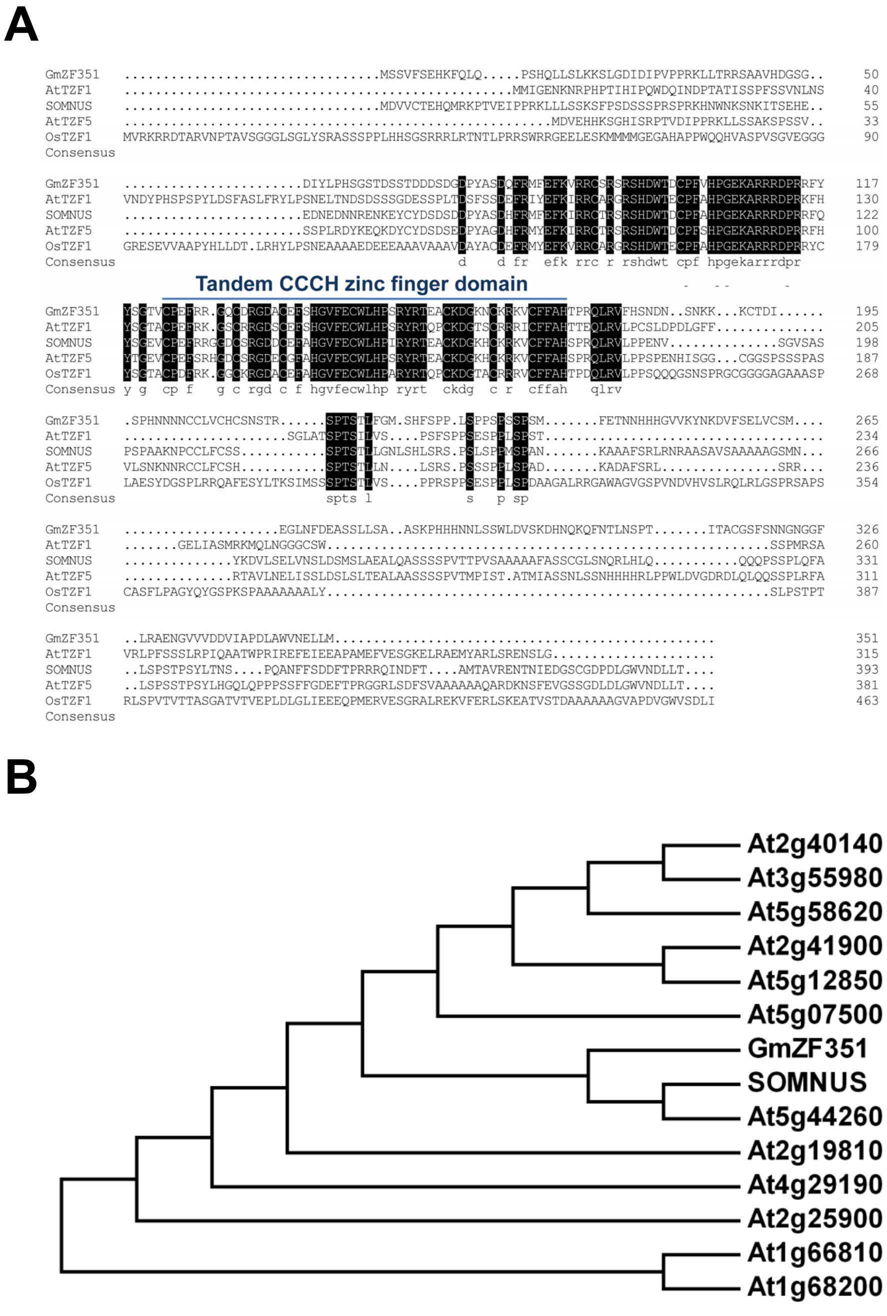
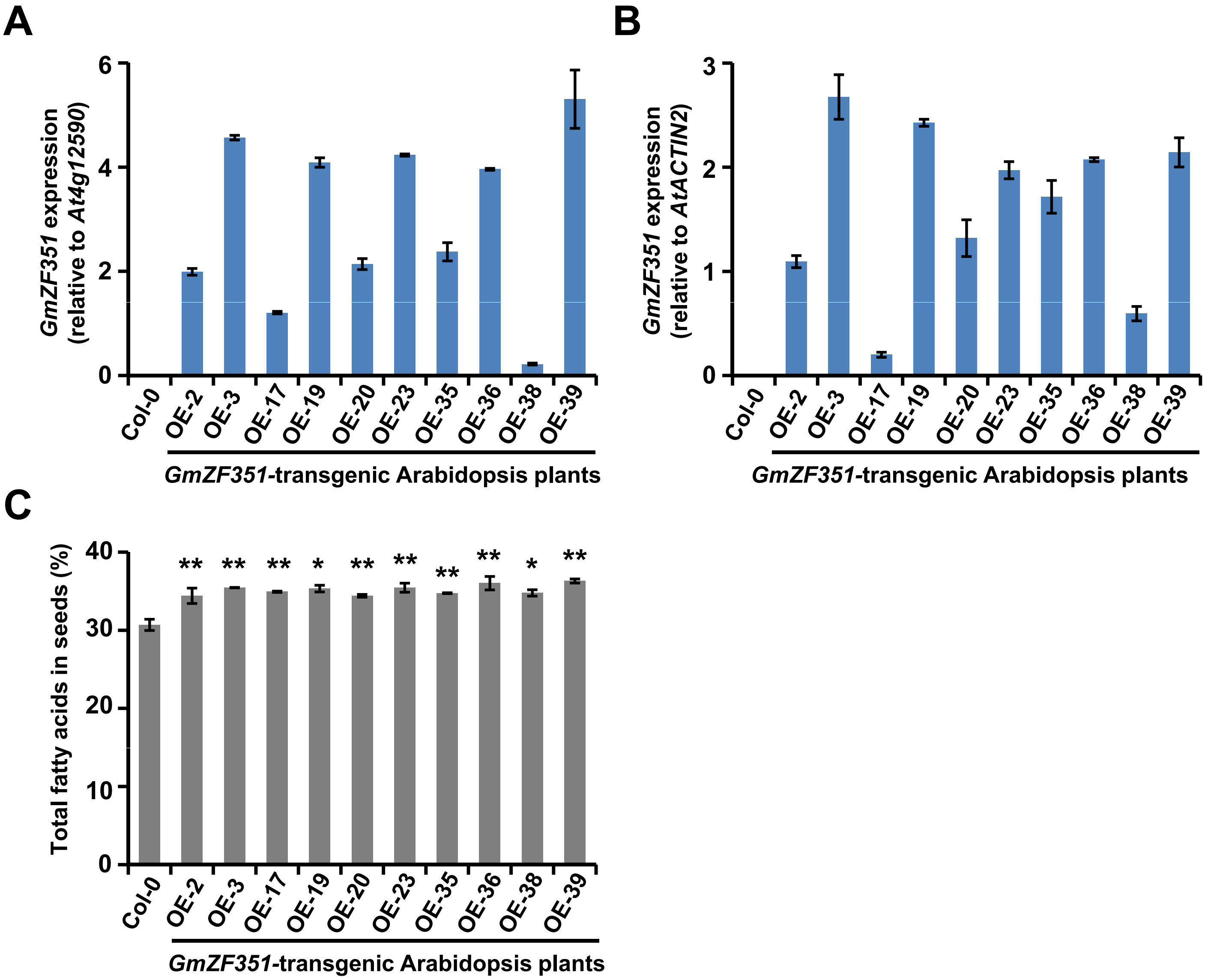
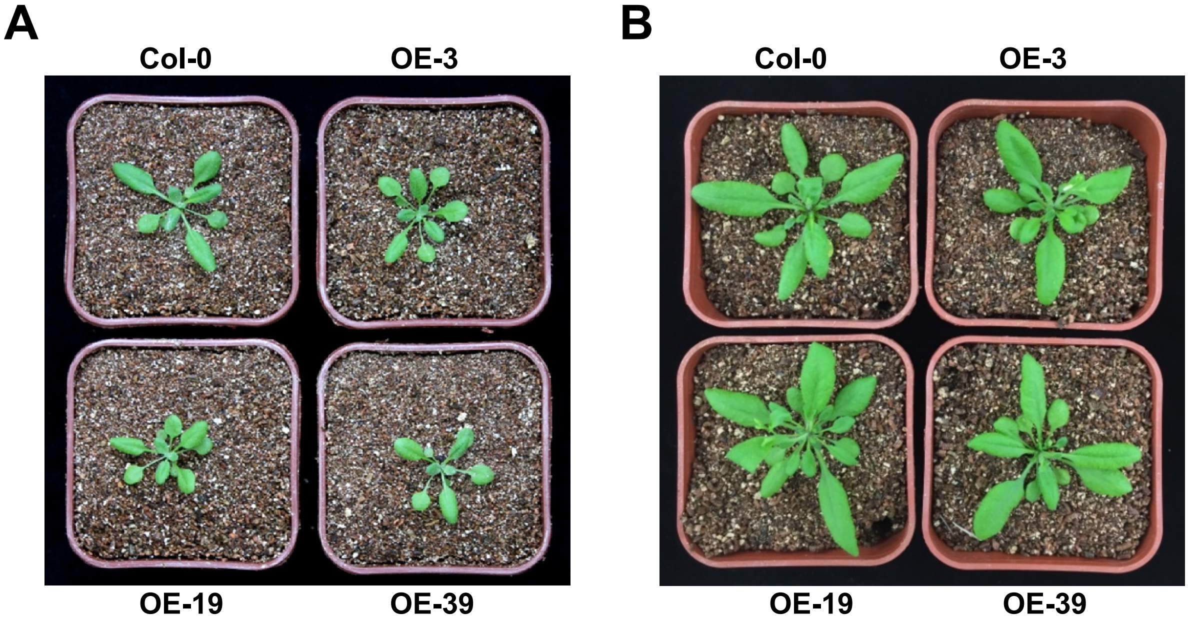


```
GmZF351   .................................MSSVFSEHKFQLQ.....PSHQLLSLKKSLGDIDIPVPPRKLLTRRSAAVHDGSG..    50
AtTZF1    ..................................................MMIGENKNRPHPTIHIPQWDQINDPTATISSPFSSVNLNS    40
SOMNUS    .................................MDVVCTEHQMRKPTVEIPPRKLLLSSKSFPSDSSSPRSPRKHNWNKSNKITSEHE..    55
AtTZF5    .......................................................MDVEHHKSGHISRPTVDIPPRKLLSSAKSPSSV..    33
OsTZF1    MVRKRRDTARVNPTAVSGGGLSGLYSRASSSPPLHHSGSRRRLRTNTLPRRSWRRGEELESKMMMMGEGAHAPPWQQHVASPVSGVEGGG    90
Consensus

GmZF351   .......................DIYLPHSGSTDSSTDDDSDGDPYASDQFRMFEFKVRRCSRSRSHDWTDCPFVHPGEKARRRDPRRFY   117
AtTZF1    VNDYPHSPSPYLDSFASLFRYLPSNELTNDSDSSSGDESSPLTDSFSSDEFRIYEFKIRRCARGRSHDWTECPFAHPGEKARRRDPRKFH   130
SOMNUS    .......................EDNEDNNRENKEYCYDSDSDDPYASDHFRMFEFKIRRCTRSRSHDWTDCPFAHPGEKARRRDPRRFQ   122
AtTZF5    .......................SSPLRDYKEQKDYCYDSDSEDPYAGDHFRMYEFKIRRCTRSRSHDWTDCPFSHPGEKARRRDPRRFH   100
OsTZF1    GRESEVVAAPYHLLDT.LRHYLPSNEAAAAEDEEEAAAVAAAVDAYACDEFRMYEFKVRRCARGRSHDWTECPFAHPGEKARRRDPRRYC   179
Consensus                                            d    d fr  efk rrc r rshdwt cpf hpgekarrrdpr
                   Tandem CCCH zinc finger domain                                  -   --       -
GmZF351   YSGTVCPEFRR.GQCDRGDACEFSHGVFECWLHPSRYRTEACKDGKNCKRKVCFFAHTPRQLRVFHSNDN..SNKK...KCTDI......   195
AtTZF1    YSGTACPEFRK.GSCRRGDSCEFSHGVFECWLHPSRYRTQPCKDGTSCRRRICFFAHTTEQLRVLPCSLDPDLGFF..............   205
SOMNUS    YSGEVCPEFRRGGDCSRGDDCEFAHGVFECWLHPIRYRTEACKDGKHCKRKVCFFAHSPRQLRVLPPENV..............SGVSAS   198
AtTZF5    YTGEVCPEFSRHGDCSRGDECGFAHGVFECWLHPSRYRTEACKDGKHCKRKVCFFAHSPRQLRVLPPSPENHISGG...CGGSPSSSPAS   187
