of Hodgkin lymphoma subsequent to FL in a patient receiving maintenance rituximab was reported.[17]

Transformation from FL to ALCL-like undifferentiated lymphoma (pattern 4) has been reported previously as neither transformation of FL nor histological change after rituximab therapy. Although two cases of FL with relapse to peripheral T-cell lymphoma after rituximab have been reported,[33,34] it was suspected that the T-cell lymphomas were another clone, thus differing from the present case. Cohen *et al.* reported large-cell transformation of CLL and FL during or soon after treatment with a fludarabine- and rituximab-containing regimen,[35] thus resembling the present case treated with fludarabine and rituximab.

Several mechanisms of resistance to rituximab have been suggested, including selection of a CD20-negative clone as a consequence of rituximab exposure, masking of CD20 epitopes by rituximab itself, or true loss of CD20 antigen due to genetic and epigenetic changes.[12,13,15,19–24] Although the present study was not intended to address the mechanism of CD20 loss, several remarkable phenomena were evident. No relationships were detected between loss of CD20 and the interval between the last dose of rituximab and rebiopsy, frequency of rituximab administration, clinical responses, and treatment regimens. It is suspected that susceptibility to rituximab differs greatly among lymphomas. Our results also indicated that loss of CD20 immediately after rituximab therapy was not frequent. Although some cases do regain CD20 expression, loss of CD20 persisting for more than 6 months is not infrequent.

A previous case report has emphasized that early relapse of FL after rituximab therapy was related to CD20-negative transformation to DLBCL.[15] However, this was not confirmed in our study using a larger series: most of the relapses were CD20-positive DLBCL, and only two were CD20-negative FL or Hodgkin lymphoma. Our results suggested that CD20-positive relapse with histological transformation occurred most frequently in cases of early relapse, and that CD20-negative relapse was relatively rare.

In conclusion, our findings indicate that 27% of B-NHL show loss of CD20 expression with four histological patterns after rituximab therapy. As the changes in morphology and CD20 expression after rituximab therapy vary widely, including not only loss of CD20 expression but also proliferation of plasmacytoid cells or transformation to special subtypes of lymphoma, and clinical outcomes are very confused, careful follow up and rebiopsy are recommended.

Acknowledgments

The authors are grateful to Ms S. Miura and Ms C. Kina for their excellent technical assistance.

References

1 Reff ME, Carner K, Chambers KS *et al.* Depletion of B cells *in vivo* by a chimeric mouse human monoclonal antibody to CD20. *Blood* 1994; **83**: 435–45.

2 Demidem A, Lam T, Alas S *et al.* Chimeric anti-CD20 (IDEC-C2B8) monoclonal antibody sensitizes a B cell lymphoma cell line to cell killing by cytotoxic drugs. *Cancer Biother Radiopharm* 1997; **12**: 177–86.

3 Shan D, Ledbetter JA, Press OW. Apoptosis of malignant human B cells by ligation of CD20 with monoclonal antibodies. *Blood* 1998; **91**: 1644–52.

4 Cardarelli PM, Quinn M, Buckman D *et al.* Binding to CD20 by anti-B1 antibody or F(ab')(2) is sufficient for induction of apoptosis in B-cell lines. *Cancer Immunol Immunother* 2002; **51**: 15–24.

5 Meeker T, Lewder J, Cteary ML *et al.* Emergence of idiotype variants during treatment of B-cell lymphoma with anti-idiotype antibodies. *N Eng J Med* 1985; **312**: 1658–65.

6 Kinoshita T, Nagai H, Murate T *et al.* CD20-negative relapse in B-cell lymphoma after treatment with rituximab. *J Clin Oncol* 1998; **16**: 3916.

7 Davis TA, Czerwinski DK, Levy R. Therapy of B-cell lymphoma with anti-CD20 antibodies can result in the loss of CD20 antigen expression. *Clin Cancer Res* 1999; **5**: 611–15.

8 Schmitz K, Brugger W, Weiss B *et al.* Clonal selection of CD20-negative non-Hodgkin's lymphoma cells after treatment with anti-CD20 antibody rituximab. *Br J Haematol* 1999; **106**: 571–2.

9 Foran JM, Norton AJ, Micallef IN *et al.* Loss of CD20 expression following treatment with rituximab (chimaeric monoclonal anti-CD20): a retrospective cohort analysis. *Br J Haematol* 2001; **114**: 881–3.

10 Chu PG, Chen YY, Molina A *et al.* Recurrent B-cell neoplasms after Rituximab therapy: an immunophenotypic and genotypic study. *Leuk Lymphoma* 2002; **43**: 2335–41.

11 Kennedy GA, Tey SK, Cobcroft R *et al.* Incidence and nature of CD20-negative relapses following rituximab therapy in aggressive B-cell non-Hodgkin's lymphoma: a retrospective review. *Br J Haematol* 2002; **119**: 412–16.

