

# PUBLISHED VERSION

Yoshihiko Onda, Keiichi Mochida, Takuhiro Yoshida, Tetsuya Sakurai, Roger S. Seymour, Yui Umekawa, Stergios Arg Pirintsos, Kazuo Shinozaki, Kikukatsu Ito

**Transcriptome analysis of thermogenic *Arum concinatum* reveals the molecular components of floral scent production.**

Scientific Reports, 2015; 5:8753-1-8753-12

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

Originally published at:

<http://doi.org/10.1038/srep08753>

## PERMISSIONS

<http://creativecommons.org/licenses/by/4.0/>



Attribution 4.0 International (CC BY 4.0)

This is a human-readable summary of (and not a substitute for) the [license](#).

[Disclaimer](#)



### You are free to:

**Share** — copy and redistribute the material in any medium or format

**Adapt** — remix, transform, and build upon the material

for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

### Under the following terms:



**Attribution** — You must give **appropriate credit**, provide a link to the license, and **indicate if changes were made**. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

**No additional restrictions** — You may not apply legal terms or **technological measures** that legally restrict others from doing anything the license permits.

<http://hdl.handle.net/2440/90036>



OPEN

SUBJECT AREAS:

PLANT MOLECULAR  
BIOLOGY

POLLINATION

Received  
21 November 2014Accepted  
2 February 2015Published  
4 March 2015

Correspondence and requests for materials should be addressed to K.M. (keiichi.mochida@riken.jp) or K.I. (kikuito@iwate-u.ac.jp)

# Transcriptome analysis of thermogenic *Arum concinatum* reveals the molecular components of floral scent production

Yoshihiko Onda<sup>1,2</sup>, Keiichi Mochida<sup>1,2,3</sup>, Takuhiro Yoshida<sup>4</sup>, Tetsuya Sakurai<sup>4</sup>, Roger S. Seymour<sup>5</sup>, Yui Umekawa<sup>6</sup>, Stergios Arg Pirtinos<sup>7</sup>, Kazuo Shinozaki<sup>1,3</sup> & Kikukatsu Ito<sup>8</sup>

<sup>1</sup>Biomass Research Platform Team, Biomass Engineering Program Cooperation Division, RIKEN Center for Sustainable Resource Science, Kanagawa, Japan, <sup>2</sup>Kihara Institute for Biological Research, Yokohama City University, Kanagawa, Japan, <sup>3</sup>Gene Discovery Research Group, RIKEN Center for Sustainable Resource Science, Kanagawa, Japan, <sup>4</sup>Integrated Genome Informatics Research Unit, RIKEN Center for Sustainable Resource Science, Kanagawa, Japan, <sup>5</sup>School of Biological Sciences, University of Adelaide, Australia, <sup>6</sup>United Graduate School of Agricultural Science, Iwate University, Morioka, Japan, <sup>7</sup>Department of Biology and Botanical Garden, University of Crete, Heraklion, Greece, <sup>8</sup>Cryobiofrontier Research Center, Iwate University, Morioka, Japan.

Several plant species can generate enough heat to increase their internal floral temperature above ambient temperature. Among thermogenic plants, *Arum concinatum* shows the highest respiration activity during thermogenesis. However, an overall understanding of the genes related to plant thermogenesis has not yet been achieved. In this study, we performed *de novo* transcriptome analysis of flower organs in *A. concinatum*. The *de novo* transcriptome assembly represented, in total, 158,490 non-redundant transcripts, and 53,315 of those showed significant homology with known genes. To explore genes associated with thermogenesis, we filtered 1266 transcripts that showed a significant correlation between expression pattern and the temperature trend of each sample. We confirmed five putative alternative oxidase transcripts were included in filtered transcripts as expected. An enrichment analysis of the Gene Ontology terms for the filtered transcripts suggested over-representation of genes involved in 1-deoxy-D-xylulose-5-phosphate synthase (DXS) activity. The expression profiles of DXS transcripts in the methyl-D-erythritol 4-phosphate (MEP) pathway were significantly correlated with thermogenic levels. Our results suggest that the MEP pathway is the main biosynthesis route for producing scent monoterpenes. To our knowledge, this is the first report describing the candidate pathway and the key enzyme for floral scent production in thermogenic plants.

Among the large number of poikilothermic plant species, several can raise their internal body temperature to a level higher than the ambient temperature. In fact, descriptions of thermogenic plants have a long history extending back to the 18th century; for example, thermogenesis of the inflorescence in the European arum lily was described by de Lamarck (1778; cited in Ref. 1) and the true flower of *Nelumbo nucifera* was described by Miyake<sup>2</sup> to warm appreciably; however, this extraordinary phenomenon in plants was not studied extensively for a long time. Ever since the 1970s, *Philodendron selloum*<sup>3</sup>, *Symplocarpus foetidus*<sup>4</sup>, *Arum maculatum* and *A. italicum*<sup>5</sup>, *Symplocarpus renifolius*<sup>6</sup>, *Nelumbo nucifera*<sup>7</sup>, *Dracunculus vulgaris*<sup>8</sup>, *Helicodictyon muscivorus*<sup>9</sup>, *Amorphophallus titanum*<sup>10</sup>, and *Arum concinatum*<sup>11</sup> were analyzed and reported as thermogenic plants. The thermogenic plants are divided into two types based on the pattern of heat production, i.e., thermoregulatory or not. Thermoregulatory plant species can regulate their internal body temperature during a few days of thermogenesis even when ambient temperature fluctuates, however, non-thermoregulatory plant species cannot regulate and thermogenesis occurs transiently for only a few hours. Thus, it is apparent that the rate of heat production must be opposite to flower temperature change in thermoregulatory plants (e.g. the rate of heat production must increase as flower temperature decreases)<sup>12</sup>. Only *S. foetidus*<sup>4,13,14</sup>, *S. renifolius*<sup>15</sup>, *D. vulgaris*<sup>8</sup>, *P. selloum*<sup>3,16</sup>, and *N. nucifera*<sup>7</sup> are definitely thermoregulatory species. There are likely to be others, particularly among the many *Philodendron* species.

There are three major benefits of heat production in thermogenic plants: (1) to protect the reproductive process from low temperatures<sup>4,17</sup>, (2) to reward insect pollinators in the flower by providing a warm environment<sup>18</sup>, and (3) to attract insect pollinators by volatilizing floral scent compounds<sup>19</sup>. All these roles are important for under-



standing the biological significance of thermogenesis in plants, which involves a substantial energy cost. However, little is known about the molecular mechanisms underlying these phenomena.

Previous studies in thermogenic plants were focused on enzymatic regulation to generate heat in plants. A well-known key player of thermogenesis in plants is mitochondrial alternative oxidase (AOX), which is also found among fungi and nematodes<sup>20–24</sup>. AOX accepts electrons from the ubiquinone pool and uses them to reduce oxygen to water<sup>25</sup>. In contrast to cytochrome *c* mitochondrial terminal oxidase, AOX bypasses complexes III and IV and conserves energy by not pumping protons. Thus, the free energy generated by the flow of electrons from the ubiquinone pool to AOX is generally believed not to lead to ATP synthesis, but to be released as heat<sup>26</sup>. In addition to AOX, plants have unique energy-dissipating systems: rotenone insensitive type II NAD(P)H dehydrogenases (NDA and NDB) in both the inner and outer surfaces of the mitochondrial inner membrane<sup>27</sup>. NDA and NDB are classified as type II, whereas rotenone sensitive and proton pumping complex I are classified as type I. NDA and NDB oxidize NAD(P)H and transport the electrons to ubiquinone without pumping protons, bypassing complex I<sup>27</sup>; however, the involvement of these molecules in thermogenesis has only been recently investigated in *A. maculatum*<sup>28</sup>. In mammals, the mitochondrial uncoupling protein (UCP) has been shown to play a crucial role in thermogenesis<sup>29</sup>. Although mammalian mitochondria contain neither AOX nor type II NAD(P)H dehydrogenases, UCP resides in the inner mitochondrial membrane. The presence of UCPs in plants gained widespread acceptance with the identification of the first plant UCP-encoding gene from potato<sup>30</sup>; since then, UCP-encoding genes have been identified from a wide range of plant species<sup>31,32</sup>. Overall, previous reports indicate that key players of plant thermogenesis differ depending on the plant species and include not only AOX but also NDA, NDB, and UCP<sup>28,33–37</sup>. Detailed molecular mechanisms (e.g., expression of protein and regulation of activity) of AOX and UCP have been reported<sup>128,33,36–40</sup>; however, an overall understanding of the genes related to plant thermogenesis has not yet been achieved.

