Large-scale miRNA-Target Data Analysis to Discover miRNA Co-regulation Network of Abiotic Stress Tolerance in Soybeans

Haowu Chang[1,2#]，Tianyue Zhang[1#]，Hao Zhang[1,2*]，Lingtao Su[2]，Qing-Ming Qin[3]，Guihua Li[3]，Xueqing Li[1]，Li Wang[1]，Tianheng Zhao[1]，Enshuang Zhao[1], Hengyi Zhao[1], Yuanning Liu[1,2]，Gary Stacey[4]，Dong Xu[2*]

[1] College of Computer Science and Technology & Symbol Computation and Knowledge Engineering, Ministry of Education, Jilin University, Jilin, China

[2] Department of Computer Science and Christopher S. Bond Life Sciences Center, University of Missouri, Missouri, USA

[3] College of Plant Sciences and Key Laboratory of Zoonosis Research, Ministry of Education, Jilin University, Jilin, China

[4] Division of Plant Sciences and Technology, Christopher S. Bond Life Sciences Center, University of Missouri, Columbia, MO 65211, USA.

[*] Corresponding author: Hao Zhang zhangh@jlu.edu.cn, Dong Xu xudong@missouri.edu

[#] These authors contributed equally to this work

Abstract

Although growing evidence shows that microRNA (miRNA) regulates plant growth and development, miRNA regulatory networks in plants are not well understood. Current experimental studies cannot characterize miRNA regulatory networks on a large scale. This information gap provides a good opportunity to employ computational methods for global analysis and to generate useful models and hypotheses. To address this opportunity, we collected miRNA-target interactions (MTIs) and used MTIs from Arabidopsis thaliana and Medicago truncatula to predict homologous MTIs in soybeans, resulting in 80,235 soybean MTIs in total. A multi-level iterative bi-clustering method was developed to identify 483 soybean miRNA-target regulatory modules (MTRMs). Furthermore, we collected soybean miRNA expression data and corresponding gene expression data in response to abiotic stresses. By clustering these data, 37 MTRMs related to abiotic stresses were identified including stress-specific MTRMs and shared MTRMs. These MTRMs have gene ontology (GO) enrichment in resistance response, iron transport, positive growth regulation, etc. Our study predicts soybean miRNA-target regulatory modules with high confidence under different stresses, constructs miRNA-GO regulatory networks for MTRMs under different stresses and provides miRNA targeting hypotheses for experimental study. The method can be applied to other biological processes and other plants to elucidate miRNA co-regulation mechanisms.

Introduction

The growth and development of crops are often restricted due to various environmental stresses, which can lead to poor harvests and yields below their genetic potential [1, 2]. In the past decade, microRNAs (miRNAs) have been identified as important gene expression regulatory factors that play an essential role in plant growth

49   and development [3]. miRNA can target multiple genes, and multiple miRNAs can also

50   target the same gene. miRNAs are involved in the expression of stress-responsive genes

51   and the plant's ability to adapt to environmental change [4]. Different stresses can induce

52   differential expressions of corresponding miRNAs in plants, while some miRNAs can

53   respond to several abiotic stresses simultaneously [5-7]. Therefore, studying the

54   cooperative relationship among miRNAs and the interactions with their target genes is

55   important for understanding the role of miRNAs in controlling plant growth and

56   development.

57   miRNAs may respond to adverse effects on plant growth and development, such as

58   drought, salinity, temperature and other abiotic environmental factors. It was shown that

59   willow leaves exposed to drought or high temperature induce differential expressions of

60   some miRNAs [8]. For example, miR169c plays a negative regulatory role under drought

61   stress by inhibiting the expression of its target gene nuclear factor Y-A (NF-YA) [9].

62   miR172a [10] and miR172c [11] endow plants with a tolerance to salt stress and water

63   deficiency. Meanwhile, miRNAs also indirectly respond to abiotic stress by regulating

64   other biological macromolecules. For example, miR398c can negatively regulate multiple

65   peroxisome-related genes (GmCSD1a/b, GmCSD2a/b/c and GmCCS) and affect the

66   drought tolerance of the soybean [12]. miR399a/b participates in developing resistance to

67   salt stress [13]. miR5036, miR862a and miR398a/b respond to phosphorus starvation and

68   salt stress [14]. Furthermore, miR166k/o, miR390g, and miR396c/k mediate BX10

69   (Al-tolerant genotype) root elongation, and miR169r triggers the BD2 (Al-sensitive B

70   genotype) oxidative stress, which in turn triggers a different type of plant aluminum

71   tolerance between BX10 and BD2 [15]. This indicates that miRNA may regulate plant

72   growth under abiotic stress through a complex network. However, current studies

73   typically explore the role of a single miRNA in response to abiotic stresses. How multiple

74   miRNAs work together as a co-regulatory mechanism has not been significantly

75   explored.

76   Several studies have uncovered interesting miRNA interactions. For example, it was

4

77 shown that miRNA lin-4 and let-7 act cooperatively in Drosophila [16], while miR375,

78 miR124 and let7b coordinately regulate Mtpn expression in mammals [17]. Similarly, the

79 miRNAs, miR125a, miR125b and miR20, work together to down-regulate erbB2/erbB3

80 in breast cancer cells [18,19]. Furthermore, compared with individual miRNAs alone, the

81 combination of miR16, miR34a and miR106b results in a stronger immortalized human

82 mammary epithelial cell cycle arrest [20]. It was also found that miR21, miR23a and

83 miR27a are three cooperative repressors in a network of tumor suppressors [21]. The

84 same combination effects also exist in plants. For example, miR160 and miR167 are

85 involved in the adventitious root program of Arabidopsis [23]. miR156 and miR172 play

86 a role in the transition of soybean nutrition [24], and miR172 also plays a role in

87 flowering by inhibiting AP2 [25]. Transgenic studies of miR482, miR1512 and miR1515

88 showed that their over-expression may lead to a substantial increase in the number of

89 soybean nodules [26]. Another study verified networks of 365 tissue-specific

90 miRNA-target interactions (MTIs) [27]. The synergistic effects of miRNAs provide a

91 new systematic perspective for the entire microRNome [23], which calls for a global

92 analysis of miRNA-target interactions.

93 With the growing miRNA-target data, several methods have been developed and

94 applied to explore this field. Shalgi et al. first constructed a miRNA network from the

95 target genes predicted by PicTar and TargetScan [28]. Xu et al. constructed a human

96 miRNA-miRNA functional synergy network through co-regulation functional modules

97 [29]. Meanwhile, biclustering was also applied for two different types of objects (gene

98 and miRNA in this case) belonging to the same cluster. In the past two decades, various

99 bi-clustering methods have been developed [30,44-46]. SAMBA [31], ISA [32], BIMAX

100 [33], QUBIC [34] and FABIA [35] are some commonly used general algorithms.

101 Contiguous column coherent (CCC) biclustering [36-41] and LateBiccluster [39] are

102 designed for temporal data analysis. BicPAM [40, 41], BicNET [42] and MCbiclust [43]

103   are the latest tools. Gianvito et al. applied the biclustering algorithm to predict human

104   miRNA-mRNA modules [47]. The application of biclustering algorithms and

105   miRNA-target regulation module (MTRM) mining is feasible and important for

106   analyzing miRNA regulation mechanisms. Compared with traditional clustering methods,

107   such as Bimax [103] and BiBit [69], CUBiBit [57] shortened the computing time and

108   provided an optimized method for finding modules in larger data. However, the result

109   obtained by CUBiBit was mostly a fully-connected bipartite graph, and the relationship

110   between miRNA and the target gene is complex and interactive.

111   In this study, we analyzed the relationship between the miRNA regulatory modules

112   in response to abiotic stresses in the soybean as a means for extending our understanding

113   of soybean resistance mechanisms. Previously, Xu et al. provided a soybean

114   miRNA-gene network, SoyFN, based on predicted miRNA targets [49]. However, this

115   work was based only on sequence comparisons, which may result in a high false

116   discovery rate. In contrast, in our work, we collected experimentally proven

117   miRNA-target relationships based on degradome sequencing in the soybean, as well as

118   the stringent homologs of miRNA-target pairs in *Arabidopsis thaliana* and *Medicago*

119   *truncatula*. Based on these reliable miRNA-target data, we performed a biclustering

120   analysis, and iteratively fused the overlapping biclusters based on the SoyNet network to

121   obtain the soybean miRNA-target regulatory modules in response to abiotic stresses. We

122   provide soybean MTRMs with high confidence relevant to various stresses and present

123   the miRNA-GO regulatory networks of these modules. Capturing these miRNA-target

124   modules with biological significance expands our understanding of the complex

125   regulatory mechanisms of miRNA. The methods used should be easily applicable to other

126   plant and animal systems where sufficient data exists to perform the analyses.

127

128

Results

130 We obtained 90,064 confirmed soybean MTIs based on multiple experimental data
131 sources and 1,189 potential soybean MTIs based on homology to experimental data from
132 *Arabidopsis thaliana* and *Medicago truncatula*. A multi-level iterative bi-clustering
133 analysis resulted in 483 soybean miRNA-target regulatory modules. We analyzed the
134 enrichment results of the top five biclusters. In addition, we identified 37 abiotic
135 stress-related modules and predicted the underlying miRNA regulatory pathway
136 networks.

137

Identification of MTIs

139 We collected soybean miRNA-target data based on databases and related
140 publications. First, we gathered all the soybean miRNA-target interactions (MTIs)
141 verified by degradome sequencing and biological experiments by mining published data,
142 including 14,958 pairs of MTIs from DPMIND [50], 86,427 pairs of MTIs from starBase
143 [51], and 38 pairs of MTIs from TarBase [52]. Additional data came from fourteen
144 publications [70-83], as shown in Supplemental Table 1 (Supplemental Table S1). A total
145 of 111,650 pairs of soybean MTIs were obtained. After removing 21,586 redundant pairs
146 of MTIs, 90,064 pairs remained.

147 To expand MTIs, we predicted the target relationship between potential miRNAs and
148 targeting genes from the MTIs of *Arabidopsis thaliana* and *Medicago truncatula* based
149 on homology. For Arabidopsis MTIs, 43 pairs were collected from TarBase, 66,212 pairs
150 from starBase, 4682 pairs from DPMIND, 106 pairs from miRTarBase, and 369 pairs
151 from 5 related publications [84-88], resulting in a total of 71,412 pairs of Arabidopsis
152 MTIs, in which 12,094 MTIs became unique after removing 59,318 pairs of redundant

153 MTIs. For Medicago MTIs, there were 302 pairs from TarBase, 22,010 pairs from

154 starBase, 781 pairs from DPMIND, and 1,349 pairs from three publications [89-91],

155 resulting in a total of 24,442 pairs of MTIs, in which 4,394 MTIs became unique after

156 removing 20,048 pairs of redundant MTIs. There are 33 miRNAs with identical

157 sequences between Arabidopsis and the soybean, covering 3,415 targeting genes in

158 Arabidopsis and 4,190 genes in the soybean. There are also 49 miRNAs with identical

159 sequences between Medicago and the soybean, covering 1,180 targeting genes in

160 Medicago and 5,096 genes in the soybean. We removed any redundant MTIs resulting

161 from these. We further validated homology-based MTIs using three miRNA target

162 prediction tools that performed well in general plants, i.e., psRNAtarget [53, 54], TAPIR

163 [55], and Targetfinder [56]. The relationship between a specific miRNA and its target was

164 validated by at least one of the prediction tools. In the Arabidopsis derived MTIs, 943

165 pairs of MTIs were validated in the psRNAtarget, 152 pairs in TAPIR, and 158 pairs in

166 Targetfinder, confirming a total of 961 unique pairs of MTIs. In the Medicago-derived

167 MTIs, 968 pairs of MTIs were validated in the psRNAtarget, 144 in TAPIR, and 134 in

168 Targetfinder, confirming a total of 986 unique pairs of MTIs, as shown in Supplemental

169 Figure S1. There is a high overlap between the two sets of MTIs (Supplemental Table S2).

170 After removing the redundant ones, a total of 1,189 pairs were used to expand soybean

171 MTIs.

172

173 **miRNA-target regulatory modules**

174 We integrated the 90,064 soybean MTIs with the 1,189 MTIs based on homology.

175 We removed MTIs involving genes that do not have the glyma2 ID. A total of 11,018

176 MTIs were removed, and the remaining 80,235 MTIs were used for analysis in the

177 following tasks.