12 Jilani I, O'Brien S, Manshuri T *et al.* Transient down-modulation of CD20 by rituximab in patients with chronic lymphocytic leukemia. *Blood* 2003; **102**: 3514–20.

13 Goteri G, Olivieri A, Ranaldi M *et al.* Bone marrow histopathological and molecular changes of small B-cell lymphomas after rituximab therapy: comparison with clinical response and patients' outcome. *Int J Immunopathol Pharmacol* 2006; **19**: 421–31.

14 Seliem RM, Freeman JK, Steingart RH *et al.* Immunophenotypic changes and clinical outcome in B-cell lymphomas treated with rituximab. *Appl Immunohistochem Mol Morphol* 2006; **14**: 18–23.

15 Alovaro-Naranjo TA, Jaen-Martinez JJ, Guma-Padro JG *et al.* CD20-negative DLBCL transformation after rituximab treatment in follicular lymphoma: a new case report and review of the literature. *Ann Hematol* 2003; **82**: 585–8.

16 Woehrer S, Streubel B, Chott A *et al.* Transformation of MALT lymphoma to pure plasma cell histology following treatment with the anti-CD20 antibody rituximab. *Leuk Lymphoma* 2005; **46**: 1645–9.

17 Scaramucci L, Perrotti A, Niscola P *et al.* Hodgkin disease subsequent to follicular lymphoma on maintenance rituximab. *Leuk Lymphoma* 2007; **48**: 1878–80.

18 Pijuan L, Vicioso L, Bellosillo B *et al.* CD20-negative T-cell-rich B-cell lymphoma as a progression of a nodular lymphocyte-predominant Hodgkin's lymphoma treated with rituximab: a molecular analysis using laser capture microdissection. *Am J Surg Pathol* 2005; **29**: 1399–403.

19 Cragg MS, Bayne MC, Illidge TM *et al.* Apparent modulation of CD20 by rituximab: an alternative explanation. *Blood* 2004; **103**: 3989–90.

20 Rawal YB, Nuovo GJ, Frambach GE *et al.* The absence of CD20 messenger RNA in recurrent cutaneous B-cell lymphoma following rituximab therapy. *J Cutan Pathol* 2005; **32**: 616–21.

21 Takei K, Yamazaki T, Sawada U *et al.* Analysis of changes in CD20, CD55, and CD59 expression on established rituximab-resistant B-lymphoma cell lines. *Leuk Res* 2006; **30**: 625–31.

22 Tomita A, Hiraga J, Kiyoi H *et al.* Epigenetic regulation of CD20 protein expression in a novel B-cell lymphoma cell line, RRBL1, established from a patient treated repeatedly with rituximab-containing chemotherapy. *Int J Hematol* 2007; **86**: 49–57.

23 Jazirehi AR, Vega MI, Bonavida B. Development of rituximab-resistant lymphoma clones with altered cell signaling and cross-resistance to chemotherapy. *Cancer Res* 2007; **67**: 1270–81.

24 Czuczman MS, Olejniczak S, Gowda A *et al.* Acquirement of rituximab resistance in lymphoma cell lines is associated with both global CD20 gene and protein down-regulation regulated at the pretranscriptional and posttranscriptional levels. *Clin Cancer Res* 2008; **14**: 1561–70.

25 Jaffe ES, Harris NL, Stein H *et al. World Health Organization Classification of Tumours Pathology & Genetics: Tumours of Haematopoietic and Lymphoid Tissues*. Lyon: IARC Press, 2001.

26 Sekiguchi N, Kobayashi Y, Yokota Y *et al.* Follicular lymphoma subgrouping by fluorescence *in situ* hybridization analysis. *Cancer Sci* 2005; **96**: 77–82.

27 Keith TA, Cousar JB, Glick AD *et al.* Plasmacytic differentiation in follicular center cell (FCC) lymphomas. *Am J Clin Pathol* 1985; **84**: 283–90.

28 Vago JF, Hurtubise PE, Redden-Borowski MM *et al.* Follicular center-cell lymphoma with plasmacytic differentiation, monoclonal paraprotein, and peripheral blood involvement: recapitulation of normal B-cell development. *Am J Surg Pathol* 1985; **9**: 764–70.

29 Maeshima AM, Omatsu M, Nomoto J *et al.* Diffuse large B-cell lymphoma after transformation from low-grade follicular lymphoma: morphological, immunohistochemical and FISH analyses. *Cancer Sci* 2008; **99**: 1760–8.

