1 **Fine-mapping identifies *NAD-ME1* as a candidate underlying**
2 **a major locus controlling temporal variation in primary and**
3 **specialized metabolism in Arabidopsis**

4

5 **Marta Francisco¹, Daniel J. Kliebenstein^{2,3}, Víctor M. Rodríguez¹, Pilar Soengas¹,**
6 **Rosaura Abilleira¹, María E. Cartea¹**

7 1Misión Biológica de Galicia, (MBG-CSIC), P.O. Box 28, 36080, Pontevedra, Spain

8 2Department of Plant Sciences, University of California at Davis, Davis, CA 95616,
9 USA.

10 3DynaMo Center of Excellence, University of Copenhagen, Thorvaldsensvej 40, DK-
11 1871 Frederiksberg C, Denmark.

12

13 **Summary**

14 Plant metabolism is modulated by a complex interplay between internal signals and
 15 external cues. A major goal of all quantitative metabolomic studies is to clone the
 16 underlying genes to understand the mechanistic basis of this variation. Using fine-scale
 17 genetic mapping, in this work we report the identification and initial characterization of
 18 *NAD-DEPENDENT MALIC ENZYME 1* (*NAD-MEI*) as the candidate gene underlying
 19 the pleiotropic network Met.II.15 QTL controlling variation in plant metabolism and
 20 circadian clock outputs in the Bay × Sha Arabidopsis population. Transcript abundance
 21 and promoter analysis in *NAD-MEI*^{Bay-0} and *NAD-MEI*^{Sha} alleles confirmed allele-
 22 specific expression that appears to be due a polymorphism disrupting a putative
 23 circadian cis-element binding site. Analysis of T-DNA insertion lines and
 24 heterogeneous inbred families (HIFs) showed that transcript variation of the *NAD-MEI*
 25 gene led to temporal shifts of tricarboxylic acid cycle (TCA) intermediates,
 26 glucosinolate (GSL) accumulation and altered regulation of several GSL biosynthesis
 27 pathway genes. Untargeted metabolomics analyses reveal complex regulatory networks
 28 of *NAD-MEI* dependent upon the day-time. The mutant lead to shifts in plant primary
 29 metabolites, cell-wall components, isoprenoids, fatty acids and plant immunity
 30 phytochemicals, among others. Our findings suggest that *NAD-MEI* may act as a key
 31 component to coordinate plant primary and secondary metabolism in a time-dependent
 32 manner.

33

34

35 **Keywords:** fine-mapping, glucosinolates, heterogeneous inbred family, *NAD-*
 36 *DEPENDENT MALIC ENZYME 1*, plant metabolism, diurnal rhythms, Arabidopsis.

37 **Introduction**

38 Plants produce large arsenals of structurally and diverse metabolic pathways with
 39 metabolites being generally categorized as primary metabolites or secondary
 40 metabolites (1). Primary metabolites, including those of the tricarboxylic acid cycle
 41 (TCA), sugars, amino acids and lipids are involved in central processes of growth and
 42 development producing all the necessary building blocks for cells and the resulting
 43 biomass. Secondary metabolites such as phenolics, terpenoids, alkaloids, and
 44 glucosinolates (GSLs) are specialized compounds that play critical roles in plant
 45 adaptation under stressful environmental events (2, 3). Given the central importance of
 46 metabolism in any physiological or developmental process, metabolic networks must be
 47 fine-tuned to make the most efficient use of available resources. The plant metabolome,
 48 however, is not a stable entity. Fluxes and metabolite concentrations are variables while
 49 enzyme activities are under genetic control (4). The precise coordination of metabolic
 50 fluxes in response to diverse endogenous signals and environmental conditions has led
 51 to strong interest in understanding how plant metabolism is regulated (5).

52 One of the key components in integrating environmental signals and regulating
 53 metabolite levels is the circadian clock (6). This coordination is critical for local
 54 adaptation because it allows plants to anticipate predictable daily environmental
 55 changes like sunrise and sunset and adjust development and physiology (7). In the
 56 model plant, *Arabidopsis*, several works revealed that transcript amounts of key
 57 enzymes from central metabolism are regulated by the circadian clock (8, 9). This
 58 regulation coordinates primary metabolite levels, reserve turnover, and growth with the
 59 daily cycle of light and darkness (10-15). In addition, evidence that specialized
 60 metabolites levels can oscillate rhythmically during a day is beginning to accumulate
 61 (16-20). An example of this are the GSLs, a major class of sulphur-containing
 62 secondary metabolites present in the members of family *Brassicaceae* that provide
 63 resistance against pathogens and insects (21). These compounds show a circadian
 64 pattern of accumulation in *Arabidopsis* and in crops from the same family (20, 22, 23).
 65 This suggests a complex and intricate biological timing mechanism that governs
 66 metabolism daily behavior. Thus, unraveling the genetic basis of this variation can
 67 remarkably enhance our understanding of plant integral regulatory systems for
 68 metabolic traits (24).

Genetic connections between both primary and secondary metabolism and circadian outputs begin to be elucidated in Arabidopsis. The study of metabolomic variation in the Bayreuth-0 (Bay-0, CS954) × Shahdara (Sha, CS929) recombinant inbred line (RIL) population identified a number of genomic regions associated with metabolite variation, namely metabolic quantitative trait loci (Met.QTL) (25, 26). Further transcriptomic analysis of the same RIL population revealed that several QTLs for natural variation altering primary metabolism were linked to the expression of the circadian clock output networks (Kerwin *et al.*, 2011). Interestingly, seven of those QTL clusters (Met.chromosome.centimorgan [cM]), Met.II.15, Met.II.47, Met.III.04, Met.IV.65, Met.V.67, AOP and ELONG found for the circadian clock and metabolomics collocated with previously identified cis-eQTL (natural variation in gene expression that map to their physical position) hot spots for GSL pathway related transcripts (West *et al.*, 2007; Wentzell *et al.*, 2007). Molecular cloning of major metabolic/clock QTLs are beginning to show that metabolomics variation may often identify key physiological regulators that are important control nodes for the metabolome with pleiotropic effects across plant circadian rhythms. In support of this, the AOP and ELONG loci play a crucial role in GSL biosynthesis determining aliphatic GSL accumulation and structure but also, AOP is involved in control of onset of flowering and circadian period alteration (27, 28). Similarly, Met. II.47 QTL caused by natural variation between the Bay and Sha alleles at *EARLY FLOWERING3 (ELF3)*, has been linked to altering flowering time, shade-avoidance, hypocotyl elongation, circadian clock oscillation and metabolic variation (29-31). Besides, the Met.V.67 locus, controlled by natural variation in *METABOLIC NETWORK MODULATOR 1 (MNMI)*, affected primary metabolism accumulation within Arabidopsis in a time-dependent fashion (32). Thus, cloned genes are illuminating new and unexpected mechanistic aspects of plant biology.

To provide new insights into the modular behavior of metabolic processes and their regulation, the objective of the present work was to identify and clone candidate genes in the genomic region of the pleiotropic network Met.II.15 QTL controlling variation in plant metabolism and circadian clock outputs within Arabidopsis. To accomplish this, we utilized a combination of fine-scale genetic mapping, using heterogeneous inbred families (HIFs), and *in silico* candidate gene identification. These analyses detected *NAD-DEPENDENT MALIC ENZYME 1 (NAD-MEI)* as the most likely candidate gene underlying the Met.II.15 QTL. Transcript abundance and promoter analysis in *NAD-*

102 *MEI*^{Bay-0} and *NAD-MEI*^{Sha} alleles confirmed allele-specific expression that appears to
 103 be due a polymorphism disrupting a putative circadian cis-element binding site.
 104 Analysis of T-DNA insertion lines and HIFs showed that transcripts variation of the
 105 *NAD-MEI* gene led to temporal shifts of tricarboxylic acid (TCA) intermediates, GSL
 106 accumulation and altered regulation of several GSL biosynthesis pathway genes.
 107 Untargeted metabolomic analyses reveal complex regulatory networks of *NAD-MEI*
 108 dependent upon the day-time. The mutant lead to shifts in plant primary metabolites,
 109 cell-wall components, isoprenoids, fatty acids and plant immunity phytochemicals,
 110 among others. Together, these data show that *NAD-MEI* may act as a key component to
 111 coordinate plant primary and secondary metabolism in a time dependent manner.