8

178   We applied CUBiBit [57] for bi-clustering analysis, with the smallest scale 2×2 or

179   6×2 for miRNA-target modules (i.e., at least two or six target genes and at least two

180   miRNAs in each module), resulting in 15,380 (2×2) miRNA-target modules or 2,461

181   (6×2) miRNA-target modules. We contracted the overlapping modules using a multi-level

182   iterative fusion method based on the soybean gene relationship network (See Methods),

183   yielding 6,577 (2×2) and 812 (6×2) soybean miRNA-target regulatory modules after

184   removing the modules that were completely included in the preliminary clustering

185   module.

186   We next merged MTRMs according to the set threshold until the level converged

187   stably (level represents the number of iterations). The iterative fusion of each level is

188   shown in Figure 1. We compared the iterative results at different scales. Soybean

189   MTRMs at the 2×2 scale showed better results at level 10, which contains 2,715 MTRMs.

190   Soybean MTRMs at the 6×2 scale showed a better effect at level 7, which contains 483

191   MTRMs. Comparing the cluster score based on the GO calculation between the two

192   scales of stable convergence (Figure 1c) shows that the cluster score quality at the 6×2

193   scale is higher than that at the 2×2 level (Supplemental Table S3). Hence, we used the

194   GO enrichment analysis result on 483 soybean MTRMs obtained at the 6×2 level 7.

195   To compare the MTRMs before and after fusion, we extracted an MTRM bicluster,

196   as shown in Figure 1(e) from the level 7 clustering results of the 6×2 scale and plotted it

197   with the corresponding MTRMs under level 1 before the fusion, as shown in Figure 1(e)

198   and after Figure 1(d), which is a level-7 fusion. The module (1534) is at level 1 before the

199   fusion has 2 miRNAs and 22 targeted genes. At level 7, the module (1534) fused an

200   additional three modules 1539, 622, and 1537, and each contains miR396. From the

201   perspective of targeting, the module at level 7 has more miRNA-target interactions than
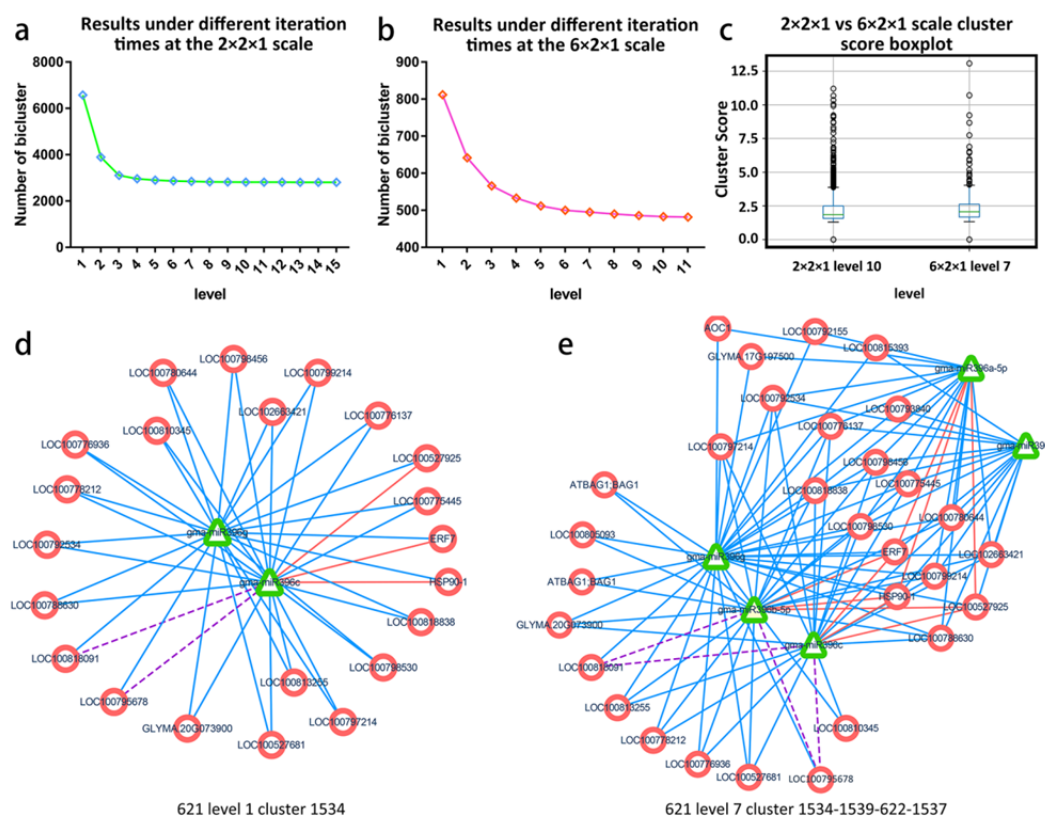
202   the one at level 1.

203

Figure 1. Multi-level iterative biclustering results of soybean MTRMs. (a) Results under different iteration times at the 2×2 scale, (b) results under different iteration times at the 6×2 scale, and (c) the boxplot of the cluster score is calculated based on the gene ontology (GO) under the two scales when converging to a stable level, where based on the overall distribution, the results at the 6×2 scale are better; (d) shows the MTRM bicluster at level 1 before the 6×2 scale fusion, and (e) shows the corresponding MTRM bicluster at level 7 after the 6×2 scale fusion.

Gene ontology (GO) analysis of MTRMs

We screened 254 GO pathways whose GO biological processes (BP) satisfied the p-value < 0.00001 for the GO enrichment and obtained 483 soybean MTRMs at the 6×2 scale at level 7. We analyzed the relationship among the enriched GO terms through

10

217   REVIGO [122] with a parameter of 0.5. These GO pathways have a specific aggregation

218   (Figure 2a). MTRMs obtained from a global perspective have several concentrated

219   distributions of GO functions, such as cellular processes, primary metabolism, and cell

220   adhesion, as well as hormone response and negative regulation of biological processes. In

221   addition, there are metachronous positive growth regulations and chalcone biosynthesis.

222   Chalcone plays an important role in soybeans and is involved in the multi-branch

223   pathway of flavonoids and isoflavone biosynthesis [92]. The enrichment results mainly

224   involve positive regulation of development, heterochronic, chalcone biosynthesis,

225   defense response, mitochondrial mRNA modification, sulfate transport, plant-type

226   primary cell wall biogenesis, and cofactor biosynthesis, as shown in Figure 2(b).

227   In addition, we extracted the enrichment results of the top biclusters in terms of

228   cluster score among the 483 MTRMs and selected the top five GO terms of each module,

229   as shown in Figure 2(c) and Supplemental Table S4).

230

231   **Abiotic stress-related modules**

232   To explore the biological significance of soybean MTRMs, we collected related

233   soybean miRNAs, which revealed 10 types of abiotic stresses, involving 1) drought, 2)

234   salt, 3) cold, 4) nodulation, 5) Pi, 6) oil, 7) rust, 8) soybean isoflavone, 9) pollen

235   development, and 10) phosphorus deficiency based on publications [22, 48, 105-121].

236   The function annotations of these miRNAs are shown in Supplemental Dataset S1. In

237   most processes, soybean miRNA responses are involved in multiple abiotic responses, as

238   shown in Figure 3(a). Therefore, mining the potential cooperative regulatory modules of

239   these miRNAs is important to understand their role in modulating soybean stress

240   responses.

10
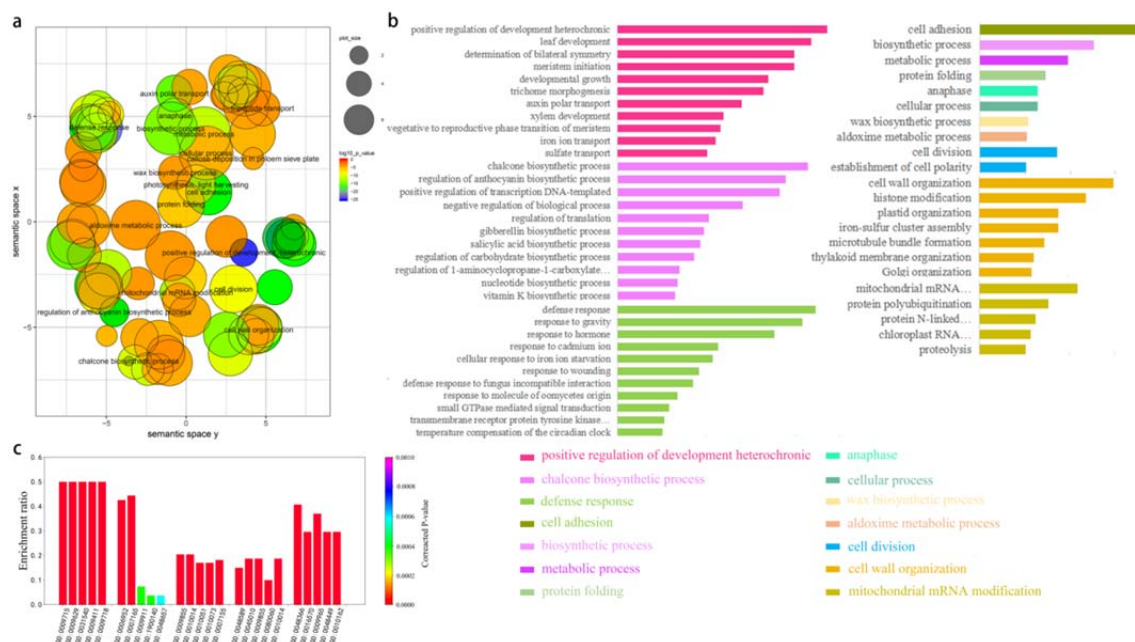
Figure 2. GO analysis of soybean MTRMs. (a) Semantic relevance of GO terms wherein the GO pathway has a certain concentration. (b) GO annotation enriched with 483 soybean MTRMs with enrichment results, which mainly involve positive regulation of development, heterochronic, chalcone biosynthesis, defense responses, and mitochondrial mRNA modification, and (c) GO enrichment of the top five soybean MTRMs. The listed GO terms were enriched with significant p-values < 0.00001
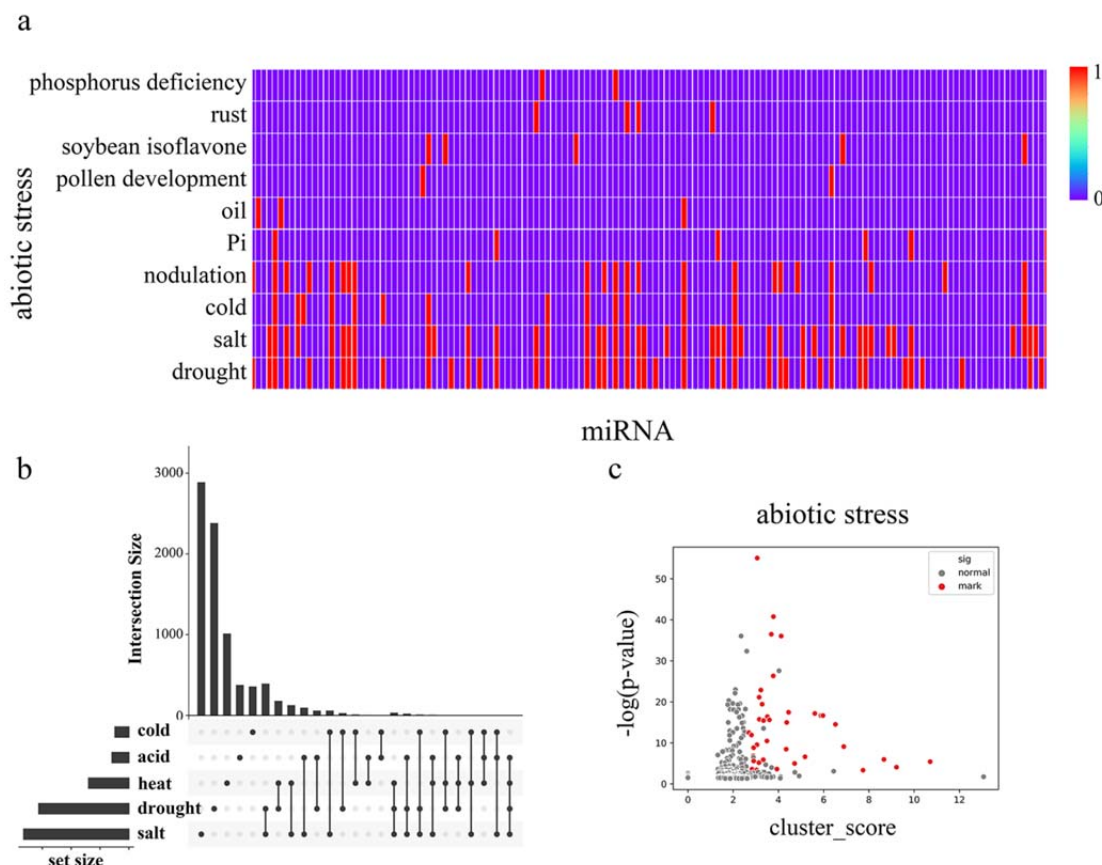
12