30 Lenz G, Dreyling M, Hoster E *et al.* Immunochemotherapy with rituximab and cyclophosphamide, doxorubicin, vincristine, and prednisone significantly improves response and time to treatment failure, but not long-term outcome in patients with previously untreated mantle cell lymphoma: results of a

doi: 10.1111/j.1349-7006.2008.01005.x

prospective randomized trial of the German Low Grade Lymphoma Study Group (GLSG). *J Clin Oncol* 2005; **23**: 1984–92.
31 Brauninger A, Hansmann ML, Strickler JG *et al.* Identification of common germinal-center B-cell precursors in two patients with both Hodgkin's disease and non-Hodgkin's lymphoma. *N Eng J Med* 1999; **340**: 1239–47.
32 Marafioti T, Hummel M, Anagnostopoulos I *et al.* Classical Hodgkin's disease and follicular lymphoma originating from the same germinal center B cell. *J Clin Oncol* 1999; **17**: 3804–9.
33 Micallef IN, Kirk A, Norton A *et al.* Peripheral T-cell lymphoma following rituximab therapy for B-cell lymphoma. *Blood* 1999; **93**: 2427–8.
34 Tetreault S, Abler SL, Robbins B *et al.* Peripheral T-cell lymphoma after anti-CD20 antibody therapy. *J Clin Oncol* 1998; **16**: 1635–7.
35 Cohen Y, Da'as N, Libstet D *et al.* Large-cell transformation of chronic lymphocytic leukemia and follicular lymphoma during or soon after treatment with fludarabine-rituximab-containing regimens: natural history- or therapy-related complication? *Eur J Haematol* 2002; **68**: 80–3.

JOURNAL OF BIOSCIENCE AND BIOENGINEERING
Vol. 106, No. 6, 587–593. 2008
DOI: 10.1263/jbb.106.587

Knowledge-based Fuzzy Adaptive Resonance Theory and Its Application to the Analysis of Gene Expression in Plants

Hiro Takahashi,[1,2,3]* Hidekazu Iwakawa,[1] Sachiko Nakao,[2] Takahiro Ojio,[2] Ryo Morishita,[2] Satomi Morikawa,[2] Yasunori Machida,[4] Chiyoko Machida,[1,2] and Takeshi Kobayashi[1,2]

Plant Biology Research Center, Chubu University, Matsumoto-cho 1200, Kasugai, Aichi 487-8501, Japan,[1] College of Bioscience and Biotechnology, Chubu University, Matsumoto-cho 1200, Kasugai, Aichi 487-8501, Japan,[2] Genetics Division, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan,[3] and Division of Biological Science, Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8602, Japan[4]

Received 30 April 2008/Accepted 25 August 2008

Gene expression data obtained from DNA microarrays are very useful in revealing the mechanisms that drive life. It is necessary to analyze these data through the use of algorithms, as in clustering and machine-learning. In a previous study, we developed fuzzy adaptive resonance theory (FuzzyART) and applied it to gene expression data, to identify genetic networks. FuzzyART was used as a clustering algorithm that is very suitable for the analysis of biological data; however, although FuzzyART is very useful in the analysis of dozens of gene expression profiles, it is difficult to apply this method to thousands of gene expression profiles, owing to inherent category proliferation and long calculation time. In the present study, we developed a knowledge-based FuzzyART (KB-FuzzyART) to mitigate these problems. We first constructed a gene list-1 from the gene database of *Arabidopsis thaliana* as knowledge for KB-FuzzyART, because KB-FuzzyART requires any knowledge as input. This method was applied to gene expression data obtained via the microarray analysis of *A. thaliana*, to identify the downstream genes of *ASYMMETRIC LEAVES1* (*AS1*) and *ASYMMETRIC LEAVES2* (*AS2*), both of which are involved in leaf development. The results of the analysis using KB-FuzzyART showed that the *KNAT6* and *YABBY5* (*YAB5*) genes are candidates for downstream factors, after a short calculation time for analysis. These results suggest that our gene list-1 is a very useful database for analyzing the expression profiles of genes that are related to the development of *A. thaliana*; they also suggest that the KB-FuzzyART has the high potential to function as a new method by which one can select candidate genes from thousands of genes, using gene expression data on mutant strains.

[**Key words:** gene expression, fuzzy adaptive resonance theory, bioinformatics, knowledge-based systems]

Rapid advances in DNA microarray technologies over the past several years have made it possible to measure the expression levels of thousands of genes simultaneously and under different conditions. The data obtained via microarray analysis are called expression profile data. Many researchers have tried to extract correlated genes from these data by simply clustering them, without using a priori knowledge. If genetic networks can be drawn from these data, we would be able to prioritize target genes that will assist in drug discovery, understanding life mechanisms, and so on. It is thus self-evident, that the identification of genetic networks is significant and important; nonetheless, the candidates for gene interaction are too numerous, and many need to be identified by experimental methods. For the selection of gene interactions, computational methods are now being investigated.