112 **Materials and Methods**

113 **Mapping population and fine-mapping strategy**

114 To confirm the effects of the locus in an isogenic background Near Isogenic Lines
 115 (NILs) were developed using a HIF following the idea published by Tuintra *et al.* (33).
 116 This powerful approach uses the residual heterozygosity in early generations of RILs.
 117 The Bay-0 × Sha RIL population at F6 showed approximately 97% homozygosity in
 118 each line (420 lines in total). This resulted in the presence of residual heterozygosity in
 119 at least a single RIL at almost all genome positions. In the present study homozygous
 120 NILs were created from the progeny of the line RIL364 (from the Bay-0 × Sha
 121 population) that show a single residual heterozygous region around marker MSAT2.28
 122 where Met.II.15 QTL was originally mapped (25-27, 33). Seeds of RIL364 at F6 stage
 123 were kindly provided by Dr. Olivier Loudet (INRA, Versailles, France). For fine
 124 mapping, a homozygous RIL364 [Bay-0] and RIL364 [Sha] pair was crossed to recreate
 125 heterozygosity (HIF364) at Met.II.15 QTL in an isogenic background and the F1 selfed
 126 to make F2. PCR-based markers flanking Met.II.15 were developed and used to screen
 127 1,600 F2 progeny to isolate individuals fixed for the parental alleles in the interval as
 128 well as possible recombinants within the heterozygous region between markers
 129 MSAT200897, MSAT2.28 and IND268. All recombinant progeny were then self-
 130 fertilized to generate F3 seeds to validate the F2 measurements. To more precisely
 131 determine the recombination breakpoints we used data sequence from 1001 Genomes
 132 project (34) to computationally predict INDELs between Bay-0 and Sha. With this
 133 approach, additional markers between markers MSAT2.28 and IND268 were developed
 134 and used to genotype all recombinants. Primers were designed using software Primer

135 3.0 (<http://bioinfo.ut.ee/primer3-0.4.0/>). DNA was extracted using the Promega DNA
136 Purification System. PCR products were run in 2% agarose gel stained with EtBr and
137 the genotype of the plant was assessed. The primers used for the PCRs are listed in
138 Table S1.

139 **T-DNA insertion lines**

140 The T-DNA insertion line Sail-374-A02 (At2g13560), was kindly provided by Dr.
141 Tronconi (Universidad Nacional de Rosario, Rosario, Argentina). Additional T-DNA
142 insertion lines SALK_024285 (AT2g13540), SALK_090811 (At2g13560) and
143 SALK_126034 (At2g13610) were obtained from the Nottingham Arabidopsis Stock
144 Center ([http:// www.arabidopsis.info/](http://www.arabidopsis.info/)). Lines with homozygous insertions were
145 identified by PCR-based genotyping and corroborated by RT-qPCR. *NAD-MEI*
146 mutant lines do not showed altered expression of the neighboring At2g13550 gene.
147 Either *NAD-MEI* mutants not showed visual phenotypic alterations under normal
148 growth conditions (Fig. S1). Since both *NAD-MEI* mutants tested displayed similar
149 phenotypes, results presented here are based on Sail-374-A02 (*NAD-MEI*) mutant.
150 Analysis of variance (ANOVA) was used to compare GSL variation between T-DNA
151 lines from the three candidate genes (At2g13560, At2g13540 and AT2g13610) and Col-
152 0 (wild-type). Each genotype account with 24 independent measurements conducted
153 across two independent experiments.

154 **Plant growth and harvest conditions**

155 All seeds were planted in potting soil stratified at 4°C in the dark for 3 days to optimize
156 germination using 3.80 x 3.80 cm pots. For fine-mapping analysis, all plants from the
157 appropriate genotypes were grown in long-day (LD) conditions (16/8h light/dark)
158 photoperiods in a growth chamber under controlled conditions (100–120 µE light
159 intensity and 22°C of temperature and 60% of humidity). Rosette leaves were used for
160 all experiments. To test different *NAD-MEI* allele's effects upon cycling within the
161 GSL network the appropriate genotypes were tested under LD photoperiod and
162 harvested after three weeks of germination. In all the experiments, plants were sown in a
163 randomized block design with two complete independent replicates of the whole
164 experiment. Sampling was performed at various Zeitgeber times (ZT), starting just
165 before dawn (ZT0). To study the effect of *NAD-MEI* alleles on transcript levels of GSL
166 related genes as well as for metabolomics analysis, samples were harvested at ZT4 and

167 ZT8. In each experiment, three individual plants per genotype were harvested at each
168 time point for sampling. This provides six total samples per genotype per time point
169 across the whole dataset. For harvesting, leaves were cut placed in plastic scintillation
170 vials and immediately frozen in liquid nitrogen. During the dark phase, the plants were
171 sampled in the presence of low-intensity green lamp. Plant material was stored at -80
172 °C.

173 **Real time PCR**

174 Total RNA was extracted using the RNeasy kit (QIAGEN, Germany). To remove any
175 traces of genomic DNA, the RNA was treated with DNase by following the
176 manufacturer instructions. Samples of RNA were reverse transcribed using the
177 GoScript™ reverse transcription system and oligo (dT20) (Promega). RT-qPCR was
178 performed in a 20 µl reaction with the Fast Start Universe SYBR Green Master (ROX)
179 mix (Roche, IN, USA). The expression levels were normalized to glyceraldehyde-3-
180 phosphate- dehydrogenase (GADPH). RT-qPCRs were carried out on a 7500 Real Time
181 PCR System (Applied Biosystem, Forster City, CA, USA) and primer efficiency was
182 calculated using the LingRegPCR software (35). Statistical significance was calculated
183 using a Student's t test to compare the relative gene expression among genotypes. All
184 the primer pairs used are listed in the Table S1.

185 **GSL identification and quantification.**

186 Sample extraction and desulfation, were performed according to Kliebenstein *et al.* (36)
187 with minor modifications. Ten microliters of the desulfo-GLS extract were used to
188 identify and quantify the GSLs. The chromatographic analyses were carried out on an
189 Ultra-High-Performance Liquid-Chromatograph (UHPLC Nexera LC-30AD;
190 Shimadzu) equipped with a Nexera SIL-30AC injector and one SPD-M20A UV/VIS
191 photodiode array detector. The UHPLC column was aXSelect HSS T3 XP ColumnC18
192 protected with a C18 guard cartridge. The oven temperature was set at 30°C.
193 Compounds were separated using the following method in aqueous acetonitrile, with a
194 flow of 0.5 mL min⁻¹: 1.5 min at 100% H₂O, an 11 min gradient from 0% to 25% (v/v)
195 acetonitrile, 1.5 min at 25% (v/v) acetonitrile, a minute gradient from 25% to 0% (v/v)
196 acetonitrile, and a final 3 min at 100% H₂O. Data was recorded on a computer with the
197 LabSolutions software (Shimadzu). All GSLs were quantified at 229 nm by using
198 glucotropaeolin (GTP, monohydrate from Phytoflan, Diehm & Neuberger GmbH,
199 Heidelberg, Germany) as internal standard and quantified by comparison to purified

standards. We reported the concentration ($\mu\text{mol cm}^2$ of FW) of individual GSL compounds as well as the sums of total aliphatic and indolic GSLs, two classes of GSLs based on the amino acid from which they have derived. ANOVA was used to compare individual and total GSL variation between Bay-0 and Sha alleles of each molecular marker at Met.II.15 QTL region. Multiple comparisons comparing GSL traits between genotypes were made post-hoc using Tukey's t-test with $P \leq 0.05$ within the model.

Co-expression network analysis

Co-expression data for each studied gene was obtained from ATTED-II version c4.1 that was calculated including 1388 microarray experiments (37). To filter candidate genes in the final interval QTL we search for significant correlations between expression variation of candidate genes, GSL-related genes and circadian core clock genes. To construct the network, we filtered the first 30 genes correlated with *NAD-ME1*. Among them, we selected only the genes with a cis-eQTL in the Bay-0 $\times$ Sha RIL population (26). The selected genes were used as bait genes. Then, co-expression relationship between the bait genes and their directly-connected genes were used to construct a network visualized using NetworkDrawer (38). Genes were subsequently interpreted by functional enrichment analysis based on Gene Ontology (GO) to identify and rank overrepresented functional categories in the network.