Figure 3. Collected miRNA data on soybeans involved in various abiotic stress responses based on the data statistics from the literature. (a) The distribution of stress types in miRNAs where each vertical line represents one miRNA, and red is marked as relevant, (b) UpSet diagram [123] of modular genes under various abiotic stresses within the horizontal correspondence, where dots are used to refer to the corresponding cold stress, acid stress, heat stress, drought stress, and salt stress on the left. The point-to-point connection is realized longitudinally to indicate the intersection between the corresponding data sets, and the upper bar graph indicates the number of genes in the intersection. In (c), the differentially expressed genes in each MTRM under abiotic stress are shown after screening. We used three indicators to filter the candidate clusters. According to the p-value, the related miRNA purity, and the cluster score of each MTRM gene are placed under the corresponding stress. We selected the corresponding threshold

261   and obtained the stress-related MTRMs with higher reliability and marked them as red

262   dots in Figure 3(c). Supplemental Figure S2 shows MTRMs under other types of stress.

263

264   We correlated the 483 soybean MTRMs obtained by clustering with the functional

265   annotations. We selected miRNAs that responded to drought resistance, salt resistance,

266   heat stress, cold stress, acid stress, and performed a statistical analysis on the miRNAs in

267   each of the 483 biclusters. We collected data on the differential expression of soybean

268   genes in the MTRMs under drought, salt, low temperature, cold, and acid stress. Among

269   them, the drought stress data are from GSE76636, the cold stress data are from

270   GSE117686, the acid stress data are from GSE75575, and the salt stress expression data

271   are from the data under 0-h and 12-h treatments of salt stress in GSE57252

272   (Supplemental Dataset S2). The conditions for screening differentially expressed genes

273   are $log2FC > 1$, $p < 0.05$. We obtained 2,145 differentially expressed genes under

274   soybean drought and 1,752 differentially expressed genes under salt treatment. Figure 3(b)

275   shows the genes in the module together with an abiotic stress diagram. At the same time,

276   we calculated the p-values and FDR. We used the Benjamin Graham formula to correct

277   the p-value of the genes in each MTRM for the differentially expressed genes under

278   abiotic stress scenarios through the hypergeometric distribution, as shown in Figure 3(c).

279   Subsequently, we screened MTRMs related to abiotic stress, drought, and salt stress

280   according to the p-value of differentially expressed genes corresponding to the stress in

281   the MTRMs ($p < 0.001$, single adversity 0.01), the proportion of the corresponding

282   miRNA family function (miR function ratio), and the cluster score (cluster score >

283   median). The screening results are shown in Supplemental Dataset S2. We obtained 37

284   MTRMs related to abiotic stress, including 34 MTRMs related to drought stress, 27

285   MTRMs related to salt stress, 3 MTRMs related to cold stress, and 21 MTRMs related to

286   heat stress. Figure 4(a) shows the set relationship of MTRMs involved in a variety of

14

287 stresses. The data suggest that soybean miRNAs have basic and universal functional

288 modules in their response mechanisms to drought, high salt, high temperature, low

289 temperature, and other abiotic stresses. There are two shared modules (31 and 493),

290 involving 6 miRNAs and 11 miRNAs, respectively. The 6 miRNAs of module 31 belong

291 to the miR156 family. The regulated gene-enriched GO pathway is a transcription

292 regulation, DNA-dependent (p-value 4.24 e-10), and a vegetative phase change

293 regulation with a p-value of 9.24 e-07. The 11 miRNAs of module 493 are mainly in the

294 miR172 family, in addition to miR156, miR1533, miR4374, miR5782, miR3939. The

295 regulated gene-enriched GO pathway involves an oxidation-reduction process (p-value =

296 4.66 e-12) and a root hair elongation (p-value = 1.63e-08). Moreover, we also found

297 stress-specific regulatory modules in our results, including 14 drought-specific MTRMs,

298 seven salt-specific MTRMs, and two heat-specific MTRMs (Supplemental Dataset S3).

299 The functions of related miRNA regulatory modules under abiotic stress are mainly

300 concentrated in positive regulation of developmental heterochrony, defense responses,

301 cell wall organization, and other biological processes, as shown in Figure (4b). The data

302 of the top-5 modules are shown in Supplemental Table S5.

303

304 miRNA regulatory pathway network under abiotic stress

305 We explored the regulatory pathway network corresponding to the miRNA of the

306 miRNA-target regulatory module in soybean abiotic stress and analyzed the GO terms of

307 MTRM genes under various abiotic stresses. Stringent screening conditions were used,

308 i.e., the p-value of MTRM stress is 0.001, and the GO BP pathway was selected with a

309 p-value of less than 10^{-5}. The REVIGO-based GO language correlation analysis is shown

310 in Figure 5. GO channels with similar functions are closer in distance in the figure. The

311 small RNAs targeting a certain cluster of GO are functionally close, and the 37 MTRMs

312 of soybeans in abiotic stresses identified in this study mainly focus on resistance response,

313   iron transport, positive growth regulation, and cell wall organization. Under abiotic stress,

314   the cooperating miRNA regulatory modules of the soybean mainly regulate these

315   pathways to respond to the stress environment. The correlation analysis between drought

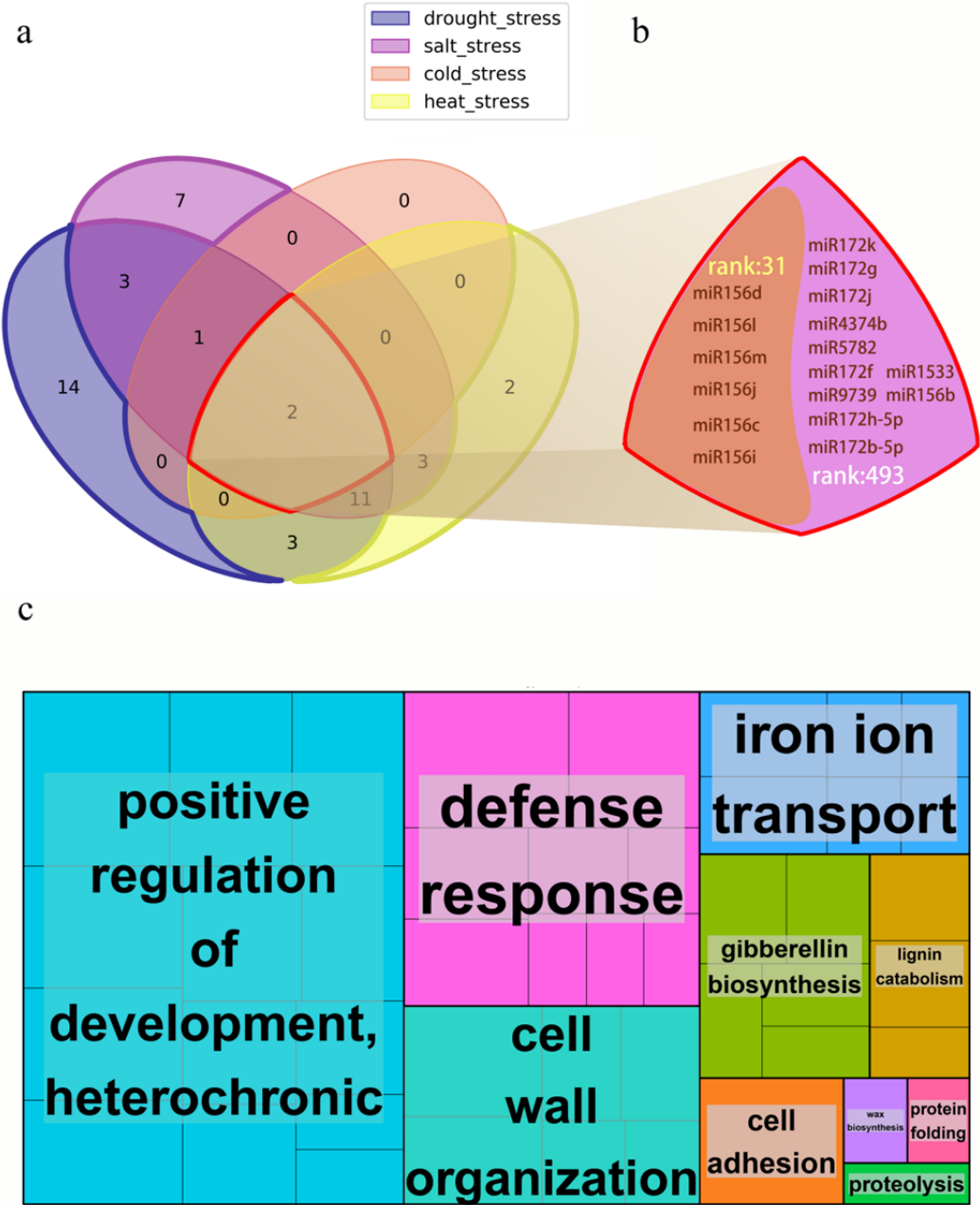316   stress and salt stress with specificity is shown in Figure 5 (b) and (c).

317

318   Figure 4. Soybean MTRMs under various abiotic stresses (a) Venn diagram of 37 kinds

319 of soybean MTRMs under various abiotic stresses, including 14 drought-specific

320 MTRMs, seven salt-specific MTRMs, two heat-specific MTRMs, and two shared

321 MTRMs, (b) the miRNAs in the two shared MTRMs 31 and 493, and (c) a GO Treemap

322 of 37 MTRMs under abiotic stress.