Among thermogenic species, *A. concinatum* is non-thermoregulatory, but shows the highest mass-specific respiration rates during heat production<sup>11,41</sup>. *A. concinatum* produces an inflorescence, which consists of a spadix with male florets and female florets in the floral chamber surrounded by the spathe. Above the floral chamber a large appendix becomes exposed when the spathe opens in the afternoon and evening of the pistillate stage of blooming (called D-day). This is accompanied by a single bout of intense warming of the appendix to over 30°C. When the appendix temperature reaches its peak, accompanied by spathe opening and scent volatilization, several insect families of Diptera (e.g., Sphaeroceridae, Sciaridae, and Drosophilidae) and a family of Coleoptera (Ceratopogonidae) are attracted to the inflorescence<sup>42</sup>. The insects remain in the floral chamber where the thermogenic male florets provide warmth overnight. About 24 hours after the pistillate stage, the staminate stage results in pollen production after which the male florets cease thermogenesis. Although male florets are more thermogenic on a mass-specific basis than the appendix, they weigh much less (0.4 g vs. 14.5 g), and therefore do not reach a temperature as high as the appendix (e.g., 34°C for the appendix, 25°C for male florets, and 22°C for female florets under an ambient temperature condition of 22°C)<sup>11</sup>. The mean rate of heat production of the whole appendix is 1076 mW, the male florets 157 mW and female florets negligible heat. These differences in heat production levels among the floral parts are convenient for comparing the molecules related to heat production in an individual.

In this study, to develop a global view of the transcriptome underlying thermogenesis and to elucidate the genes involved in heat production in plants, we performed *de novo* transcriptome analysis of the flower organs in *A. concinatum*. Because the genome size of this plant was reported as 23,445 Mbp<sup>43</sup>, it is not practical to perform

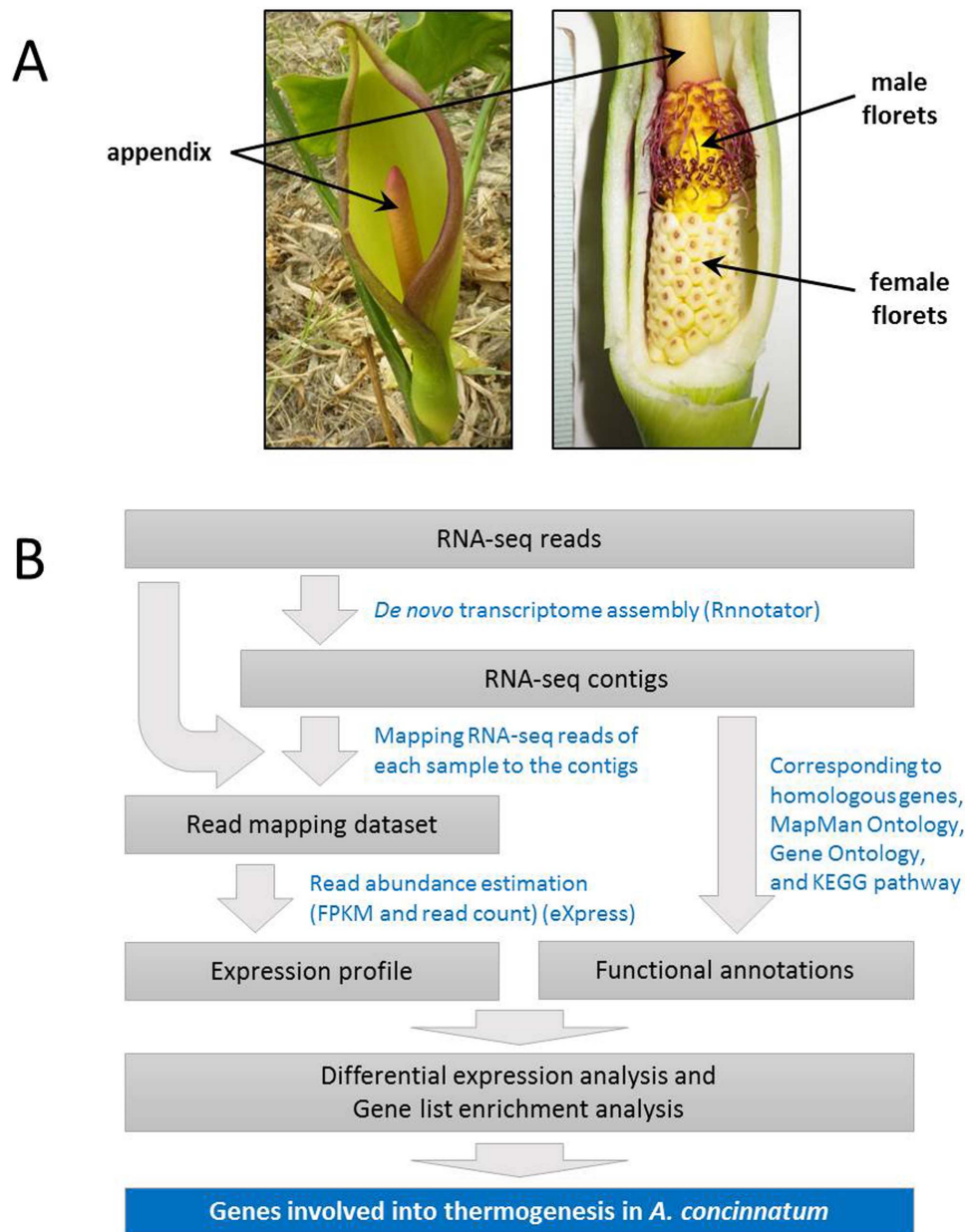
whole-genome sequencing. Thus, RNA-seq with a reference sequence could not be applied. *De novo* transcriptome analysis is becoming a useful tool to gain sequence information regarding expressed genes and expression patterns without a genome sequence as a reference<sup>44–46</sup>. Based on the RNA-seq dataset of *A. concinatum*, we present comparative properties of the transcriptome among stages (on pre-D-day and on D-day) and tissues (male florets for intermediate heating level, female florets for never heating, and the appendix for most thermogenic) with different thermogenic properties. We also predict putative genes that are associated with thermogenic levels of the flower organs based on a co-expression analysis. Herein, we present the first comprehensive view of the transcriptome to provide new insights into plant thermogenesis, especially with respect to a metabolic process of volatilizing floral scent compounds.

## Results

### *De novo* transcriptome assembly of flower organs in *A. concinatum*.

Transcriptome analysis was performed using RNAs extracted from three organs of the inflorescence (male florets, female florets, and the appendix) at two stages (pre-D-day and D-day) (Figure 1A). These organs and stages exhibit different temperatures<sup>11</sup>. In this sampling, we confirmed the heating level of the appendix in comparison with air temperature; the appendix temperature was elevated to more than 31°C at D-day. By using an Illumina HiSeq2000 platform, 26 to 32 million paired-end reads were obtained in each sample, and these passed a filtering process as cleaned reads (Supplementary Table S1). In total, 180 million paired-end reads representing 34.2 Gbase cleaned nucleotides were used to the assembly process of sequence by using the Rnnotator program<sup>47</sup> to generate a non-redundant dataset of contig sequences. The forward reads were mapped to the dataset of contig sequences as the reference transcripts using a bowtie program<sup>48</sup> to estimate the abundance of reads corresponding to each transcript in each sample. In the mapping, 44% to 59% of reads were mapped to the reference transcripts (Supplementary Table S1). Based on the mapping results, the fragments per kilobase of transcript per million mapped reads (FPKM) value of each gene in each sample was computed as a dataset of global expression profiles by using the eXpress program (Figure 1B).

**Overview of the transcriptome in thermogenic inflorescences.** The *de novo* transcriptome assembly represents, in total, 158,490 non-redundant transcripts identified in the *A. concinatum* inflorescence (Figure 2A). To classify the transcripts to putative homologs of known genes or others, we performed a sequence similarity search against known protein sequence datasets (NCBI nr and *Arabidopsis* proteome in TAIR10) by using the BLASTx program ( $E < 10^{-5}$ ). The BLASTx search showed 53,315 (33.6%) transcripts with significant sequence similarity with entries of NCBI nr and/or those of TAIR10 (known), and the others 105,175 (66.4%) transcripts did not provide any significant hits (unknown) (Figure 2A). The sequence length distribution differed between the known and unknown transcripts. The sequence length of the known transcripts ranged to more than 4 kbp. On the other hand, the unknown transcripts were less than 2.2 kbp long, and most transcripts were less than 600 bp long (Supplementary Figure S1). Throughout this study, we defined transcripts with FPKM values  $\geq 1$  in at least one sample as the threshold of an expressed gene. Figure 2B shows numbers and proportions of transcripts of known genes and those of unknown genes that are expressed or not expressed in each sample. In every sample, more than 35,000 transcripts of putative homologs of known genes were identified with FPKM  $\geq 1$ . Most sequences without any significant hits to known proteins also showed significant expression, with FPKM  $\geq 1$  in each sample (Figure 2B). Organ-wise comparisons of expressed transcripts between pre-D-day and D-day in male florets, female florets, and the appendix showed commonly and