In our previous paper (1), we constructed FuzzyART, based on ART, and applied it to the clustering of gene expression profiles. Cluster construction obtained by FuzzyART showed high reproducibility, and a very high clustering robustness was obtained, even when adding random noise corresponding to a twofold change. This means that ART can select genes with a similar expression pattern, and do so with high robustness. For this reason, we applied FuzzyART to gene expression profiles, to construct genetic networks in our previous paper (2). Although FuzzyART is a very useful clustering algorithm for dozens of gene expression profiles, it is difficult to apply this method to thousands of gene expression profiles, owing to inherent category proliferation and long calculation time. To mitigate these problems, we developed knowledge-based FuzzyART (KB-FuzzyART), and this method was applied in the present study to gene ex-

* Corresponding author. e-mail: hirtakah@isc.chubu.ac.jp
phone/fax: +81-(0)568-51-6369

pression data regarding *Arabidopsis thaliana*. This method is a knowledge-based system that requires knowledge in the form of input; therefore, we constructed a gene database for *A. thaliana* by referring to data from our previous paper (3). This gene database consists of gene lists related in terms of cell fate determination, cell division, and plant hormones. A total of 322 genes are listed in this database.

The *ASYMMETRIC LEAVES2* (*AS2*) gene, a member of the *AS2/LOB* gene family, and the *ASYMMETRIC LEAVES1* (*AS1*) gene of *A. thaliana* are involved in the development of a symmetrical expanded lamina, with the rachis as the axis (4). Mutations in the *AS2* and the *AS1* genes have pleiotropic effects, and the phenotypes of the corresponding mutants are similar to one another, with asymmetrical lobes along the leaf margin, asymmetrically and downwardly curled leaves, malformed vein systems with a less prominent mid-vein, ectopic accumulation of transcripts of class-1 knotted-like homeobox (*KNOX*) genes in mature leaves, an increase in shoot-regeneration ability, and a slight decrease in the adaxialization of leaves (4–10). Thus, the *AS2* and *AS1* genes are involved in the establishment of the entire venation system, which includes the prominent mid-vein as the structural axis of the left-right symmetry of the leaf, the adaxial-abaxial polarity, and the development of a symmetrical expanded lamina. Both *AS2* and *AS1* depress levels of transcripts of class-1 *KNOX* genes, such as *BREVIPEDICELLUS* (*BP*)/*KNAT1* in leaves (11). *AS2* and *AS1* function to maintain leaf cells in a developmentally determinate state. As the phenotypes described above are common to both *as2-1* and *as1-1* mutant plants, the AS2 and AS1 proteins might function in the same pathway. Indeed, a molecular-level association between the two proteins has been proposed as a possible explanation for their actions (8, 12). In addition, both *AS2* and *AS1* are involved in the repression of the expression of abaxial determinant genes, such as *ETTIN(ETT)/AUXINRESPONSE FACTOR3* (*ARF3*) and members of *KANADY* and *YABBY* families (11, 13). In the present study, *as1-1*, *as2-1* mutant, and the expression of *pAS1::AS2* in *as2-1* plants (*as2-1/pAS1::AS2*) were used.

Finally, we could find *KNAT6* and *YABBY5* (*YAB5*) by applying KB-FuzzyART to the gene expression data of *A. thaliana*. These two genes were not included in gene list-1, but they were candidates for downstream genes of *AS1* and *AS2*, as reported by Semiarti *et al.* (4), Garcia *et al.* (14), and Iwakawa *et al.* (13). These results imply that novel genes can be extracted by KB-FuzzyART for gene expression analysis of mutant strains of various species.

MATERIALS AND METHODS

Plant strains and growth conditions *Arabidopsis thaliana* ecotype Col-0 (CS1092) and mutants *as1-1* (CS3374) and *as2-1* (CS3117) were obtained from the Arabidopsis Biological Resource Center (ABRC) (Columbus, OH, USA). The transgenic plant *as2-1/pAS1::AS2* was obtained by introducing plasmids *pGpAS1::AS2* to *as2-1* plants (13). For the analysis of plants, seeds were sown on soil or on Murashige and Skoog (MS) medium. After two days at 4°C in darkness, plants were transferred to a regimen of white light at 50 μmol m^{-2} S^{-1} for 16 h daily at 22°C, as described previously (4). Ages of plants are given in terms of numbers of days after sowing.

Microarray analysis Shoot apices of *Arabidopsis* plants were harvested 15 d after sowing, and immediately frozen in liquid nitrogen and stored at −80°C. Total RNA extracted from each sample of shoot apices containing leaf primordia and young leaves were biotin-labeled and hybridized to high-density oligonucleotide microarrays (Affymetrix Full Genome Array ATH1; Affymetrix, Santa Clara, CA, USA) containing 22,746 probe sets representing approximately 24,000 genes, according to the manufacturer's instructions. The scanned array data were processed by Affymetrix GeneChip Operating Software (GCOS), which scaled the average intensity of all the genes on each array to a target signal of 200.