323

324 Subsequently, we constructed the GO BP regulatory network of cooperative miRNAs

325 under soybean abiotic stress for the above main regulatory GO categories and miRNAs,

326 as shown in Figure 5(d). Gene expressions related to abiotic stress responses are mainly

327 regulated by multi-component miRNA families. For example, the miR167 family

328 regulates the resistance response pathway; the miR171 family regulates the gibberellin

329 biosynthesis pathway, while the miR395 family participates in regulating the iron uptake.

330 Moreover, some miRNAs have multiple GO functional partitions, such as miR156b,

331 which regulates developmental growth, the timing of developmental events, the response

332 to hormones, and the response to heavy metal cadmium. The miRNA families and

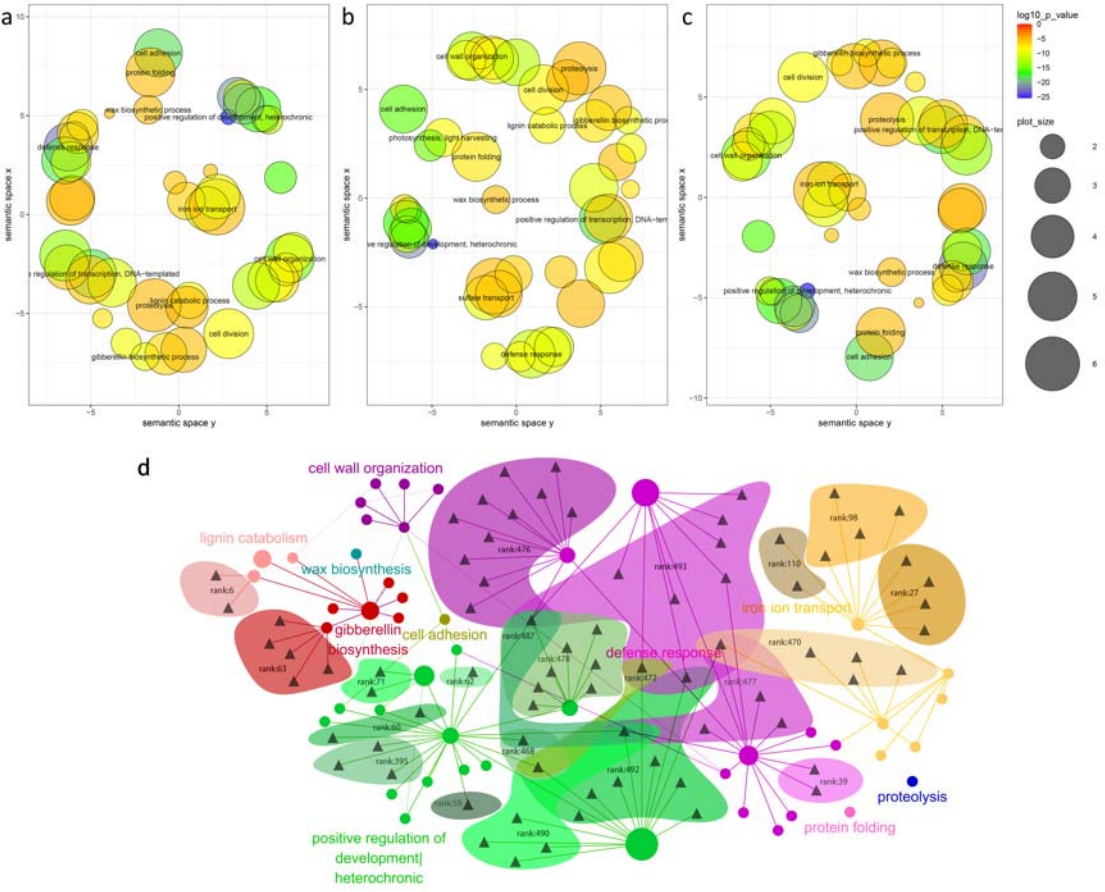333 regulatory pathways involved in MTRM are detailed in Supplemental Dataset S4.

Figure 5. GO term analysis of MTRM genes under various abiotic stresses (a-c). (a) GO semantic correlation analysis of abiotic stress, (b) drought stress, and (c) salt stress, and (d) the GO BP regulatory network of cooperative miRNAs under abiotic stresses. Triangles represent different miRNAs and circles represent different GOs. The size of the circle is determined by the number of genes contained in the GO in this article. The color of the circle depends on the representative GO. The areas with different colors show the modules obtained by our method.

Discussions

miRNAs are major regulators of plant growth and development. They can also

18

345 regulate environmental responses [58-67]. Hence, the study of the role of miRNAs is

346 crucial—not only to understanding the basic events of plant biology but to improve

347 breeding for higher yields and more resilient crop plants. While a variety of papers have

348 noted the role of one or a few miRNAs in regulating plant stress responses, a global

349 analysis of the cooperative interactions is lacking. To study miRNA regulation in

350 response to abiotic response in the soybean, we collected a large number of soybean

351 MTIs. In addition, we collected MTIs from the model plant *Arabidopsis thaliana* and the

352 legume *Medicago truncatula* and used them to predict potential targeting relationships in

353 soybeans through homologous gene analysis. Meanwhile, psRNAtarget, TAPIR, and

354 Targetfinder were used to support the targeting relationships. To ensure the reliability of

355 the data, we mainly chose drought stress, salt stress, cold stress, and heat stress.

356 In the process of bi-clustering soybean MTIs, CUBiBit obtained many MTRMs

357 based on targeting relationships between miRNAs and their target genes. To merge

358 related MTRMs, we proposed a multi-level iterative fusion method of soybean MTRMs

359 based on soybean gene networks. The method determines whether the miRNAs of the

360 two modules have intersections, and the distribution of their genes in the soybean gene

361 network determines whether they can be fused into a unified class.

362 We mined 483 soybean MTRMs, which provide a data reference for the analysis of

363 the cooperative miRNA mechanism of the soybean. Some MTRMs are involved in the

364 biosynthesis of chalcone, which is derived from the general phenylpropanoid pathway

365 that plays a wide variety of roles in soybeans, as well as other plants. In most cases, gene

366 regulation in each MTRMs involved a multi-component miRNA gene family. In some

367 cases, these families were predicted to act cooperatively, which is consistent with the

368 conclusion of Wang et al. [27]. The functions of a specific miRNA family can represent

369 the main functions of miRNAs in the family, such as miR396 and its growth regulatory

370 factors (GRFs) [68]; moreover, these regulatory factors include drought stress, high salt

371   stress, low-temperature stress, and ultraviolet radiation stress. Interestingly, we found that

372   miRNAs from different families are also involved in the same regulatory gene clusters,

373   which indicates that different miRNA families may have cross-family cooperative

374   regulatory mechanisms in regulating certain functions. In contrast, miRNAs in the same

375   family can be in different MTRMs; for example, the miR171 family (miR172b-5p,

376   miR172h-5p, miR172f, miR172g, miR172j, and miR172k) are in multiple regulatory

377   modules during drought and salt stress. Such hub miRNAs may be useful research targets

378   for exploring soybean resistance mechanisms and resistance to breeding research under

379   different stresses. After further combining the analysis of differentially expressed genes

380   in soybeans under various stresses, we obtained the miRNA-GO regulatory network

381   under abiotic stress. The GO BP contains a variety of important related pathways for

382   understanding the common mechanisms in stress response. The research covering the

383   plant miRNA regulation module can analyze the coordination mechanism of miRNA

384   from a global perspective and determine the regulation relationship between modules,

385   which may help explore the regulation mechanism of soybean miRNAs.

386

387

Methods

388

389   We collected soybean MTIs from *Arabidopsis thaliana* and *Medicago truncatula*

390   databases and publications on miRNAs and genes of soybean response to several abiotic

391   stresses. Subsequently, we used homology prediction on the collected MTIs to expand

392   the soybean MTIs. We used the biclustering method to mine the soybean MTRMs to

393   perform overlap analysis on the results and practiced redundancy removal. Then, based

394   on the soybean gene interaction network, biclustering was applied through multi-level

395   iteration. Finally, based on soybean abiotic stress-related miRNAs and genes, the fusion

396   regulatory module was screened to obtain soybean abiotic stress-related MTRMs. Figure

397   6 shows a flowchart of our tasks and results.

398
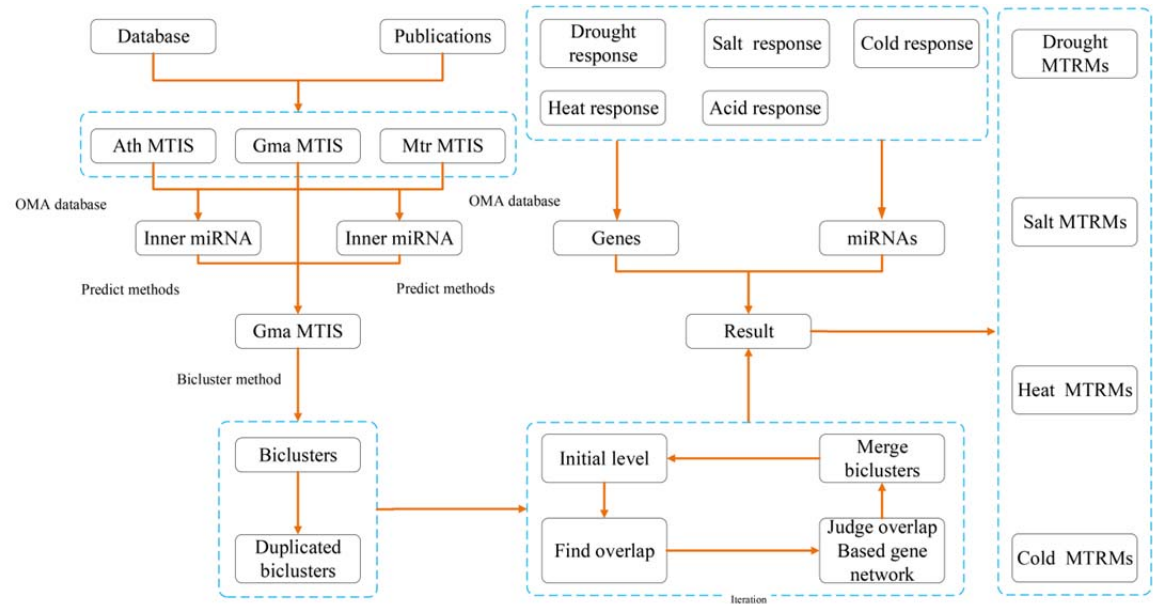
399   Figure 6. Flowchart of the authors' research method.

400

401 **Data collection**

402   We collected miRNA-target data of *Arabidopsis thaliana*, soybean and *Medicago*

403   *truncatula* based on experimentally verified degradome sequencing results from

404　databases (DPMIND, (c) Tarbase, mirTarbase, and Starbase [93-96]) and publications

405　(Supplemental Table 1). In addition, we collected the miRNA information of the three

406　species from the miRbase [97], the gene annotation of the species in the NCBI,

407　EnsemblPlants, and the Phytozome [98, 99]. We also downloaded the homologous genes

408　of *Arabidopsis thaliana* and *Medicago truncatula* in Orthologous MAtrix (OMA) [100].

409　Besides, we downloaded the soybean cDNA sequence and soybean gene GO annotations

410　from SoyBase [101], and obtained soybean gene network data from SoyNet [102].