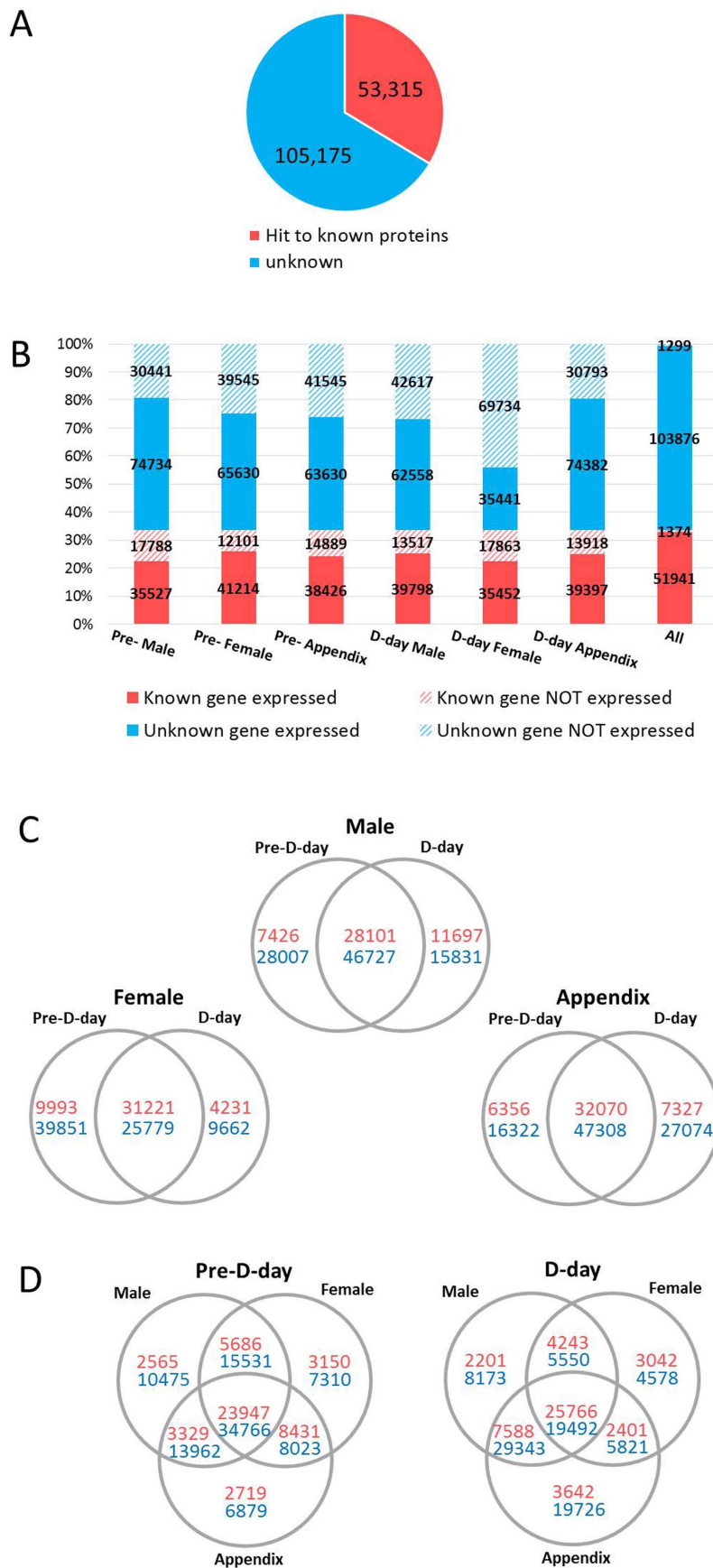


**Figure 1 | Schematic overview of the transcriptome analysis of thermogenic inflorescences of *A. concinatum*.** (A) Anatomical representation of the thermogenic flower organs of *A. concinatum*. (B) Analytical workflow of a *de novo* RNA-seq analysis used in this study.

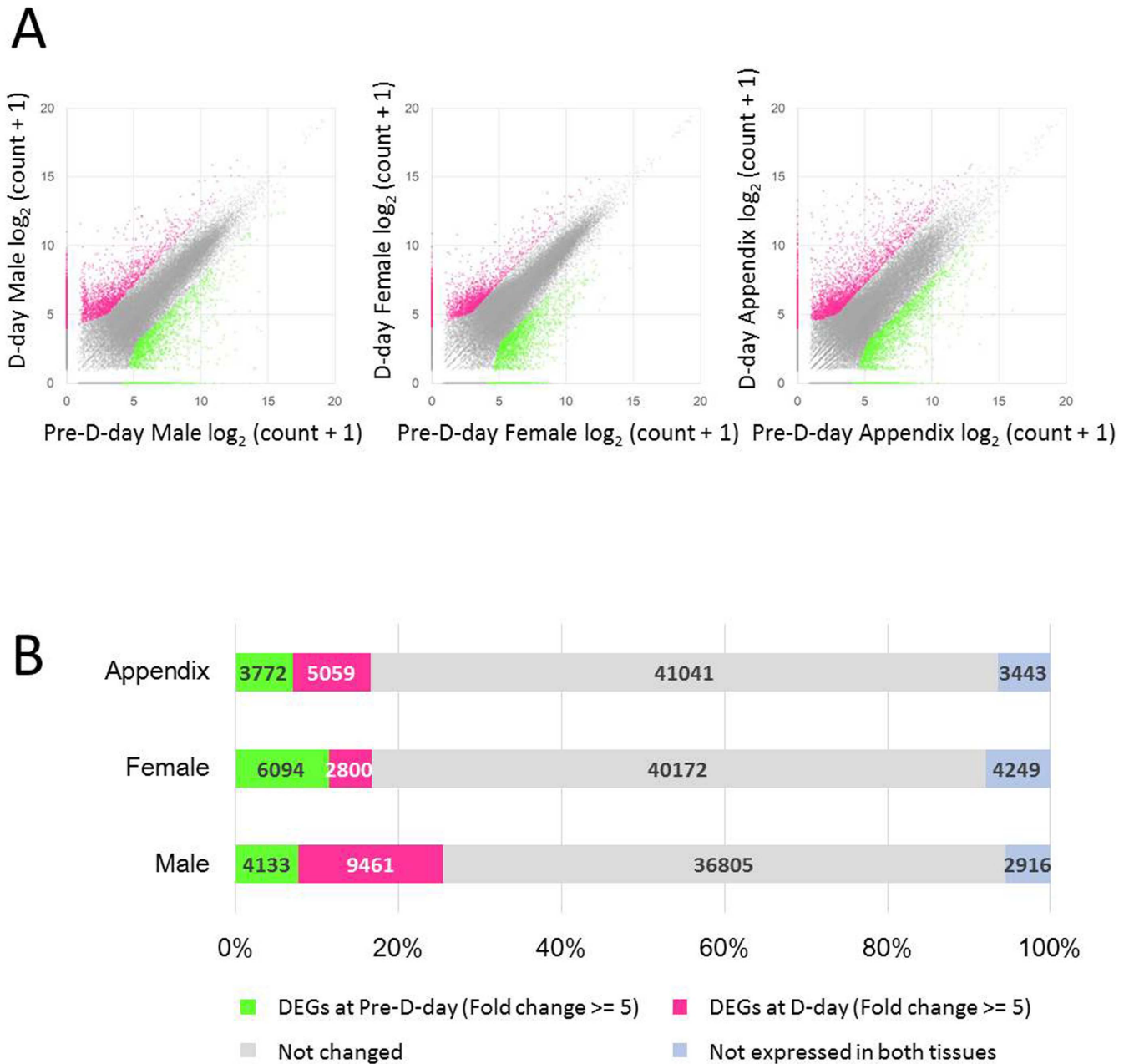
specifically expressed transcripts in each organ at stages with different thermal states (Figure 2C). For instance, in thermogenic male florets and the appendix at D-day, 11,697 and 7327 transcripts of known genes were specifically expressed, respectively. The distribution of expressed transcripts among the three organs showed specific expression of transcripts for each organ at pre-D-day and D-day (Figure 2D). More than 8400 transcripts of known genes were commonly expressed in the female florets and the appendix only at pre-D-day. On the other hand, 7500 transcripts of known genes were commonly expressed in the male florets and the appendix only at the D-day. These results of the *de novo* transcriptome assembly represent a number of genes that are specifically co-expressed in each of the thermogenic organs (male florets and appendix) of the inflorescence at D-day in *A. concinatum* (Figure 2D). In addition to the sequence similarity searches, we performed further functional annotation of the *A. concinatum* transcripts to collect clues as much as possible to infer gene function (Supplementary Figure S2).

**Differentially expressed genes at pre-D-Day and D-day.** A normalized RNA-seq read count dataset from the eXpress program<sup>49</sup> of transcripts of known genes was used to identify genes that are differentially expressed in each organ between pre-D-day and D-day. Differentially expressed genes (DEGs) were determined using a false discovery rate (FDR) threshold corrected  $P$  value  $< 10^{-5}$  using Fisher's exact test and a fold change  $\geq 5$  with the normalized read count mapped to each transcript. These were computed using the DEGseq program (Figure 3A). Based on the DEG analysis, 5059 transcripts were up-regulated in the appendix at D-day, whereas 3772 transcripts were up-regulated in the appendix at pre-D-day. In the female florets, 2800 and 6094 transcripts were up-regulated at D-day and pre-D-day, respectively (Figure 3B). In the male florets, 9461 and 4123 transcripts were up-regulated at D-day and pre-D-day, respectively.

