

Mapping and Functional Characterization of the Golden Fruit 1 (gf1) in Melon (*Cucumis melo* L.)

Shuai Li

Chinese Academy of Agricultural Sciences

Huihui Wang

Qingdao Agricultural University

Yang Li

Qingdao Agricultural University

Feng Jing

Qingdao Agricultural University

Yuanchao Xu

Chinese Academy of Agricultural Sciences

Shijun Deng

Qingdao Agricultural University

Naonao Wang

Qingdao Agricultural University

Zhonghua Zhang

Qingdao Agricultural University

Sen Chai

chaisen@qau.edu.cn

Qingdao Agricultural University <https://orcid.org/0009-0001-0142-4652>

Research Article

Keywords: Melon, Fruit color, EGY1, Chloroplast development

Posted Date: November 6th, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-5295132/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Theoretical and Applied Genetics on February 26th, 2025. See the published version at <https://doi.org/10.1007/s00122-025-04849-4>.

Abstract

Melon (*Cucumis melo* L.) is a globally cultivated vegetable crop with substantial economic importance. Fruit color is an important factor that affects the appearance quality and economic value of crops, which is determined mainly by the type and concentration of pigments such as chlorophyll, carotenoids, anthocyanins, and flavonoids. Identifying and utilizing the genetic loci that govern melon fruit color is crucial for guiding breeding efforts aimed at enhancing melon rind coloration. In this study, we identified an EMS-induced mutant that produces golden-colored fruit which named gf1. Through MUTmap and map-based cloning, we narrowed the gf1 locus to an 862 kb region containing 28 SNPs, consistent with typical alkylation-induced mutations. Of these SNPs, only one was located in a coding region, resulting in a stop-gained mutation in the gene MELO3C008847, which exhibits the highest sequence similarity to AtEGY1 in Arabidopsis. Genome editing of CsEGY1, the homolog of CmEGY1 in cucumber, led to the golden fruit. Transcriptome and metabolome analysis reveal that flavonoid and carotenoid biosynthesis pathways were changed, which may affect the appearance quality and flavor of gf1 fruit. The identification and characterization of egy1 offer new insights into modifying fruit color in melons, providing valuable implications for future breeding programs.

Introduction

Melon (*Cucumis melo* L., $2n = 2x = 24$) is an annual crop from the *Cucurbit*s, which is widely cultivated all over the world. It holds significant economic value due to its diverse edible and medicinal properties. In addition to yield and flavor, appearance quality is also an important factor affecting the economic value of crops. In order to meet the market demand, it is urgent to improve the edible quality and appearance quality of melon.

Fruit color is an important commercial trait, which is primarily influenced by the content of pigments such as chlorophyll, carotenoids, and flavonoids (Freilich et al. 2015). Green, white and yellow are three major skin colors in melon fruit, which were primarily influenced by chlorophyll and flavonoids accumulation. Studies have shown that green skin is dominant over white skin, while white skin is dominant over yellow skins (Zhao et al. 2019). Previous studies showed that *APRR2* had a conserved function in regulating chloroplast development, which in turn affected the green-white color of *Cucurbit* crops (Gebretsadik et al. 2024; Liu et al. 2016; Oren et al. 2019). The accumulation of naringenin chalcone is responsible for the yellow skin color, which was reported to be negatively regulated by the F-box protein encoded by *CmKFB* (Feder et al. 2015; Gur et al. 2017; Tadmor et al. 2010). Additionally, *CmOR* and *CmGLK* were reported to influence fruit color by regulate β -carotene accumulation and chloroplast development respectively (Tzuri et al. 2015; Yang et al. 2023). Although several genes related to melon skin color has been reported, the specific regulatory mechanisms of skin color remain unclear.

Chloroplasts are the primary carriers of photosynthetic pigments and the main sites of photosynthesis in plants. According to the endosymbiotic theory, chloroplasts originated from the evolution of a eukaryotic host acquiring a cyanobacterial endosymbiont (McFadden 2001). Chloroplasts are typical semi-

autonomous organelles, with their genomes usually containing only 120–130 genes in a nucleoid-like structure (Kessler and Schnell 2009). Most other chloroplast proteins are encoded by the nuclear genome and are imported into the chloroplasts after translation. Thus, chloroplast development is co-regulated by both nuclear and chloroplast genes. Any disruption in this process can interfere with chloroplast development, leading to changes in color phenotypes. For example, genes such as *PFKB1* (Zhu et al. 2020), *WLP1* (Song et al. 2014), and *GARS* (Cao et al. 2019) identified in rice have been found to affect chloroplast development to varying degrees, resulting in changes in leaf color. Chloroplast division is also crucial for maintaining its physiological activity. In the tomato *ACR6* mutant, only one or two giant chloroplasts are present in mesophyll cells, causing an albino phenotype in stems and developing fruits, although the red color and carotenoid content of the fruit remain largely unchanged (Chang et al. 2021; Forth and Pyke 2006); a similar phenotype is observed in cucumber (Sun et al. 2023). Meanwhile, the *ACR5* mutation also inhibits chloroplast division, causing light green peel in cucumber (Zhou et al. 2015). Furthermore, premature termination of *CsAPRR2* translation reduces the number of chloroplasts in the peel, leading to white fruit skin (Tang et al. 2018).

EGY1 is a member of the S2P protease gene family and is a nuclear-encoded, plastid-targeted, membrane-bound, ATP-independent metalloprotease. EGY1 was initially identified in an Arabidopsis mutant (Chen et al. 2005), and the Arabidopsis genome encodes four homologs of EGY1: AtS2P1/AraSP (At2g32480); AtS2P2 (At1g05140); AtAMOS1/AtEGY1 (At5g35220); and AtEGY2 (At5g05740) (Adam 2015). Loss-of-function of EGY1 results in similar chloroplast impairment in Arabidopsis, tomato (*Solanum lycopersicum*), and monocot foxtail millet (*Setaria italica*), suggesting a conserved role of EGY1 homologs (Barry et al. 2012; Zhang et al. 2018). The *Ategy1* mutant exhibits a dual phenotype characterized by reduced chlorophyll accumulation and abnormal gravitropic bending of the hypocotyl. Moreover, the expression level of *EGY1* in the mutant can be upregulated under light and ethylene treatment. Additionally, EGY1 is essential for the accumulation of chlorophyll and chlorophyll a/b binding (CAB) proteins in the chloroplast membrane and is involved in the response to ammonium and phosphate stress (Chen et al. 2005; Guo et al. 2008; Li et al. 2012; Yu et al. 2016).

In this study, an EMS-induced melon mutant shows golden-colored fruit was identified and designated as *gf1*. TEM observation and chlorophyll content determination showed that the golden fruit in *gf1* was mainly resulted from impaired chloroplast development and reduced chlorophyll content. To identify the gene responsible for the golden fruit in *gf1*, MUTmap analysis and map-based cloning were performed, narrowing down the *gf1* locus to an 862 kb region. Of all the variations in this region, an SNP resulting in a stop-gained mutation in the gene *MELO3C008847* which predicted to encode an ATP-independent metalloprotease (EGY1), was predicted to be the causative mutation. Knock-out of *CsEGY1* in cucumber led to yellow leaves, stem and also fruit peel. Transcriptome and metabolome analysis showed that in addition to affecting chloroplast development, EGY1 also affected flavonoid and carotenoid synthesis. This study provides a novel genetic resource and new insights for improving appearance traits in melon breeding.

Materials and methods

Plant materials and growth conditions

The cultivated melon inbred line ssp. *agrestis* 13C was used to construct an EMS-mutagenized library. The seeds of inbred line 13C were treated with EMS (Sigma M0880, USA) for 12 h. The M₁ plants were self-pollinated, and the golden-fruit line S134 (*gf1* hereafter) was identified in the M₂ population. *gf1* was self-pollinated for over two generations, and then planted for phenotypic data collection with wild-type 13C. The M₂ population was used for inheritance analysis and map-based cloning. All plants were grown in the greenhouse at Dagou River National Science and Technology Park, Qingdao.

All materials were planted at Agricultural Hi-tech Industry Zone, Jimo, Shandong (36°56'N, 120°21'E) in plastic green house in 2020, 2021 and 2022.

