated with early seedling development in a wild and a domesticated common bean (*Phaseolus vulgaris* L.) accession. *Plant Sci.* **2018**, *274*, 153–162. [CrossRef]

28. Lippman, Z.; Tanksley, S.D. Dissecting the genetic pathway to extreme fruit size in tomato using a cross between the small-fruited wild species *Lycopersicon pimpinellifolium* and *L. esculentum* var. Giant Heirloom. *Genetics* **2001**, *158*, 413–422.

29. Müller, N.A.; Wijnen, C.L.; Srinivasan, A.; Ryngajllo, M.; Ofner, I.; Lin, T.; Ranjan, A.; West, D.; Maloof, J.N.; Sinha, N.R.; et al. Domestication selected for deceleration of the circadian clock in cultivated tomato. *Nat. Genet.* **2016**, *48*, 89–93. [CrossRef] [PubMed]

30. Sauvage, C.; Rau, A.; Aichholz, C.; Chadoeuf, J.; Sarah, G.; Ruiz, M.; Santoni, S.; Causse, M.; David, J.; Glémin, S. Domestication rewired gene expression and nucleotide diversity patterns in tomato. *Plant J.* **2017**, *91*, 631–645. [CrossRef]

31. Razifard, H.; Ramos, A.; Della Valle, A.L.; Bodary, C.; Goetz, E.; Manser, E.J.; Li, X.; Zhang, L.; Visa, S.; Tieman, D.; et al. Genomic evidence for complex domestication history of the cultivated tomato in Latin America. *Mol. Biol. Evol.* **2020**, *37*, 1118–1132. [CrossRef]

32. Hernández-Verdugo, S.; Luna-Reyes, R.; Oyama, K. Genetic structure and differentiation of wild and domesticated populations of *Capsicum annuum* (Solanaceae) from Mexico. *Plant Syst. Evol.* **2001**, *226*, 129–142. [CrossRef]

33. Paran, I.; Van Der Knaap, E. Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. *J. Exp. Bot.* **2007**, *58*, 3841–3852. [CrossRef]

34. Carvalho, S.; Ragassi, C.; Bianchetti, L.; Reifschneider, F.; Buso, G.; Faleiro, F. Morphological and genetic relationships between wild and domesticated forms of peppers (*Capsicum frutescens* L. and *C. chinense* Jacquin). *Genet. Mol. Res.* **2014**, *13*, 7447–7464. [CrossRef] [PubMed]

35. Taitano, N.; Bernau, V.; Jardón-Barbolla, L.; Leckie, B.; Mazourek, M.; Mercer, K.; McHale, L.; Michel, A.; Baumler, D.; Kantar, M.; et al. Genome-wide genotyping of a novel Mexican Chile Pepper collection illuminates the history of landrace differentiation after *Capsicum annuum* L. domestication. *Evol. Appl.* **2019**, *12*, 78–92. [CrossRef]

36. Qin, C.; Yu, C.; Shen, Y.; Fang, X.; Chen, L.; Min, J.; Cheng, J.; Zhao, S.; Xu, M.; Luo, Y.; et al. Whole-genome sequencing of cultivated and wild peppers provides insights into Capsicum domestication and specialization. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 5135–5140. [CrossRef]

37. Kim, S.; Park, M.; Yeom, S.I.; Kim, Y.M.; Lee, J.M.; Lee, H.A.; Seo, E.; Choi, J.; Cheong, K.; Kim, K.T.; et al. Genome sequence of the hot pepper provides insights into the evolution of pungency in Capsicum species. *Nat. Genet.* **2014**, *46*, 270–278. [CrossRef]

38. Ahn, Y.K.; Karna, S.; Jun, T.H.; Yang, E.Y.; Lee, H.E.; Kim, J.H.; Kim, J.H. Complete genome sequencing and analysis of *Capsicum annuum* varieties. *Mol. Breed.* **2016**, *36*, 140. [CrossRef]

39. Hulse-Kemp, A.M.; Maheshwari, S.; Stoffel, K.; Hill, T.A.; Jaffe, D.; Williams, S.R.; Weisenfeld, N.; Ramakrishnan, S.; Kumar, V.; Shah, P.; et al. Reference quality assembly of the 3.5-Gb genome of *Capsicum annuum* from a single linked-read library. *Hortic. D* **2018**, *5*, 4. [CrossRef] [PubMed]

40. Albert, V.A.; Chang, T.H. Evolution of a hot genome. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 5069–5070. [CrossRef] [PubMed]