OsTZF1    YSGTACPDFRK.GGCKRGDACEFAHGVFECWLHPARYRTQPCKDGTACRRRVCFFAHTPDQLRVLPPSQQQGSNSPRGCGGGGAGAAASP   268
Consensus y g  cp f   g c rgd c f hgvfecwlhp ryrt  ckdg  c r  cffah   qlrv

GmZF351   .SPHNNNNCCLVCHCSNSTR......SPTSTLFGM.SHFSPP.LSPPSPSSPSM.......FETNNHHHGVVKYNKDVFSELVCSM....   265
AtTZF1    .....................SGLATSPTSILVS....PSFSPPSESPPLSPST....................................   234
SOMNUS    PSPAAKNPCCLFCSS...........SPTSTLLGNLSHLSRS.PSLSPPMSPAN.......KAAAFSRLRNRAASAVSAAAAAGSMN...   266
AtTZF5    VLSNKNNRCCLFCSH...........SPTSTLLN....LSRS.PSSSPPLSPAD.......KADAFSRL...............SRR...   236
OsTZF1    LAESYDGSPLRRQAFESYLTKSIMSSSPTSTLVS....PPRSPPSESPPLSPDAAGALRRGAWAGVGSPVNDVHVSLRQLRLGSPRSAPS   354
Consensus                           spts l            s   p sp

GmZF351   ....................EGLNFDEASSLLSA..ASKPHHHNNLSSWLDVSKDHNQKQFNTLNSPT.......ITACGSFSNNGNGGF   326
AtTZF1    .......GELIASMRKMQLNGGGCSW.........................................................SSPMRSA   260
SOMNUS    ..............YKDVLSELVNSLDSMSLAEALQASSSSPVTTPVSAAAAAFASSCGLSNQRLHLQ...........QQQPSSPLQFA   331
AtTZF5    ..............RTAVLNELISSLDSLSLTEALAASSSSPVTMPIST.ATMIASSNLSSNHHHHRLPPWLDVGDRDLQLQQSSPLRFA   311
OsTZF1    CASFLPAGYQYGSPKSPAAAAAAALY.........................................................SLPSTPT   387
Consensus

GmZF351   ..LRAENGVVVDDVIAPDLAWVNELLM.................................................                351
AtTZF1    VRLPFSSSLRPIQAATWPRIREFEIEEAPAMEFVESGKELRAEMYARLSRENSLG.....................                315
SOMNUS    ..LSPSTPSYLTNS....PQANFFSDDFTPRRRQINDFT....AMTAVRENTNIEDGSCGDPDLGWVNDLLT....                393
AtTZF5    ..LSPSSTPSYLHGQLQPPPSSFFGDEFTPRGGRLSDFSVAAAAAAQARDKNSFEVGSSGDLDLGWVNDLLT....                381
OsTZF1    RLSPVTVTTASGATVTVEPLDLGLIEEEQPMERVESGRALREKVFERLSKEATVSTDAAAAAAGVAPDVGWVSDLI                463
Consensus
```