Mapping strategy and identification of the candidate gene of *gf1*

MUTmap analysis was used for mapping the candidate gene of *gf1*. The young leaves of 15 wild-type and 15 golden-fruit individuals were sampled respectively for genomic DNA extracting with CTAB method (Murray and Thompson 1980). Pair-end sequencing libraries with a read length of 150 bp and insert sizes of ~ 350 bp were subjected to whole genome re-sequencing with Illumina NovaSeq 6000, and roughly 30× genome sequences for each pool were generated. Clean reads were mapped to Cme13C reference genome using BWA (0.7.17-r1188) with default parameters (Li and Durbin 2009). SAMtools (v1.11) (Li et al. 2009) was used to transform the mapping results from sam to bam format, sort the bam file, remove mark duplicates and index the bam file. The SNPs and InDels were identified using Genome Analysis Toolkit (4.1.4.0) (McKenna et al. 2010). The annotation and effects of mutations on gene function were predicted using SnpEff software (v4.3) (Cingolani et al. 2012). For the SNPs in two bulks, we further filtered SNPs with two more steps: (i) extraction of SNPs that exhibit G-A or C-T transitions, which are the most frequent changes caused by ethyl methanesulfonate mutagenesis. (ii) retention of SNPs which SNP index ≤ 0.5 in dominant pool and SNP index ≥ 0.3 in mutant pool. SNP index plot regression lines were obtained by averaging SNP indices from a moving window of five consecutive SNPs and shifting the window one SNP at a time. The x-axis value of each averaged SNP index was set at a midpoint between the first and fifth SNP.

According to the re-sequencing data, variants with ΔSNP values above 0.6 and read depths over 10 were selected and developed into CAPS marker with dCAPS Finder 2.0 (<http://biology4.wustl.edu/dcaps/>). All primers were synthesized by Sangon Biological Company (primer information were listed in Supplementary Table S1). The polymorphism CAPS markers were screened with *gf1* and wild type, then used for genotyping in M₂ population.

Chlorophyll and carotene content

The total chlorophyll in the wild-type and *gf1* fruit was extracted with 95% ethanol and analyzed using a spectrophotometer (Shimadzu, Japan). The total chlorophyll, chlorophyll a, chlorophyll b and carotene

contents were estimated with light absorption values at 649, 665 and 470 nm, respectively (Sartory and Grobbelaar, 1984).

Transmission Electron Microscope

The 30 DAP fruit skin of WT and *gf1* were cut into $\sim 0.5 \text{ cm}^2$ pieces and fixed in 2.5% glutaraldehyde for 12 hours, then washed three times with PBS. Afterwards, the samples were dehydrated through different concentrations of alcohol, dried, coated with gold particles and examined with S-3500N scanning electron microscope (Hitachi, Japan).

Phylogenetic Analysis

Multiple sequence (Gene ID shown in Supplemental Table 2) alignment were performed using the Muscle program with default parameters (Edgar, 2004). The sequence alignments were used for the subsequent phylogenetic analysis. Phylogenetic trees were generated with MEGA7.0 software (<https://megasoftware.net/>) using the neighbor join algorithms (Kumar et al., 2016) and then beautified with ITOL website (<http://itol.embl.de/>).

Stable genetic transformation in cucumber

In order to obtain the CRISPR/Cas9 mediated gene editing of *CsEGY1*, specific sgRNA target sites were designed with the online CRISPR-GE (<http://skl.scau.edu.cn/home/>). A PCR fragment were amplified by using vector *pCBC-DT1T2* as template, and then inserted into the vector *pBSE402* with In-Fusion method (Hu et al., 2017; Xing et al., 2014). Transformation of cucumber (XTMC) was carried out as previously described (Xin et al., 2022).

RNA-seq analysis and co-expression analysis

The fruit of 13C and *gf1* (30 DPA) were used for RNA extraction and sequenced with Illumina HiSeq2000 platform. Using Hisat2 (v2.2.1) (Kim et al. 2015) with the default parameters, we mapped the clean reads of six samples to the melon reference genome (Cme13C). For each gene expression, the FPKM (Fragments Per Kilobase of Transcript Per Million Mapped Reads) value was calculated using StringTie (v2.1.5) (Pertea et al. 2015) to estimate the gene expression level. For further analysis, WGCNA (Langfelder and Horvath 2008) was performed on genes with a coefficient of variation (CV) exceeding 0.3 in the transcriptome of the entire melon tissues to conduct co-expression analysis, and EGY1 related genes were subsequently extracted from the corresponding module. DESeq2 (Love et al. 2014) was used to normalize read counts and to test for differential expression. The DEGs between the bulks were identified with $P\text{-adjust} < 0.05$ and $|\log_2(\text{fc})| \geq 1$ as the significance cut-off. Gene Ontology terms were determined with the Omicshare (<https://www.omicshare.com/tools/Home/Soft/gogsea>) website.

Measurements of metabolites

Fruit sample of 13C and *gf1* were subjected to freeze-drying, followed by weighting and grounding into a fine powder. Subsequently, 50 mg of the resulting dry powder was dissolved in 1.2 mL of 70% methanol and then vortexed for 30 s with a 30-min interval, and this process was repeated for six times. After refrigerating the suspension overnight at 4°C, it was then centrifugated at 12,000 rpm for 10 min. The resulting supernatant was filtered using a microporous member with a pore size of 0.22 µm. The filtrate obtained was used for measurements of broadly targeted metabolites. For the relative quantification of these metabolites, an ultra-performance liquid chromatography-electrospray ionization-tandem mass spectrometry system (UPLC-MS/MS) was employed, using a scheduled multiple reaction monitoring method (Chen et al. 2013; Chen et al. 2016). The monitored signal intensities were normalized by the internal standard (lidocaine, at a concentration of 0.1 mg l⁻¹). Confirmation of the broadly targeted metabolites was achieved through the use of authentic standards. Two biological replicates were performed for each sample.

Results

The *gf1* mutant has golden fruit and reduced chlorophyll accumulation

A golden-fruit mutant was identified in the previously constructed EMS-mutagenized library, which was designated as *gf1*. Compared with wild-type 13C, the *gf1* mutant fruit showed golden peel and flesh at the mature stage of fruit development (Fig. 1A-B). Actually, the ovary of *gf1* exhibited yellow from the very early stage of fruit development, and then turned light-green gradually, which became golden eventually (Supplemental Fig. S1A-D). Chlorophyll and carotenoid measurement found that *gf1* mutant showed severely reducing chlorophyll a, chlorophyll b and carotenoid (Fig. 1C). Considering that reduced chlorophyll content often affects plant development and assimilate accumulation by reducing photosynthetic efficiency, fruit weight and soluble solids content (SSC) of wild-type 13C and *gf1* were measured in the mature stage. Compared with 13C, *gf1* showed no difference in fruit weight (Supplemental Fig. 1A), while SSC content in *gf1* showed a slightly decrease with no significance (Supplemental Fig. 1B).

Thylakoid development was severely impaired in *gf1*

To explore whether the golden-fruit phenotype in *gf1* was associated with the number or development of chloroplast, transmission electron microscopic analysis was conducted with fruit peels from *gf1* and 13C. The chloroplast number in *gf1* was significantly lower than that in 13C (Fig. 2A, D). Moreover, *gf1* chloroplasts had reduced granal thylakoids and poorly developed lamellae networks. There were neither a conspicuous sheet membrane nor grana structures, but instead ubiquitous plastoglobuli in mutant chloroplasts (Fig. 2B-C, E-F). These results indicate that the causative mutation severely inhibited chloroplast development which resulted in golden-fruit phenotype.

Genetic inheritance analysis, map-based cloning and candidate gene isolation in *gf1*

Previous re-sequencing data from EMS-induced library of 13C showed that limited mutations were induced in each line, thus M₂ population which without extra backcrossing process with wild-type was used for inheritance analysis and map-based cloning. All F₁ individuals displayed same fruit color as WT, indicating the dominance of normal fruit color over the golden-fruit phenotype. Among the 783 plants in M₂ population, 589 individuals exhibited normal fruit color and 194 individuals displayed golden-fruit phenotype, which showed a segregation ratio of 3:1 ($\chi^2 = 0.021$, $P > 0.05$), implies that the golden-fruit phenotype in *gf1* was controlled by a single gene (Table 1).