41. Díaz-Valenzuela, E.; Sawers, R.H.; Cibrián-Jaramillo, A. Cis-and Trans-Regulatory Variations in the Domestication of the Chili Pepper Fruit. *Mol. Biol. Evol.* **2020**, *37*, 1593–1603. [CrossRef]

42. Holland, M.J. Transcript abundance in yeast varies over six orders of magnitude. *J. Biol. Chem.* **2002**, *277*, 14363–14366. [CrossRef] [PubMed]

43. Wu, Z.; Liu, W.; Jin, X.; Ji, H.; Wang, H.; Glusman, G.; Robinson, M.; Liu, L.; Ruan, J.; Gao, S. NormExpression: An R package to normalize gene expression data using evaluated methods. *Front. Genet.* **2019**, *10*, 400. [CrossRef] [PubMed]

44. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140. [CrossRef]

45. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **1995**, *57*, 289–300. [CrossRef]

46. Kothari, S.; Joshi, A.; Kachhwaha, S.; Ochoa-Alejo, N. Chilli peppers? A review on tissue culture and transgenesis. *Biotechnol. Adv.* **2010**, *28*, 35–48. [CrossRef] [PubMed]

47. Tian, S.L.; Li, L.; Shah, S.; Gong, Z.H. The relationship between red fruit colour formation and key genes of capsanthin biosynthesis pathway in Capsicum annuum. *Biol. Plant.* **2015**, *59*, 507–513. [CrossRef]

48. Liu, W.; Chen, L.; Zhang, S.; Hu, F.; Wang, Z.; Lyu, J.; Wang, B.; Xiang, H.; Zhao, R.; Tian, Z.; et al. Decrease of gene expression diversity during domestication of animals and plants. *BMC Evol. Biol.* **2019**, *19*, 19. [CrossRef]

49. Santaguida, S.; Musacchio, A. The life and miracles of kinetochores. *EMBO J.* **2009**, *28*, 2511–2531. [CrossRef]

50. D?Archivio, S.; Wickstead, B. Trypanosome outer kinetochore proteins suggest conservation of chromosome segregation machinery across eukaryotes. *J. Cell Biol.* **2017**, *216*, 379–391. [CrossRef]

51. Filkov, V. Identifying gene regulatory networks from gene expression data. In *Handbook of Computational Molecular Biology*; CRC Press: Boca Raton, FL, USA, **2005**; pp. 101–132.

52. Allocco, D.J.; Kohane, I.S.; Butte, A.J. Quantifying the relationship between co-expression, co-regulation and gene function. *BMC Bioinform.* **2004**, *5*, 18. [CrossRef]

53. Spies, D.; Ciaudo, C. Dynamics in transcriptomics: Advancements in RNA-seq time course and downstream analysis. *Comput. Struct. Biotechnol. J.* **2015**, *13*, 469–477. [CrossRef]

54. Darwin, C. *The Variation of Animals and Plants Under Domestication*; Cambridge University Press: Cambridge, UK, 1868; Volume 2.

55. Swain, P.S.; Elowitz, M.B.; Siggia, E.D. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12795–12800. [CrossRef]

56. Chapman, M.A.; Pashley, C.H.; Wenzler, J.; Hvala, J.; Tang, S.; Knapp, S.J.; Burke, J.M. A genomic scan for selection reveals candidates for genes involved in the evolution of cultivated sunflower (*Helianthus annuus*). *Plant Cell* **2008**, *20*, 2931–2945. [CrossRef]

57. Li, Y.h.; Zhao, S.c.; Ma, J.x.; Li, D.; Yan, L.; Li, J.; Qi, X.t.; Guo, X.s.; Zhang, L.; He, W.m.; et al. Molecular footprints of domestication and improvement in soybean revealed by whole genome re-sequencing. *BMC Genom.* **2013**, *14*, 579. [CrossRef]

58. Huang, C.L.; Hung, C.Y.; Chiang, Y.C.; Hwang, C.C.; Hsu, T.W.; Huang, C.C.; Hung, K.H.; Tsai, K.C.; Wang, K.H.; Osada, N.; et al. Footprints of natural and artificial selection for photoperiod pathway genes in Oryza. *Plant J.* **2012**, *70*, 769–782. [CrossRef] [PubMed]

59. Pollinger, J.P.; Bustamante, C.D.; Fledel-Alon, A.; Schmutz, S.; Gray, M.M.; Wayne, R.K. Selective sweep mapping of genes with large phenotypic effects. *Genome Res.* **2005**, *15*, 1809–1819. [CrossRef]