Quality check of gene expression data Gene expression data were obtained by using an ATH1 chip for Col-0, *as1-1*, *as2-1*, and *as2-1/pAS1::AS2*. The experiments were repeated in triplicate, and all obtained similar results (Spearman's rank correlation coefficients ~0.9). To determine the experiment with the highest-quality data, we applied Kadota's method (15). The data of Col-0, *as1-1*, *as2-1*,

TABLE 1. The gene list-1 (cell fate determination, cell division, and plant hormones)

A. Categories and gene numbers

Category	Subcategory	Number of genes
Cell fate determination related genes	Meristem related genes	10
	Adaxial determinants	6
	Abaxial determinants	9
	Organ differentiation related genes	13
	Small RNA biogenesis related genes	7
	AS2/LOB family	41
	TCP family	17
Cell division related genes	Cell division related genes	61
Plant hormone related genes	TIR1 gene	1
	AUX/IAA family	29
	ARF family	21
	GH3 family	17
	SAUR family	61
	IPT genes	7
	AHK genes	3
	ARR family	18
	PIN gene	1
Total		322

TABLE 1. Continued

B. Member genes

Gene name	AGI code	Gene name	AGI code	Gene name	AGI code
Meristem related genes		Cell division related genes		GH3 family	
BLR	AT5G02030	Arath;CDKB1;2	AT2G38620	GH3-1	AT2G14960
BP	AT4G08150	Arath;CDKB2;1	AT1G76540	GH3-2	AT4G37390
CLV1	AT1G75820	Arath;CDKB2;2	AT1G20930	GH3-3	AT2G23170
CLV3	AT2G27250	Arath;CDKC;1	AT5G10270	GH3-4	AT1G59500
CUC1	AT3G15170	Arath;CDKC;2	AT5G64960	GH3-5	AT4G27260
CUC2	AT5G53950	Arath;CDKD;1	AT1G73690	GH3-6	AT5G54510
CUC3	AT1G76420	Arath;CDKD;2	AT1G66750	GH3-7	AT1G23160
KNAT2	AT1G70510	Arath;CDKD;3	AT1G18040	GH3-8	AT5G51470
STM	AT1G62360	Arath;CDKE;1	AT5G63610	GH3-9	AT2G47750
WUS	AT2G17950	Arath;CDKF;1	AT4G28980	GH3-10	AT4G03400
Adaxial determinants		Arath;CKS1	AT2G27960	GH3-11	AT2G46370
AS1	AT2G37630	Arath;CKS2	AT2G27970	GH3-12	AT5G13320
AS2	AT1G65620	Arath;CYCA1;1	AT1G44110	GH3-13	AT5G13350
CNA	AT1G52150	Arath;CYCA1;2	AT1G77390	GH3-14	AT5G13360
PHB	AT2G34710	Arath;CYCA2;1	AT5G25380	GH3-15	AT5G13370
PHV	AT1G30490	Arath;CYCA2;2	AT5G11300	GH3-16	AT5G13380
REV	AT5G60690	Arath;CYCA2;3	AT1G15570	GH3-17	AT1G28130
Abaxial determinants		Arath;CYCA2;4	AT1G80370	SAUR family	
ARF3/ETT	AT2G33860	Arath;CYCA3;1	AT5G43080	SAUR1	AT4G34770
ARF4	AT5G60450	Arath;CYCA3;2	AT1G47210	SAUR2	AT4G34780
CRC	AT1G69180	Arath;CYCA3;3	AT1G47220	SAUR3	AT4G34790