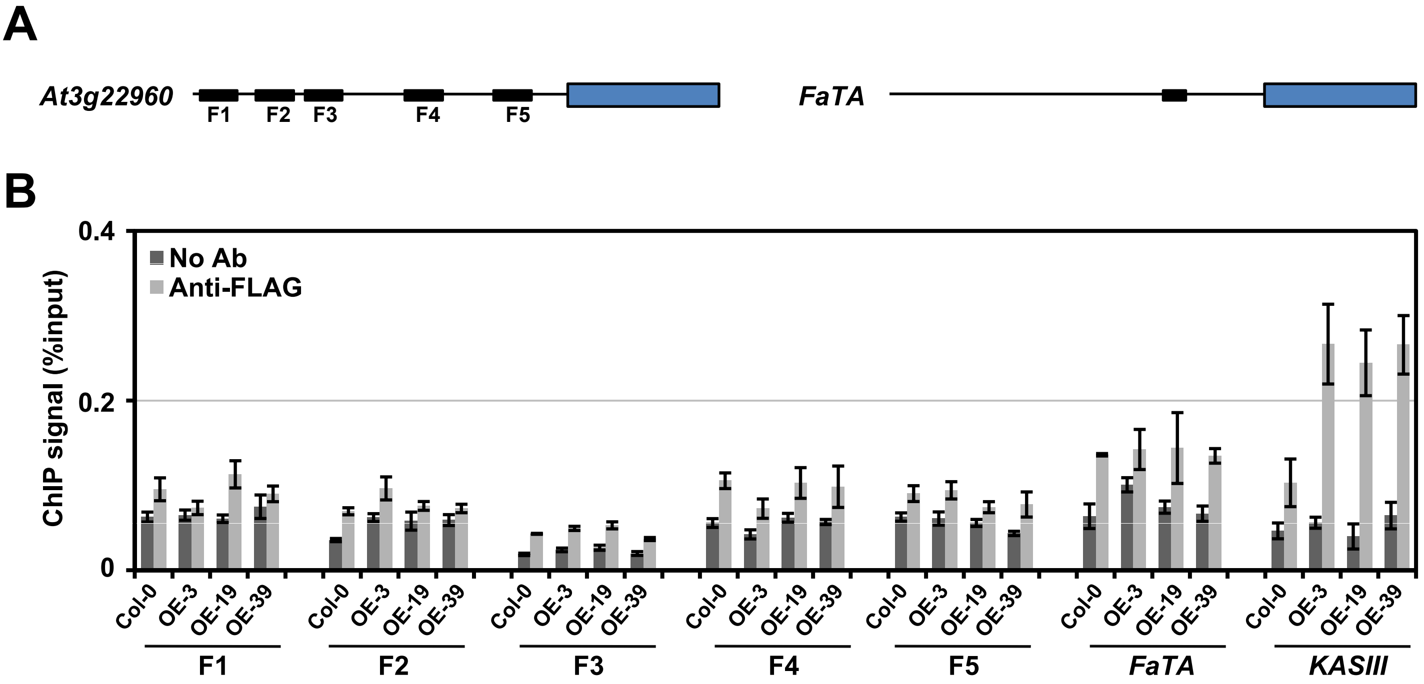
Supplemental Figure S1. Sequence feature and phylogenetic analysis of GmZF351.
(A) Amino acid sequence alignment of GmZF351, AtTZF1, SOMNUS, AtTZF5, and OsTZF1. Characters with black background indicate conserved amino acids. (B) Phylogenetic analysis of GmZF351 and Arabidopsis TZF proteins.