411

412

413　**Data processing**

414　　Due to the differences in the annotation methods in various databases and the

415　differences in the data formats in the publications, we unified the miRNA and gene

416　formats in the data and put the data of the same species together. We annotated the

417　miRNA-target data based on the collected and processed miRNA details and the gene

418　annotations derived from the data of three species, including miRNA target data, related

419　notes and data sources. After processing the duplicated data, we obtained the

420　miRNA-target data of the three species.

421

422　**Homologous extension**

423　　We chose *Arabidopsis thaliana* and *Medicago truncatula* to explore the potential

424　targets. *Arabidopsis thaliana* as a model plant has rich high-quality data. *Medicago*

425　*truncatula* and soybean are closely related and have many similar biological

426　characteristics. We extracted the miRNA sequence and removed redundant miRNAs with

427　the same sequence in the soybean and *Arabidopsis thaliana*. Subsequently, we extracted

428 the target gene corresponding to the miRNA ID. Based on these targeted genes, we

429 obtained soybean genes homologous to these genes from the *Arabidopsis*

430 *thaliana*-soybean homologous genes downloaded by OMA. We assumed that if the

431 sequences coexist and the genes are homologous, targeting relationships may exist.

432 Therefore, these homologous genes may be targeted by these miRNAs in soybeans.

433 Targets obtained only based on homology information may not exist; so, we

434 extracted these miRNA sequences and the cDNA sequence of target genes (SoyBase) and

435 used miRNA-target prediction tools to predict potential relationships. Most of the

436 existing plant miRNA target gene prediction tools showed high accuracy in *Arabidopsis*

437 *thaliana*, but the prediction results in non-*Arabidopsis thaliana* plant data were different.

438 Therefore, we chose psRNAtarget, TAPIR, and Targetfinder, whose results were better in

439 non-Arabidopsis plants to predict potential soybean miRNA-target relationships [104].

440 The three prediction software tools have different scoring methods. We analyzed their

441 respective scores and merged them. The homology extension method for *Medicago*

442 *truncatula*-soybean is the same as above.

443

Clustering method

445 The current research on miRNA targeting relationships is mostly based on

446 one-to-one relative targeting. However, the miRNA targeting relationship is a complex

447 interaction. The traditional clustering method is to cluster the same type of data, such as

448 k-means, whose mining results in the miRNA-target regulatory module are poor because

449 the targeting of miRNAs is sparse. The relationship between miRNA and the target gene

450 is a bipartite graph structure; thus, the miRNA-target regulatory group can be found by

451 analyzing the bipartite graph. Bimax has been widely used in gene expression data [103].

452 However, for more data, the calculation cost is higher. Domingo et al. proposed the BiBit

453 [69], whose results are almost the same as that of Bimax, but the calculation time is still

454 longer. Subsequently, ParBiBit and CUBiBit were proposed [57], which shortened the

455 computing time and provided an optimized method for finding modules in larger data.

456 We added the miRNA-target data based on the homology expansion predictions obtained

457 from *Arabidopsis thaliana* and *Medicago truncatula* into the collected soybean

458 miRNA-target data and extracted the miRNA-Target data with GO annotations and

459 glyma2ID based on the soybean gene annotations of SoyBase. We used the CUBiBit to

460 perform bi-clustering to obtain the results.

461

462

463 **Overlap and iterative fusion**

464 The primary result obtained by CUBiBit was mostly a fully-connected bipartite

465 graph. However, the relationship between miRNA and target gene is complex and

466 interactive. Therefore, we proposed a method of iterative fusion for MTRM modules

467 based on a gene interaction network (Figure 7).

468 We detected completely included classes in the clustering results and removed the

469 included classes as the initial level result. First, for each class of this level containing

470 miRNAs and genes, we judged the degree of overlap with other classes of miRNA and

471 genes to form alpha and beta matrices, both of which are upper triangular matrices. After

472 that, we set two thresholds of miRNA and genes that can be potentially merged for the

473 two classes. We then recorded the two classes that met the potential fusion class-class

474 table requirements to form a Boolean matrix. The initial threshold of alpha was 0.3, and

475 each iteration increased at a pace of 0.05 to conservatively determine the fusionable

476 module and to keep this value unchanged after increasing to 0.8. It was sufficient if the

477 beta threshold was greater than 0. Next, we extracted the union of two class genes, and

478   the network blocks of this pair of genes with a depth of 2 layers based on the SoyNet

479   network, for each pair of classes in the potentially merged class-class table. Subsequently,

480   from the obtained block set, network blocks containing smaller classes were extracted.

481   We assumed that the network block with the most smaller class genes represented the

482   function of the genes in the smaller class. Therefore, judging the number of genes in the

483   major category of this network block can determine whether the genes of the two

484   categories are similar in function. If the genes of the two classes were concentrated on a

485   network block, which means that their genes interact closely and meet the conditions of

486   potential fusion, the two classes can be merged. We compared the number of genes in the

487   major category with the number of genes in all major categories in the sub-category

488   function module to obtain scores and determine the correlation. Finally, we compared the

489   number of genes in the major category with the number of genes in all major categories

490   in the sub-category function module to obtain scores to determine the correlation. The

491   threshold was recorded as gamma. When gamma > 0.3 was satisfied, the two classes are

492   merged; otherwise, they would not be merged. For the class pairs that meet the fusion

493   condition, we arranged them in descending order of alpha value and performed top-down

494   non-repetitive fusion. Each class can only merge at most one class in one iteration. A new

495   class set was formed as the new level, and the fusion result was the output. The next

496   iteration would be performed and then another iteration until no fusion class pair could
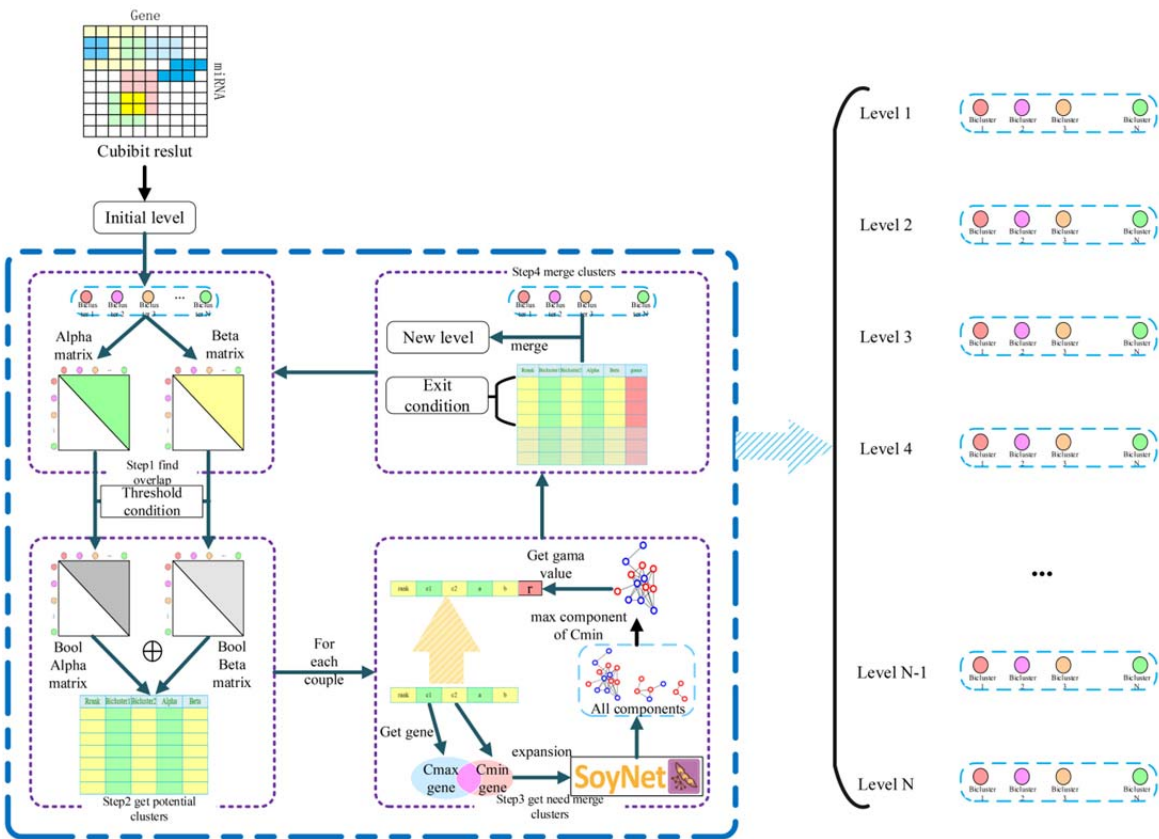
497   meet the two conditions.

498

Figure 7. MTRM iterative merge algorithm flowchart used
to derive a gene interaction network.

499
500
501

502

Result analysis

503

For the results of the above iterative fusion, the enrichment of the classes in each

504

level were separately analyzed. For a bicluster, we extracted its genes, used SoyBase's

505

GO BP and GO MF for enrichment analysis, and used the corrected GO ID with the

506

smallest p-value as the best enrichment result for this type of cluster. When evaluating

507

each class, only the smallest p-value was not enough to judge the importance of the class.

508

Here, we should observe the situation in all the GO IDs enriched by this class to make a

509

reasonable judgment on this class. Therefore, we used the cluster score to evaluate the

510

26

511 enrichment of the class. For all the enriched GO IDs of this class, we screened all the

512 results with a p-value of less than 0.05 and then used Eq. (1) to calculate the cluster score

513 of the class.

514 $$cluster\ score = \frac{\sum_1^n (x_i * -log(correct\ P_i))}{\sum_1^n x_i} \qquad (1)$$

515 Among them, n is the number of gene ontologies enriched in the module, x_i is the

516 number of genes enriched in the i-th GO, and $correct\ P_i$ is the adjusted p-value of the

517 i-th enriched GO.

518

519 **Abiotic Stress response MTRM**

520 We collected the miRNAs of soybeans that respond to drought, salt, acid, and low

521 temperature based on our studies of publications [22, 48, 105-121]. At the same time, we

522 collected the differentially expressed soybean genes under various stresses and screened

523 these genes with foldchange ≥ 2 and t-test p-value less than 0.05 as related genes under

524 abiotic stress. Then we marked the genes in the module and calculated the p-value related

525 to abiotic stress based on the hypergeometric distribution. Finally, we screened based on

526 the cluster score calculated by the module, the p-value related to stress, and the

527 proportion of miRNA related to stress. In addition, the screening procedures related to

528 drought and salt stress were consistent with the screening steps of the abiotic stress

529 module.

530

531 **Construction of miRNA-GO network under abiotic stress**

532 Based on the results of MTRM mining under stress, we first screened the GO of the

533 enrichment results in the screened module by p-value to remove the GO with a p-value

534 less than 10^{-5}; then, we performed a REVIGO semantic relevance analysis and extraction

535 of concentrated representative GO channels. Based on the MTI data, the miRNA-GO

536 relationship data was constructed through the gene pointed to by the miRNA in the

537 module, and the enriched go pathway to which the gene belongs. The relationship

538 between GO is based on the results of REVIGO and the GO similarity calculation. The

539 relationship is presented by setting a threshold to remove some weaker relationships.

540 More detailed parameters are provided here or in the location of the specific figure.

541

542 **Supplemental data**

543 **Supplemental Table S1.** Summary of ATH, GMA and MTR MTIs sources.

544 **Supplemental Table S2.** Potential targets found by homology expansion.

545 **Supplemental Table S3.** Comparison of the biclustering results of soybean MTRMs.

546 **Supplemental Table S4.** The top 5 GO enrichment of each top 5 soybean MTRMs.

547 **Supplemental Table S5.** Top 5 soybean abiotic response MTRMs.

548 **Supplemental Dataset S1.** miRNA function annotation.