To identify the candidate region contributing to the golden-fruit phenotype in *gf1*, 15 normal and 15 golden-fruit M₂ plants were sampled respectively to construct W-pool and M-pool. Genome-wide re-sequencing and BSA analysis were subsequently performed. This resulted in a peak region between 11.42 to 24.75 Mb with a cluster of SNPs harboring high SNP-index resided on chromosome 8 (Fig. 3A). Among the high-quality variations between W-pool and M-pool, 7 SNPs in predicted interval were selected and developed into CAPS marker. By applying 7 CAPS/dCAPS markers to 783 plants in M₂ population, *gf1* locus was delimited to an 862 kb region between *CAPS-2* and *CAPS-4* (Fig. 3B).

According to the whole-genome resequencing data of W-pool and M-pool, 133 SNPs in total were found in 862 kb region, while only 28 of them displayed G/C to A/T transitions (Fig. 3C), which was consistent with canonical alkylation-induced mutation (Drake and Baltz 1976). Among the 28 SNPs, 21 of them were located in intergenic region. The remaining 7 SNPs consist of 4 intron-located SNPs, 2 3' UTR-located SNPs, 2 coding-region-located SNP which resulted in a synonymous mutation and a stop-gained mutation respectively (Fig. 3C). The gene *MELO3C008847*, which shares the highest sequence similarity to AtEGY1, encodes a S2P-like chloroplast membrane-located protein and reported to participate in chloroplast development, harbored the stop-gained SNP. The G-A transition in the 7th exon in *gf1* leads to a 116-aa-long residue truncation in the C-terminal of *CmEGY1*, which consist a part of highly conserved sequence in S2P-M50-like domain (Fig. 3D-E; Supplemental Fig. S3).

Above all, the stop-gained mutation in *CmEGY1* was identified as the causative variation which leads to golden-fruit phenotype and for further analyzed.

Temporal-spatial expression patterns and phylogenetic analysis of CmEGY1

As the *Cmegy1* mutant showed yellow leaves, stems and fruits, the expression pattern of *CmEGY1* was investigated in different organs at different developmental stages with previous RNA-seq data. As expected, *CmEGY1* showed no organ-specific pattern, which express in all organs at different stage (Fig. 4A). Among all detected samples, the expression of *CmEGY1* was higher in leaves and mature fruit while lower in root. CmEGY1 encodes a 560-aa-long protein with an estimated molecular mass of 60.33 kDa, while C-terminal (aa: 259–484) predicted to be S2P-M50_like_2 domain which consist of catalytic center required for protease function (Supplemental Fig. S3). There are 8 transmembrane regions in C-

terminal of CmEGY1 (aa: 249–545). Premature termination in *egy1* result in loss of 5 transmembrane regions and part of S2P-M50_like_2 domain (Fig. 4B). Phylogenetic analysis revealed that EGY proteins grouped into three clades, while CmEGY1 belongs to EGY1 clade (Fig. 4C).

CRISPR-mediated knock-out of CsEGY1 results in yellow plant and fruit in cucumber

To further determine the function of *CmEGY1*, the homolog of *CmEGY1* in cucumber, *CsEGY1* (*CsaV3_4G012000*), was knock-out with CRISPR/Cas9 system. We selected two target sites on the first and second exon of the *CsEGY1* as small guide ribonucleic acids (sgRNAs) to construct a vector for transform in cucumber inbred line 'XTMC' (Fig. 5A). Two null *CsEGY1* alleles with a complicated editing type (10 bases deletion, 6 bases insertion and 4 bases substitution) and a 1158-bp deletion were obtained, which named *cr-1* and *-2* respectively (Fig. 5B). The *Csegy1-cr* mutants exhibited yellow leaves, stem, sepal of male flowers and also fruit peel (Fig. 5C-F). These results confirmed that *CsEGY1* is essential for chloroplast development, which demonstrates that *Cmegy1* was the gene leads to golden-fruit phenotype in *gf1*.

Transcriptome profiling and metabolome analysis reveals the regulated network of CmEGY1

To explore the molecular function of *CmEGY1* in regulating of fruit development, we performed RNA-seq analysis of the fruit samples from 13C and *gf1* harvested at 35 DAP (days after pollination). Using a *P*-adjust < 0.05 and fold change ≥ 1 as thresholds, 1261 genes were identified differentially expressed between 13C and *gf1*, while 591 genes down-regulated and 670 genes up-regulated (Supplementary Table 3). GO and KEGG pathway enrichment analysis were performed with these DEGs. Not surprisingly, the GO terms 'photosynthesis' (GO:0015979), 'photosystem II' (GO:0009523), 'chloroplast' (GO:0009507), 'thylakoid' (GO:0009579), and KEGG pathways 'photosynthesis proteins' (ko00194), 'photosynthesis - antenna proteins' (ko00196), 'carbon fixation in photosynthetic organisms' (ko00710) were significantly enriched (Supplemental Fig. S4).

Additionally, to further investigate the regulated network which consist of higher correlation genes with *EGY1* in fruit development, previously generated transcriptome profiling with different part of fruit were used for gene co-expression analysis. In total, a module contains 402 genes which exhibited co-expression tendency with *EGY1* was identified. Among 402 co-expressed genes, 123 genes showed differentially expressed between 13C and *gf1* (Fig. 6A). KEGG pathway enrichment analysis was performed with the 123 genes, and the top 20 pathways were shown in Fig. 6B. Apart from pathways related to chloroplast development and photosynthesis, 'carotenoid biosynthesis' (ko00906) was also enriched significantly, which implies that changes of carotenoid may also contribute to golden fruit phenotype in some extent.

Given to the apparent phenotype and differentially expressed genes in *gf1*, we wanted to understand the metabolic changes in *gf1* fruit. Fruits of 13C and *gf1* were collected for metabolites measurements, and 162 differential metabolites were identified, including 65 down-regulated metabolites and 97 up-regulated metabolites in *gf1*. KEGG analysis showed that pathways including 'plant hormone signal

transduction', 'flavonoid biosynthesis', 'carotenoid biosynthesis' and 'flavone and flavonol biosynthesis' were enriched significantly (Fig. 6C). To better study the potential important metabolic pathways in the content of substances in different groups of change, differential metabolites in top 5 pathways were selected for clustering analysis. The result showed that organic acids, phenolic acids and flavonoids were the main components of the differential metabolites (Fig. 6D), which may affect the quality and flavor of melon fruit.

Discussion

Identification of a novel fruit color mutant in melon

Melon is an important vegetable crop in *Cucurbitaceae* and is widely planted worldwide. As a major character of appearance quality, melon fruit color greatly affects the commercial and economic value. Mining the genes related to fruit color can not only fully understand the formation mechanism of skin color diversity, but also of great practical significance for melon appearance diversity breeding. In general, composition and content of chlorophyll, carotenoid, anthocyanins and flavonoids are major factors that affect leaf and fruit color in melon. For example, the GOLDEN2-LIKE (GLK2) transcription factor regulating chlorophyll levels in melon (Yang et al. 2023); the *Or* gene (*CmOr*) increasing carotenoids accumulation by inducing the biogenesis of chromoplasts with an enhanced sink strength (Tzuri et al. 2015); the Kelch domain-containing F-box protein coding (*CmKFB*) gene functions as a posttranscriptional regulator that negative regulate flavonoid accumulation (Feder et al. 2015).

In the current study, we discovered a golden fruit mutant in an EMS-mutagenized library of the melon inbred line 13C. We show that the golden fruit phenotype was mainly from strikingly reduced chlorophyll content and impaired chloroplasts development (Fig. 1B; Fig. 2). Using BSA-seq method combined with map-based cloning strategy, the *gf1* locus was delimited to an 862 kb region (Fig. 3). Of all the variants in this interval, only one mutation resulted in premature termination of CmEGY1 protein translation and was considered to be a functional variant leading to phenotypic generation. The knock-out of *CsEGY1* resulted in yellow fruit and leaf in cucumber, indicating that *CmEGY1* regulate fruit color in melon (Fig. 5). To sum up, *gf1* is a novel fruit color mutant which due to impaired chloroplasts development in melon.