Supplemental Figure S2. Expression level of *GmZF351* and total fatty acids contents of seeds in *GmZF351*-transgenic Arabidopsis plants.
(A) Expression level of *GmZF351* in Col-0 and *GmZF351*-transgenic Arabidopsis plants. *At4g12590* is used as internal control. Error bars indicate SD (n = 3). (B) Expression level of *GmZF351* in Col-0 and *GmZF351*-transgenic Arabidopsis plants. *AtACTIN2* is used as internal control. Error bars indicate SD (n = 3). (C) Total fatty acids contents in seeds of Col-0 and *GmZF351*-transgenic Arabidopsis plants. Asterisks indicate a significant difference compared with Col-0 (*P<0.05 and **P<0.01). Error bars indicate SD (n = 4).

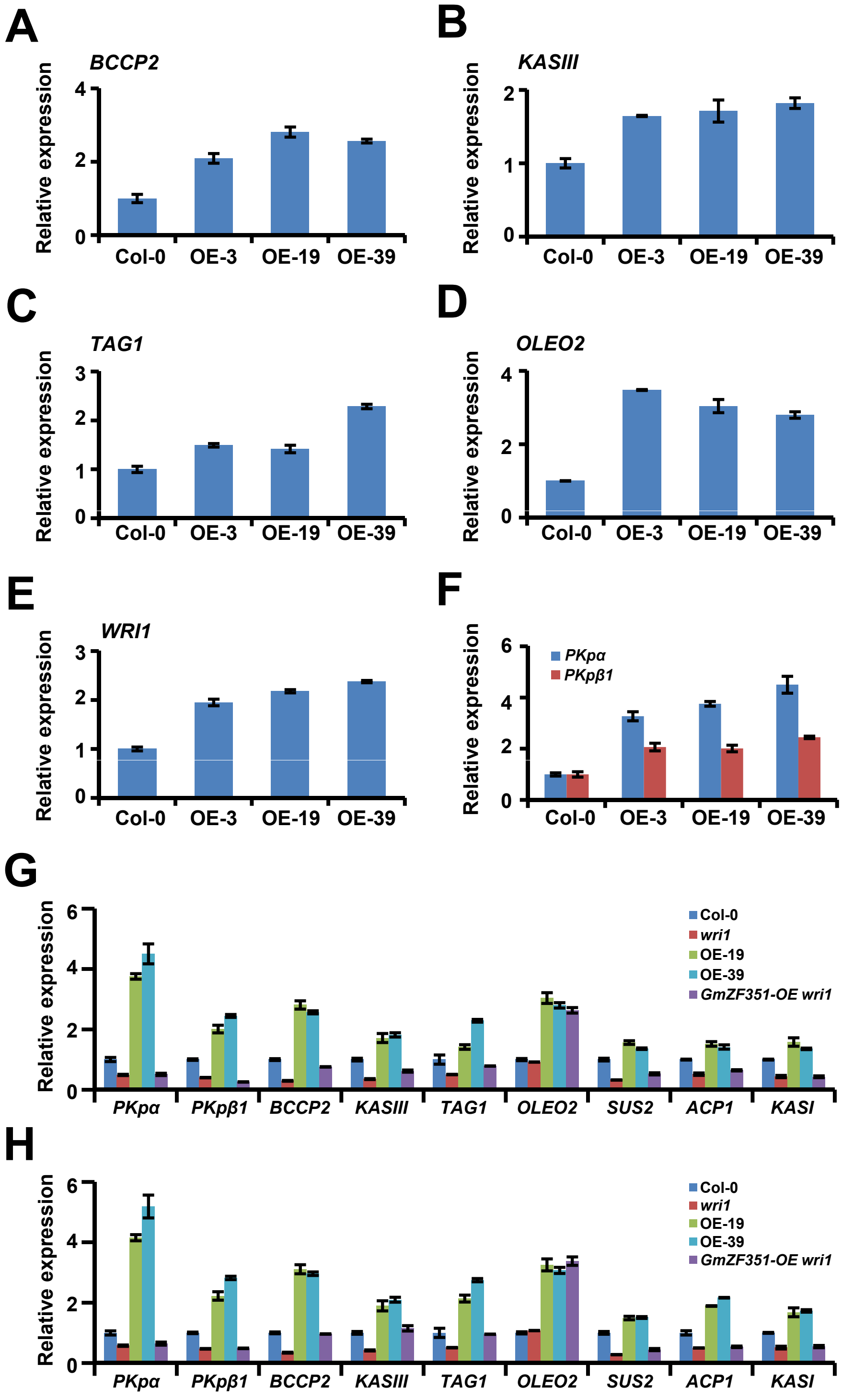


Supplemental Figure S3. Morphology of *GmZF351*-transgenic Arabidopsis.
(A) Rosettes of 3-week-old Col-0 and *GmZF351*-transgenic Arabidopsis plants (OE-3, OE-19 and OE-39).
(B) Rosettes of 5-week-old Col-0 and *GmZF351*-transgenic Arabidopsis plants (OE-3, OE-19 and OE-39).