Metabolomic analysis

For metabolomic composition analysis we used an ultra-performance liquid chromatography coupled with electrospray ionization quadrupole (Thermo Dionex Ultimate 3000 LC) time-of-flight mass spectrometry (UPLC-Q-TOF-MS/MS) (Bruker Compact™) with a heated electrospray ionization (ESI) source. Chromatographic separation was performed in a Kinetex™ C18 LC Column (2.1 $\times$ 100 mm 1.7 μm pore size) using a binary gradient solvent mode consisting of 0.1% formic acid in water (solvent A) and acetonitrile (solvent B). The following gradient was used: 3 % B (0-4 min), from 3% to 25 % B (4-16 min), from 25 to 80% B (16-25min), from 80 to 100% B (25-30 min), hold 100% B until 32 min, from 100% to 3% B (32-33 min), hold 3% B until 36 min. The injection volume was 5 μL , the flow rate was established at 0.4 ml/min and column temperature was controlled at 35 °C. MS analysis was operated in spectra acquisition range from 50 to 1200 m/z. Both polarities ($\pm$) of ESI mode were used under the following specific conditions: gas flow 9 l/min, nebulizer pressure 38 psi, dry gas 9 l/min, and dry temperature 220 °C. Capillary and

end plate offset were set to 4500 and 500 V, respectively. MS/MS analysis was performed based on the previously determined accurate mass and RT and fragmented by using different collision energy ramps to cover a range from 15 to 50 eV. The algorithm T-Rex 3D from the MetaboScape 4.0 software (Bruker Daltoniks, Germany) was used for peak alignment and detection. The generated dataset was imported into Metaboanalyst (39) to perform statistical analyses. In order to remove non-informative variables, data was filtered using the interquartile range filter (IQR). Moreover, Pareto variance scaling was used to remove the offsets and adjust the importance of high- and low-abundance ions to an equal level. The resulting three-dimensional matrix (peak indices, samples and variables) was further subjected to multivariate data analysis. Partial least squares discriminant analysis (PLS-DA) was carried out to investigate and visualize the pattern of metabolite changes. This analysis was applied to obtain an overview of the complete dataset and discriminate those variables that are responsible for variation between groups. The PLS-DA model was evaluated through a cross-validation (R^2 and Q^2 parameters). The differential features were selected according to the VIP (variable importance in projection) list obtained from PLS-DA analysis. The PLS-DA was combined with univariate analysis (one-way ANOVA) with a P-value ≤ 0.05 to find differentially expressed metabolites. Using the Volcano Plot (VP) approach, which measure differentially accumulated metabolites based on t statistics and fold changes simultaneously, we also highlighted the metabolites with a $-1 \leq \log FC \leq 1$ and statistically significant difference (P-value < 0.05) between genotypes at each time point. Identification of putative metabolites was performed using accurate mass metabolites reported in different publicly available databases such as METLIN, KEGG, Pubchem, HMDB and Plant Metabolic Network. Additionally, further partial identification of the most significant metabolites was made by comparison of MS/MS fragmentation patterns against reference compounds found in databases such as METLIN.

Results

Validation of Met.II.15 QTL and high-resolution mapping using HIFs

We initially focused on the GSL related phenotype of the Met.II.15 QTL to begin identifying the causal gene. However, the Met.II.15 QTL does not overlap any known biosynthetic or regulatory GSL genes. Thus, to facilitate map-based cloning of

265 Met.II.15, we first validated it using HIF focusing on GSL phenotype (Fig.1A). To
 266 avoid temporal metabolite variation, in all the phenotype experiments, samples for
 267 GSLs analysis were harvested at mid-day. An initial screen of 200 plants from the
 268 HIF364 [Bay-0 × Sha] progeny, comparing plants homozygous for the Sha allele with
 269 plants homozygous for the Bay-0 allele and heterozygous at the QTL region, confirmed
 270 the phenotypic impact of Met.II.15 QTL on GSLs accumulation (Fig. 1B; Table S2). As
 271 previously found in the QTL mapping, total and individual aliphatic and indolic GSLs
 272 content of Met-HIF364 [Sha] lines was consistently higher than those of Met-HIF364
 273 [Bay] at marker MSAT2.28 where the QTL was mapped originally. Lines heterozygous
 274 for HIF364 showed intermediate aliphatic and total GSL phenotypes, suggesting a
 275 semi-dominant relationship between the Bay-0 and Sha alleles (Fig. 1B). We then
 276 screened 1600 F₂ plants and identified 32 recombination events in the initial
 277 heterozygous region of ~6 Mb between markers MSAT2008, MSAT2.28 and IND628.
 278 By recurrent genotypic selection and phenotypic analysis of recombinants within the
 279 QTL interval we selected eight lines since they all had a recombination break point
 280 between MSAT2.28 and IND628 and segregated for the GSL genotype (Fig. 1C; Table
 281 S3). Fine-mapping of Met.II.15 by adding new markers in the target region and
 282 progeny testing of lines homozygous for the different recombinant alleles, HIF364 [Bay
 283 x Sha] F₃, narrowed the Met.II.15 locus to the region between markers M.II.D and
 284 M.II.E (Fig. 1C, D). Five recombinant lines (R4–R8) and control parental lines
 285 homozygous for Sha in the target region showed increased GSL content. In contrast,
 286 three recombinant line (R1–R3) and control parental line homozygous for Bay in the
 287 target region showed decreased GSL content (Fig. 1D).

288 **Identification of a candidate gene for the Met.II.15 QTL**

289 The chromosomal region delimited by the M.II.D and M.II.E interval in the Met.II.15
 290 region confidence interval includes 14 predicted protein coding genes (Fig. 2E; Table
 291 1). Accordingly to previous Bay-0 × Sha eQTL analysis (26) we anticipate natural
 292 variation in gene expression for Met.II.15 causative gene. Thus, we focused only on the
 293 genes with evidence for a cis-eQTL. This analysis reduced the candidates to three
 294 genes: *ABA HYPERSENSITIVE 1* (At2g13540; *ABH1*), *NAD-DEPENDENT MALIC*
 295 *ENZYME 1* (At2g13560; *NAD-ME1*) and *ABC TRANSPORTER G FAMILY MEMBER*
 296 *5* (AT2g13610; *ABCG5*). From them, in agreement with the overlapping circadian
 297 network, *NAD-ME1* and *ABCG5* had also a diel expression pattern (40). To confirm the

phenotypic effects, GSL levels were evaluated in T-DNA insertion lines from each of the candidate genes with cis-eQTL (Fig. 1F; Table S4). Results evidenced that only *NAD-MEI* knockout mutant showed significant effects on GSLs levels. Simultaneously, we constructed co-expression networks by searching for connections between candidates and genes that are affiliated with GSL metabolism and circadian clock. Results corroborated that only *NAD-MEI* knockout mutant tightly co-expressed with SUR1 and SUR2, central components of the GSLs biosynthetic pathway. Moreover, *NAD-MEI* co-expressed with CRY2 (Cryptochrome 2) and PHYA (Phytochrome A), key components of the circadian oscillator complex (41). Thus, we hypothesize that *NAD-MEI* is the most likely candidate gene underlying the Met.II.15 QTL.

Our candidate, *NAD-MEI* gene, encodes a malic enzyme (ME) (EC 1.1.1.39) found in mitochondria of all cells (42). MEs are widespread in all kinds of organisms and catalyzes the reversible oxidative decarboxylation of malate to pyruvate, CO₂, and NAD(P)H in the presence of a divalent metal ion (43). It has been reported that MEs play a central role in the management of flux through the TCA cycle (44). However, to our knowledge *NAD-MEI* gene was not related before with GSL metabolism and/or circadian rhythmicity. Hence, in the next sections, several experiments were carried out in order to elucidate the interplay between plant metabolome and daily rhythms mediated by *NAD-MEI*.

***NAD-MEI* alleles show altered diel regulation**

To measure the potential influence of diurnal variation in conjunction with genetic variation, we studied the oscillation patterns of *NAD-MEI* gene expression in the progeny of the recombinant lines R3 (*NAD-MEI*^{Bay-0}) and R4 (*NAD-MEI*^{Sha}) carrying either parental allele in the final target region as well as in the Col-0 natural accession (*NAD-MEI*^{Col-0}). RT-qPCR showed that the *NAD-MEI*^{Sha} and *NAD-MEI*^{Col-0} transcript levels are differently expressed between day and night (Fig. 2A). *NAD-MEI* transcript levels from Col-0 and Sha alleles exhibited their peaks from evening to early night and troughs at dusk suggesting circadian clock regulation. In contrast *NAD-MEI*^{Bay-0} transcripts were significantly lower than those of *NAD-MEI*^{Sha} and *NAD-MEI*^{Col-0} and lost the time-of-day-specific peak expression.