Potential for increasing production and enhancing the purity of hybrid seed production

To meet the growing global demand for food, increasing crop yields has become a key objective for biotechnological innovation. Enhancing crop architecture and photosynthesis through genetic engineering and synthetic biology has proven to be effective in boosting yields. Reducing chlorophyll (Chl) content may improve the conversion efficiency of absorbed photosynthetically active radiation into biomass, thereby increasing yield in dense monoculture crops by enhancing light penetration and distribution within the canopy (Pettigrew et al. 1989; Sakowska et al. 2018; Slattery et al. 2017). Moreover, improving light distribution within the canopy can help narrow the gap between theoretical and

actual yields (Cutolo et al. 2023). In addition to the golden-fruit phenotype, the cotyledon, true leaf, and stem color of *gf1* are markedly different from those of 13C at very early stage (Supplemental Fig. S1). Thus, *gf1* holds the potential to increase yield by optimizing planting density.

Additionally, the instability of hybrid seed purity is a key factor limiting hybrid breeding production. Traditionally, seed purity detection has relied on field identification; however, the use of seedling marker traits offers a more intuitive, accurate, simple, fast, and cost-effective method for hybrid breeding and seed purity identification. Genetic dissections of chloroplast development using leaf coloration as the phenotypic readout have been proven to be extremely fruitful as a spectrum of leaf color mutants ranging from albino, yellow, pale green, virescent, and variegation can be easily identified through genetic screening (Yu et al. 2007). By introducing the recessive *egy1* mutation into the female parent, hybrid purity can be verified before seeds are used in production, allowing for the elimination of self-cross progeny of the female parent at the seedling stage. This approach would significantly streamline hybrid production.

Conserved function in chloroplast development suggests its potential for reducing fertilizer use

Chloroplast thylakoid membrane protein complexes, including PSII, PSI, cytochrome b6f, and ATP synthase, play a crucial role in converting light energy into chemical energy (Eberhard et al. 2008). Mutations in chloroplast proteases often lead to arrested or delayed chloroplast development, underscoring the importance of proteostasis within chloroplasts. The function of EGY1 in chloroplast development is highly conserved across various species. A previous study demonstrated that EGY1 is essential for the proper accumulation of PSI complexes and the homeostasis of PSII complexes in *Arabidopsis* (Qi et al. 2020). *egy1* mutant in tomato exhibit pleiotropic chlorophyll-deficient phenotypes, including reduced chlorophyll accumulation in leaves, flowers and also unmaturing fruit (Barry et al. 2012). The loss-of-function of Seita.9G060500 (SiYGL2) showed chlorotic leaves, stems and panicles in millet (Zhang et al., 2018). As expected, GO and KEGG analysis showed that EGY1 is closely related to photosynthesis, carbon fixation and thylakoid development (Supplemental Fig. S4). In land plants, the GOLDEN2-LIKE (GLK) family of transcription factors are master regulators of chloroplast biogenesis. A recent study found two RR-MYBs acts synergistically with GLK to control chloroplast biogenesis (Frangedakis et al. 2024). The GLK proteins and RR-MYBs have a broader set of targets that extends to genes allowing CO₂ fixation, photorespiration, photosystem assembly, and repair. Whether EGY1 is a target gene of GLK and RR-MYBs needs further investigation.

Furthermore, loss-of-function of EGY1 in *Arabidopsis* confer resistance to ammonium and phosphate stress via ABA signaling pathway (Li et al. 2012; Yu et al. 2016). Given the challenges posed by over-fertilization and unbalanced fertilization in our country, it is crucial to evaluate the effects of the *egy1* mutant on reducing fertilizer application and improving efficiency in melons and other crops in the following study.

Declarations

Funding

This work was supported by the National Natural Science Foundation of China (32225044 to Z.Z., 32130093 to Z.Z., 32472767 to S.C., 32102404 to S.C.). This work was also supported by the “Taishan Scholar” Foundation of the People’s Government of Shandong Province.

Conflict of interest

All authors declare that they have no conflict of interest.

Author contribution statement

Z.Z., S.C. and S.L.: conceived and designed the experiments; S.C., S.L. and H.W. performed most of the experiments and analyzed the data; J.F., S.D. and N.W. helped in planting melon population; Y.L. and Y.X. performed the bioinformatics analyses; Z.Z., S.C. and S.L.: wrote the paper. All authors have read and approved the final manuscript.

Data availability

The materials presented in this article will be freely available to any researcher wishing to use them for non-commercial purposes. The data that support the findings of this study are available on request from the corresponding author upon reasonable request. The author responsible for distribution of materials and sequencing data in accordance with the policy described in the Instructions for Authors (www.springer.com) is: Sen Chai (chaisen@qau.edu.cn).

References

1. Adam Z (2015) Plastid intramembrane proteolysis. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1847:910–914
2. Barry CS, Aldridge GM, Herzog G, Ma Q, McQuinn RP, Hirschberg J, Giovannoni JJ (2012) Altered Chloroplast Development and Delayed Fruit Ripening Caused by Mutations in a Zinc Metalloprotease at the *lutescent2* Locus of Tomato. *Plant Physiol* 159:1086–1098
3. Cao P, Ren Y, Liu X, Zhang T, Zhang P, Xiao L, Zhang F, Liu S, Jiang L, Wan J (2019) Purine nucleotide biosynthetic gene *GARS* controls early chloroplast development in rice (*Oryza sativa* L). *Plant Cell Rep* 38:183–194
4. Chang J, Zhang F, Qin H, Liu P, Wang J, Wu S (2021) Mutation of *SIARC6* leads to tissue-specific defects in chloroplast development in tomato. *Hortic Res* 8:127

5. Chen G, Bi YR, Li N (2005) *EGY1* encodes a membrane-associated and ATP-independent metalloprotease that is required for chloroplast development. *Plant J* 41:364–375
6. Chen W, Gong L, Guo Z, Wang W, Zhang H, Liu X, Yu S, Xiong L, Luo J (2013) A Novel Integrated Method for Large-Scale Detection, Identification, and Quantification of Widely Targeted Metabolites: Application in the Study of Rice Metabolomics. *Mol Plant* 6:1769–1780
7. Chen W, Wang W, Peng M, Gong L, Gao Y, Wan J, Wang S, Shi L, Zhou B, Li Z, Peng X, Yang C, Qu L, Liu X, Luo J (2016) Comparative and parallel genome-wide association studies for metabolic and agronomic traits in cereals. *Nat Commun* 7:12767
8. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms. *SnEff Fly* 6:80–92
9. Cutolo EA, Guardini Z, Dall'Osto L, Bassi R (2023) A paler shade of green: engineering cellular chlorophyll content to enhance photosynthesis in crowded environments. *New Phytol* 239:1567–1583
10. Drake JW, Baltz RH (1976) The Biochemistry of Mutagenesis. *Annu Rev Biochem* 45:11–37
11. Eberhard S, Finazzi G, Wollman F-A (2008) The Dynamics of Photosynthesis. *Annu Rev Genet* 42:463–515
12. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797
13. Feder A, Burger J, Gao S, Lewinsohn E, Katzir N, Schaffer AA, Meir A, Davidovich-Rikanati R, Portnoy V, Gal-On A, Fei Z, Kashi Y, Tadmor Y (2015) A Kelch Domain-Containing F-Box Coding Gene Negatively Regulates Flavonoid Accumulation in Muskmelon. *Plant Physiol* 169:1714–1726
14. Forth D, Pyke KA (2006) The suffulta mutation in tomato reveals a novel method of plastid replication during fruit ripening. *J Exp Bot* 57:1971–1979
15. Frangedakis E, Yelina NE, Billakurthi K, Hua L, Schreier T, Dickinson PJ, Tomaselli M, Haseloff J, Hibberd JM (2024) MYB-related transcription factors control chloroplast biogenesis. *Cell* 187:4859–4876e4822
16. Freilich S, Lev S, Gonda I, Reuveni E, Portnoy V, Oren E, Lohse M, Galpaz N, Bar E, Tzuri G, Wissotsky G, Meir A, Burger J, Tadmor Y, Schaffer A, Fei Z, Giovannoni J, Lewinsohn E, Katzir N (2015) Systems approach for exploring the intricate associations between sweetness, color and aroma in melon fruits. *BMC Plant Biol* 15:71
17. Gebretsadik K, Chen W, Duan Y, Sun Y, He Y, Liao Q, Wang C, Bo K (2024) Map-based cloning reveals *Cpgp* gene encoding an APRR2 protein to regulate the green fruit peel formation in *Cucurbita pepo*. *Mol Breeding* 44:53
18. Guo D, Gao X, Li H, Zhang T, Chen G, Huang P, An L, Li N (2008) *EGY1* plays a role in regulation of endodermal plastid size and number that are involved in ethylene-dependent gravitropism of light-grown *Arabidopsis* hypocotyls. *Plant Mol Biol* 66:345–360