549 **Supplemental Dataset S2.** Screening results of soybean MTRMs related to abiotic stress,

550 drought and salt stress.

551 **Supplemental Dataset S3.** Stress-specific MTRMs and shared MTRMs.

552 **Supplemental Dataset S4.** GO terms of miRNAs in the soybean abiotic response

553 MTRMs.

554 **Supplemental Figure S1.** Potential soybean MTIs predicted based on three prediction

555 tools. (a) 961 pairs of MTIs obtained based on the Arabidopsis data, (b) 986 pairs of

556 MTIs obtained based on the Medicago data, and (c) 1,189 soybean MTIs from the union

557 between (a) and (b).

28

558   **Supplemental Figure S2.** The differentially expressed genes in each MTRM under other

559   stress after the screening.

560

561   **Funding**

565

566   **Acknowledgments**

568

569   **Conflict of interest statement.** None declared.

570

571

Parsed Citations

1. Ku, Y.S., Wong, J.W., Mui, Z., Liu, X., Hui, J.H., et al. (2015) Small RNAs in Plant Responses to Abiotic Stresses: Regulatory Roles and Study Methods. Int J Mol Sci, 16(10), 24532-24554.
Google Scholar: Author Only Title Only Author and Title

2. Li, S., Castillo-Gonzalez, C., Yu, B., Zhang, X. (2017) The functions of plant small RNAs in development and in stress responses. Plant J, 90(4), 654-670.
Google Scholar: Author Only Title Only Author and Title

3. Ruiz-Ferrer, V., Voinnet, O. (2009) Roles of plant small RNAs in biotic stress responses. Annu Rev Plant Biol, 60, 485-510.
Google Scholar: Author Only Title Only Author and Title

4. Sunkar, R., Chinnusamy, V., Zhu, J., Zhu, J.K. (2007) Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. Trends Plant Sci, 12(7), 301-309.
Google Scholar: Author Only Title Only Author and Title

5. Shukla, L.I., Chinnusamy, V., Sunkar, R. (2008) The role of microRNAs and other endogenous small RNAs in plant stress responses. Biochim Biophys Acta, 1779(11), 743-748.
Google Scholar: Author Only Title Only Author and Title

6. Song, X., Li, Y., Cao, X., Qi, Y. (2019) MicroRNAs and Their Regulatory Roles in Plant-Environment Interactions. Annu Rev Plant Biol, 70, 489-525.
Google Scholar: Author Only Title Only Author and Title

7. Sun, X., Lin, L., Sui, N. (2019) Regulation mechanism of microRNA in plant response to abiotic stress and breeding. Mol Biol Rep, 46(1), 1447-1457.
Google Scholar: Author Only Title Only Author and Title

8. Hivrale, V., Zheng, Y., Puli, C.O.R., Jagadeeswaran, G., et al. (2016) Characterization of drought- and heat-responsive microRNAs in switchgrass. Plant Sci, 242, 214-223.
Google Scholar: Author Only Title Only Author and Title

9. Yu, Y., Ni, Z., Wang, Y., Wan, H., Hu, Z., et al. (2019) Overexpression of soybean miR169c confers increased drought stress sensitivity in transgenic Arabidopsis thaliana. Plant Sci, 285, 68-78.
Google Scholar: Author Only Title Only Author and Title

10. Pan, W.J., Tao, J.J., Cheng, T., Bian, X.H., Wei, W., et al. (2016) Soybean miR172a Improves Salt Tolerance and Can Function as a Long-Distance Signal. Mol Plant, 9(9), 1337-1340.
Google Scholar: Author Only Title Only Author and Title

11. Li, W., Wang, T., Zhang, Y., Li, Y. (2016) Overexpression of soybean miR172c confers tolerance to water deficit and salt stress, but increases ABA sensitivity in transgenic Arabidopsis thaliana. J Exp Bot, 67(1), 175-194.
Google Scholar: Author Only Title Only Author and Title

12. Zhou, Y., Liu, W., Li, X., Sun, D., Xu, K., et al. (2020) Integration of sRNA, degradome, transcriptome analysis and functional investigation reveals gma-miR398c negatively regulates drought tolerance via GmCSDs and GmCCS in transgenic Arabidopsis and soybean. BMC Plant Biol, 20(1), 190.
Google Scholar: Author Only Title Only Author and Title

13. Sun, Z, Wang, Y, Mou, F, Tian, Y, Chen, L, Zhang, S, Jiang, Q, Li, X (2015) Genome-Wide Small RNA Analysis of Soybean Reveals Auxin-Responsive microRNAs that are Differentially Expressed in Response to Salt Stress in Root Apex. Front Plant Sci 6, 1273.
Google Scholar: Author Only Title Only Author and Title

14. Ning, L.H., Du,W.K., Song, H.N., Shao, H.B., Qi, W.C., et al. (2019) Identification of responsive miRNAs involved in combination stresses of phosphate starvation and salt stress in soybean root. Environmental and Experimental Botany, 167, 103823.
Google Scholar: Author Only Title Only Author and Title

15. Huang, S.C,. Lu, G.H., Tang, C.Y., Ji, Y.J., Tan, G.S., et al. (2018) Identification and comparative analysis of aluminum-induced microRNAs conferring plant tolerance to aluminum stress in soybean. Biologia Plantarum, 62(1), 97-108.
Google Scholar: Author Only Title Only Author and Title

16. Enright, A.J., John, B., Gaul, U., Tuschl, T., Sander, C., Marks, D.S. (2003) MicroRNA targets in Drosophila. Genome Biol, 5(1), R1.
Google Scholar: Author Only Title Only Author and Title

17. Azra, K., Dominic, G., Matthew, N.P., Rachel, W., Lauren, R., et al. (2005) Combinatorial microRNA target predictions. Nat Genet, 37(5), 495-500.
Google Scholar: Author Only Title Only Author and Title

18. Wahdan-Alaswad, R., Liu, B. (2013) "Sister" miRNAs in cancers. Cell Cycle, 12(24),3703-3704.
Google Scholar: Author Only Title Only Author and Title

19. Wang, S., Huang, J., Lyu, H., Lee, C.K., Tan, J., Wang, J., Liu, B. (2013) Functional cooperation of miR-125a, miR-125b, and miR-205 in

entinostat-induced downregulation of erbB2/erbB3 and apoptosis in breast cancer cells. Cell Death Dis, 4, e556.
Google Scholar: Author Only Title Only Author and Title

20. Ivanovska I., Cleary M.A. (2009) Combinatorial microRNAs: working together to make a difference, 7(20), 3137-3142.
Google Scholar: Author Only Title Only Author and Title

21. Frampton, A.E., Castellano, L., Colombo, T., Giovannetti, E., Krell, J., et al. (2014) MicroRNAs cooperatively inhibit a network of tumor suppressor genes to promote pancreatic tumor growth and progression. Gastroenterology, 146(1), 268-277 e218.
Google Scholar: Author Only Title Only Author and Title

22. Ramanjulu, S., Li, Y.F., Guru, J. (2012) Functions of microRNAs in plant stress responses. Trends in plant science, 17(4), 196-203.
Google Scholar: Author Only Title Only Author and Title

23. Xu, Y., Guo, M., Liu, X., Wang, C., Liu, Y. (2014) Inferring the soybean (Glycine max) microRNA functional network based on target gene network. Bioinformatics, 30(1), 94-103.
Google Scholar: Author Only Title Only Author and Title

24. Yoshikawa, T., Ozawa, S., Sentoku, N., Itoh, J., Nagato, Y., Yokoi, S. (2013) Change of shoot architecture during juvenile-to-adult phase transition in soybean. Planta, 238(1), 229-237.
Google Scholar: Author Only Title Only Author and Title

25. Wang, T., Sun, M.Y., Wang, X.S., Li, W.B., Li, Y.G. (2016) Over-Expression of GmGla-Regulated Soybean miR172a Confers Early Flowering in Transgenic Arabidopsis thaliana. Int J Mol Sci, 17(5).
Google Scholar: Author Only Title Only Author and Title

26. Li, H., Deng, Y., Wu, T., Subramanian, S., Yu, O. (2010) Misexpression of miR482, miR1512, and miR1515 increases soybean nodulation. Plant Physiol, 153(4), 1759-1770.
Google Scholar: Author Only Title Only Author and Title

27. Wang, R., Yang, Z, Fei, Y., Feng, J., Zhu, H., et al. (2019) Construction and analysis of degradome-dependent microRNA regulatory networks in soybean. BMC Genomics, 20(1), 534.
Google Scholar: Author Only Title Only Author and Title

28. Shalgi, R., Lieber, D., Oren, M., Pilpel, Y. (2007) Global and local architecture of the mammalian microRNA-transcription factor regulatory network. PLoS Comput Biol, 3(7), e131.
Google Scholar: Author Only Title Only Author and Title

29. Xu, J., Li, C.X., Li, Y.S., Lv, J.Y., Ma, Y., et al. (2011) MiRNA-miRNA synergistic network: construction via co-regulating functional modules and disease miRNA topological features. Nucleic Acids Res, 39(3), 825-836.
Google Scholar: Author Only Title Only Author and Title

30. Xie, J., Ma, A, Fennell, A, Ma, Q., Zhao, J. (2019) It is time to apply biclustering: a comprehensive review of biclustering applications in biological and biomedical data. Brief Bioinform, 20(4), 1449-1464.
Google Scholar: Author Only Title Only Author and Title

31. Tanay, A, Sharan, R., Shamir, R. (2002) Discovering statistically significant biclusters in gene expression data. Bioinformatics, 18 Suppl 1, S136-144.
Google Scholar: Author Only Title Only Author and Title

32. Bergmann, S., Ihmels, J., Barkai, N. (2003) Iterative signature algorithm for the analysis of large-scale gene expression data. Phys Rev E Stat Nonlin Soft Matter Phys, 67(3 Pt 1), 031902.
Google Scholar: Author Only Title Only Author and Title

33. Prelic, A, Bleuler. S., Zimmermann, P., Wille, A, Buhlmann, P., et al. (2006) A systematic comparison and evaluation of biclustering methods for gene expression data. Bioinformatics, 22(9), 1122-1129.
Google Scholar: Author Only Title Only Author and Title

34. Li, G., Ma, Q., Tang, H., Paterson, A.H., Xu, Y. (2009) QUBIC: a qualitative biclustering algorithm for analyses of gene expression data. Nucleic Acids Res, 37(15), e101.
Google Scholar: Author Only Title Only Author and Title

35. Hochreiter, S., Bodenhofer, U., Heusel, M., Mayr, A, Mitterecker, A, et al. (2010) FABIA: factor analysis for bicluster acquisition. Bioinformatics, 26(12), 1520-1527.
Google Scholar: Author Only Title Only Author and Title

36. Medina, I., Carbonell, J., Pulido, L., Madeira, S.C., Goetz, S., et al. (2010) Babelomics: an integrative platform for the analysis of transcriptomics, proteomics and genomic data with advanced functional profiling. Nucleic Acids Res, 38(Web Server issue), W210-213.
Google Scholar: Author Only Title Only Author and Title

37. Goncalves, J.P., Madeira, S.C., Oliveira, A.L. (2009) BiGGEsTS: integrated environment for biclustering analysis of time series gene expression data. BMC Res Notes, 2, 124.
Google Scholar: Author Only Title Only Author and Title

38. Madeira, S.C., Teixeira, M.C., Sa-Correia, I., Oliveira, A.L. (2010) Identification of regulatory modules in time series gene expression

data using a linear time biclustering algorithm. IEEE/ACM Trans Comput Biol Bioinform, 7(1), 153-165.
  Google Scholar: Author Only Title Only Author and Title