FIL	AT2G45190	Arath;CYCA3;4	AT1G47230	SAUR4	AT4G34800
KAN1	AT5G16560	Arath;CYCB1;1	AT4G37490	SAUR5	AT4G34810
KAN2	AT1G32240	Arath;CYCB1;2	AT5G06150	SAUR6	AT2G21210
KAN3	AT4G17695	Arath;CYCB1;3	AT3G11520	SAUR7	AT2G21200
KAN4	AT5G42630	Arath;CYCB1;4	AT2G26760	SAUR8	AT2G16580
YAB3	AT4G00180	Arath;CYCB2;1	AT2G17620	SAUR9	AT4G36110
Organ differentiation related genes		Arath;CYCB2;2	AT4G35620	SAUR10	AT2G18010
AN	AT1G01510	Arath;CYCB2;3	AT1G20610	SAUR11	AT5G66260
ANT	AT4G37750	Arath;CYCB2;4	AT1G76310	SAUR12	AT2G21220
AP2	AT4G36920	Arath;CYCB3;1	AT1G16330	SAUR14	AT4G38840
AP3	AT3G54340	Arath;CYCD1;1	AT1G70210	SAUR15	AT4G38850
ARGOS	AT3G59900	Arath;CYCD2;1	AT2G22490	SAUR16	AT4G38860
BOP1	AT3G57130	Arath;CYCD3;1	AT4G34160	SAUR17	AT4G09530
CLF	AT2G23380	Arath;CYCD3;2	AT5G67260	SAUR18	AT3G51200
FWA	AT4G25530	Arath;CYCD3;3	AT3G50070	SAUR19	AT5G18010
JAG	AT1G68480	Arath;CYCD4;1	AT5G65420	SAUR20	AT5G18020
LFY	AT5G61850	Arath;CYCD4;2	AT5G10440	SAUR21	AT5G18030
MYB33	AT5G06100	Arath;CYCD5;1	AT4G37630	SAUR22	AT5G18050
ROT3	AT4G36380	Arath;CYCD6;1	AT4G03270	SAUR23	AT5G18060
SUP	AT3G23130	Arath;CYCD7;1	AT5G02110	SAUR24	AT5G18080
Small RNA biogenesis related genes		Arath;CYCH;1	AT5G27620	SAUR25	AT4G13790
AGO1	AT1G48410	Arath;DEL1	AT3G48160	SAUR26	AT3G03850
AGO7	AT1G69440	Arath;DEL2	AT5G14960	SAUR27	AT3G03840
DCL1	AT1G01040	Arath;DEL3	AT3G01330	SAUR28	AT3G03830
HEN1	AT4G20910	Arath;DPa	AT5G02470	SAUR29	AT3G03820
HST	AT3G05040	Arath;DPb	AT5G03415	SAUR30	AT5G53590
HYL1	AT1G09700	Arath;E2Fa	AT2G36010	SAUR32	AT2G46690
PNH	AT5G43810	Arath;E2Fb	AT5G22220	SAUR33	AT3G61900
AS2/LOB family		Arath;E2Fc	AT1G47870	SAUR35	AT4G12410
ASL1	AT5G66870	Arath;KRP1	AT2G23430	SAUR36	AT2G45210
ASL2	AT2G23660	Arath;KRP2	AT3G50630	SAUR37	AT4G31320
ASL3	AT3G27650	Arath;KRP3	AT5G48820	SAUR38	AT2G24400
ASL4	AT5G63090	Arath;KRP4	AT2G32710	SAUR39	AT3G43120
ASL5	AT2G30130	Arath;KRP5	AT3G24810	SAUR41	AT1G16510
ASL6	AT1G31320	Arath;KRP6	AT3G19150	SAUR43	AT5G42410
ASL7	AT2G28500	Arath;KRP7	AT1G49620	SAUR45	AT2G36210
AS2/LOB family		Arath;Rb	AT3G12280	SAUR46	AT2G37030
ASL8	AT1G07900	Arath;WEE1	AT1G02970	SAUR47	AT3G20220
ASL9	AT1G16530	TIR1 gene		SAUR49	AT4G34750
ASL10	AT2G30340	TIR1	AT3G62980	SAUR50	AT4G34760
ASL11	AT2G40470	AUX/IAA family		SAUR51	AT1G75580
ASL12	AT3G11090	IAA1	AT4G14560	SAUR52	AT1G75590
ASL13	AT3G26660	IAA2	AT3G23030	SAUR53	AT1G19840