19. Gur A, Tzuri G, Meir A, Sa'ar U, Portnoy V, Katzir N, Schaffer AA, Li L, Burger J, Tadmor Y (2017) Genome-Wide Linkage-Disequilibrium Mapping to the Candidate Gene Level in Melon (*Cucumis melo*). *Sci Rep* 7:9770
20. Hu B, Li D, Liu X, Qi J, Gao D, Zhao S, Huang S, Sun J, Yang L (2017) Engineering Non-transgenic Gynoecious Cucumber Using an Improved Transformation Protocol and Optimized CRISPR/Cas9 System. *Mol Plant* 10:1575–1578
21. Kessler F, Schnell D (2009) Chloroplast biogenesis: diversity and regulation of the protein import apparatus. *Curr Opin Cell Biol* 21:494–500
22. Kim D, Langmead B, Salzberg SL (2015) HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 12:357–360
23. Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33:1870–1874
24. Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:1–13
25. Li B, Li Q, Xiong L, Kronzucker HJ, Krämer U, Shi W (2012) Arabidopsis Plastid AMOS1/EGY1 Integrates Abscissic Acid Signaling to Regulate Global Gene Expression Response to Ammonium Stress. *Plant Physiol* 160:2040–2051
26. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25:1754–1760
27. 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
28. Liu H, Jiao J, Liang X, Liu J, Meng H, Chen S, Li Y, Cheng Z (2016) Map-based cloning, identification and characterization of the *w* gene controlling white immature fruit color in cucumber (*Cucumis sativus* L). *Theor Appl Genet* 129:1247–1256
29. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550
30. McFadden GI (2001) Chloroplast Origin and Integration1. *Plant Physiol* 125:50–53
31. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–1303
32. Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res* 8:4321–4326
33. Oren E, Tzuri G, Vexler L, Dafna A, Meir A, Faigenboim A, Kenigswald M, Portnoy V, Schaffer AA, Levi A, Buckler ES, Katzir N, Burger J, Tadmor Y, Gur A (2019) The multi-allelic *APRR2* gene is associated with fruit pigment accumulation in melon and watermelon. *J Exp Bot* 70:3781–3794

34. Perteau M, Perteau 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
35. Pettigrew WT, Hesketh JD, Peters DB, Woolley JT (1989) Characterization of Canopy Photosynthesis of Chlorophyll-Deficient Soybean Isolines. *Crop Sci* 29:1025–1029
36. Qi Y, Wang X, Lei P, Li H, Yan L, Zhao J, Meng J, Shao J, An L, Yu F, Liu X (2020) The chloroplast metalloproteases VAR2 and EGY1 act synergistically to regulate chloroplast development in *Arabidopsis*. *J Biol Chem* 295:1036–1046
37. Sakowska K, Alberti G, Genesio L, Peressotti A, Delle Vedove G, Gianelle D, Colombo R, Rodeghiero M, Panigada C, Juszczak R, Celesti M, Rossini M, Haworth M, Campbell BW, Mevy J-P, Vescovo L, Cendrero-Mateo MP, Rascher U, Miglietta F (2018) Leaf and canopy photosynthesis of a chlorophyll deficient soybean mutant. *Plant Cell Environ* 41:1427–1437
38. Sartory DP, Grobbelaar JU (1984) Extraction of chlorophyll a from freshwater phytoplankton for spectrophotometric analysis. *Hydrobiologia* 114:177–187
39. Slattery RA, VanLoocke A, Bernacchi CJ, Zhu X-G, Ort DR (2017) Photosynthesis, Light Use Efficiency, and Yield of Reduced-Chlorophyll Soybean Mutants in Field Conditions. *Frontiers in Plant Science* 8
40. Song J, Wei X, Shao G, Sheng Z, Chen D, Liu C, Jiao G, Xie L, Tang S, Hu P (2014) The rice nuclear gene *WLP1* encoding a chloroplast ribosome L13 protein is needed for chloroplast development in rice grown under low temperature conditions. *Plant Mol Biol* 84:301–314
41. Sun W, Li X, Huang H, Wei J, Zeng F, Huang Y, Sun Q, Miao W, Tian Y, Li Y, Gao L, Li X, Gao H (2023) Mutation of *CsARC6* affects fruit color and increases fruit nutrition in cucumber. *Theor Appl Genet* 136:111
42. Tadmor Y, Burger J, Yaakov I, Feder A, Libhaber SE, Portnoy V, Meir A, Tzuri G, Sa'ar U, Rogachev I, Aharoni A, Abeliovich H, Schaffer AA, Lewinsohn E, Katzir N (2010) Genetics of Flavonoid, Carotenoid, and Chlorophyll Pigments in Melon Fruit Rinds. *J Agric Food Chem* 58:10722–10728
43. Tang H-Y, Dong X, Wang J-K, Xia J-H, Xie F, Zhang Y, Yao X, Xu Y-J, Wan Z-J (2018) Fine Mapping and Candidate Gene Prediction for White Immature Fruit Skin in Cucumber (*Cucumis sativus* L). *International Journal of Molecular Sciences*
44. Tzuri G, Zhou X, Chayut N, Yuan H, Portnoy V, Meir A, Sa'ar U, Baumkoler F, Mazourek M, Lewinsohn E, Fei Z, Schaffer AA, Li L, Burger J, Katzir N, Tadmor Y (2015) A 'golden' SNP in *CmOr* governs the fruit flesh color of melon (*Cucumis melo*). *Plant J* 82:267–279
45. Xin T, Tian H, Ma Y, Wang S, Yang L, Li X, Zhang M, Chen C, Wang H, Li H, Xu J, Huang S, Yang X (2022) Targeted creation of new mutants with compact plant architecture using CRISPR/Cas9 genome editing by an optimized genetic transformation procedure in cucurbit plants. *Hortic Res* 9:uhab086
46. Yang S, Wang X, Yan W, Zhang Y, Song P, Guo Y, Xie K, Hu J, Hou J, Wu Y, Zhu H, Sun S, Yang L (2023) Melon yellow-green plant (*Cmygp*) encodes a Golden2-like transcription factor regulating chlorophyll synthesis and chloroplast development. *Theor Appl Genet* 136:66

47. Yu FEI, Fu A, Aluru M, Park S, Xu Y, Liu H, Liu X, Foudree A, Nambogga M, Rodermeel S (2007) Variegation mutants and mechanisms of chloroplast biogenesis. *Plant Cell Environ* 30:350–365
48. Yu FW, Zhu XF, Li GJ, Kronzucker HJ, Shi WM (2016) The Chloroplast Protease AMOS1/EGY1 Affects Phosphate Homeostasis under Phosphate Stress. *Plant Physiol* 172:1200–1208
49. Zhang S, Zhi H, Li W, Shan J, Tang C, Jia G, Tang S, Diao X (2018) SiYGL2 Is Involved in the Regulation of Leaf Senescence and Photosystem II Efficiency in *Setaria italica* (L.) P. Beauv. *Frontiers in Plant Science* 9
50. Zhao G, Lian Q, Zhang Z, Fu Q, He Y, Ma S, Ruggieri V, Monforte AJ, Wang P, Julca I, Wang H, Liu J, Xu Y, Wang R, Ji J, Xu Z, Kong W, Zhong Y, Shang J, Pereira L, Argyris J, Zhang J, Mayobre C, Pujol M, Oren E, Ou D, Wang J, Sun D, Zhao S, Zhu Y, Li N, Katzir N, Gur A, Dogimont C, Schaefer H, Fan W, Bendahmane A, Fei Z, Pitrat M, Gabaldón T, Lin T, Garcia-Mas J, Xu Y, Huang S (2019) A comprehensive genome variation map of melon identifies multiple domestication events and loci influencing agronomic traits. *Nat Genet* 51:1607–1615
51. Zhou Q, Wang S, Hu B, Chen H, Zhang Z, Huang S (2015) An *ACCUMULATION AND REPLICATION OF CHLOROPLASTS 5* gene mutation confers light green peel in cucumber. *J Integr Plant Biol* 57:936–942
52. Zhu X, Ze M, Yin J, Chern M, Wang M, Zhang X, Deng R, Li Y, Liao H, Wang L, Tu B, Song L, He M, Li S, Wang W-M, Chen X, Wang J, Li W (2020) A phosphofructokinase B-type carbohydrate kinase family protein, PFKB1, is essential for chloroplast development at early seedling stage in rice. *Plant Sci* 290:110295

Tables

Tables 1 is not available with this version.