60. Rothammer, S.; Seichter, D.; Förster, M.; Medugorac, I. A genome-wide scan for signatures of differential artificial selection in ten cattle breeds. *BMC Genom.* **2013**, *14*, 908. [CrossRef]

61. Dai, Q.; Geng, L.; Lu, M.; Jin, W.; Nan, X.; He, P.a.; Yao, Y. Comparative transcriptome analysis of the different tissues between the cultivated and wild tomato. *PLoS ONE* **2017**, *12*, e0172411. [CrossRef]

62. Rong, J.; Lammers, Y.; Strasburg, J.L.; Schidlo, N.S.; Ariyurek, Y.; De Jong, T.J.; Klinkhamer, P.G.; Smulders, M.J.; Vrieling, K. New insights into domestication of carrot from root transcriptome analyses. *BMC Genom.* **2014**, *15*, 895. [CrossRef] [PubMed]

63. Liu, T.; Tang, S.; Zhu, S.; Tang, Q.; Zheng, X. Transcriptome comparison reveals the patterns of selection in domesticated and wild ramie (*Boehmeria nivea* L. Gaud). *Plant Mol. Biol.* **2014**, *86*, 85–92. [CrossRef]

64. Chaudhary, B.; Hovav, R.; Rapp, R.; Verma, N.; Udall, J.A.; Wendel, J.F. Global analysis of gene expression in cotton fibers from wild and domesticated Gossypium barbadense. *Evol. Dev.* **2008**, *10*, 567–582. [CrossRef] [PubMed]

65. Monson, M.S.; Cardona, C.J.; Coulombe, R.A.; Reed, K.M. Hepatic transcriptome responses of domesticated and wild turkey embryos to aflatoxin B1. *Toxins* **2016**, *8*, 16. [CrossRef]

66. Christie, M.R.; Marine, M.L.; Fox, S.E.; French, R.A.; Blouin, M.S. A single generation of domestication heritably alters the expression of hundreds of genes. *Nat. Commun.* **2016**, *7*, 10676. [CrossRef]

67. Albert, F.W.; Somel, M.; Carneiro, M.; Aximu-Petri, A.; Halbwax, M.; Thalmann, O.; Blanco-Aguiar, J.A.; Plyusnina, I.Z.; Trut, L.; Villafuerte, R.; et al. A comparison of brain gene expression levels in domesticated and wild animals. *PLoS Genet.* **2012**, *8*, e1002962. [CrossRef]

68. Azzi, L.; Deluche, C.; Gévaudant, F.; Frangne, N.; Delmas, F.; Hernould, M.; Chevalier, C. Fruit growth-related genes in tomato. *J. Exp. Bot.* **2015**, *66*, 1075–1086. [CrossRef]

69. Bourdon, M.; Frangne, N.; Mathieu-Rivet, E.; Nafati, M.; Cheniclet, C.; Renaudin, J.P.; Chevalier, C. Endoreduplication and growth of fleshy fruits. In *Progress in Botany 71*; Springer: Berlin/Heidelberg, Germany, **2010**; pp. 101–132.

70. Chevalier, C. Cell cycle control and fruit development. *Annu. Plant Rev. Online* **2018**, *32*, 269–293.

71. Ogawa, D.; Ishikawa, K.; Nunomura, O.; Mii, M. Correlation between fruit characters and degree of polysomaty in fruit tissues of *Capsicum*. *J. Jpn. Soc. Hortic. Sci.* **2010**, *79*, 168–173. [CrossRef]

72. Ogawa, D.; Ishikawa, K.; Mii, M. Difference in the polysomaty degree during fruit development among plants with different ploidy levels produced by artificial chromosome doubling of a pepper (*Capsicum annuum*) cultivar 'Shishitou No. 562'. *Sci. Hortic.* **2012**, *134*, 121–126. [CrossRef]

73. Vanneste, S.; Coppens, F.; Lee, E.; Donner, T.J.; Xie, Z.; Van Isterdael, G.; Dhondt, S.; De Winter, F.; De Rybel, B.; Vuylsteke, M.; et al. Developmental regulation of CYCA2s contributes to tissue-specific proliferation in Arabidopsis. *EMBO J.* **2011**, *30*, 3430–3441. [CrossRef]

74. Guo, M.; Simmons, C.R. Cell number counts—The fw2. 2 and CNR genes and implications for controlling plant fruit and organ size. *Plant Sci.* **2011**, *181*, 1–7. [CrossRef]