TABLE 1. Continued

B. Member genes

Gene name	AGI code	Gene name	AGI code	Gene name	AGI code
ASL14	AT3G26620	IAA4	AT5G43700	SAUR54	AT1G19830
ASL15	AT2G42440	IAA5	AT1G15580	SAUR56	AT1G76190
ASL16	AT3G58190	IAA6	AT1G52830	SAUR57	AT3G53250
ASL17	AT2G31310	IAA7/AXR7	AT3G23050	SAUR59	AT3G60690
ASL18	AT2G42430	IAA8	AT2G22670	SAUR61	AT1G29420
ASL19	AT4G00220	IAA9	AT5G65670	SAUR62	AT1G29430
ASL20	AT2G45420	IAA10	AT1G04100	SAUR63	AT1G29440
ASL21	AT3G03760	IAA11	AT4G28640	SAUR64	AT1G29450
ASL22	AT4G00210	IAA12/BDL	AT1G04550	SAUR65	AT1G29460
ASL23	AT2G45410	IAA13	AT2G33310	SAUR66	AT1G29500
ASL24	AT5G06080	IAA14/SLR1	AT4G14550	SAUR67	AT1G29510
ASL25	AT3G50510	IAA15	AT1G80390	SAUR68	AT1G29490
ASL26	AT4G22700	IAA16	AT3G04730	SAUR69	AT5G10990
ASL27	AT5G35900	IAA17/AXR3	AT1G04250	SAUR71	AT1G56150
ASL28	AT3G27940	IAA18	AT1G51950	SAUR72	AT3G12830
ASL29	AT3G47870	IAA19/MSG2	AT3G15540	IPT genes	
ASL30	AT3G13850	IAA20	AT2G46990	IPT1	AT1G68460
ASL31	AT1G72980	IAA26/PAP1	AT3G16500	IPT3	AT3G63110
ASL32	AT1G06280	IAA27/PAP2	AT4G29080	IPT4	AT4G24650
ASL33	AT1G36000	IAA28	AT5G25890	IPT5	AT1G19040
ASL34	AT2G19510	IAA29	AT4G32280	IPT6	AT1G25410
ASL35	AT2G19820	IAA3/SHY2	AT1G04240	IPT7	AT3G23630
ASL36	AT1G68510	IAA30	AT3G62100	IPT8	AT3G19160
ASL37	AT1G67100	IAA31	AT3G17600	AHK genes	
ASL38	AT3G02550	IAA32	AT2G01200	AHK2	AT5G35750
ASL39	AT5G67420	IAA33	AT5G57420	AHK3	AT1G27320
ASL40	AT3G49940	IAA34	AT1G15050	AHK4	AT2G01830
ASL41	AT4G37540	ARF family		ARR family	
TCP family		ARF1	AT1G59750	ARR1	AT3G16857
TCP2	AT4G18390	ARF2	AT5G62000	ARR3	AT1G59940
TCP3	AT1G53230	ARF5/IAA24/MP	AT1G19850	ARR4	AT1G10470
TCP4	AT3G15030	ARF6	AT1G30330	ARR5	AT3G48100
TCP5	AT5G60970	ARF7/NPH4	AT5G20730	ARR6	AT5G62920
TCP8	AT1G58100	ARF8	AT5G37020	ARR7	AT1G19050
TCP9	AT2G45680	ARF9	AT4G23980	ARR11	AT1G67710
TCP10	AT2G31070	ARF10	AT2G28350	ARR12	AT2G25180
TCP11	AT2G37000	ARF11	AT2G46530	ARR13	AT2G27070
TCP13	AT3G02150	ARF12	AT1G34310	ARR14	AT2G01760
TCP14	AT3G47620	ARF13	AT1G34170	ARR15	AT1G74890
TCP15	AT1G69690	ARF14	AT1G35540	ARR16	AT2G40670
TCP17	AT5G08070	ARF15	AT1G35520	ARR17	AT3G56380
TCP19	AT5G51910	ARF16	AT4G30080	ARR18	AT5G58080
TCP20	AT3G27010	ARF17	AT1G77850	ARR19	AT1G49190
TCP21	AT5G08330	ARF18	AT3G61830	ARR20	AT3G62670
TCP23	AT1G35560	ARF19/IAA22	AT1G19220	ARR21	AT5G07210
TCP24	AT1G30210	ARF20	AT1G35240	ARR23	AT5G62120
Cell division related genes		ARF21	AT1G34410	PIN gene	
Arath;CDKA;1	AT3G48750	ARF22	AT1G34390	PIN1	AT1G73590
Arath;CDKB1;1	AT3G54180	ARF23	AT1G43950		

and *as2-1/pAS1::AS2* in the first experiment and *as2-1/pAS1::AS2* in the second experiment were selected as the lower-quality data. Therefore, we used the third experiment in the present study for analysis.

Data processing We used four gene expression data sets obtained from each strain (Col-0, *as1-1*, *as2-1*, and *as2-1/pAS1::AS2*). Each data set comprised 22,746 plant genes (probe sets). First of all, we calculated the expression signal for all strains and the log2 (ratio) for three strains (*as1-1*, *as2-1*, and *as2-1/pAS1::AS2*) against Col-0 by GCOS. In this experiment, we excluded those genes for which all strain data sets showed an absent call (*i.e.*, detection call determined by GCOS), because absent indicates that the expression signal was undetectable. We also excluded those genes for which all strain data sets showed a no change call (*i.e.*, change call determined by GCOS), because no change indicates that the expression signal is almost equal to one of Col-0. Thus, 5065 genes were selected. In these data, the log2(ratio) values (*as1-1*, *as2-1*, and *as2-1/pAS1::AS2*) that were over twofold or under half-fold were rounded to twofold and half-fold, respectively; the log2 (ratio) values with no change were rounded to 0, to avoid category proliferation in clustering.

Gene list-1 for *Arabidopsis thaliana* The KB-FuzzyART is a knowledge-based system that requires any knowledge. Therefore, we constructed a gene database for *A. thaliana* by referring to our previous paper (3), as shown in Table 1. The *AS1* and *AS2* genes are involved in the early stage of leaf development in *A. thaliana*; both genes repress the expression of meristem-related homeobox genes in leaves (4–6) and genes that specify abaxial cell

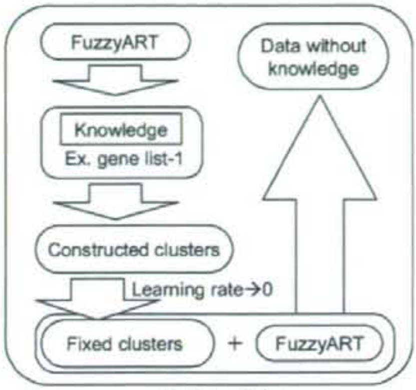

FIG. 1. Concept of KB-FuzzyART. FuzzyART can be applied to any knowledge, such as the gene list-1. The learning rate is then set to 0, to fix constructed clusters. Secondly, FuzzyART with 17 fixed clusters was applied to data (gene set), for analysis. All of the genes were assigned to each cluster or outliers.