Table 1 Genetic inheritance analysis of gf1.

Figures

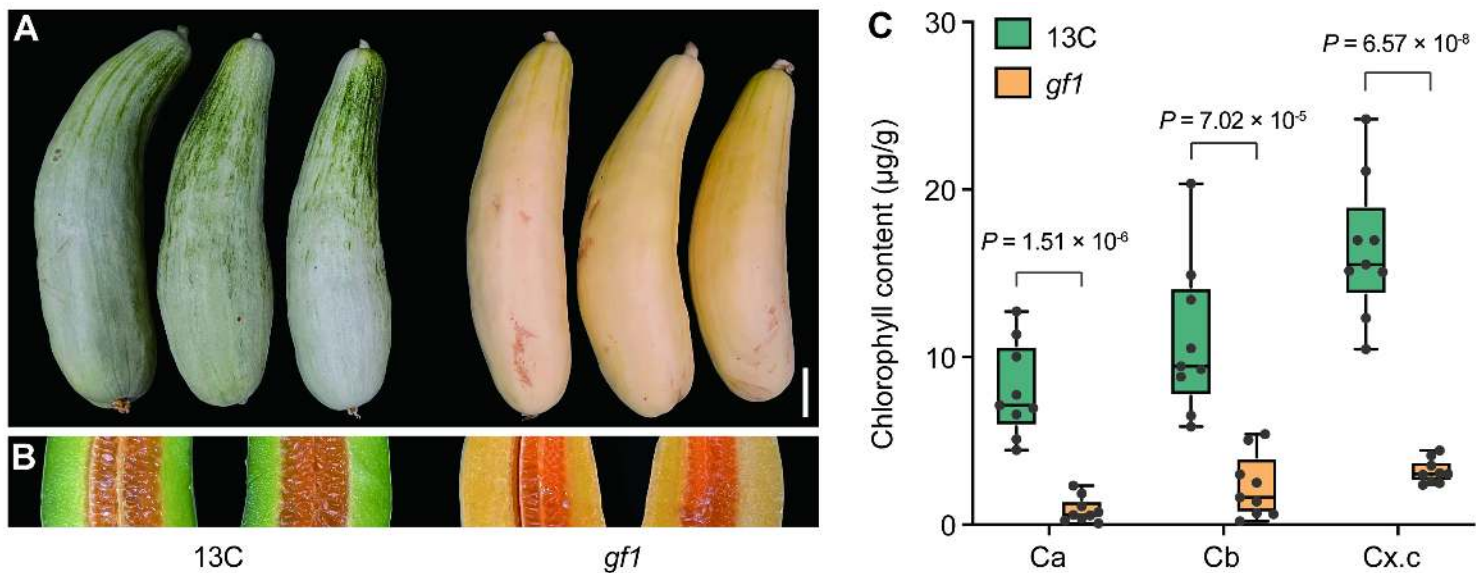


Figure 1

Phenotype analysis of *gf1*. The peel (A) and flesh (B) of 13C and *gf1* fruit in 30 DAP. (C) Measurement of chlorophyll concentration in 13C and *gf1*. Scale bars, 5 cm.

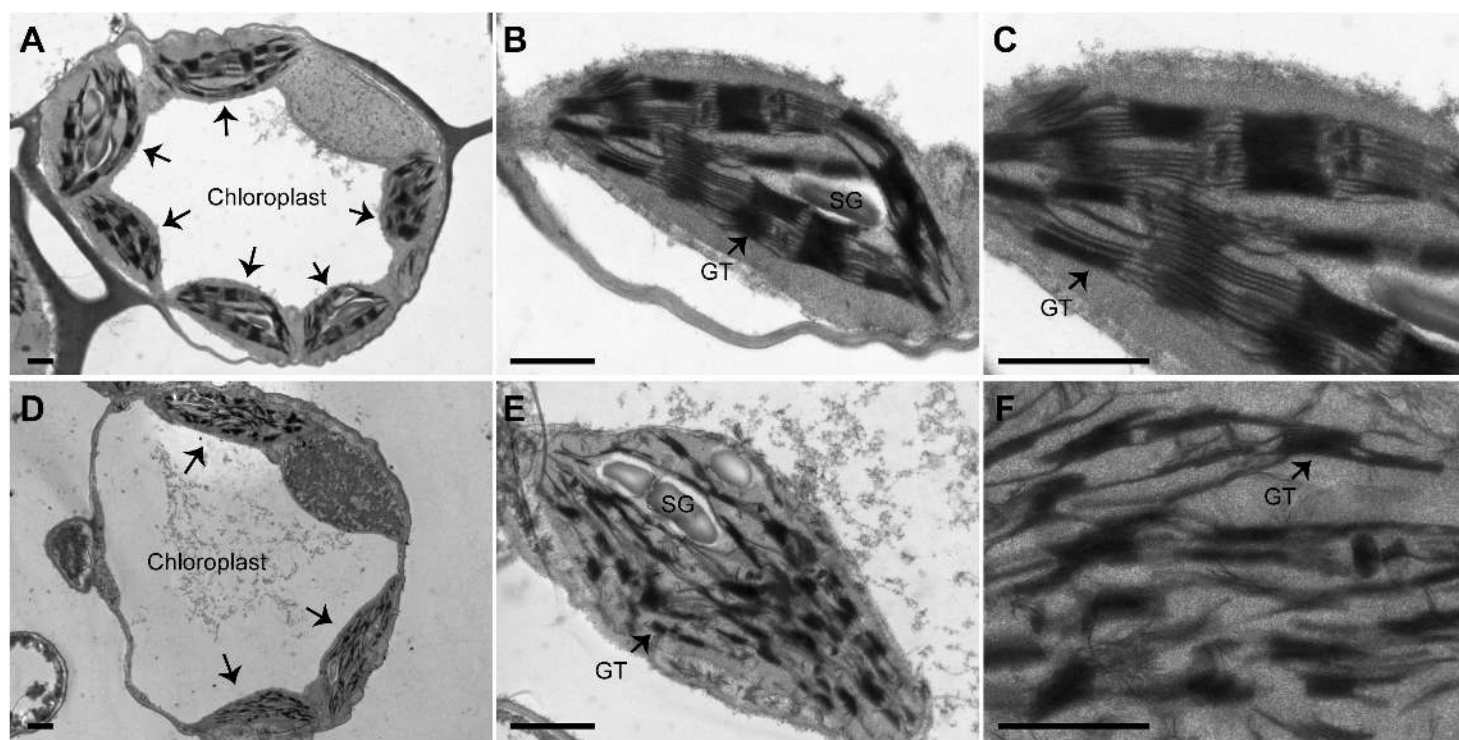


Figure 2

Transmission electron microscopic images of chloroplasts in 13C (A-C) and *gf1* (D-F). GT, grana thylakoid stacks; SG, starch granules. Scale bars, 1 µm.