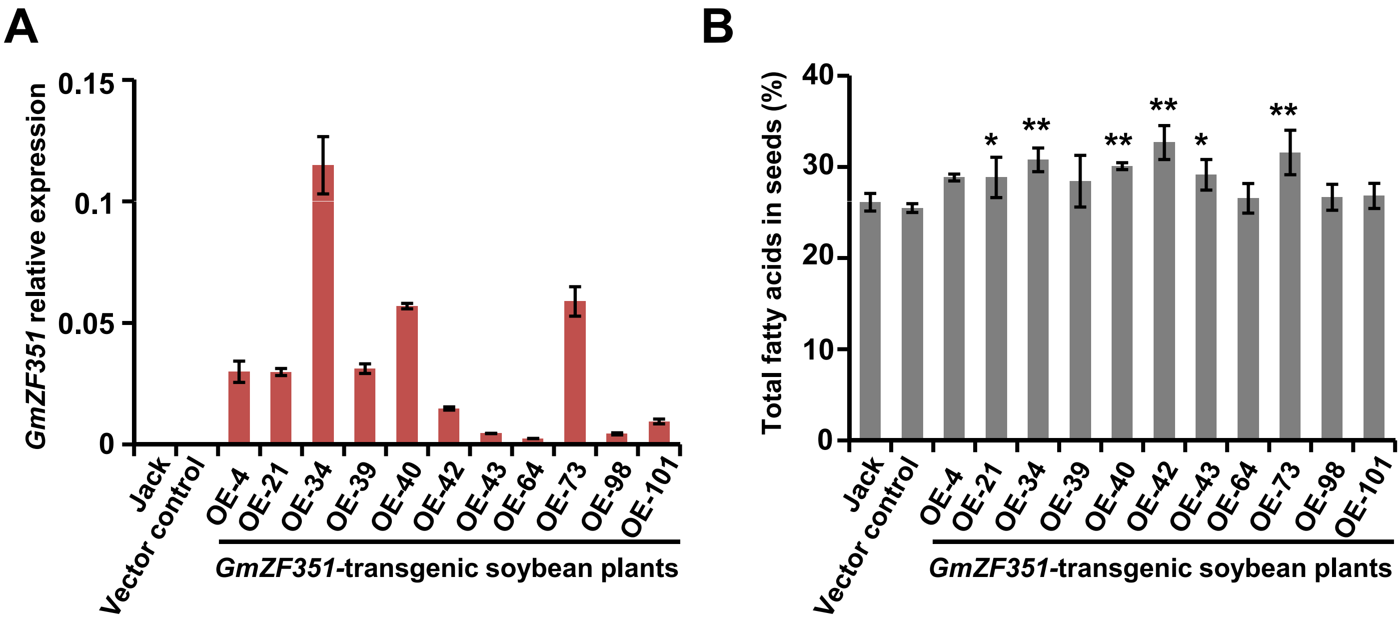


Supplemental Figure S4. GmZF351 ChIP analysis of the *At3g22960* promoter and *FaTA* promoter.

(A) The schematic diagrams depict the promoters of *At3g22960* and *FaTA*. *At3g22960* encoding Pkpα is regulated by *WRI1,* and its promoter is not enriched in GmZF351 ChIP-Seq analysis. The right panel represents various fragments (F1 to F5, black rectangles) used for qPCR in 2 kb promoter region (black line) of *At3g22960.* The left panel represents the location of the enriched interval (black rectangle) identified by ChIP-Seq in 2 kb promoter region (black line) of *FaTA.* The blue rectangles indicate CDS regions of *At3g22960* and *FaTA*. The primer sequences are shown in Supplemental table S4. (B) The graph indicates the percentages of DNA fragments that coimmunoprecipitated with FLAG antibody relative to the input DNAs in Col-0 and *GmZF351*-transgenic Arabidopsis plants (OE-3, OE-19, OE-39). The "no Ab" samples are coimmunoprecipitated without antibody and served as negative control. The *KASIII* promoter fragment identified as a GmZF351 binding site is used as a positive control. Error bars indicate SD (n = 3).