Gene Ontology (GO) enrichment analysis was performed to infer particular functional categories of genes over-represented in up-regulated DEGs in each sample. Various GO terms were significantly



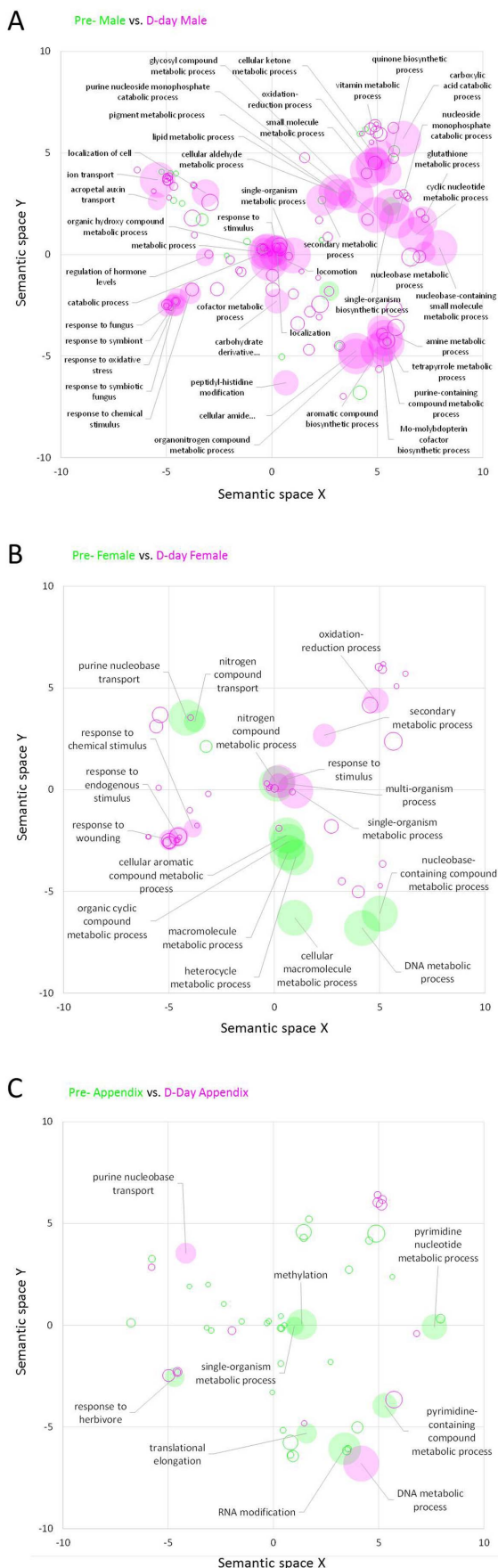
**Figure 2 | Overview of expressed transcripts in the thermogenic inflorescences.** (A) Proportion of known transcripts that showed sequence similarity with NCBI nr entries or proteins of *Arabidopsis* (TAIR10). (B) Number of expressed transcripts (FPKM  $\geq 1$ ) in each tissue and at least one tissue RNA sequenced in this study. (C) Organ-wise comparison of specificity and commonality of expressed transcripts. (D) Specificity and commonality of expressed transcripts among organs in pre-thermogenic and thermogenic stages.



**Figure 3 | Differentially expressed genes (DEG) in the flower organs.** In total, 53,315 transcripts that showed significant homology with known protein sequences were analyzed. (A) Scatter plots of pairwise combinations of tissues with selected DEGs (colored) in the male and female florets and the appendix. DEGs were selected using DEGseq with Fishers exact test ( $P < 10^{-5}$ , fold change  $\geq 5$ ) (B) Numbers of DEGs in each combinations of organs.

enriched in each group of DEGs with a false discovery rate (FDR) threshold corrected  $P$  value  $< 0.05$  using Fisher's exact test on the BLAST2GO program (Supplementary Table S2). In particular, the number of GO terms that were enriched are 120 in pre-D-day male, 52 in pre-D-day female, 189 in pre-D-day appendix, 890 in D-day male, 184 in D-day female, and 92 in D-day appendix tissues (Supplementary Table S2). These GO terms were enriched from DEGs as shown in Figure 3B. To summarize the significantly enriched GO terms along with semantic similarities, the enriched GO terms of the biological process ontology were represented on the 2D semantic space (Figure 4). In the male florets, 19 GO terms at pre-D-day and 120 GO terms at D-day were significantly enriched. In the male florets at D-day, genes involved in a number of biological processes were particularly up-regulated. The summarization on the 2D semantic

space presents semantic clusters of GO terms related to the response to some biological stimulus, and the metabolic processing of various primary and secondary metabolites in the male florets at D-day (Figure 4A). In the female florets, 11 GO terms at pre-D-day and 42 GO terms at D-day were significantly enriched. In the female organs, GO terms of some metabolic processes were particularly over-represented at pre-D-day, whereas GO terms related to response to stimulus were weakly over-represented at D-day (Figure 4B). In the appendix, 39 GO terms at pre-D-day and 14 GO terms at D-day were significantly enriched. Although not many GO terms were particularly over-represented in the appendix, GO terms of pyrimidine nucleotides or compound metabolic processes were enriched at pre-D-day, whereas those of purine nucleobases or compound metabolic processes were enriched at D-day (Figure 4C). The comparative GO



**Figure 4** | Significantly enriched gene functions in DEGs found in combinations of pre-D-day and D-day tissues. Enriched gene functions were projected to a 2D semantic space of the gene ontology, biological

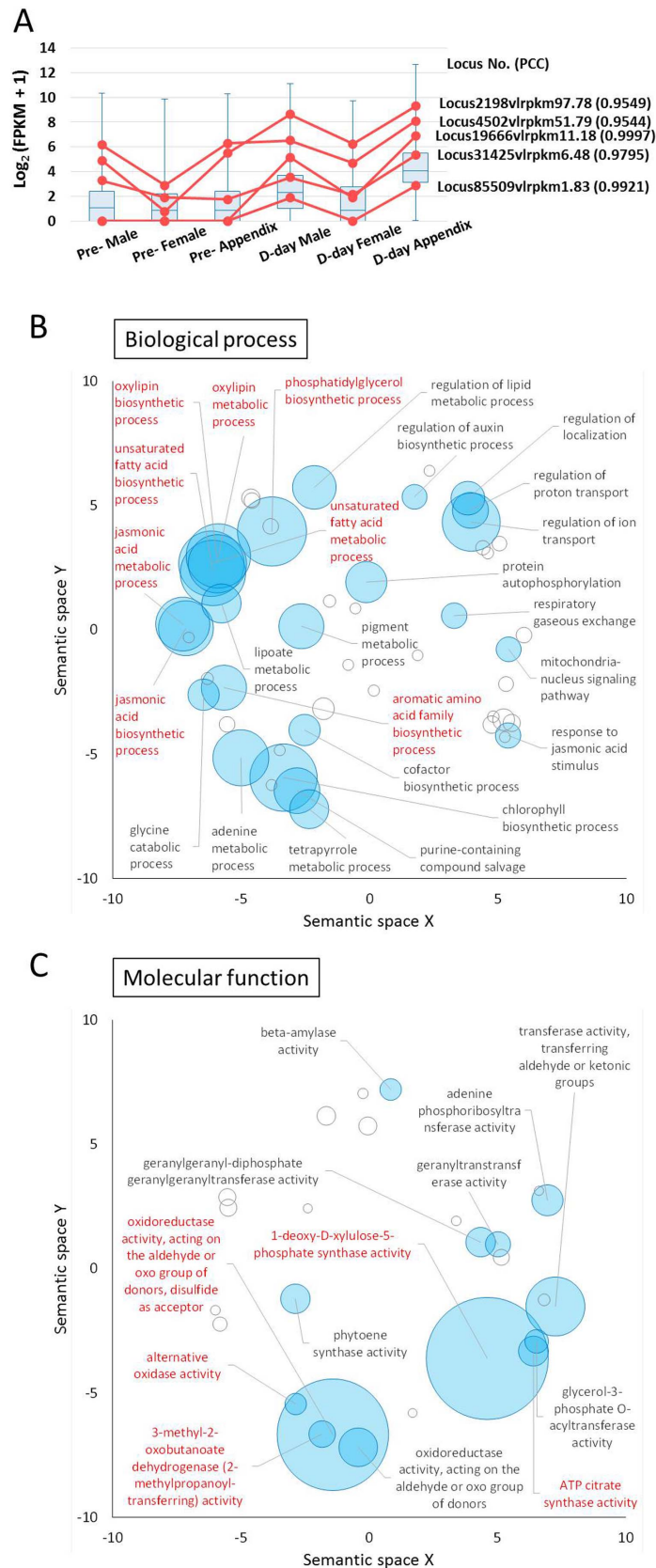
process. Green circles and pink circles correspond to significantly enriched GO terms found in DEG genes at pre-D-day and D-day, respectively. The circle size represents the  $-\log_{10}$  transformed FDR in REVIGO analysis. Circles depicted by filled color show significantly enriched GO terms with  $FDR < 10^{-5}$ . (A) Male florets at pre-D-day and D-day, (B) Female florets at pre-D-day and D-day, and (C) Appendix at pre-D-day and D-day. The semantic space was generated by the REVIGO web service with all enriched GO terms found in the DEG gene sets.

enrichment analyses between pre-D-day and D-day in each organ represent differential up-regulation of genes involved in various processes behind the flowering of *A. concinatum*.