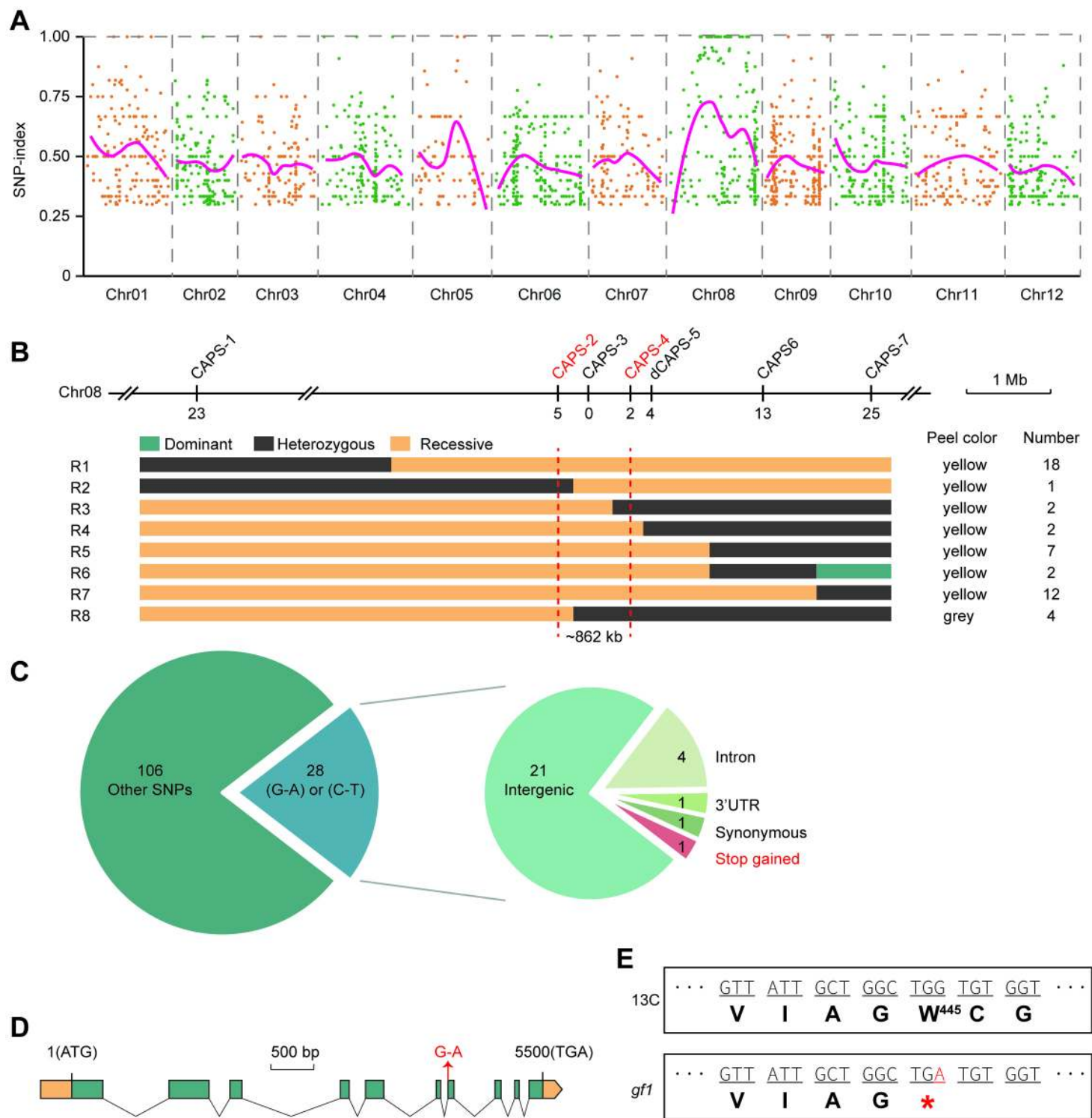


Figure 3

Genetic inheritance analysis, map-based cloning and candidate gene isolation in *gf1*. MUTmap analysis (A) and map-based cloning (B) of *gf1*. (C) Variations analysis between *CAPS-2* and *CAPS-4* in 13C and *gf1*. (D) Gene structure of EGY1. (E) The W⁴⁴⁵ mutation identified in *CmEGY1*.

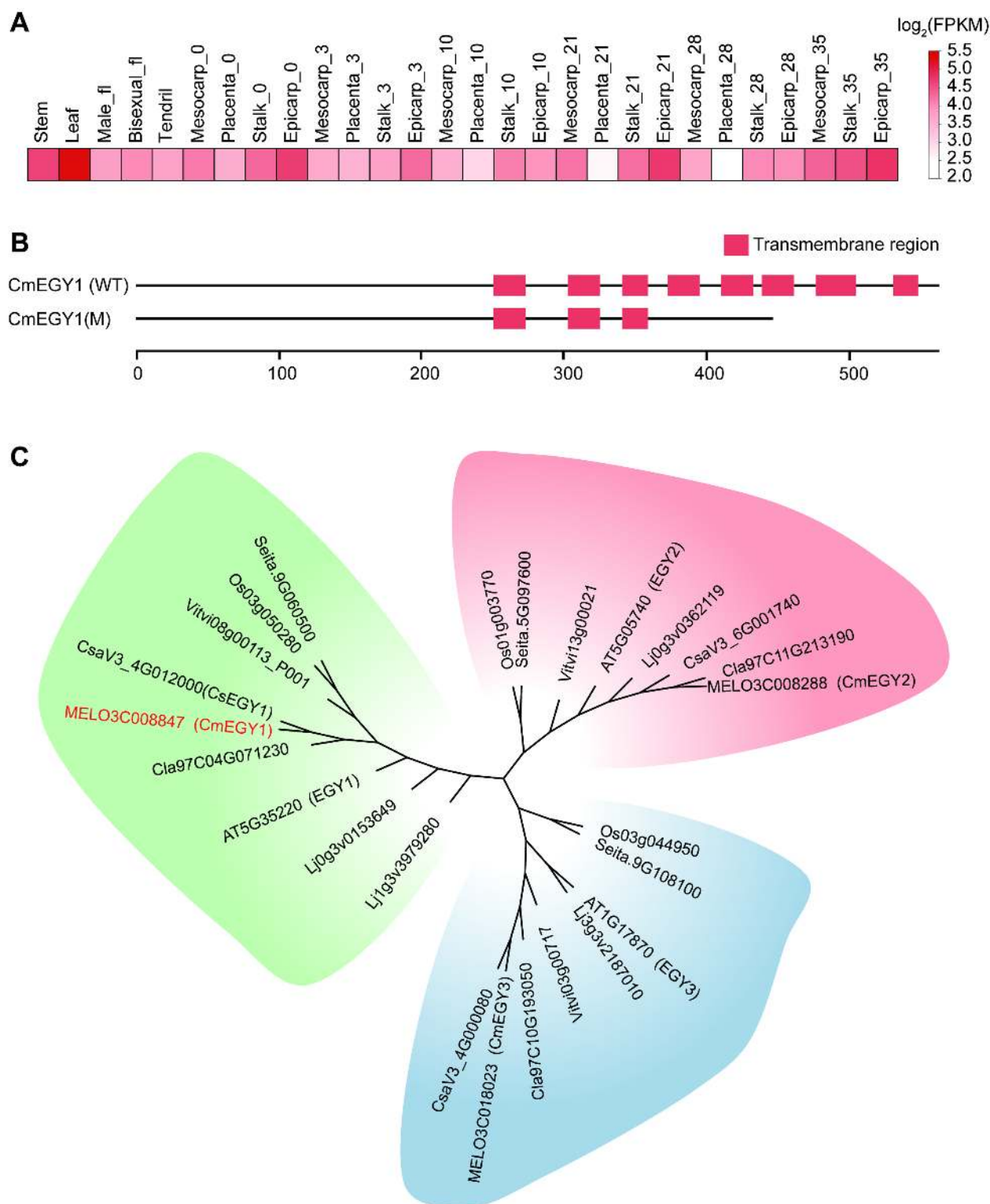


Figure 4

Temporal-spatial expression patterns and phylogenetic analysis of CmEGY1. (A) *CmEGY1* expression analysis among different tissues. (B) Transmembrane region analysis of *CmEGY1*. (C) Phylogenetic analysis of EGY1 proteins among 7 species (*Oryza sativa*; *Setaria italica*; *Arabidopsis thaliana*; *Lotus japonicus*; *Cucumis sativus*; *Cucumis melo*; *Citrullus lanatus*).