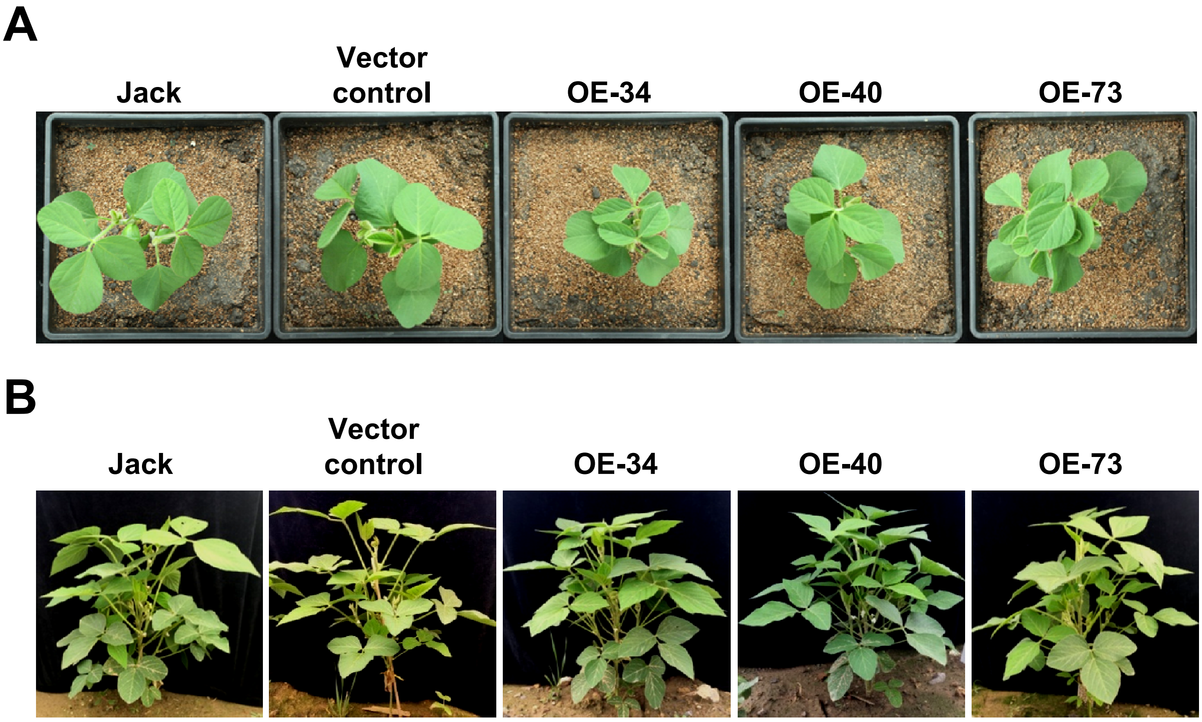


Supplemental Figure S5. Expression levels of *GmZF351* downstream genes in various Arabidopsis plants. (A-F) Expression levels of *BCCP2* (A), *KASIII* (B), *TAG1* (C), *OLEO2* (D), *WRI1* (E), *PKpα* and *PKpβ1* (F) in 5-day-old seedlings of Col-0 and *GmZF351*-transgenic Arabidopsis plants. *AtACTIN2* is used as internal control. Error bars indicate SD (n = 3). (G) Expression levels of *PKpα* , *PKpβ1*, *BCCP2*, *KASIII, TAG1*, *OLEO2*, *SUS2*, *ACP1* and *KASI* in 5-day-old seedlings of Col-0, *wri1*, *GmZF351*-OE (OE-19 and OE-39) and *GmZF351*-OE *wri1* plants. *AtACTIN2* is used as internal control. Error bars indicate SD (n = 3). For all panels, relative mRNA levels of each gene in Col-0 are set to 1. (H) Expression levels of *PKpα* , *PKpβ1*, *BCCP2*, *KASIII, TAG1*, *OLEO2*, *SUS2*, *ACP1* and *KASI* in 5-day-old seedlings of Col-0, *wri1*, *GmZF351*-OE (OE-19 and OE-39) and *GmZF351*-OE *wri1* plants. *At4g12590* is used as internal control. Error bars indicate SD (n = 3). For all panels, relative mRNA levels of each gene in Col-0 are set to 1.

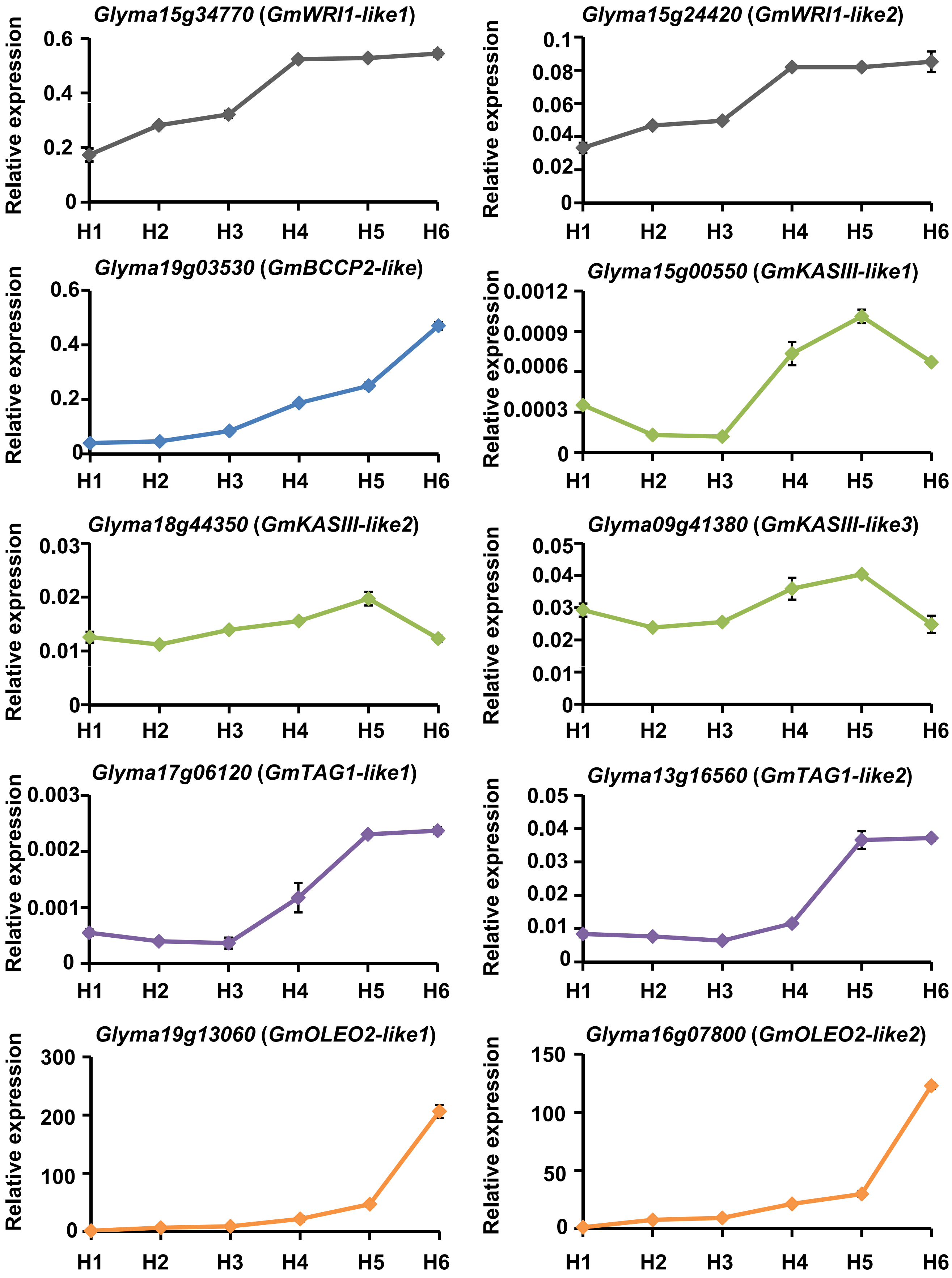


Supplemental Figure S6. Expression level of *GmZF351* and total fatty acids contents of seeds in *GmZF351*-transgenic soybean plants.

(A) Expression level of *GmZF351* in leaves of Jack, vector control and *GmZF351*-transgenic soybean plants. Error bars indicate SD (n = 3). (B) Total fatty acids contents in seeds of Jack, vector control and *GmZF351*-transgenic soybean plants. Asterisks indicate a significant difference compared with Jack (*P<0.05 and **P<0.01). Error bars indicate SD (n = 4).