39. Goncalves, J.P., Madeira, S.C. (2014) LateBiclustering: Efficient Heuristic Algorithm for Time-Lagged Bicluster Identification. IEEE/ACM Trans Comput Biol Bioinform, 11(5), 801-813.
  Google Scholar: Author Only Title Only Author and Title

40. Henriques, R., Ferreira, F.L., Madeira, S.C. (2017) BicPAMS: software for biological data analysis with pattern-based biclustering. BMC Bioinformatics, 18(1), 82.
  Google Scholar: Author Only Title Only Author and Title

41. Henriques, R., Madeira, S.C. (2014) BicPAM: Pattern-based biclustering for biomedical data analysis. Algorithms Mol Biol, 9(1), 27.
  Google Scholar: Author Only Title Only Author and Title

42. Henriques, R., Madeira, S.C. (2016) BicNET: Flexible module discovery in large-scale biological networks using biclustering. Algorithms Mol Biol, 11, 14.
  Google Scholar: Author Only Title Only Author and Title

43. Bentham, R.B., Bryson, K., Szabadkai, G. (2017) MCbiclust: a novel algorithm to discover large-scale functionally related gene sets from massive transcriptomics data collections. Nucleic Acids Res, 45(15), 8712-8730.
  Google Scholar: Author Only Title Only Author and Title

44. Caldas, J., Kaski, S. (2011) Hierarchical generative biclustering for microRNA expression analysis. J Comput Biol, 18(3), 251-261.
  Google Scholar: Author Only Title Only Author and Title

45. Yoon, S., De, M.G. (2005) Prediction of regulatory modules comprising microRNAs and target genes. Bioinformatics, 21(Suppl 2), ii93-100.
  Google Scholar: Author Only Title Only Author and Title

46. Huang, J.Y., Brutlag, D.L. (2001) The EMOTIF database. Nucleic Acids Res, 29(1), 202-204.
  Google Scholar: Author Only Title Only Author and Title

47. Pio, G., Ceci, M., D'Elia, D., Loglisci, C., Malerba, D. (2013) A novel biclustering algorithm for the discovery of meaningful biological correlations between microRNAs and their target genes. BMC Bioinformatics, 14(Suppl 7), S8.
  Google Scholar: Author Only Title Only Author and Title

48. Muhammad, D., Muhammad Y.K.B., Iftekhar A.B. (2016) Profiling and annotation of microRNAs and their putative target genes in chilli (Capsicum annuum L.) using ESTs. Gene Reports, 5, 62-69.
  Google Scholar: Author Only Title Only Author and Title

49. Xu, Y.G., Guo, M.Z., Liu, X.Y., Wang, C.Y., Liu, Y. (2014) SoyFN: a knowledge database of soybean functional networks. Database, 2014(1), bau019.
  Google Scholar: Author Only Title Only Author and Title

50. Fei, Y., Wang, R., Li, H., Liu, S., Zhang, H., Huang, J. (2018) DPMIND: degradome-based plant miRNA-target interaction and network database. Bioinformatics, 34(9), 1618-1620.
  Google Scholar: Author Only Title Only Author and Title

51. Yang, J.H., Li, J.H., Shao, P., Zhou, H., Chen, Y.Q., Qu, L.H. (2011) starBase: a database for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data. Nucleic Acids Res, 39(Database issue), D202-209.
  Google Scholar: Author Only Title Only Author and Title

52. Sethupathy, P., Corda, B., Hatzigeorgiou, A.G. (2006) TarBase: A comprehensive database of experimentally supported animal microRNA targets. RNA, 12(2), 192-197.
  Google Scholar: Author Only Title Only Author and Title

53. Dai, X., Zhuang, Z., Zhao, P.X. (2018) psRNATarget: a plant small RNA target analysis server (2017 release). Nucleic Acids Res, 46(W1), W49-W54.
  Google Scholar: Author Only Title Only Author and Title

54. Dai, X., Zhao, P.X. (2011) psRNATarget: a plant small RNA target analysis server. Nucleic Acids Res, 39(Web Server issue), W155-159.
  Google Scholar: Author Only Title Only Author and Title

55. Bonnet, E., He, Y., Billiau, K., Van, de, P.Y. (2010) TAPIR, a web server for the prediction of plant microRNA targets, including target mimics. Bioinformatics, 26(12), 1566-1568.
  Google Scholar: Author Only Title Only Author and Title

56. Bo, X.C., Wang, S.Q. (2005) TargetFinder: a software for antisense oligonucleotide target site selection based on MAST and secondary structures of target mRNA Bioinformatics, 21(8), 1401-1402.
  Google Scholar: Author Only Title Only Author and Title

57. Jorge González-Domínguez, R.R.E. (2019) Accelerating binary biclustering on platforms with CUDA-enabled GPUs. Information Sciences, 496, 317-325.

Google Scholar: Author Only Title Only Author and Title

58. Khraiwesh, B., Zhu, J.K., Zhu, J. (2012) Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. Biochim Biophys Acta, 1819(2), 137-148.
   Google Scholar: Author Only Title Only Author and Title

59. Kulcheski, F.R., Molina, L.G., da, Fonseca, G.C., de, Morais, G.L., de, Oliveira, L.F., Margis, R. (2016) Novel and conserved microRNAs in soybean floral whorls. Gen, 575(2 Pt 1), 213-223.
   Google Scholar: Author Only Title Only Author and Title

60. Zhu, Q.H., Helliwell, C.A. (2011) Regulation of flowering time and floral patterning by miR172. J Exp Bot, 62(2), 487-495.
   Google Scholar: Author Only Title Only Author and Title

61. Chen, X. (2001) A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. Science, 303(5666), 2022-2025.
   Google Scholar: Author Only Title Only Author and Title

62. Aukerman, M.J., Sakai, H. (2003) Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. Plant Cell, 15(11), 2730-2741.
   Google Scholar: Author Only Title Only Author and Title

63. Wong, J., Gao, L., Yang, Y., Zhai, J., Arikit, S., et al. (2014) Roles of small RNAs in soybean defense against Phytophthora sojae infection. Plant J, 79(6), 928-940.
   Google Scholar: Author Only Title Only Author and Title

64. Mao, G., Turner, M., Yu, O., Subramanian, S. (2013) miR393 and miR164 influence indeterminate but not determinate nodule development. Plant Signal Behav, 8(10), doi: 10 4161/psb 26753.
   Google Scholar: Author Only Title Only Author and Title

65. Turner, M., Nizampatnam, N.R., Baron, M., Coppin, S., Damodaran, S., et al. (2013) Ectopic expression of miR160 results in auxin hypersensitivity, cytokinin hyposensitivity, and inhibition of symbiotic nodule development in soybean. Plant Physiol, 162(4), 2042-2055.
   Google Scholar: Author Only Title Only Author and Title

66. Wang, Y., Li, K., Chen, L., Zou Y., Liu, H., et al. (2015) MicroRNA167-Directed Regulation of the Auxin Response Factors GmARF8a and GmARF8b Is Required for Soybean Nodulation and Lateral Root Development. Plant Physiol, 168(3), 984-999.
   Google Scholar: Author Only Title Only Author and Title

67. Yan, Z., Hossain, M.S., Wang, J., Valdes-Lopez, O., Liang, Y., et al. (2013) miR172 regulates soybean nodulation. Mol Plant Microbe Interact, 26(12), 1371-1377.
   Google Scholar: Author Only Title Only Author and Title

68. Omidbakhshfard, M.A, Proost, S., Fujikura, U., Mueller-Roeber, B. (2015) Growth-Regulating Factors (GRFs): A Small Transcription Factor Family with Important Functions in Plant Biology. Mol Plant, 8(7), 998-1010.
   Google Scholar: Author Only Title Only Author and Title

69. Rodriguez-Baena, D.S., Perez-Pulido, A.J., Aguilar-Ruiz, J.S. (2011) A biclustering algorithm for extracting bit-patterns from binary datasets. Bioinformatics, 27(19), 2738-2745.
   Google Scholar: Author Only Title Only Author and Title

70. Yan, Z, Hossain, M., Arikit, S., et al. (2015) Identification of micro RNAs and their mRNA targets during soybean nodule development: functional analysis of the role of miR393j-3p in soybean nodulation. New Phytologist, 207(3), 748-759.
   Google Scholar: Author Only Title Only Author and Title

71. Yan, Z, Hossain, M., Lopez, O., et al. (2016) Identification and functional characterization of soybean root hair microRNAs expressed in response to Bradyrhizobium japonicum infection. Plant Biotechnology Journal, 14(1), 332-341.
   Google Scholar: Author Only Title Only Author and Title

72. Liu, T., Fang, C., Ma, Y., et al. (2016) Global investigation of the co-evolution of MIRNA genes and micro RNA targets during soybean domestication. Plant Journal, 85(3), 396-409.
   Google Scholar: Author Only Title Only Author and Title

73. Ding, X., Li, J., Zhang, H., et al. (2016) Identification of miRNAs and their targets by high-throughput sequencing and degradome analysis in cytoplasmic male-sterile line NJCMS1A and its maintainer NJCMS1B of soybean. BMC Genomics, 17, 24.
   Google Scholar: Author Only Title Only Author and Title

74. Chen, H., Adam, Arsovski, A, Yu, K., et al. (2017) Deep sequencing leads to the identification of eukaryotic translation initiation factor 5A as a key element in Rsv1-mediated lethal systemic hypersensitive response to Soybean mosaic virus infection in soybean. Molecular Plant Pathology, 18(3), 391-404.
   Google Scholar: Author Only Title Only Author and Title

75. Xu, S., Liu, N., Mao, W., et al. (2016) Identification of chilling-responsive microRNAs and their targets in vegetable soybean (Glycine max L.). SCIENTIFIC REPORTS, 6.
   Google Scholar: Author Only Title Only Author and Title

76. Chen, H., Arsovski, A.A., Yu, K., et al. (2017) Correction: Genome-Wide Investigation Using sRNA-Seq, Degradome-Seq and Transcriptome-Seq Reveals Regulatory Networks of microRNAs and Their Target Genes in Soybean during Soybean mosaic virus Infection. PLoS ONE, 12(11), 1.
Google Scholar: Author Only Title Only Author and Title

77. Turner, M., Yu, O., Subramanian, S. (2012) Genome organization and characteristics of soybean microRNAs. BMC Genomics, 13(1), 169.
Google Scholar: Author Only Title Only Author and Title

78. Shamimuzzaman, M., Vodkin, L. (2012) Identification of soybean seed developmental stage-specific and tissue-specific miRNA targets by degradome sequencing. BMC Genomics, 13(1).
Google Scholar: Author Only Title Only Author and Title

79. Xu, F., Liu, Q., Chen, L., et al. (2013) Genome-wide identification of soybean microRNAs and their targets reveals their organ-specificity and responses to phosphate starvation. BMC Genomics, 14(1), 66.
Google Scholar: Author Only Title Only Author and Title

80. Fang, X., Zhao, Y., Ma, Q., et al. (2013) Identification and comparative analysis of cadmium tolerance- associated miRNAs and their targets in two soybean genotypes. Plos One, 8(12), e81471.
Google Scholar: Author Only Title Only Author and Title

81. Xu, M., Li, Y., Zhang, Q., et al. (2014) Novel MiRNA and PhasiRNA Biogenesis Networks in Soybean Roots from Two Sister Lines That Are Resistant and Susceptible to SCN Race 4. PLoS ONE, 9(10), 1-20.
Google Scholar: Author Only Title Only Author and Title

82. Ye, C., Xu, H., Shen, E., et al. (2014) Genome-wide identification of non-coding RNAs interacted with microRNAs in soybean. FRONTIERS IN PLANT SCIENCE, 5, 743.
Google Scholar: Author Only Title Only Author and Title