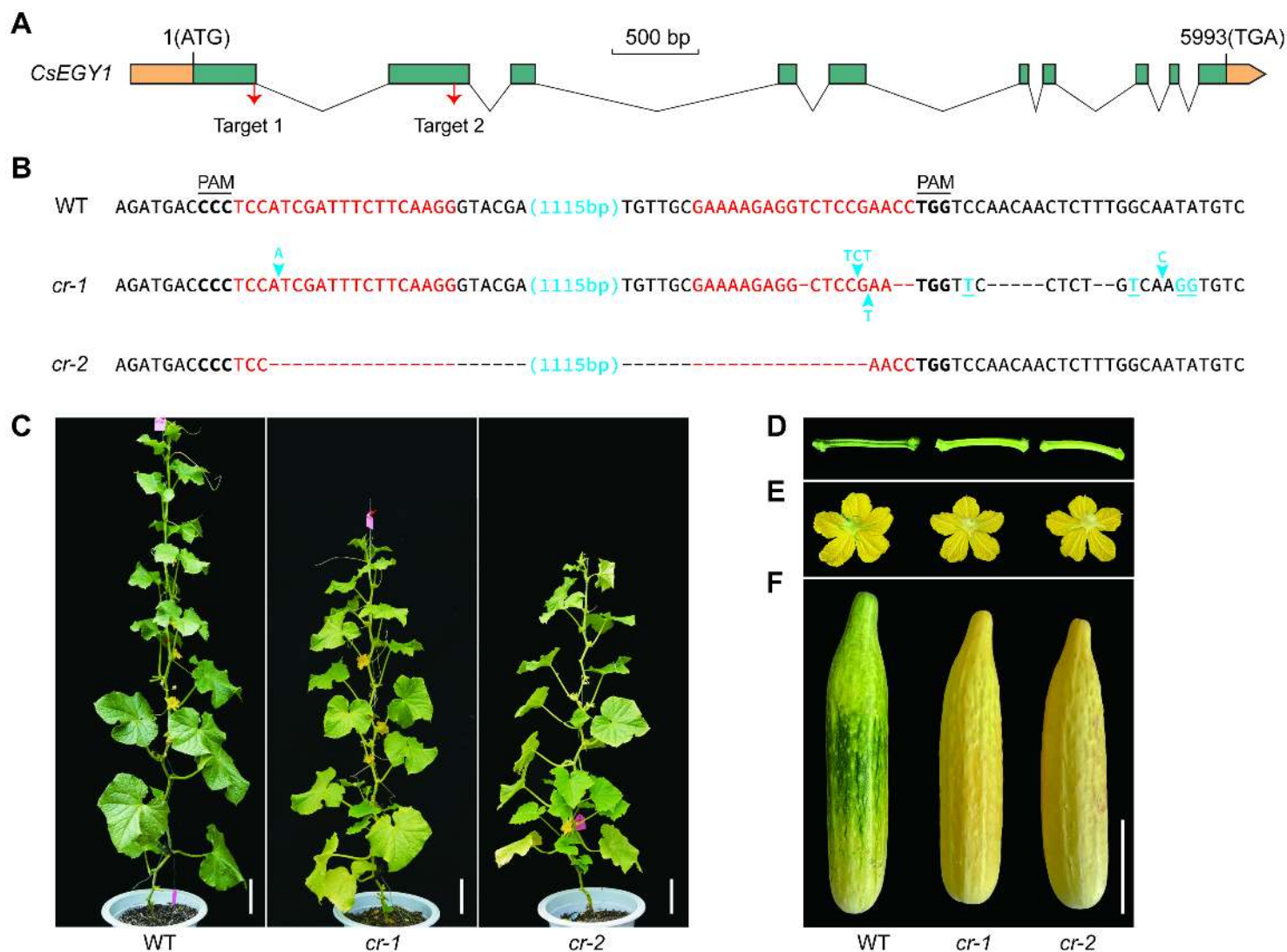


Figure 5

CRISPR-mediated knock-out of *CsEGY1* results in yellow plant and fruit in cucumber. (A) Gene structure of *CsEGY1*. Yellow rectangle: untranslated region; green rectangle: exon; black line: intron. **(B)** *CsEGY1* mutations generated by CRISPR/Cas9 using two single-guide RNAs. Phenotype observation of WT and edited mutants, include plant height **(C)**, stem **(D)**, male flower **(E)** and fruit color **(F)**. Scale bars, 10 cm.

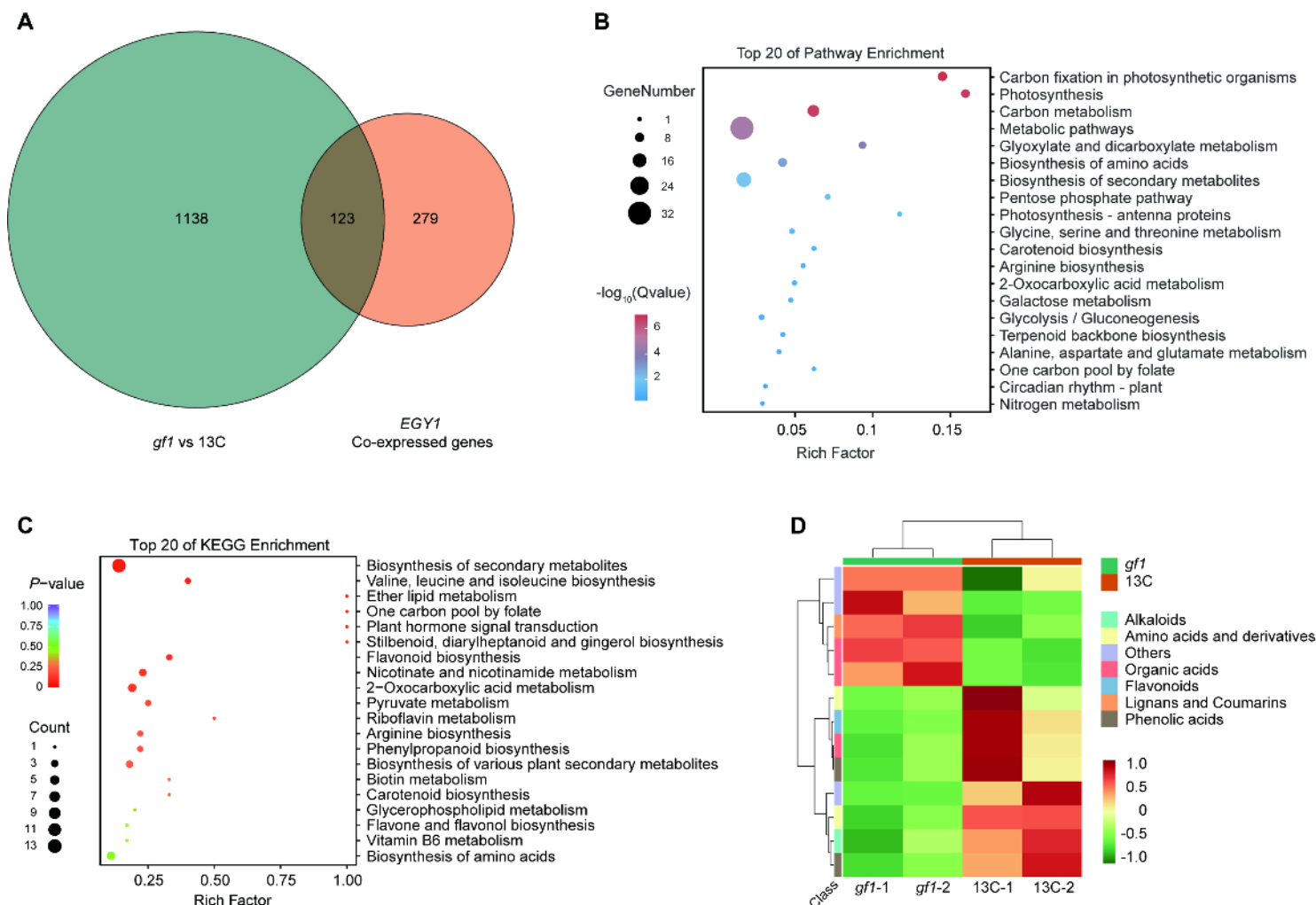


Figure 6

Transcriptome and metabolome analysis of *gf1* mutant. (A) Venn diagram showing the DEGs (*gf1* vs 13C) and EGY1-coexpressed genes in melon fruit respectively. (B) The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of shared genes between DEGs and EGY1-coexpressed genes in melon fruit. (C) KEGG analysis of differential metabolites between *gf1* and 13C. (D) Heatmap of differential metabolites in top 5 KEGG pathways.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFiguresS1S4TablesS1S2.docx](#)
- [SupplementaryTablesS3S4.xlsx](#)