75. Gonzalez, N.; Gévaudant, F.; Hernould, M.; Chevalier, C.; Mouras, A. The cell cycle-associated protein kinase WEE1 regulates cell size in relation to endoreduplication in developing tomato fruit. *Plant J.* **2007**, *51*, 642–655. [CrossRef] [PubMed]

76. Ariizumi, T.; Shinozaki, Y.; Ezura, H. Genes that influence yield in tomato. *Breed. Sci.* **2013**, *63*, 3–13. [CrossRef] [PubMed]

77. Giovannoni, J.J. Genetic regulation of fruit development and ripening. *Plant Cell* **2004**, *16*, S170–S180. [CrossRef] [PubMed]

78. Ma, B.; Ding, Y.; Li, C.; Li, M.; Ma, F.; Yuan, Y. Comparative Proteomic Analysis Reveals Key Proteins Linked to the Accumulation of Soluble Sugars and Organic Acids in the Mature Fruits of the Wild Malus Species. *Plants* **2019**, *8*, 488. [CrossRef]

79. Guo, J.; Wang, P.; Cheng, Q.; Sun, L.; Wang, H.; Wang, Y.; Kao, L.; Li, Y.; Qiu, T.; Yang, W.; et al. Proteomic analysis reveals strong mitochondrial involvement in cytoplasmic male sterility of pepper (*Capsicum annuum* L.). *J. Proteom.* **2017**, *168*, 15–27. [CrossRef]

80. Voelckel, C.; Borevitz, J.O.; Kramer, E.M.; Hodges, S.A. Within and between whorls: Comparative transcriptional profiling of *Aquilegia* and *Arabidopsis*. *PLoS ONE* **2010**, *5*, e9735. [CrossRef]

81. Mizushima, N.; Komatsu, M. Autophagy: Renovation of cells and tissues. *Cell* **2011**, *147*, 728–741. [CrossRef]

82. Lee, T.A.; Wetering, S.W.V.; Brusslan, J.A. Stromal protein degradation is incomplete in Arabidopsis thaliana autophagy mutants undergoing natural senescence. *BMC Res. Notes* **2013**, *6*, 17. [CrossRef]

83. Launholt, D.; Merkle, T.; Houben, A.; Schulz, A.; Grasser, K.D. Arabidopsis chromatin-associated HMGA and HMGB use different nuclear targeting signals and display highly dynamic localization within the nucleus. *Plant Cell* **2006**, *18*, 2904–2918. [CrossRef]

84. Kobayashi, K.; Suzuki, T.; Iwata, E.; Nakamichi, N.; Suzuki, T.; Chen, P.; Ohtani, M.; Ishida, T.; Hosoya, H.; Müller, S.; et al. Transcriptional repression by MYB3R proteins regulates plant organ growth. *EMBO J.* **2015**, *34*, 1992–2007. [CrossRef]

85. Shin, J.; Jeong, G.; Park, J.Y.; Kim, H.; Lee, I. MUN (MERISTEM UNSTRUCTURED), encoding a SPC24 homolog of NDC80 kinetochore complex, affects development through cell division in arabidopsis thaliana. *Plant J.* **2018**, *93*, 977–991. [CrossRef]

86. Yao, M.; Wakamatsu, Y.; Itoh, T.J.; Shoji, T.; Hashimoto, T. Arabidopsis SPIRAL2 promotes uninterrupted microtubule growth by suppressing the pause state of microtubule dynamics. *J. Cell Sci.* **2008**, *121*, 2372–2381. [CrossRef] [PubMed]

87. Wightman, R.; Chomicki, G.; Kumar, M.; Carr, P.; Turner, S.R. SPIRAL2 determines plant microtubule organization by modulating microtubule severing. *Curr. Biol.* **2013**, *23*, 1902–1907. [CrossRef] [PubMed]

88. Buschmann, H.; Fabri, C.O.; Hauptmann, M.; Hutzler, P.; Laux, T.; Lloyd, C.W.; Schäffner, A.R. Helical growth of the Arabidopsis mutant tortifolia1 reveals a plant-specific microtubule-associated protein. *Curr. Biol.* **2004**, *14*, 1515–1521. [CrossRef]

89. Petrovská, B.; Jeřábková, H.; Kohoutová, L.; Cenklová, V.; Pochylová, Ž.; Gelová, Z.; Kočárová, G.; Váchová, L.; Kurejová, M.; Tomaštíková, E.; et al. Overexpressed TPX2 causes ectopic formation of microtubular arrays in the nuclei of acentrosomal plant cells. *J. Exp. Bot.* **2013**, *64*, 4575–4587. [CrossRef]