Further sequencing of the Bay-0 and Sha promoter and body gene of *NAD-MEI* revealed 51 SNPs and one deletion between Bay and Sha allele (Fig. S2). There were

three SNPs within the predicted body of the RNA but these did not affect any key splicing sites or amino acid change. Thus, it is more likely that cis-eQTL is caused by the variation between the Bay and Sha promoters. To investigate that, *NAD-MEI* promoter sequences were analyzed through online program PLACE (www.dna.affrc.go.jp/PLACE/). Results showed that the 1.6 kb sequence upstream of the translation initiation codon ATG contains 20 types of putative cis-acting elements (Table 2). Supporting the idea that the clock regulates *NAD-MEI* transcripts, we found cis-regulatory elements known as the evening element (EE; AAAATATCT), which was identified in the promoters of Arabidopsis clock-controlled genes that are co-expressed late in the day (9), one putative circadian regulation element (CAANNNNATC) and two *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA-1*) binding sites (CBS; AAATCT). Interestingly, we found that within 546 bp of the transcription start site a SNP in *NAD-MEI* Bay-0 allele promoter sequence may disrupt the putative CBS/MYB motif (Fig. 2B). It is remarkable that in close proximity to this sequence, is localized the G-box core element (Fig. 2B; Table 2). Typically, G-box and CBS motifs are part of light/circadian regulated gene promoters and usually cooperate in defining the transcript oscillation properties (45). In addition, disrupted sequence in Bay allele may be also affecting the binding of C2H2 zinc-finger proteins which are involved in participate in various aspects of normal plant growth and development, as well as in environmental stress tolerance regulation (46). Although further work is required to identify the specific transcription factor that leads to the alterations in *NAD-MEI* gene expression, we hypothesized that difference of *NAD-MEI* expression among alleles is likely to be caused by polymorphisms in cis-acting elements at the promoter.

Amplitude of the GSLs diel variation was affected by *NAD-MEI* alleles

The finding that *NAD-MEI* transcripts could be regulated in a diel manner prompted us to test if this also leads to an alteration in the temporal variation in GSL accumulation during the day. To test if *NAD-MEI* controls diel variation in GSL metabolism, we carried out a time course under LD cycle conditions in three-weeks-old plants, sampling every 4 hours over 1 day in the *NAD-MEI*^{Bay-0} and *NAD-MEI*^{Sha} genotypes, as well as in the knockout mutant line of the candidate gene *NAD-MEI* (*nad-mei*) and in the wild type, *NAD-MEI*^{Col-0}. Results showed that the amplitude of the GSLs diel variation was higher in the *nad-mei* and *NAD-MEI*^{Bay-0} genotypes compared with *NAD-MEI*^{Col-0} and *NAD-MEI*^{Sha} alleles respectively (Fig. 2C, D). In general, during

the first hours of light, with a maximum at ZT4, the *nad-me1* mutants and *NAD-MEI*^{Bay-0} accumulated significantly more GSLs than *NAD-MEI*^{Col-0} and *NAD-MEI*^{Sha} lines respectively. On the contrary, at mid-day (ZT8) GSL content in *nad-me1* mutant and *NAD-MEI*^{Bay-0} lines showed a clear reduction in GSL content. Therefore, different alleles of *NAD-MEI* affect the amplitude of GSLs diel variation.

To further investigate the influence of *NAD-MEI* on GSLs accumulation, we studied transcript levels of GSL related genes at ZT4 and ZT8 (Fig. 2E; Table S5). Regulation of numerous GSL biosynthesis genes was strongly affected in the *nad-me1* mutant and in the recombinant line carrying Bay allele. In general, down-regulation of *NAD-MEI* transcript levels was related with up-regulation of GSL related transcript genes during morning while dampening during the evening. Thus, expression of *NAD-MEI* is correlated with the GSL metabolism suggesting that *NAD-MEI* is a strong candidate for Met.II.15 cis-eQTL.

Untargeted metabolomics revealed that *NAD-MEI* loss of function induces time-specific metabolome changes in Arabidopsis

Since *NAD-MEI* was assumed to play a central role in the management of flux through the TCA cycle by coordination of the carbon and nitrogen metabolisms in Arabidopsis (47) and Met.II.15 co-localizes with a major QTL controlling central metabolism associated with the TCA cycle (25), we suggest that *NAD-MEI* may be also affecting broad metabolomics changes. Thus, we carried out an untargeted metabolomics analysis using UPLC-Q-TOF-MS/MS to evaluate the effect of *NAD-MEI* loss of function on plant metabolome. We analyzed global trends of metabolite variations by searching for changes that occur in leaves of *nad-me1* mutants compared with Col-0 at ZT4 and ZT8. These time points were chosen because wider amplitude on GSLs variation was found in previous sections. PLS-DA score plots showed a clear separation between Col-0 and *nad-me1* group plant genotypes (Fig. 3A). We detected 57 and 81 differentially expressed metabolites among *nad-me1* mutant and Col-0 at ZT4 and ZT8, respectively (Table S6). Among the 138 ions significantly altered, only 15 were coincident at both time points studied (Fig. 3B). Thus, close to 90% of the significant altered ions among genotypes, including TCA intermediates, were dependent upon the day-time, which evidenced complex temporal metabolomics changes in Arabidopsis *NAD-MEI* network (Fig. 3C). Metabolomics comparison of *NAD-MEI*^{Bay-0} and *NAD-MEI*^{Sha} genotypes corroborated the temporal shift on primary metabolites related with TCA cycle (Fig.

397 S3). At ZT4, the time point equivalent to that used for the QTL mapping sample, the
398 direction of effect agreed with the previously published analysis from Met.II.15 in the
399 Bay $\times$ Sha RIL population (25). This suggests that *NAD-MEI* plays a role in mediating
400 the temporal partitioning of TCA intermediates within the Arabidopsis leaf.

401 Rather than TCA intermediates itself, lack of *NAD-MEI* caused alterations in many
402 other metabolites involved in a wide a wide range of metabolic processes (Table S6). At
403 ZT4, in terms of fold change, the major differences among genotypes were found for
404 isoprenoids potentially involved in plant photosynthesis and cell-wall precursors. At
405 ZT8, amino acids and their derivatives were the most altered compounds. In addition,
406 several compounds related with phospholipid metabolism, fatty acids, cysteine and
407 glutathione conjugates, flavonoids, and plant immunity compounds (phytoalexins,
408 isothiocyanates and alkaloids) were altered in the mutant in a time-dependent fashion
409 (Fig. 3C; Table S6).

410 ***NAD-MEI* regulatory network construction**

411 To understand the molecular mechanisms behind the *NAD-MEI* effect on plant
412 metabolism, we worked to identify potential co-regulated modules of our candidate
413 gene which may explain metabolic changes between genotypes. To achieve that, we
414 first searched for genes closely expressed with *NAD-MEI* through the ATTEDII co-
415 expression database and then filtered for genes with a cis-eQTL in the Bay-0 $\times$ Sha RIL
416 population (26). Selected genes were used as bait genes to construct the co-expression
417 network (Fig. 4; Table S7).

418 Studying Gene Ontology (GO) terms in the ‘biological process’ annotation, revealed
419 over-representation of genes involved in primary carbon metabolism, response to
420 stimulus, amino acid metabolism, N-glycan biosynthesis, cell-wall remodeling and
421 phospholipid metabolism . Remarkably, 73% of the genes from the *NAD-MEI* co-
422 expressed network with an eQTL were assigned to plant response to different stimulus.
423 This set included several genes shown to be involved in cell redox homeostasis, plant
424 hormone signal transduction, response to sucrose stimulus and systemic acquired
425 resistance related genes (Table S7). The largest hubs within the network were the
426 *PROLYL 4-HYDROXYLASE 5* (*P4H5*; At2g17720), the *ISOCITRATE*
427 *DEHYDROGENASE 1* (*IDH1*; AT4g35260) and *HAP6* (At4g21150). The *P4H5*
428 catalyze an important post-translational modification required for proper cell-wall self-
429 assembly and protein stability of transcription factors in response to biotic and abiotic

plant stresses and hence involved in plant defense (48). The *IDH1* performs an essential role in cellular defense against oxidative stress-induced damage of the TCA cycle (49). Lastly, the *HAP6* gene is a member of the oligosaccharyltransferase (OST) multi-membrane protein complex, with a role in protein N-linked glycosylation and involved in cellulose biosynthesis in response to abiotic stress (50). All together, these results suggest that impaired *NAD-MEI* expression is correlated with various plant cell stress-sensing mechanisms affecting a combination of different signal transduction pathways.