**Co-expressed genes with their associated temperature profiles.** To explore genes associated with *A. concinatum* thermogenesis, we filtered transcripts from 53,315 known transcripts that showed a significant correlation between expression pattern and the temperature trend of each sample. As a result, 1266 known transcripts (2.37%) that showed a strong correlation (Pearson's correlation coefficient (PCC)  $\geq 0.95$ ) with the temperature trend were labeled with an ID that begins with "Locus" (Supplementary Table S3). A box plot summarizing the expression of these transcripts agreed with the temperature trend, which are slightly heated in male florets and well heated in the appendix at D-day (Figure 5A). On the other hand, among 105,175 unknown transcripts, 3195 transcripts (3.72%) showed a significant correlation between expression pattern and the temperature trend of each organ. There is a possibility that a number of unknown transcripts from related gene families are associated with the heat production.

**AOX-encoding genes.** The 1266 transcripts that correlated with temperature include those of five genes (Locus19666v1rpkm11.18, Locus2198v1rpkm97.78, Locus31425v1rpkm6.48, Locus4502v1rpkm51.79, and Locus85509v1rpkm1.83) putatively encoding AOX. Transcripts of Locus19666 and Locus85509 were especially correlated with the heat pattern, being not expressed in any organ at pre-D-day, and mildly expressed in male florets, and having highest expression in the appendix at D-day. Transcripts of three other genes (Locus2198, Locus31425, and Locus4502) were expressed in the organs even at pre-D-day (Figure 5A). In this correlation analysis, only AOX showed significantly higher correlation; other genes for mitochondrial energy-dissipating enzymes (namely UCP, NDA, and NDB) were not significantly correlated with the thermogenic levels of each organ (Table 1). These results suggest that gene expression for AOX plays a critical role in thermogenesis of *A. concinatum*, and UCP, NDA and NDB may act in a rather ubiquitous fashion in this non-thermoregulatory plant. In a previous study, two mRNAs encoding AOX in *A. concinatum* were reported (AB485993 for *AcoAOX1a* and AB485994 for *AcoAOX1b* in DDBJ)<sup>39</sup>. In *A. maculatum*, another *Arum* species distributed around Europe, at least seven transcripts for AOX were cloned<sup>50</sup>. Our transcriptome analysis revealed at least seven independent transcripts putatively encoding AOX. This result suggests that there are unidentified AOX homologs other than *AcoAOX1a* and *1b* in *A. concinatum*.

**Enriched GO terms.** GO enrichment analysis was performed to infer particular functional categories of genes over-represented in the 1266 selected transcripts. In total, 131 GO terms were significantly enriched in this group of transcripts (Supplementary Table S4). The summarization on the biological process ontology of the 2D semantic space presents semantic clusters related to metabolic processes of the unsaturated fatty acid, oxylipin, and jasmonic acid, particularly (Figure 5B). On the molecular function ontology, we observed particularly enriched categories of oxidoreductase and alternative oxidase activity because of putative AOX-encoding genes. In addition, we observed 1-deoxy-D-xylulose-5-phosphate synthase



**Figure 5** | Genes with expression patterns that correlated with the temperature pattern. (A) Expression trend of 1266 transcripts whose expression patterns were highly correlated with temperature patterns in each organ (Pearson's correlation coefficient (PCC)  $\geq 0.95$ ). Red colored line plots are expression patterns of homologs of AOX encoding genes. (B) and (C) Enriched gene functions of the correlated genes with the temperature pattern projected onto a 2D semantic space. The circle size represents the  $-\log_{10}$  transformed FDR in REVIGO analysis. Circles depicted by filled color show significantly enriched GO terms with FDR  $< 10^{-5}$ . (B); GO categories related to biological process. (C); GO categories related to molecular function.





Table 1 | The results of expression and correlation analysis of energy dissipating system found in this study.

ID	Expression (FPKM)						
	Pre-D-day Male	Pre-D-day Female	Pre-D-day Appendix	D-day Male	D-day Female	D-day Appendix	PCC
<b>alternative oxidase (AOX)</b>							
Locus104409v1rpkm1.43	2.92E-01	1.73E+00	2.01E+00	2.06E+00	0.00E+00	6.94E-01	-0.086
Locus113826v1rpkm1.27	1.08E+01	7.67E+00	1.32E+01	0.00E+00	0.00E+00	4.61E+00	-0.294
Locus19666v1rpkm11.18	0.00E+00	0.00E+00	0.00E+00	3.44E+01	2.66E+00	1.16E+02	1.000
Locus2198v1rpkm97.78	2.85E+01	7.10E-01	4.42E+01	3.97E+02	7.46E+01	6.35E+02	0.955
Locus2221v1rpkm96.81	5.54E+01	2.34E+01	8.18E+01	1.06E+02	3.13E+01	9.09E+01	0.573
Locus31425v1rpkm6.48	8.65E+00	2.72E+00	2.33E+00	1.06E+01	3.36E+00	3.91E+01	0.980
Locus3229v1rpkm69.53	6.88E+01	2.75E+01	1.21E+02	8.68E+01	6.81E+01	1.29E+02	0.617
Locus4502v1rpkm51.79	7.08E+01	6.33E+00	7.73E+01	9.12E+01	2.45E+01	2.72E+02	0.954
Locus52656v1rpkm3.43	5.29E+00	3.51E+00	0.00E+00	9.58E+00	0.00E+00	8.26E+00	0.658
Locus85509v1rpkm1.83	0.00E+00	0.00E+00	0.00E+00	2.68E+00	0.00E+00	6.28E+00	0.992
Locus87119v1rpkm1.79	0.00E+00	0.00E+00	0.00E+00	0.00E+00	5.75E+00	6.76E+00	0.621
<b>uncoupling protein (UCP)</b>							
Locus1928v1rpkm109.42	1.70E+01	2.96E+00	2.09E+01	1.82E+02	1.27E+01	7.97E+01	0.476
Locus56578v1rpkm3.14	2.19E+01	1.02E+01	1.06E+01	3.20E+00	2.64E+01	3.86E+00	-0.611
Locus8198v1rpkm28.98	5.31E+01	5.29E+01	1.13E+01	2.80E+01	3.86E+01	9.98E+00	-0.610
<b>rotenone insensitive type II NAD(P)H dehydrogenases (NDA and NDB)</b>							
Locus110607v1rpkm1.32	0.00E+00	8.52E-01	3.43E+00	4.05E+00	4.16E+00	3.81E+00	0.410
Locus13660v1rpkm16.85	1.06E+01	5.54E+00	1.21E+01	6.09E+00	0.00E+00	9.88E+00	0.236
Locus1612v1rpkm125.79	9.22E+01	5.81E+00	1.50E+02	1.60E+02	2.79E+01	9.54E+01	0.225
Locus18090v1rpkm12.32	5.07E+00	1.87E+01	1.22E+01	1.39E+01	2.52E+01	6.41E+00	-0.465
Locus23870v1rpkm8.93	1.09E+01	0.00E+00	9.13E+00	2.82E+01	3.61E+00	1.86E+01	0.569
Locus2753v1rpkm80.23	7.20E+01	9.26E+00	1.29E+02	2.69E+02	3.68E+01	1.23E+02	0.352
Locus60938v1rpkm2.86	2.58E+00	4.46E-01	8.56E+00	6.57E+00	6.32E-01	5.24E+00	0.301
Locus611v1rpkm271.69	4.19E+02	4.06E+01	6.62E+02	1.12E+03	1.39E+02	5.98E+02	0.366
Locus611v2rpkm105.37	2.12E+02	1.92E+01	3.59E+02	4.10E+02	7.64E+01	2.83E+02	0.356
Locus614v1rpkm269.67	3.73E+01	2.41E+00	6.04E+01	4.90E+02	7.65E+01	2.93E+02	0.603

(DXS) activity was especially enriched, with the lowest corrected *P* value ( $5.85 \times 10^{-19}$ ) in the tested group (Figure 5C).

**DXS-encoding genes.** In the transcripts with known protein hits, we identified 21 putative homologs of the *Arabidopsis thaliana* DXS gene (At4g15550). In the 1266 selected transcripts, 24 entries were classified into the GO category of DXS activity, which includes 17 putative DXS encoding transcripts. This means expression of most of the putative DXS genes was correlated with the thermogenic level (heating pattern) of the flower organs. DXS is the first enzyme of the methyl-D-erythritol 4-phosphate (MEP) pathway, and plays a major role in the regulation of this pathway<sup>51</sup>. It is well known that there are two metabolic pathways for terpenoids biosynthesis: the MEP pathway (non-mevalonate pathway) and the mevalonate pathway. Therefore, we examined the expression profiles of transcripts putatively encoding enzymes involved in both these pathways based on the RNA-seq dataset (Figure 6). The expression pattern of each of 17 transcripts encoding putative DXS showed strong correlation with the temperature pattern of the flower organs and stages, and showed extremely elevated expression in the appendix at D-day compared with pre-D-day. For example, the expression Locus2538 was more than a thousand-fold increased from pre-D-day to D-day in the appendix, with strong correlation (Pearson's correlation coefficient (PCC) = 0.963). On the other hand, only one transcript of the mevalonate pathway (Locus 1825, encoding the first step enzyme ACAA) showed strong correlation between its expression and temperature (PCC = 0.939). Previous studies reported that DXS catalyze one of the late-limiting reaction of the MEP pathway for isoprenoid biosynthesis in *Arabidopsis thaliana*<sup>52</sup> and *Solanum lycopersicum*<sup>53,54</sup>. Our *de novo* transcriptome analysis revealed that the expression

patterns and particular up-regulation of DXS were strongly correlated with thermogenesis. These results also suggest that this step might be a rate-limiting step in biosynthesis of terpenoids leading to volatile compounds for the scent from this plant.