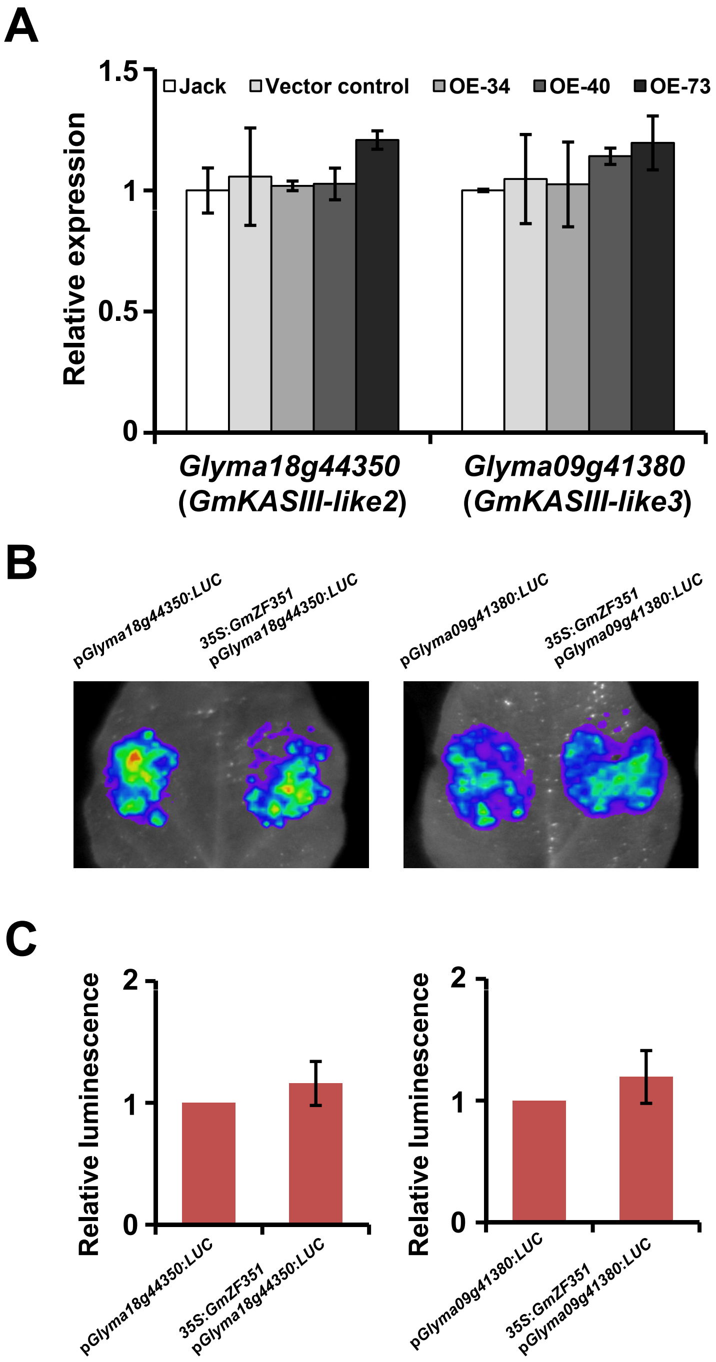


Supplemental Figure S7. Morphology of *GmZF351*-transgenic soybean plants.
(A) Morphology of transgenic soybean plants in V1 stage. Jack, vector control and *GmZF351*-transgenic soybean plants (OE-34, OE-40 and OE-73) are observed. (B) Morphology of transgenic soybean plants in R3 stage. Jack, vector control and *GmZF351*-transgenic soybean plants (OE-34, OE-40 and OE-73) are observed.



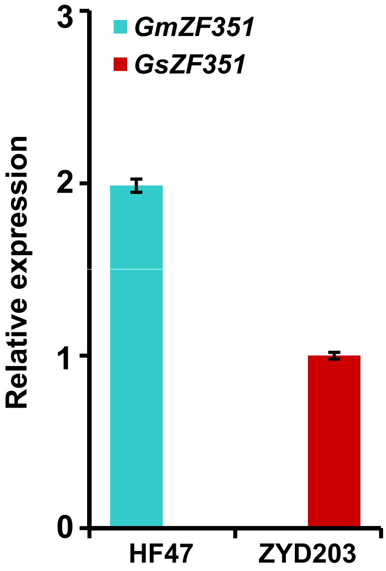
Supplemental Figure S8. Expression pattern of potential downstream genes in six developmental stages of soybean seeds.

The gray lines indicate *WRI1* homologs, the blue line indicates *BCCP2* homolog, the green lines indicate *KASIII* homologs, the purple lines indicate *TAG1* homologs, and the orange lines indicate *OLEO2* homologs. Error bars indicate SD (n = 3).

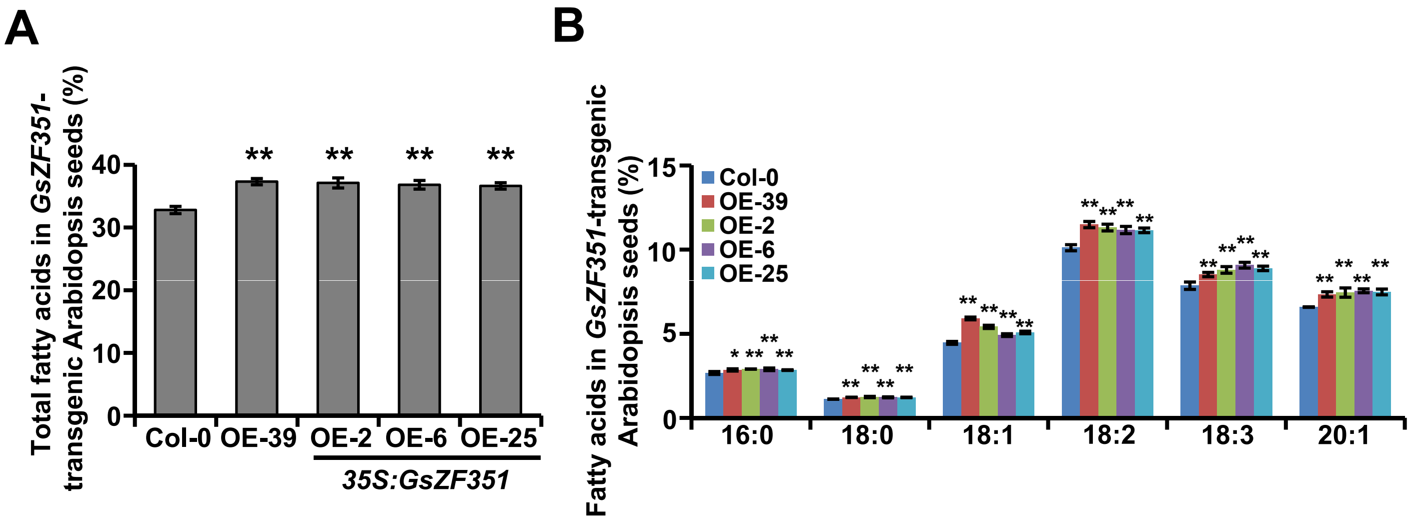


Supplemental Figure S9. Overexpression of *GmZF351* does not influence the expression of *Glyma18g44350* and *Glyma09g41380.*