Discussion

During the last decade, increasing evidence suggest the existence of a complex interplay between plant metabolic pathways and daily accumulation of a large range of plant metabolites, including GSLs (16-20, 27, 32). However, identification and cloning of new loci controlling quantitative traits related to both plant metabolism and diel patterns have been less studied. In this work, we created NILs developed from HIFs that only segregates at a small region around the Met.II.15 QTL controlling temporal variation in primary and specialized metabolism in the Arabidopsis Bay-0 × Sha population. This approach facilitated map-based cloning and successfully identified *NAD-MEI*, involved in the decarboxylation of malate to pyruvate, as the most likely candidate gene underlying the Met.II.15 QTL leading to temporal shifts of TCA intermediates, GSL accumulation and altered regulation of several GSL biosynthesis pathway genes in a time-dependent manner.

The availability of specific recombinants of RILs containing *NAD-MEI* allelic variation allowed us to confirm the light/circadian dependent cis-eQTL on transcript abundance (Fig.2A). Transcripts of *NAD-MEI* from Col-0 and Sha alleles peaks from evening to early night, corroborating a major role of *NAD-MEI* during night mitochondrial metabolism (51). Transcripts from Bay-0 allele were significantly down-regulated and lost the time-of-day-specific peak expression, indicating that these differences in expression among alleles are likely to be caused by polymorphisms in light/clock-controlled cis-acting elements at the promoter. Interestingly, a polymorphism disrupting a putative CBS motif in *NAD-MEI* promoter region of Bay allele, may be potentially affecting the binding of the core clock elements *CCA1* and *LHY*. Supporting our findings, Graf et al. (12) found that *NAD-MEI* transcripts and protein levels are significantly reduced in *cca1/lhy* double mutants. These evidences suggested that *CCA1* and *LHY* are highly likely to regulate *NAD-MEI* expression.

Intriguingly, the loss of *NAD-ME1* transcript regulation was linked to temporal shifts of several TCA intermediates (Fig. 3; Fig S3; Table S6). Although the biological meaning of the linkage between circadian rhythms and TCA cycle remains unclear, Arabidopsis clock mutants showed deregulation of TCA metabolic homeostasis within plant mitochondria (11, 52). More recently, it has been reported that many TCA intermediates drove the diel changes under different natural and artificial light regimes in Arabidopsis (15). In general, these metabolites are low at dawn and high towards the end of the light period. However, we found that plants lacking *NAD-ME1* displayed enhanced levels of several TCA cycle intermediates after dawn and decreased at evening (Fig.3C; Fig. S3; Table S6). Interestingly, neither malate nor pyruvate, the substrate and product of the *NAD-ME1* reaction, were altered in plants lacking *NAD-ME1* at the time points tested. This can be explained because *NAD-ME1* is not the sole source of pyruvate in mitochondria. In accordance with this, the other NAD-dependent malic enzyme *NAD-ME2* (At4g00570) and two genes encoding *MITOCHONDRIAL MALATE DEHYDROGENASE* (*MMDH*; At1g53240 and At3g47520) can metabolize malate in the mitochondria. It has been suggested that MMDHs and NAD-MEs manage the flux of malate through the TCA cycle differentially during a diurnal cycle. MMDHs would have a prevalent role during the light period, while NAD-MEs would be more important during the night period (51). This is consistent with the suggested role of *NAD-ME1* during night mitochondrial metabolism based on our results in time-of-day-specific peak expression (Fig. 2). When *NAD-ME1* is not present, like in *nad-me1* mutant plants used in the present study, Arabidopsis leaves display half of total NAD-ME enzyme activity compared with wild-type, presumably associated with *NAD-ME2* activity (51). Therefore, we speculated that reduced NAD-ME activity during the dark period in Bay-0 and *nad-me1* produces an excess of mitochondrial TCA intermediates at the end of the night and, in consequence, many primary and secondary plant metabolites derived from TCA will be affected in a time-dependent fashion. Thus, *NAD-ME1* seems to be key factor involved in to an intricate system to modulate the flux into and out of central metabolism.

TCA cycle intermediates provide the carbon skeletons to support the biosynthesis of the majority of amino acids (53, 54). In the current work, we found that amino acids and related metabolites were altered in plants lacking *NAD-ME1*. Indeed, mutant plants displayed temporal shifts of several plant defensive compounds derived from amino

acids including GSLs, flavonoids, phytoalexins and alkaloids (Fig. 3, Table S6). To our knowledge, *NAD-MEI* has not previously been implicated in the production of plant defensive compounds. Thus, potential explanations for temporal shift of these compounds may be related with altered cysteine and glutathione conjugates pool sizes during the light hours due to the lack of *NAD-MEI* during nocturnal metabolism (Fig.3C, Table S6). Cysteine occupies a central position in plant metabolism because it is a reduced sulfur donor molecule involved in the synthesis of essential biomolecules such as the other proteinogenic amino acid methionine, vitamins, cofactors, and the antioxidant glutathione (Romero et al., 2014). Methionine is the amino acid precursor of aliphatic GSLs and glutathione plays a role as sulfur donor in GSLs formation. In addition, several intermediates of tryptophan derived metabolites including indole GSLs as well as intermediates from phytoalexins biosynthesis were altered in the *nad-mei* mutant. These findings are in line with our data from the co-expressed *NAD-MEI* network (Fig.4; Table S7). One of the most significant GO categories identified from the *NAD-MEI* network were genes involved in amino acid metabolism, including cysteine, glutathione and tryptophan related genes. Remarkably, transcript variation of most of those genes was related with plant response to plant hormones and plant-pathogen interactions with a role in the plant systemic acquired resistance response (Fig. 4, Table S7). This agrees with the fact that the regulation of several GSL biosynthesis pathway genes, including major transcriptional regulators of the aliphatic and indolic GSLs biosynthesis pathways in response to biotic and abiotic stress (*MYB28*, *MYB29*, *MYB34*) were strongly time-dependent affected in plants lacking *NAD-MEI*. These results may indicate that *NAD-MEI* down-regulation might be a decisive factor for basal and inducible levels of chemical defenses and thereby plant resistance. Further research is necessary to test whether temporal alterations in plant phytochemicals under biotic and abiotic stress might affect plant defense response in *nad-mei* mutant plants.

A closer examination of untargeted metabolomics analyses revealed that rather than TCA intermediates itself and plant defensive compounds, the lack of *NAD-MEI* caused alterations in many other metabolites involved in a wide a wide range of metabolic processes (Fig.3C; Table S6). Several precursors of cell-wall components and phospholipids, critical components of cell membranes, appear to be temporary altered in the mutant. Moreover, down-regulation of *NAD-MEI* produced reduction in photosynthetic pigments while chlorophyll degradation products increased during the

first part of the day–light hours. Similar modifications on metabolite composition that occurred in *nad–me1* mutant plants were before associated with plant responses against high light, cold or nutrient stresses among others (55-57). In concordance with that, genes related with phospholipid biosynthetic process and cell–wall organization in response to cold, osmotic stress, cell redox homeostasis and sucrose stimulus were over–represented in the *NAD–ME1* co–regulated network (Fig.4; Table S7). Since no effect on growth or development are found for *NAD–ME1* mutant plants under the tested conditions (Fig. S1) it seems that *NAD–ME1* down–regulation may initiate several mechanisms of acclimation and adaptation establishing a new metabolic homeostasis that mimics plant responses to environmental stress conditions. Results from this work will facilitate future research investigating how *NAD–ME1* would respond to diverse and dynamic environmental cues.