## Discussion

**De novo assembly of the *A. concinatum* transcriptome.** *De novo* transcriptome analysis for thermogenic plants provides an opportunity to discover genes involved in floral thermogenesis and scent production. It has become a useful approach to rapidly build comprehensive sequence resources of expressed genes without whole genome sequencing and to examine genome-scale gene expression patterns of RNA samples of interest. Thus, *de novo* transcriptome analysis has been applied to a wide range of plant species such as medicinal plants, crops, and plants with particular biological systems<sup>55–57</sup>. For organisms with large, complex genomes containing repetitive sequences and polyploidy, including thermogenic plants, RNA-seq based *de novo* transcriptome analysis is a promising procedure to examine properties of a transcriptome as a proxy for the whole genome<sup>44</sup>.

*De novo* transcriptome assembly contained more than 150,000 transcripts, even though only non-redundant transcripts were included in this study. Functional prediction of these transcripts resulted in only 33.6% of the transcripts being associated with at least one homologous protein; no proteins were identified for the remaining transcripts. Possibly, the dataset of assembled transcripts includes many noisy sequences, or some non-coding transcripts such as functional non-coding RNAs and expressed transposable elements. With mapping of RNA-seq reads and estimation of gene expression, most of the unknown transcripts also showed a significant level of gene expression, with FPKM values  $\geq 1$ , as well as expression in several tissues





(Figure 2). The large-scale RNA-seq reads should become a significant sequence resource that should facilitate identifying homologous counterparts in *A. concinatum* and could shed light on the largely unexplored transcriptome in thermogenic plants.

**Differential gene expression analysis.** Differential gene expression analysis in this study identified a number of transcripts that were differentially expressed between pre-D-day and D-day. In total, 139 (19 at pre-D-day and 120 at D-day), 53 (11 at pre-D-day and 42 at D-day), and 53 (39 at pre-D-day and 14 at D-day) summarized GO terms in biological process were significantly enriched in the male florets, female florets, and the appendix, respectively (Figure 4). Interestingly, although more GO terms were enriched at D-day in the thermogenic male and non-thermogenic female florets compared with that at pre-D-day, fewer GO terms were enriched in the thermogenic appendix at D-day. This result suggests that a number of transcripts categorized into the same GO terms have been sufficiently prepared before thermogenesis, or less significance of over-representation of functional category of GOs due to a wider variety of transcripts in the appendix which organ produces remarkable heat in this plant.

Although we could not identify prominent DGEs in the appendix by a simple comparison between pre-D-day and D-day samples, we did identify genes that were significantly correlated with thermogenic levels among samples (Figure 5). First, we focused on AOX, a well-known energy-dissipating mitochondrial enzyme in thermogenic plants, as a positive control. Five putative AOX genes were significantly correlated with the temperature patterns of each sample. This expression pattern was consistent with that of a previous study in *A. concinatum*<sup>39</sup> and other thermogenic species such as *D. vulgaris*<sup>34</sup> and *S. renifolius*<sup>38</sup>.

The GO categories that were most enriched related to the biological process involved in metabolism of the unsaturated fatty acids, oxylipin, and jasmonic acid (Figure 5B). Jasmonic acid, an oxylipin, is derived from the unsaturated fatty acid, linolenic acid, in the chloroplast membrane<sup>58–60</sup>. The jasmonic acid signaling pathway is involved in defense responses to injury and in programmed cell death and senescence<sup>61</sup>. When the appendix temperature rises and reaches its peak, accompanied by floral scent volatilization, several species of pollinator insects are attracted to the inflorescence<sup>42</sup>. In the only other study of the transcriptome in thermogenic plants, Ito-Inaba et al. demonstrated that genes involved in stress responses and protein degradation are up-regulated during the post-thermogenic stages in skunk cabbage (*S. renifolius*)<sup>62</sup>. Accordingly, the enriched gene expression for jasmonic acid metabolism seems to be related to programmed cell death and senescence, followed by energy-consuming thermogenesis, rather than to injury caused by attracted pollinator insects.

For molecular function, the most outstanding enriched GO category was 1-deoxy-D-xylulose-5-phosphate synthase activity (Figure 5C). The expression of *DXS* genes in heated organs, especially in the appendix, might be involved in the synthesis of terpenoids, depending on the temperature increases in the flower organs (Figure 6). As a distinctive phenomenon associated with thermogenesis in *A. concinatum*, the flowers volatilize a strong dung/urine-like smell at D-day<sup>42</sup>. Metabolome analysis by gas chromatography/mass spectrometry (GC-MS) on the volatile compounds showed that terpenoids (3,7-dimethyl-1,6-octadine, 2,7-dimethyl-1,7-octadiene, and 3-7-dimethyl-1-octane) mainly account for the smell of *A. concinatum*<sup>42,63</sup>. Because the transcripts of the first step of the mevalonate pathway showed lower correlation with the thermogenic level compared with transcripts of the first step of the MEP pathway (Fig. 6), the MEP pathway seems to be the main biosynthetic route for producing these three basic scent monoterpenes in this plant. To the best of our knowledge, this is the first report describing the candidate pathway and the key enzyme for floral scents in thermogenic plants.

**Evolution of thermogenic plants.** Heat is a byproduct of metabolism in all flowering plants. In most plants, metabolic reactions are too slow to elevate tissue temperature<sup>64</sup>, or heat escapes very quickly because the tissue is too small to retain heat<sup>41</sup>. Irrespective of whether plants produce heat, the reproductive process is one of the most pivotal life events for all seed plants. In thermogenic plants, diverse relationships exist between the types of odor and the insect pollinators they attract. For example, *P. sellowii* produces 4-methoxystyrene and 3,4-dimethoxystyrene and attracts the dynastid scarab beetle (*Cyclocephala variolosa*)<sup>65</sup>. *H. muscivorus* (dead-horse arum) produces oligosulphides, dimethyl mono-, di- and trisulphide and attracts flies<sup>66</sup>. *A. maculatum* produces indole and 2-heptanone<sup>67</sup> and attracts drain flies (*Psychoda phalaenoides*)<sup>68</sup>. These observations suggest that heat production and the accompanying floral scent volatilization may have coevolved with the olfactory receptor of insect pollinators. As described above, thermogenic plants produce different floral scents and use different combinations of energy-dissipating systems for heat production to attract insect pollinators. Thermogenic plants have undergone a variety of evolutionary changes and has diversified with regard to reproduction. In conclusion, the present study firstly establishes a comprehensive and detailed relationship between global gene expression and reproductive processes, in particular, floral thermogenesis and scent production in *A. concinatum*. With *de novo* transcriptome analysis, it is possible to shed light on the unknown molecular components of the metabolic pathway involved in plant thermogenesis and pollinating insect attraction.

## Methods

**Plant materials and temperature measurements.** *A. concinatum* plants were collected in May 2013 from a field situated near the village of Panormos, Crete Island, Greece. The appendix temperature was measured by thermistors placed into the tissue. Air temperature was measured in the same way. Three individuals were randomly selected and collected for following RNA extraction. Mean temperature of both appendix and air were  $23.0 \pm 0.0^\circ\text{C}$  at pre-D-day, and mean temperature of appendix and air were  $31.3 \pm 0.5^\circ\text{C}$  and  $23.0 \pm 0.0^\circ\text{C}$  at D-day, respectively.

**RNA-seq analysis.** Total RNA was extracted using NucleoSpin RNA Plant (MACHEREY-NAGEL) from flash-frozen male florets, female florets, and the appendix of *A. concinatum* plants at pre-D-day and D-day from the three independent individuals. Quality and quantity of all total RNA samples were checked by using NanoDrop (Thermo Fisher Scientific) and a 2100 Bioanalyzer (Agilent). Before library construction, total RNAs from the three individuals were combined and analyzed as a sample. First-Strand cDNA synthesis was performed using SMARTer® Ultra™ Low Input RNA for Illumina® Sequencing – HV (Clontech), according to the manufacturer's instructions. Libraries for RNA-seq analysis were constructed using a Low Input Library Prep Kit (Clontech) with Indexing Reagents to be barcoded according to the manufacturer's instructions. Prepared libraries were sequenced with HiSeq2000 by using TruSeq SBS Kit v3-HS (Illumina). The dataset of the Illumina reads has been submitted to DDBJ Sequence Read Archive under the accession number DRA002584.