90. Vos, J.W.; Pieuchot, L.; Evrard, J.L.; Janski, N.; Bergdoll, M.; de Ronde, D.; Perez, L.H.; Sardon, T.; Vernos, I.; Schmit, A.C. The plant TPX2 protein regulates prospindle assembly before nuclear envelope breakdown. *Plant Cell* **2008**, *20*, 2783–2797. [CrossRef]

91. Nashima, K.; Shimizu, T.; Nishitani, C.; Yamamoto, T.; Takahashi, H.; Nakazono, M.; Itai, A.; Isuzugawa, K.; Hanada, T.; Takashina, T.; et al. Microarray analysis of gene expression patterns during fruit development in European pear (Pyrus communis). *Sci. Hortic.* **2013**, *164*, 466–473. [CrossRef]

92. Steiner, A.; Rybak, K.; Altmann, M.; McFarlane, H.E.; Klaeger, S.; Nguyen, N.; Facher, E.; Ivakov, A.; Wanner, G.; Kuster, B.; et al. Cell cycle-regulated PLEIADE/At MAP 65–3 links membrane and microtubule dynamics during plant cytokinesis. *Plant J.* **2016**, *88*, 531–541. [CrossRef] [PubMed]

93. Pignocchi, C.; Minns, G.E.; Nesi, N.; Koumproglou, R.; Kitsios, G.; Benning, C.; Lloyd, C.W.; Doonan, J.H.; Hills, M.J. ENDOSPERM DEFECTIVE1 is a novel microtubule-associated protein essential for seed development in Arabidopsis. *Plant Cell* **2009**, *21*, 90–105. [CrossRef] [PubMed]

94. Chen, C.; Marcus, A.; Li, W.; Hu, Y.; Calzada, J.P.V.; Grossniklaus, U.; Cyr, R.J.; Ma, H. The Arabidopsis ATK1 gene is required for spindle morphogenesis in male meiosis. *Development* **2002**, *129*, 2401–2409. [PubMed]

95. Cifuentes, M.; Jolivet, S.; Cromer, L.; Harashima, H.; Bulankova, P.; Renne, C.; Crismani, W.; Nomura, Y.; Nakagami, H.; Sugimoto, K.; et al. TDM1 regulation determines the number of meiotic divisions. *PLoS Genet.* **2016**, *12*, e1005856. [CrossRef] [PubMed]

96. McGuinness, B.E.; Hirota, T.; Kudo, N.R.; Peters, J.M.; Nasmyth, K. Shugoshin prevents dissociation of cohesin from centromeres during mitosis in vertebrate cells. *PLoS Biol.* **2005**, *3*, e86. [CrossRef] [PubMed]

97. Kitajima, T.S.; Kawashima, S.A.; Watanabe, Y. The conserved kinetochore protein shugoshin protects centromeric cohesion during meiosis. *Nature* **2004**, *427*, 510–517. [CrossRef] [PubMed]

98. Zhu, Z.; Sun, B.; Cai, W.; Zhou, X.; Mao, Y.; Chen, C.; Wei, J.; Cao, B.; Chen, C.; Chen, G.; et al. Natural variations in the MYB transcription factor MYB31 determine the evolution of extremely pungent peppers. *New Phytol.* **2019**, *223*, 922–938. [CrossRef]

99. Sun, B.; Zhu, Z.; Chen, C.; Chen, G.; Cao, B.; Chen, C.; Lei, J. Jasmonate-inducible R2R3-MYB transcription factor regulates capsaicinoid biosynthesis and stamen development in Capsicum. *J. Agric. Food Chem.* **2019**, *67*, 10891–10903. [CrossRef]

100. Pickersgill, B. Domestication of plants in the Americas: Insights from Mendelian and molecular genetics. *Ann. Bot.* **2007**, *100*, 925–940. [CrossRef]

101. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2013.

102. Escoto-Sandoval, C.; Flores-Díaz, A.; Reyes-Valdés, M.H.; Ochoa-Alejo, N.; Martinez, O. An R package for data mining chili pepper fruit transcriptomes. *Evaluation* **2020**. [CrossRef]

103. Edgar, R.; Domrachev, M.; Lash, A.E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **2002**, *30*, 207–210. [CrossRef]

104. Martínez, O.; Escoto-Sandoval, C. *Salsa: An R Package of Data Mining Facilities for Capsicum Gene Expression Profiles*; **2021**. [CrossRef]