As a conclusion, this work identifies *NAD–ME1* as a novel actor of plant metabolomics network and provides new insights into the modular behavior of biochemical processes and their diel regulation within Arabidopsis. Transcript variation of the *NAD–ME1* gene led to temporal shifts of TCA intermediates, GSLs accumulation and altered regulation of several GSL biosynthesis pathway genes. In addition, lack of *NAD–ME1* induced time–dependent changes of plant phytochemicals related to plant immunity, initiated a complicated re–organization of plant cell–wall components and it seems to be involved in photosynthesis optimization. These findings suggest that *NAD–ME1* may coordinate plant primary and secondary metabolism in a diel dependent manner. Because some metabolites modulate the plant clock, it is needed to elucidate the interaction mechanism between the biological clock system and metabolism such as in the TCA cycle or GSLs pathways. A key future step is to investigate whether diurnal regulation of plant metabolome mediated by *NAD–ME1* may affect plant physiology, adaptation and survival under environmental stress conditions.

Acknowledgements

This effort was funded by the grant RTI2018-094650-J-I00 by the Spanish Ministry of Science, Innovation and Universities to Marta Francisco.

Author contributions

M.F., D.K., and E.C. conceived and designed the experiments. M.F., V.M.R., P.S. and R.A. assisted with the plant trials setup; M.F., and R.A. conducted laboratory work;

561 V.M.R. did the metabolomics analysis; M.F. analyzed the data and wrote the
562 manuscript. All authors read and approved the final manuscript. The authors declare that
563 they have no competing interests.

Table 1. Shown is the list of candidate genes in the interval between markers M.II.D and M.II.E for the Met.II.15 QTL (At2g13500 to At2g13650) that have either a cis-eQTL or are regulated in a circadian manner with the phase reported.

AGI	cis_eQTL	Circadian phase	Gene name	GO Biological Process
AT2G13500	no	16	Uncharacterized protein	Unknown
AT2G13510	no	no	Uncharacterized protein	Unknown
AT2G13540	yes	no	ABH1 (ABA HYPERSENSITIVE 1)	ABA signaling and flowering
AT2G13542	no	no	DEFENSIN-LIKE (DEFL) FAMILY PROTEIN	Defense response to fungus
AT2G13547	no	no	Uncharacterized protein	Unknown
AT2G13550	no	no	Uncharacterized protein	Unknown
AT2G13560	yes	8	NAD-DEPENDENT MALIC ENZYME (NAD-ME1)	Carbohydrate metabolism
AT2G13570	no	no	NF-YB7 (NUCLEAR FACTOR Y, SUBUNIT B7)	Regulation of transcription at pollen stage
AT2G13600	no	no	SLOW GROWTH 2 (SLO2)	Sugar mediated signaling pathway
AT2G13610	yes	16	ATP-BINDING CASSETTE G5 (ABCG5)	Transmembrane transport
AT2G13620	no	no	CATION/HYDROGEN EXCHANGER 15 (ATCHX15)	Cation exchanger
AT2G13630	no	21	F-BOX FAMILY PROTEIN-RELATED	Unknown
AT2G13640	no	no	Uncharacterized protein	Unknown
AT2G13650	no	no	GOLGI NUCLEOTIDE SUGAR TRANSPORTER 1 (GONST1)	Carbohydrate transport

568 **Table 2.** List of known cis-element motifs found in an interval of 1.6 kb upstream from the transcription start site (TSS) of NAD-ME1 promoter
569 in the Bay-0 and Sha alleles. Motifs are listed alphabetically including their motif name, sequence, position from the TSS and specific function.

Motif name	Motif sequence	Position from the TSS	Specific function of cis-elements
ABRE	ACGT	188, 883	ABA responsive elements
ARF	TGTCTC	394, 1007	Auxine response factor
ARR1AT	NGATT	5, 356, 385, 421, 667, 731, 829, 1058, 1062, 1233, 1277, 1674, 1988	Citokinin response regulator
CARGCW8GAT	CWWWWWWWWG	597	Binding site for AGAMOUS-like 15
CCA1ATLHCB0	AAMAATCT	1652	Circadian rhythm
CCAATBOX1	CCAAT	244, 1156, 1487, 1576	Heat shock element
CIACADIANLELHC	CAANNNNATC	1504	Circadian rhythm
DOFCOREZM	AAAG	83, 89, 116, 126, 281, 309, 605, 1321, 1435, 1562, 1583, 1613	Carbon metabolism
ERE	AWTTCAAA	288	Ethylene responsive element
ERE	AGCCGCC	1331	Ethylene responsive element
GATABOX	GATA	145, 1001, 1447	Light-dependent and tissue specific element
G-box/ABRE	CACGTG	561	Light and ABA-responsive element
GT1CONSENSUS	GRWAAW	86, 113, 317, 446, 1411	Ligth responsive element
MYB/SANT;ARR-B	AAATCT	1518, 546	Circadian rhythm
MYBCORE; E-box/ABRE	CANNTG	163, 645, 738, 951, 1031	Light and ABA-responsive element
OSE2ROOTNODULE	CTCTT	137, 488, 1159	Organ-specific element
RHERPATEXPA7	KCACGW	562, 1005	Root hair-specific cis-elements
SEBFCONSSTPR10A	YTGTCWC	395	Pathogenesis
SREATMSD	TTATCC	375	Sugar-repressive element (SRE)
SURECOREATSULTR11	GAGAC	77	Core of sulfur-responsive element (SURE)

	WBOXATNPR1	TTGAC	132, 167, 615	Salicylic acid (SA)-induced WRKY DNA binding proteins
570	<hr/> K=G/T; W=T/A; R=A/G			
571				

572 REFERENCES

- 573 1. R. A. Dixon, D. Strack, Phytochemistry meets genome analysis, and beyond.
574 *Phytochemistry* **62**, 815-816 (2003).
- 575 2. R. Akula, G. A. Ravishankar, Influence of abiotic stress signals on secondary
576 metabolites in plants. *Plant Signaling & Behavior* **6**, 1720-1731 (2011).
- 577 3. J. Kroymann, Natural diversity and adaptation in plant secondary metabolism.
578 *Current Opinion in Plant Biology* **14**, 246-251 (2011).
- 579 4. P. D. Keightley, Models of Quantitative Variation of Flux in Metabolic
580 Pathways. *Genetics* **122**, 719-719 (1989).
- 581 5. A. Gaudinier, M. Tang, D. J. Kliebenstein, Transcriptional networks governing
582 plant metabolism. *Current Plant Biology* **3-4**, 56-64 (2015).
- 583 6. J. A. Kim *et al.*, The Importance of the Circadian Clock in Regulating Plant
584 Metabolism. *International journal of molecular sciences* **18**, 2680 (2017).
- 585 7. S. L. Harmer, The Circadian System in Higher Plants. *Annual Review of Plant*
586 *Biology* **60**, 357-377 (2009).
- 587 8. K. Greenham, C. R. McClung, Integrating circadian dynamics with
588 physiological processes in plants. *Nat Rev Genet* **16**, 598-610 (2015).
- 589 9. S. L. Harmer *et al.*, Orchestrated Transcription of Key Pathways in
590 *Arabidopsis* by the Circadian Clock. *Science* **290**, 2110 (2000).
- 591 10. M. F. Covington, J. N. Maloof, M. Straume, S. A. Kay, S. L. Harmer, Global
592 transcriptome analysis reveals circadian regulation of key pathways in plant
593 growth and development. *Genome Biology* **9**, R130 (2008).
- 594 11. A. Fukushima *et al.*, Impact of clock-associated *Arabidopsis*
595 pseudo-response regulators in metabolic coordination. *Proceedings of the*
596 *National Academy of Sciences* **106**, 7251 (2009).
- 597 12. A. Graf *et al.*, Parallel analysis of *Arabidopsis* circadian clock mutants reveals
598 different scales of transcriptome and proteome regulation. *Open biology* **7**,
599 160333 (2017).
- 600 13. K. Nozue *et al.*, Rhythmic growth explained by coincidence between internal
601 and external cues. *Nature* **448**, 358-361 (2007).
- 602 14. D. Augustijn, H. J. M. de Groot, A. Alia, A robust circadian rhythm of
603 metabolites in *Arabidopsis thaliana* mutants with enhanced growth
604 characteristics. *PLOS ONE* **14**, e0218219 (2019).
- 605 15. M. G. Annunziata *et al.*, Response of *Arabidopsis* primary metabolism and
606 circadian clock to low night temperature in a natural light environment. *Journal*
607 *of Experimental Botany* **69**, 4881-4895 (2018).
- 608 16. C. J. Doherty, S. A. Kay, Circadian Control of Global Gene Expression Patterns.
609 *Annual review of genetics* **44**, 419-444 (2010).
- 610 17. G. Jander, Timely plant defenses protect against caterpillar herbivory.
611 *Proceedings of the National Academy of Sciences* **109**, 4343 (2012).
- 612 18. H. S. Atamian, S. L. Harmer, Circadian regulation of hormone signaling and
613 plant physiology. *Plant Molecular Biology* **91**, 691-702 (2016).
- 614 19. D. Goodspeed, E. W. Chehab, A. Min-Venditti, J. Braam, M. F. Covington,
615 *Arabidopsis* synchronizes jasmonate-mediated defense with insect circadian
616 behavior. *Proceedings of the National Academy of Sciences* **109**, 4674-4677
617 (2012).
- 618 20. D. Goodspeed *et al.*, Postharvest Circadian Entrainment Enhances Crop Pest
619 Resistance and Phytochemical Cycling. *Current Biology* **23**, 1235-1241 (2013).