**De novo transcriptome assembly.** The FASTX tool kit ([http://hannonlab.cshl.edu/fastx\\_toolkit](http://hannonlab.cshl.edu/fastx_toolkit)) was used to assess quality score distributions of the Illumina reads. We estimated the most appropriate k-mer size for the following *de novo* assembly by using the KmerGenie program (<http://kmergenie.bx.psu.edu/>). The Rnnotator program<sup>47</sup> was used for the *de novo* assembly of the Illumina RNA-seq reads with the estimated K-mer = 31. The significantly expressed transcripts were filtered with a threshold FPKM values  $\geq 1$ , as described below. The assembled transcripts with length  $\geq 200$  bp were used as the sequence dataset of reference transcripts.

**Functional annotations.** To predict the function of the assembled transcripts, we performed a similarity search of the assembled sequences as queries by using the NCBI BLASTx program<sup>69</sup>, with a threshold E value less than  $10^{-5}$  against protein databases of the NCBI nr and the *A. thaliana* proteome dataset of TAIR10<sup>70</sup>. Functional classification of the assembled transcripts with the MapMan ontology was performed using the web-based Mercator pipeline<sup>71</sup> (<http://mapman.gabipd.org/web/guest/app/Mercator>). Functional classification of the assembled transcripts with Gene Ontology was performed by integrated use of the BLAST2GO program with BLASTx search (against the Viridiplantae subset of NCBI nr, with a threshold E value less than  $10^{-6}$ ) and InterProScan (default settings on the BLAST2GO) and TAIR GO assignment based on BLASTx search to the proteome dataset of TAIR10 described above. Functional classification of the assembled transcripts with KEGG enzymes was performed using the KAAS web service (<http://www.genome.jp/tools/kaas/>) with



bidirectional mode and using the KEGG assignment function provided by BLAST2GO.

**Differential gene expression analysis.** The Illumina reads were mapped to the assembled transcript sequence set using the bowtie program (ver. 1.0.0)<sup>48</sup> with a default setting in the single end mode. Expression of each transcript was quantified using eXpress<sup>49</sup>. We defined transcripts with FPKM  $\geq 1$  in at least one tissue as significantly expressed. Based on a normalized read count mapped to each transcript, the DEGseq, an R package<sup>72</sup> was used to test and identify differentially expressed transcripts with a FDR-corrected *P* value from Fisher's exact test  $< 10^{-5}$  and fold change  $\geq 5$ .

**GO Enrichment analysis.** GO enrichment analyses of selected transcripts were performed using a GO enrichment analysis function provided by BLAST2GO with a threshold FDR-corrected *P* value from Fisher's exact test  $< 0.05$ . Summarization of GO terms based on their semantic similarities and representation of summarized GOs on the 2D semantic space was performed on the REVIGO web service (<http://revigo.irb.hr/>)<sup>73</sup>.

- Raskin, I. Salicylate, a new plant hormone. *Plant Physiol.* **99**, 799–803 (1992).
- Miyake, K. Some physiological observations on *Nelumbo nucifera*, Gaertn. *Bot. Mag. Tokyo* **12**, 85–101 (1898).
- Nagy, K. A., Odell, D. K. & Seymour, R. S. Temperature regulation by the inflorescence of *Philodendron*. *Science* **178**, 1195–1197 (1972).
- Knutson, R. M. Heat production and temperature regulation in eastern skunk cabbage. *Science* **186**, 746–747 (1974).
- Skubatz, H., Nelson, T. A., Dong, A. M., Meeuse, B. J. & Bendich, A. J. Infrared thermography of *Arum lily* inflorescences. *Planta* **182**, 432–436 (1990).
- Uemura, S., Ohkawara, K., Kudo, G., Wada, N. & Higashi, S. Heat-production and cross-pollination of the Asian skunk cabbage *Symplocarpus renifolius* (Araceae). *Am. J. Bot.* **80**, 635–640 (1993).
- Seymour, R. S. & Schultze-Motel, P. Thermoregulating lotus flowers. *Nature* **383**, 305 (1996).
- Seymour, R. S. & Schultze-Motel, P. Respiration, temperature regulation and energetics of thermogenic inflorescences of the dragon lily *Dracunculus vulgaris* (Araceae). *Proc. R. Soc. B Biol. Sci.* **266**, 1975–1983 (1999).
- Seymour, R. S., Gibernau, M. & Ito, K. Thermogenesis and respiration of inflorescences of the dead horse arum *Helicodiceros muscivorus*, a pseudo-thermoregulatory aroid associated with fly pollination. *Funct. Ecol.* **17**, 886–894 (2003).
- Barthlott, W., Szarzynski, J., Vlek, P., Lobin, W. & Korotkova, N. A torch in the rain forest: thermogenesis of the Titan arum (*Amorphophallus titanum*). *Plant Biol.* **11**, 499–505 (2009).
- Seymour, R. S., Gibernau, M. & Pirintsos, S. A. Thermogenesis of three species of *Arum* from Crete. *Plant Cell Environ.* **32**, 1467–1476 (2009).
- Seymour, R. S. Biophysics and physiology of temperature regulation in thermogenic flowers. *Biosci. Rep.* **21**, 223–236 (2001).
- Seymour, R. S. & Blaylock, A. J. Switching off the heater: influence of ambient temperature on thermoregulation by eastern skunk cabbage *Symplocarpus foetidus*. *J. Exp. Bot.* **50**, 1525–1532 (1999).
- Seymour, R. S. Dynamics and precision of thermoregulatory responses of eastern skunk cabbage *Symplocarpus foetidus*. *Plant Cell Environ.* **27**, 1014–1022 (2004).
- Ito, K., Onda, Y., Sato, T., Abe, Y. & Uemura, M. Structural requirements for the perception of ambient temperature signals in homeothermic heat production of skunk cabbage (*Symplocarpus foetidus*). *Plant Cell Environ.* **26**, 783–788 (2003).
- Seymour, R. S. Pattern of respiration by intact inflorescences of the thermogenic arum lily *Philodendron selloum*. *J. Exp. Bot.* **50**, 845–852 (1999).
- Seymour, R. S., Ito, Y., Onda, Y. & Ito, K. Effects of floral thermogenesis on pollen function in Asian skunk cabbage *Symplocarpus renifolius*. *Biol. Lett.* **5**, 568–570 (2009).
- Seymour, R. S., White, C. R. & Gibernau, M. Environmental biology: heat reward for insect pollinators. *Nature* **426**, 243–244 (2003).
- Meeuse, B. D. & Raskin, I. Sexual reproduction in the arum lily family, with emphasis on thermogenicity. *Sex. Plant Reprod.* **1**, 3–15 (1988).
- Berthold, D. A. & Siedow, J. N. Partial purification of the cyanide-resistant alternative oxidase of skunk cabbage (*Symplocarpus foetidus*) mitochondria. *Plant Physiol.* **101**, 113–119 (1993).
- McIntosh, L. Molecular biology of the alternative oxidase. *Plant Physiol.* **105**, 781–786 (1994).
- Day, D. A. & Wiskich, J. T. Regulation of alternative oxidase activity in higher plants. *J. Bioenerg. Biomembr.* **27**, 379–385 (1995).
- Wagner, A. M. & Krab, K. The alternative respiration pathway in plants: Role and regulation. *Physiol. Plant.* **95**, 318–325 (1995).
- Moore, A. L. *et al.* Unraveling the heater: new insights into the structure of the alternative oxidase. *Annu. Rev. Plant Biol.* **64**, 637–663 (2013).
- Siedow, J. N. & Umbach, A. L. Plant mitochondrial electron transfer and molecular biology. *Plant Cell* **7**, 821–831 (1995).
- Moore, A. L. & Siedow, J. N. The regulation and nature of the cyanide-resistant alternative oxidase of plant mitochondria. *Biochim. Biophys. Acta* **1059**, 121–140 (1991).
- Rasmussen, A. G., Soole, K. L. & Elthon, T. E. Alternative NAD(P)H dehydrogenases of plant mitochondria. *Annu. Rev. Plant Biol.* **55**, 23–39 (2004).
- Kakizaki, Y., Moore, A. L. & Ito, K. Different molecular bases underlie the mitochondrial respiratory activity in the homeothermic spadices of *Symplocarpus renifolius* and the transiently thermogenic appendices of *Arum maculatum*. *Biochem. J.* **445**, 237–246 (2012).
- Porter, R. K. A new look at UCP 1. *Biochim. Biophys. Acta* **1757**, 446–448 (2006).
- Laloi, M. *et al.* A plant cold-induced uncoupling protein. *Nature* **389**, 135–136 (1997).
- Vercesi, A. E. *et al.* Plant uncoupling mitochondrial proteins. *Annu. Rev. Plant Biol.* **57**, 383–404 (2006).
- Borecky, J. *et al.* The plant energy-dissipating mitochondrial systems: depicting the genomic structure and the expression profiles of the gene families of uncoupling protein and alternative oxidase in monocots and dicots. *J. Exp. Bot.* **57**, 849–864 (2006).
- Onda, Y. *et al.* Functional coexpression of the mitochondrial alternative oxidase and uncoupling protein underlies thermoregulation in the thermogenic florets of skunk cabbage. *Plant Physiol.* **146**, 636–645 (2008).
- Ito, K. & Seymour, R. S. Expression of uncoupling protein and alternative oxidase depends on lipid or carbohydrate substrates in thermogenic plants. *Biol. Lett.* **1**, 427–430 (2005).
- Watling, J. R., Robinson, S. A. & Seymour, R. S. Contribution of the alternative pathway to respiration during thermogenesis in flowers of the sacred lotus. *Plant Physiol.* **140**, 1367–1373 (2006).
- Grant, N. M., Miller, R. E., Watling, J. R. & Robinson, S. A. Synchronicity of thermogenic activity, alternative pathway respiratory flux, AOX protein content, and carbohydrates in receptacle tissues of sacred lotus during floral development. *J. Exp. Bot.* **59**, 705–714 (2008).
- Grant, N. *et al.* Two cys or not two cys? That is the question; alternative oxidase in the thermogenic plant sacred lotus. *Plant Physiol.* **150**, 987–995 (2009).
- Onda, Y. *et al.* Pyruvate-sensitive AOX exists as a non-covalently associated dimer in the homeothermic spadix of the skunk cabbage, *Symplocarpus renifolius*. *FEBS Lett.* **581**, 5852–5858 (2007).
- Kakizaki, Y., Seymour, R. S. & Ito, K. A novel functional element in the N-terminal region of *Arum concinatum* alternative oxidase is indispensable for catalytic activity of the enzyme in HeLa cells. *Biochim. Biophys. Acta* **1797**, 20–28 (2010).
- Ito, K. *et al.* Metabolite profiling reveals tissue- and temperature-specific metabolomic responses in thermoregulatory male florets of *Dracunculus vulgaris* (Araceae). *Metabolomics* **9**, 919–930 (2013).
- Seymour, R. S. Scaling of heat production by thermogenic flowers: limits to floral size and maximum rate of respiration. *Plant Cell Environ.* **33**, 1474–1485 (2010).
- Urru, I. *et al.* Pollination strategies in Cretan *Arum* lilies. *Biol. J. Linn. Soc.* **101**, 991–1001 (2010).
- Vesely, P., Bures, P., Smarda, P. & Pavlíček, T. Genome size and DNA base composition of geophytes: the mirror of phenology and ecology? *Ann. Bot.* **109**, 65–75 (2012).
- Martin, J. A. & Wang, Z. Next-generation transcriptome assembly. *Nat. Rev. Genet.* **12**, 671–682 (2011).
- Ward, J. A., Ponnala, L. & Weber, C. A. Strategies for transcriptome analysis in nonmodel plants. *Am. J. Bot.* **99**, 267–276 (2012).
- Hamilton, J. P. & Buell, C. R. Advances in plant genome sequencing. *Plant J.* **70**, 177–190 (2012).
- Martin, J. *et al.* Rnnotator: an automated de novo transcriptome assembly pipeline from stranded RNA-Seq reads. *BMC Genomics* **11**, 663 (2010).
- Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
- Roberts, A. & Pachter, L. Streaming fragment assignment for real-time analysis of sequencing experiments. *Nat. Methods* **10**, 71–73 (2013).
- Ito, K. *et al.* Identification of a gene for pyruvate-insensitive mitochondrial alternative oxidase expressed in the thermogenic appendices in *Arum maculatum*. *Plant Physiol.* **157**, 1721–1732 (2011).
- Cordoba, E., Salmi, M. & León, P. Unravelling the regulatory mechanisms that modulate the MEP pathway in higher plants. *J. Exp. Bot.* **60**, 2933–2943 (2009).
- Estévez, J. M., Cantero, A., Reindl, A., Reichler, S. & León, P. 1-Deoxy-D-xylulose-5-phosphate synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in plants. *J. Biol. Chem.* **276**, 22901–22909 (2001).
- Lois, L. M., Rodríguez-Concepción, M., Gallego, F., Campos, N. & Boronat, A. Carotenoid biosynthesis during tomato fruit development: regulatory role of 1-deoxy-D-xylulose 5-phosphate synthase. *Plant J.* **22**, 503–513 (2000).
- Enfissi, E. M. A. *et al.* Metabolic engineering of the mevalonate and non-mevalonate isopentenyl diphosphate-forming pathways for the production of health-promoting isoprenoids in tomato. *Plant Biotechnol. J.* **3**, 17–27 (2005).
- Yamazaki, M. *et al.* Coupling deep transcriptome analysis with untargeted metabolic profiling in *Ophiorrhiza pumila* to further the understanding of the biosynthesis of the anti-cancer alkaloid camptothecin and anthraquinones. *Plant Cell Physiol.* **54**, 686–696 (2013).
- Xie, F. *et al.* De novo sequencing and a comprehensive analysis of purple sweet potato (*Impomoea batatas* L.) transcriptome. *Planta* **236**, 101–113 (2012).
- Iorizzo, M. *et al.* De novo assembly and characterization of the carrot transcriptome reveals novel genes, new markers, and genetic diversity. *BMC Genomics* **12**, 389 (2011).