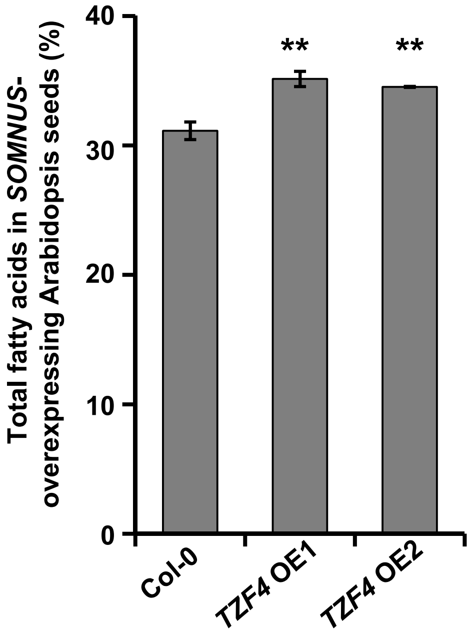
(A) The relative expression of *Glyma18g44350* and *Glyma09g41380* are unchanged in *GmZF351*-transgenic soybean plants compared with Jack and vector control plants. The mRNA levels (relative to *GmTUBULIN*) of each gene in Jack are set to 1. Error bars indicate SD (n = 3). (B) Inclusion of *35S*:*GmZF351* effector can not activate the expression of *Glyma18g44350* and *Glyma09g41380* by transient expression in tobacco leaves. The *A. tumefacien* harboring reporter vector is mixed with *A. tumefacien* harboring *35S*:*GmZF351-FLAG* and infiltrated into tobacco leaves. *A. tumefacien* harboring pGWB412 is used as negative control. (C) Quantitative analysis of luminescence intensity in (B). LUC image is taken 2 days after infiltration. Luminescence intensity is analyzed with IndiGo software. Error bars indicate SD (n = 5).



Supplemental Figure S10. Expression of *GmZF351* and *GsZF351* in seeds of HF47 and ZYD203 as determined by TaqMan RT-qPCR. *GmZF351* or *GsZF351* can not be detected in ZYD203 or HF47, respectively. The expression of *GsZF351* allele in ZYD203 seed is set to 1. *GmTUBULIN* is used as internal control. Error bars indicate SD (n = 3).

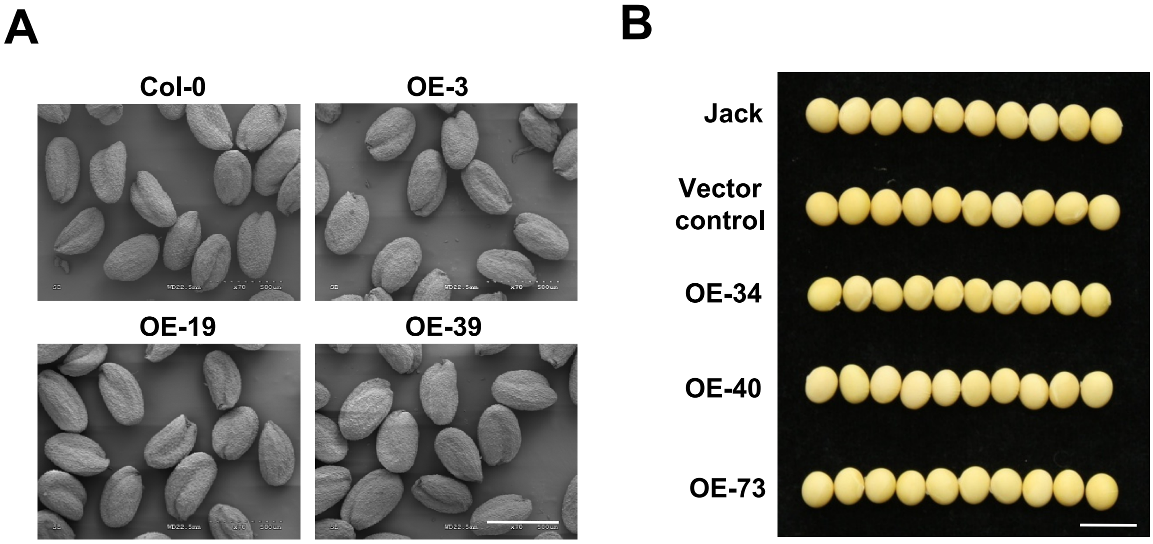


Supplemental Figure S11. Function of *GsZF351* from ZYD203 on seed oil accumulation.
(A) Total fatty acids contents in seeds of *GsZF351*-transgenic Arabidopsis plants (OE-2, OE-6 and OE-25). *GmZF351*-transgenic Arabidopsis plants (OE-39) is used as positive control. Error bars indicate SD (n = 4). (B) Contents of fatty acid compositions in seeds of *GsZF351*-transgenic plants. *GmZF351*-transgenic Arabidopsis plants (OE-39) is used as positive control. Error bars indicate SD (n = 4). For two panels, asterisks indicate a significant difference compared to Col-0 (*P<0.05 and **P<0.01).



Supplemental Figure S12. Overexpression of *SOMNUS*/*TZF4* improves seed oil accumulation in transgenic Arabidopsis seeds.

Total fatty acids contents in seeds of Col-0 and *SOMNUS*-overexpressing Arabidopsis plants (*TZF4* OE1 and *TZF4* OE2) are measured. Error bars indicate SD (n = 4). Asterisk indicates a significant difference compared to Col-0 (**P<0.01).



Supplemental Figure S13. Morphology of *GmZF351*-transgenic Arabidopsis seeds and *GmZF351*-transgenic soybean seeds.

(A) Morphology of seeds from Col-0 and *GmZF351*-transgenic Arabidopsis plants (OE-3, OE-19 and OE-39). Bar = 0.5 mm. (B) Morphology of seeds from Jack, vector control and *GmZF351*-transgenic soybean plants (OE-34, OE-40 and OE-73). 10 seeds are used for photographing. Bar = 1 cm.