- 620 21. A. Singh, in *Glucosinolates*, J.-M. Mérillon, K. G. Ramawat, Eds. (Springer
621 International Publishing, Cham, 2017), pp. 237-246.
- 622 22. S. Huseby *et al.*, Diurnal and light regulation of sulphur assimilation and
623 glucosinolate biosynthesis in Arabidopsis. *Journal of experimental botany* **64**,
624 1039-1048 (2013).
- 625 23. P. Soengas, M. E. Cartea, P. Velasco, M. Francisco, Brassica glucosinolate
626 rhythmicity in response to light-dark entrainment cycles is cultivar-dependent.
627 *Plant Science* **275**, 28-35 (2018).
- 628 24. S. Wu *et al.*, Mapping the Arabidopsis Metabolic Landscape by Untargeted
629 Metabolomics at Different Environmental Conditions. *Molecular Plant* **11**, 118-
630 134 (2018).
- 631 25. H. C. Rowe, B. G. Hansen, B. A. Halkier, D. J. Kliebenstein, Biochemical
632 Networks and Epistasis Shape the *Arabidopsis thaliana*
633 Metabolome. *The Plant Cell* **20**, 1199 (2008).
- 634 26. A. M. Wentzell *et al.*, Linking metabolic QTLs with network and cis-eQTLs
635 controlling biosynthetic pathways. *PLoS Genet* **3**, 1687-1701 (2007).
- 636 27. R. E. Kerwin *et al.*, Network quantitative trait loci mapping of circadian clock
637 outputs identifies metabolic pathway-to-clock linkages in Arabidopsis. *Plant*
638 *Cell* **23**, 471-485 (2011).
- 639 28. L. M. Jensen, H. S. K. Jepsen, B. A. Halkier, D. J. Kliebenstein, M. Burow,
640 Natural variation in cross-talk between glucosinolates and onset of flowering in
641 Arabidopsis. *Frontiers in plant science* **6**, 697-697 (2015).
- 642 29. M. P. Coluccio, S. E. Sanchez, L. Kasulin, M. J. Yanovsky, J. F. Botto, Genetic
643 mapping of natural variation in a shade avoidance response: ELF3 is the
644 candidate gene for a QTL in hypocotyl growth regulation. *Journal of*
645 *Experimental Botany* **62**, 167-176 (2010).
- 646 30. Y. Jiang *et al.*, The ELF3-PIF7 Interaction Mediates the Circadian Gating of the
647 Shade Response in Arabidopsis. *iScience* **22**, 288-298 (2019).
- 648 31. J. M. Jiménez-Gómez, A. D. Wallace, J. N. Maloof, Network Analysis Identifies
649 ELF3 as a QTL for the Shade Avoidance Response in Arabidopsis. *PLOS*
650 *Genetics* **6**, e1001100 (2010).
- 651 32. B. Li, D. J. Kliebenstein, The AT-hook motif-encoding gene METABOLIC
652 NETWORK MODULATOR 1 underlies natural variation in Arabidopsis
653 primary metabolism. *Frontiers in Plant Science* **5**, 415 (2014).
- 654 33. M. R. Tuinstra, G. Ejeta, P. B. Goldsbrough, Heterogeneous inbred family (HIF)
655 analysis: a method for developing near-isogenic lines that differ at quantitative
656 trait loci. *Theoretical and Applied Genetics* **95**, 1005-1011 (1997).
- 657 34. X. Gan *et al.*, Multiple reference genomes and transcriptomes for Arabidopsis
658 *thaliana*. *Nature* **477**, 419-423 (2011).
- 659 35. C. Ramakers, J. M. Ruijter, R. H. L. Deprez, A. F. M. Moorman, Assumption-
660 free analysis of quantitative real-time polymerase chain reaction (PCR) data.
661 *Neuroscience Letters* **339**, 62-66 (2003).
- 662 36. D. J. Kliebenstein, V. M. Lambrix, M. Reichelt, J. Gershenzon, T. Mitchell-
663 Olds, Gene Duplication in the Diversification of Secondary Metabolism:
664 Tandem 2-Oxoglutarate-Dependent Dioxygenases Control Glucosinolate
665 Biosynthesis in Arabidopsis. *The Plant Cell* **13**, 681-693 (2001).
- 666 37. T. Obayashi, Y. Aoki, S. Tadaka, Y. Kagaya, K. Kinoshita, ATTED-II in 2018:
667 A Plant Coexpression Database Based on Investigation of the Statistical
668 Property of the Mutual Rank Index. *Plant and Cell Physiology* **59**, e3-e3 (2017).

- 669 38. E. A. R. Serin, H. Nijveen, H. W. M. Hilhorst, W. Ligterink, Learning from Co-
670 expression Networks: Possibilities and Challenges. *Frontiers in plant science* **7**,
671 444-444 (2016).
- 672 39. J. Chong, D. S. Wishart, J. Xia, Using MetaboAnalyst 4.0 for Comprehensive
673 and Integrative Metabolomics Data Analysis. *Current Protocols in*
674 *Bioinformatics* **68**, e86 (2019).
- 675 40. T. C. Mockler *et al.*, The Diurnal Project: Diurnal and Circadian Expression
676 Profiling, Model-based Pattern Matching, and Promoter Analysis. *Cold Spring*
677 *Harbor Symposia on Quantitative Biology* **72**, 353-363 (2007).
- 678 41. Y. Jiao, O. S. Lau, X. W. Deng, Light-regulated transcriptional networks in
679 higher plants. *Nature Reviews Genetics* **8**, 217-230 (2007).
- 680 42. B. M. Winning, J. Bourguignon, C. J. Leaver, Plant mitochondrial NAD⁺-
681 dependent malic enzyme. cDNA cloning, deduced primary structure of the 59-
682 and 62-kDa subunits, import, gene complexity and expression analysis. *J Biol*
683 *Chem* **269**, 4780-4786 (1994).
- 684 43. M. F. Drincovich, P. Casati, C. S. Andreo, NADP-malic enzyme from plants: a
685 ubiquitous enzyme involved in different metabolic pathways. *FEBS Lett* **490**, 1-
686 6 (2001).
- 687 44. S. D. Grover, P. F. Canellas, R. T. Wedding, Purification of NAD malic enzyme
688 from potato and investigation of some physical and kinetic properties. *Arch*
689 *Biochem Biophys* **209**, 396-407 (1981).
- 690 45. T. P. Michael *et al.*, Network Discovery Pipeline Elucidates Conserved Time-of-
691 Day-Specific cis-Regulatory Modules. *PLOS Genetics* **4**, e14 (2008).
- 692 46. P. M. Sendon *et al.*, Activation of C2H2-type zinc finger genes induces
693 dwarfism in Arabidopsis thaliana. *Journal of the Korean Society for Applied*
694 *Biological Chemistry* **57**, 35-41 (2014).
- 695 47. A. Maier, M. B. Zell, V. G. Maurino, Malate decarboxylases: evolution and
696 roles of NAD(P)-ME isoforms in species performing C4 and C3 photosynthesis.
697 *Journal of Experimental Botany* **62**, 3061-3069 (2011).
- 698 48. Silvia M. Velasquez *et al.*, Complex Regulation of Prolyl-4-Hydroxylases
699 Impacts Root Hair Expansion. *Molecular Plant* **8**, 734-746 (2015).
- 700 49. Y.-F. Tan, N. Toole, N. L. Taylor, A. H. Millar, Divalent Metal Ions in Plant
701 Mitochondria and Their Role in Interactions with Proteins and Oxidative Stress-
702 Induced Damage to Respiratory Function. *Plant Physiology* **152**, 747 (2010).
- 703 50. I. S. Jeong *et al.*, Purification and characterization of Arabidopsis thaliana
704 oligosaccharyltransferase complexes from the native host: a protein super-
705 expression system for structural studies. *Plant J* **94**, 131-145 (2018).
- 706 51. M. A. Tronconi *et al.*, Arabidopsis NAD-Malic Enzyme Functions As a
707 Homodimer and Heterodimer and Has a Major Impact on Nocturnal
708 Metabolism. *Plant Physiology* **146**, 1540 (2008).
- 709 52. N. Nakamichi *et al.*, PSEUDO-RESPONSE REGULATORS 9, 7, and 5 Are
710 Transcriptional Repressors in the Arabidopsis Circadian Clock. *The*
711 *Plant Cell* **22**, 594 (2010).
- 712 53. A. Nunes-Nesi, W. L. Araújo, T. Obata, A. R. Fernie, Regulation of the
713 mitochondrial tricarboxylic acid cycle. *Current Opinion in Plant Biology* **16**,
714 335-343 (2013).
- 715 54. G. Galili, R. Amir, A. R. Fernie, The Regulation of Essential Amino Acid
716 Synthesis and Accumulation in Plants. *Annual Review of Plant Biology* **67**, 153-
717 178 (2016).