58. Santino, A. *et al.* Jasmonate signaling in plant development and defense response to multiple (a)biotic stresses. *Plant Cell Rep.* **32**, 1085–1098 (2013).
59. Ishiguro, S., Kawai-Oda, A., Ueda, J., Nishida, I. & Okada, K. The *DEFFECTIVE IN ANther DEHISCENCE1* gene encodes a novel phospholipase A1 catalyzing the initial step of jasmonic acid biosynthesis, which synchronizes pollen maturation, anther dehiscence, and flower opening in *Arabidopsis*. *Plant Cell* **13**, 2191–2209 (2001).
60. Hyun, Y. *et al.* Cooperation and functional diversification of two closely related galactolipase genes for jasmonate biosynthesis. *Dev. Cell* **14**, 183–192 (2008).
61. Balbi, V. & Devoto, A. Jasmonate signalling network in *Arabidopsis thaliana*: crucial regulatory nodes and new physiological scenarios. *New Phytol.* **177**, 301–318 (2008).
62. Ito-Inaba, Y. *et al.* The gene expression landscape of thermogenic skunk cabbage suggests critical roles for mitochondrial and vacuolar metabolic pathways in the regulation of thermogenesis. *Plant Cell Environ.* **35**, 554–566 (2012).
63. Urru, I., Stensmyr, M. C. & Hansson, B. S. Pollination by brood-site deception. *Phytochemistry* **72**, 1655–1666 (2011).
64. Seymour, R. S. & Schultze-Motel, P. Heat-producing flowers. *Endeavour* **21**, 125–129 (1997).
65. Gottsberger, G., Silberbauer-Gottsberger, I. & Dötterl, S. Pollination and floral scent differentiation in species of the *Philodendron bipinnatifidum* complex (Araceae). *Plant Syst. Evol.* **299**, 793–809 (2013).
66. Stensmyr, M. C. *et al.* Pollination: Rotting smell of dead-horse arum florets. *Nature* **420**, 625–626 (2002).
67. Kite, G. C. The floral odour of *Arum maculatum*. *Biochem. Syst. Ecol.* **23**, 343–354 (1995).
68. Albre, J., Quilichini, A. & Gibernau, M. Pollination ecology of *Arum italicum* (Araceae). *Bot. J. Linn. Soc.* **141**, 205–214 (2003).
69. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
70. Lamesch, P. *et al.* The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.* **40**, D1202–D1210 (2012).
71. Lohse, M. *et al.* Mercator: a fast and simple web server for genome scale functional annotation of plant sequence data. *Plant Cell Environ.* **37**, 1250–1258 (2014).
72. Wang, L., Feng, Z., Wang, X., Wang, X. & Zhang, X. DESeq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* **26**, 136–138 (2010).
73. Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* **6**, e21800 (2011).

## Acknowledgments

The authors thank Tadamas Sasaki and Yukiko Uehara-Yamaguchi of RIKEN Center for Sustainable Resource Science, for their valuable suggestions and comments on the manuscript. This work was supported by the Australian Research Council (Discovery Project DP0771854), the Japanese Society for the Promotion of Science (Grant No. 24380182), and the Alexander von Humboldt Foundation to R.S.S. This work was partially supported by a Grant-in-Aid for Scientific Research (B) (Grant No. 22380184) from the Japan Society for the Promotion of Science to K.I. This work was also supported by RIKEN incentive research project in fiscal year 2013 to K.M.

## Author contributions

Y.O., K.M. and K.I. designed the experiments. Y.O. and K.M. wrote the manuscript. K.M., T.Y. and T.S. performed bioinformatics analysis. Y.O., Y.U. and K.I. prepared samples for RNA-seq analysis. R.S.S., S.A.P. and K.S. contributed to biological interpretation of the results. All authors read and approved the final manuscript.

## Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Onda, Y. *et al.* Transcriptome analysis of thermogenic *Arum concinmatum* reveals the molecular components of floral scent production. *Sci. Rep.* **5**, 8753; DOI:10.1038/srep08753 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>