83. Song, Q., Liu, Y., Hu, X., et al. (2011) Identification of miRNAs and their target genes in developing soybean seeds by deep sequencing. BMC PLANT BIOLOGY, 11(1), 1-16.
Google Scholar: Author Only Title Only Author and Title

84. Xu, M., Zhang, L., Li, W., et al. (2014) Stress-induced early flowering is mediated by miR169 in Arabidopsis thaliana. Journal of Experimental Botany, 65(1), 89-101.
Google Scholar: Author Only Title Only Author and Title

85. Ding, J., Li, D., Ohler, U., et al. (2012) Genome-wide search for miRNA-target interactions in Arabidopsis thaliana with an integrated approach. BMC Genomics, 13(Suppl 3), S3.
Google Scholar: Author Only Title Only Author and Title

86. German, M.A, Pillay, M., Jeong, D-H. (2008) Global identification of microRNA–target RNA pairs by parallel analysis of RNA ends. Nature biotechnology, 26(8), 941-946.
Google Scholar: Author Only Title Only Author and Title

87. Addo-Quaye, C., Eshoo, T.W., Bartel, D.P., et al. (2008) Endogenous siRNA and miRNA targets identified by sequencing of the Arabidopsis degradome. Current biology: CB, 18(10), 758-762.
Google Scholar: Author Only Title Only Author and Title

88. Ma, X., Yu, D., Shao, W., et al. (2018) Transcriptome-wide identification and characterization of the copper and cadmium stress-responsive small RNAs and their targets in Arabidopsis thaliana. Plant & Soil, 429(1-2), 391-405.
Google Scholar: Author Only Title Only Author and Title

89. Lauressergues, D., Delaux, P.M., Formey, D., et al. (2012) The microRNA miR171h modulates arbuscular mycorrhizal colonization of Medicago truncatula by targeting NSP2. Plant J, 72(3): 512-522.
Google Scholar: Author Only Title Only Author and Title

90. Zhou, Z, Zeng, H., Liu, Z., et al. (2012) Genome-wide identification of Medicago truncatula microRNAs and their targets reveals their differential regulation by heavy metal. Plant Cell Environ, 35(1), 86-99.
Google Scholar: Author Only Title Only Author and Title

91. Devers, E.A, Branscheid, A, May, P. (2011) Stars and Symbiosis: MicroRNA- and MicroRNA*-Mediated Transcript Cleavage Involved in Arbuscular Mycorrhizal Symbiosis. PLANT PHYSIOLOGY, 156(4), 1990-2010.
Google Scholar: Author Only Title Only Author and Title

92. Senthil, S., Gary, S., Oliver, Y. (2006) Endogenous isoflavones are essential for the establishment of symbiosis between soybean and Bradyrhizobium japonicum. The Plant journal : for cell and molecular biology, 48(2), 261-273.
Google Scholar: Author Only Title Only Author and Title

93. Fei, Y.H., Wang, R., Li, H.Y., Liu, S., Zhang, H.S., Huang, J. (2018) DPMIND: degradome-based plant miRNA-target interaction and network database. Bioinformatics, 34(9), 1618-1620.
Google Scholar: Author Only Title Only Author and Title

94. Vlachos, I.S., Paraskevopoulou, M.D., Karagkouni, D., Georgakilas, G., Vergoulis, T., et al. (2015) DIANA-TarBase v7.0: indexing

more than half a million experimentally supported miRNA:mRNA interactions. Nucleic Acids Research, 43, D153-D159.
Google Scholar: Author Only Title Only Author and Title

95. Hsu, S.D., Lin, F.M., Wu, W.Y., Liang, C., Huang, W.C., et al. (2011) miRTarBase: a database curates experimentally validated microRNA–target interactions. Nucleic Acids Research, 39, D163-D169.
Google Scholar: Author Only Title Only Author and Title

96. Li, J.H., Liu, S., Zhou, H., Qu, L.H., Yang, J.H. (2014) starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. NUCLEIC ACIDS RESEARCH, 42, D92-D97.
Google Scholar: Author Only Title Only Author and Title

97. Harpreet, K.S., Stijn, V.D., Anton, J.E. (2007) miRBase: tools for microRNA genomics. Nucleic Acids Research, 36(50), 5.
Google Scholar: Author Only Title Only Author and Title

98. Howe, K.L., Contreras-Moreira, B., De, Silva, N., Maslen, G., Akanni, W., et al. (2019) Ensembl Genomes 2020-enabling non-vertebrate genomic research. Nucleic Acids Research, 48, D689–D695.
Google Scholar: Author Only Title Only Author and Title

99. Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., et al. (2012) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Research, 40, D1178-D1186.
Google Scholar: Author Only Title Only Author and Title

100. Adrian, M.A, Adrian, S., Gaston, H.G., Christophe, D. (2011) OMA 2011: orthology inference among 1000 complete genomes. Nucleic Acids Research, 39, D289-D294.
Google Scholar: Author Only Title Only Author and Title

101. Grant, D., Nelson, R.T., Cannon, S.B., Shoemaker, R.C. (2010) SoyBase, the USDA-ARS soybean genetics and genomics database. Nucleic Acids Research, 38, D843-D846.
Google Scholar: Author Only Title Only Author and Title

102. Eiru, K., Sohyun, H., Insuk, L. (2017) SoyNet: a database of co-functional networks for soybean Glycine max. Nucleic Acids Research, 45, D1082-D1089.
Google Scholar: Author Only Title Only Author and Title

103. Amela, P., Stefan, B., Philip, Z, Anja, W., Peter, B., et al. (2006) A systematic comparison and evaluation of biclustering methods for gene expression data. Bioinformatics, 22(9), 1122-1129.
Google Scholar: Author Only Title Only Author and Title

104. Prashant, K.S., Taraka, R.M., Priyanka, P., Ian, T.B., Shree, P.P. (2014) A comparison of performance of plant miRNA target prediction tools and the characterization of features for genome-wide target prediction. BMC genomics, 15(1), 348.
Google Scholar: Author Only Title Only Author and Title

105. Zhang, Z.B., Dunwell, J.M., Zhang, Y.M. (2018) An integrated omics analysis reveals molecular mechanisms that are associated with differences in seed oil content between Glycine max and Brassica napus. BMC Plant Biology, 18(1), 1-15.
Google Scholar: Author Only Title Only Author and Title

106. Ning, L.H., Du, W.k., Song, H.N., Shao, H.B., Qi, W.C., et al. (2019) Identification of responsive miRNAs involved in combination stresses of phosphate starvation and salt stress in soybean root. Environmental and Experimental Botany, 167, 103823.
Google Scholar: Author Only Title Only Author and Title

107. Hélène, P., Jérémy, M., Martin, C., Caroline, H., Christine, L.B. (2019) Small RNA diversity and roles in model legumes. The Model Legume Medicago truncatula, 948-962.
Google Scholar: Author Only Title Only Author and Title

108. Gupta, O.P., Dahuja, A, Sachdev, A, Kumari, S., Jain, P.K., et al. (2019) Conserved miRNAs modulate the expression of potential transcription factors of isoflavonoid biosynthetic pathway in soybean seeds. Molecular Biology Reports, 46(4), 3713-3730.
Google Scholar: Author Only Title Only Author and Title

109. Sha, A.H., Chen, Y.H., Ba, H.P., Shan, Z.H., Zhang, X.J., et al. (2012) Identification of Glycine Max MicroRNAs in response to phosphorus deficiency. Journal of Plant Biology, 55(4), 268-280.
Google Scholar: Author Only Title Only Author and Title

110. Yu, J.Y., Zhang, Z.G., Huang, S.Y., Han, X., Wang, X.Y., et al. (2019) Analysis of miRNAs Targeted Storage Regulatory Genes during Soybean Seed Development Based on Transcriptome Sequencing. Genes, 10(6), 408.
Google Scholar: Author Only Title Only Author and Title

111. Chen, L.F., Ding, X.L., Zhang, H., He, T.T., Li, Y.W., et al. (2018) Comparative analysis of circular RNAs between soybean cytoplasmic male-sterile line NJCMS1A and its maintainer NJCMS1B by high-throughput sequencing. BMC Genomics, 19(1), 1-14.
Google Scholar: Author Only Title Only Author and Title

112. Li, HY, Dong, YY, Yin, HL, Wang, N, Yang, J, et al. (2011) Characterization of the stress associated microRNAs in Glycine max by deep sequencing BMC PLANT BIOLOGY 11(1), 170-181.
Google Scholar: Author Only Title Only Author and Title

113. Senthil, S. (2012) microRNA Regulation of Symbiotic Nodule Development in Legumes. MicroRNAs in Plant Development and Stress Responses, 15, 177-195.
Google Scholar: Author Only Title Only Author and Title

114. Sun, Z.X., Wang, Y.N., Mou, F.P., Tian, Y.P., et al. (2016) Genome-Wide Small RNA Analysis of Soybean Reveals Auxin-Responsive microRNAs that are Differentially Expressed in Response to Salt Stress in Root Apex. FRONTIERS IN PLANT SCIENCE, 6, 1273.
Google Scholar: Author Only Title Only Author and Title

115. Zheng, Y., Hivrale, V., Zhang, X., Valliyodan, B., Lelandais-Briere, C., et al. (2016) Small RNA profiles in soybean primary root tips under water deficit. BMC systems biology, 10, 126.
Google Scholar: Author Only Title Only Author and Title

116. Xu, S.C, Liu, N., Mao, W.H., Hu, Q.Z., Wang, G.F., Gong, Y.M. (2016) Identification of chilling-responsive microRNAs and their targets in vegetable soybean (Glycine max L.). SCIENTIFIC REPORTS, 6.
Google Scholar: Author Only Title Only Author and Title

117. Dong, Z.H., Shi, L., Wang, Y.W., Chen, L., Cai, Z.M., et al. (2013) Identification and Dynamic Regulation of microRNAs Involved in Salt Stress Responses in Functional Soybean Nodules by High-Throughput Sequencing. Int J Mol Sci, 14(2), 2717-2738.
Google Scholar: Author Only Title Only Author and Title

118. Sonia, C.B., Roseeta, D.M., Shivani, K., Santosh, K., Saloni, M., Saurabh, R. (2015) Insights into the Small RNA-Mediated Networks in Response to Abiotic Stress in Plants. Elucidation of Abiotic Stress Signaling in Plants, 45-91.
Google Scholar: Author Only Title Only Author and Title

119. Senthil, S., Yan F., Ramanjulu, S., W, B.B., Zhu, J.K., Oliver, Y. (2008) Novel and nodulation-regulated microRNAs in soybean roots. BMC genomics 9,160.
Google Scholar: Author Only Title Only Author and Title

120. Zhang, S.L., Wang, Y.N., Li, K.X., Zou, Y.M., Chen, L., Li, X. (2014) Identification of Cold-Responsive miRNAs and Their Target Genes in Nitrogen-Fixing Nodules of Soybean. INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES, 15(8), 13596-13614.
Google Scholar: Author Only Title Only Author and Title

121. Kulcheski, F.R., de Oliveira, L.F., Molina, L.G. et al. (2011) Identification of novel soybean microRNAs involved in abiotic and biotic stresses. BMC Genomics, 12, 307.
Google Scholar: Author Only Title Only Author and Title

122. Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. PloS one, 6(7), e21800. 123. Lex, A, Gehlenborg, N., Strobelt, H., Vuillemot, R., & Pfister, H. (2014). UpSet: Visualization of Intersecting Sets. IEEE transactions on visualization and computer graphics, 20(12), 1983-1992.
Google Scholar: Author Only Title Only Author and Title