- 718 55. T. Engelsdorf *et al.*, The plant cell wall integrity maintenance and immune
719 signaling systems cooperate to control stress responses in *Arabidopsis*
720 *thaliana*. *Science Signaling* **11**, eaao3070 (2018).
- 721 56. M. Michaud, J. Jouhet, Lipid Trafficking at Membrane Contact Sites During
722 Plant Development and Stress Response. *Frontiers in plant science* **10**, 2-2
723 (2019).
- 724 57. T. Zhi *et al.*, Loss of fumarylacetoacetate hydrolase causes light-dependent
725 increases in protochlorophyllide and cell death in *Arabidopsis*. *The Plant*
726 *Journal* **98**, 622-638 (2019).
- 727
- 728

Figure captions

Figure 1. Fine mapping, phenotype characterization of heterogeneous inbred families (HIFs) segregating for Met.II.15 and candidate gene analysis. **A)** Map location of Met.II.15 QTL in a region (at MSAT2.38 and adjacent markers) on Arabidopsis chromosome 2 based on Bay x Sha QTL analysis ((25-27)). Molecular markers are represented with their corresponding names and positions in megabases below according to TAIR (see Supplemental Table1). **B)** Glucosinolate (GSL) ($\mu\text{mol}/\text{cm}^2$ of FW) phenotypes from F2 plants (n=200) from HIF364 [Bay x Sha] (See table S2). **C)** High-resolution linkage map of selected F2 recombinant HIF364 (R1-8) narrowed the Met.II.15 locus to the region between markers M.II.D and M.II.E. **D)** Total GSLs progeny testing of fixed recombinant plants (HIF364 [Bay x Sha] F3, n=24). Five recombinant lines (R4-R8) and control parental lines homozygous for Sha in the target region showed increased GSL content. In contrast, three recombinant line (R1-R3) and control parental line homozygous for Bay in the target region showed decreased GSL content (See Table S3). **E)** Predicted protein coding genes in the region based on Arabidopsis reference genome (TAIR). Candidate genes with variation on transcript abundance between Bay-0 and Sha alleles are highlighted in red. **F)** Average GSL accumulation ($\mu\text{mol}/\text{cm}^2$ of FW) of the evaluated T-DNA insertion lines from candidate genes (n=24) (See Table S4). Different letters indicate statistically significant differences determined by analysis of variance (ANOVA) with post hoc Tukey's honestly significant difference (HSD) test ($P < 0.05$). Error lines represent $\pm$ standard deviation of the mean.

Figure 2. NAD-ME1 allelic effect and promoter analysis. **A)** Diurnal expression pattern of NAD-ME1 mRNA in long day (LD) conditions in three-weeks-old Arabidopsis leaves from three different genotypes: HIF364_R3 (NAD-ME1^{Bay-0}), HIF364_R4 (NAD-ME1^{Sha}) and Col-0 (NAD-ME1^{Col-0}). Samples were harvested at various Zeitgeber times (ZT) every 4h in a 24h time-course experiment. **B)** Predicted cis-acting regulatory elements by PLACE in the NAD-ME1 promoter region are shown in black box. Sequence variation between Bay-0, Sha and Col-0 genotypes is highlighted in red. Numbers indicate positions from putative transcription start site. Sequence between -540 and -578 is shown, with two identified transcription factor binding motifs, one G-box/ABRE and one CBS/MYB/C2H2 binding site. **C, D)** Diurnal total glucosinolate content ($\mu\text{mol}/\text{cm}^2$ of FW) of three-weeks-old Arabidopsis leaves from four different genotypes: NAD-ME1^{Col-0} vs *nad-me1* knockout mutant and NAD-ME1^{Bay-0} vs NAD-ME1^{Sha}. Samples were harvested at various ZT every 4h in a 24h time-course experiment. **E)** Expression data of GSL aliphatic and indolic biosynthetic pathway genes from leaves of NAD-ME1^{Col-0} vs *nad-me1* knockout mutant and NAD-ME1^{Bay-0} vs NAD-ME1^{Sha} genotypes at ZT4 and ZT8. Highlighted with red squares are significantly up-regulated genes from NAD-ME1^{Bay-0} and *nad-me1* knockout mutant; Highlighted with blue squares are significantly up-regulated genes from NAD-ME1^{Sha} and NAD-ME1^{Col-0}. In all experiments three individual plants per genotype from two independent experiments were harvested each time point. Error lines represent $\pm$ standard deviation of the mean. The asterisk indicates statistically significant differences between mean values according to Student's t test (* $p < 0.05$).

775 **Figure 3. Impact of NAD-ME1 in Arabidopsis metabolome.** A) PLS-DA score plots
 776 from UPLC-Q-TOF-MS/MS analysis showing a clear group separation between Col-0
 777 and *nad-me1* at two different time points (ZT4 and ZT8). Green and red ellipsoids show
 778 the 95% confidence interval in Col-0 and *nad-me1* plants, respectively. B) Comparison
 779 of the significant altered ions at both time points analyzed. C) Schematic model of
 780 NAD-ME1 metabolomics changes in the cell. Red arrow represents the enzymatic step
 781 catalyzed by NAD-ME1. Highlighted with red squares are significantly up-regulated
 782 metabolites from *nad-me1* knockout mutant; Highlighted with blue squares are
 783 significantly up-regulated metabolites from Col-0. G3P, 3-phosphoglycerate; PEP,
 784 phosphoenolpyruvate; Fructose-6P, fructose 6-phosphate; Glucose 6-P, glucose 6-
 785 phosphate; IAA, indole-3-acetic acid; JA, jasmonic acid; GSLs, glucosinolates; ITCs,
 786 isothiocyanates; GSH, glutathione; Phe, phenylalanine; Tyr, tyrosine; Tryp, tryptophan;
 787 Cys, cysteine; Met, methionine; Gly, glycine; Asp, asparagine; Ser, serine; PSer,
 788 Phosphoserine; Val, valine; Leu, leucine; Arg, arginine; Glu, glutamic acid; Lys, lysine;
 789 PE, Phosphatidylethanolamine derivative; PS, Phosphoserine derivative; PG,
 790 Phosphatidylglycerophosphate derivative; PI: Phospho-myo-inositol derivative. For
 791 complete metabolite information see Table S6.
 792

793 **Figure 4. Co-expressed NAD-ME1 network.** To construct the network, we filtered
 794 the first 30 genes correlated with *NAD-ME1* from the ATTED-II database. Among
 795 them, we selected only the genes with a cis-eQTL in the Bay-0 × Sha RIL population
 796 (26). The selected genes were used as bait genes. Then, co-expression relationship
 797 between the bait genes and their directly-connected genes were used to construct a
 798 network visualized using NetworkDrawer. Genes were subsequently interpreted by
 799 functional enrichment analysis based on Gene Ontology (GO) to identify and rank
 800 overrepresented functional categories in the network. Red circles depict genes with cis-
 801 eQTL. For complete gene information see Table S7.

802

A

B

C