fate (10, 13, 14, 16). In the present study, we listed the genes involved in the formation and maintenance of the shoot apical meristem and organ development from the meristem. Second, since the *AS1* and *AS2* genes repress cell proliferation in the adaxial domain of leaves (13), we listed the genes that are related to cell division. Third, since the *AS1* and *AS2* genes are involved in the establishment of venation in leaves (4) and repress the *ETTIN* known as *AUXIN RESPONSE FACTOR3* (13), we listed the genes that are involved in the synthesis of plant hormones and which are regulated by plant hormones. We also added other members of gene families that include the genes described above. This database consists of gene lists related to cell fate determination, cell division, and plant hormones. There were a total of 322 genes in this list; among them, 94 genes are commonly included in the 5065 filtered genes. Therefore, data regarding these 94 genes were used as the knowledge base for KB-FuzzyART.

The algorithm of KB-FuzzyART FuzzyART is used as a clustering algorithm that is very useful in processing biological data (1, 2). However, it difficult for FuzzyART to analyze thousands of gene expression profiles, owing to inherent category proliferation and long calculation time. In the present study, we developed KB-FuzzyART to mitigate these problems, as shown in Fig. 1. First, for KB-FuzzyART, FuzzyART was applied to any knowledge, such as the 94 genes obtained from gene list-1. The FuzzyART parameters (*i.e.*, vigilance parameter and learning rate) were optimized by using an optimal clustering index (OCI) (2). Then, 17 clusters were constructed for the present study. Weight vectors (representative patterns for each cluster) for the 17 clusters were constructed by FuzzyART, and the learning rate was set to 0, to fix the constructed weight vectors. Second, FuzzyART with 17 fixed weight vectors was applied to the 5065 filtered genes. All of the 5065 genes were assigned to each cluster or outliers. The outliers indicate that the patterns of these genes did not match patterns from gene list-1.

RESULTS AND DISCUSSION

Construction of gene list-1 as knowledge for clustering Our new method is a knowledge-based system that requires any knowledge. Therefore, we constructed a gene database for *A. thaliana*. This database consists of gene lists related to cell fate determination, cell division, and plant hormones. The total number of genes in this list was 322, as shown in Table 1; the number of categories reached 17 and includes meristem-related genes, abaxial determinants, and adaxial determinants. We excluded genes that have undetectable calls or no-change calls against Col-0 (wild-type). The filtered 94 genes from among them were used for clustering.

Cluster construction by gene list-1 on the KB-FuzzyART FuzzyART was applied to the filtered 94 genes obtained from gene list-1. The parameters of FuzzyART were optimized by OCI (2) and 17 clusters were then constructed for KB-FuzzyART. For the sake of comparison, a total of 435 clusters were constructed for the original FuzzyART (data not shown). It is difficult to understand the downstream factors, owing to the many clusters constructed by the original FuzzyART.

Assignment of all filtered genes to the KB-FuzzyART FuzzyART with 17 fixed clusters was applied to 5065 filtered genes. The total calculation time for KB-FuzzyART was about 10 seconds; in comparison, FuzzyART lacking knowledge required over 7 d of calculation time. All of the 5065 genes were assigned to each cluster or outliers, as shown in Table 2. The outliers indicate that the patterns of these genes did not match the patterns of gene list-1. As shown in Table 2, outliers comprised only 87 of the 5065 genes. Many genes could be assigned as outliers, if enough kinds of gene expression patterns had not been included in gene list-1. Thus, this result implies that we have listed appropriate genes in gene list-1 that are necessary for constructing a suitable variety of clusters.

Members of the *BP* cluster *AS1* and *AS2* were suggested as repressor genes (4, 11). As shown in Table 2, only cluster 3 included genes that were repressed by *AS1* and *AS2*. The pattern of gene expression for cluster 3 was up-regulation, up-regulation, and down-regulation for *as1-1*, *as2-1*, and *as2-1/pAS1::AS2*, respectively. Cluster 3 included 48 genes, as shown in Table 3; of these, two genes, *BP/KNAT1* and *ETT/ARF3*, are gene list-1 members. *BP/KNAT1* and *ETT/ARF3* were expressed in the shoot apical meristem and reported as candidates of the downstream genes of *AS1* and *AS2* (4–6, 13). Among the remaining genes, two genes—*KNAT6* and *YAB5*—were not gene list-1 members; this is reasonable, given that these two genes have been shown to be candidates of the downstream genes of *AS1* and *AS2* (4, 13, 14). These findings suggest that KB-FuzzyART, using gene list-1, could confirm candidate genes.

Distribution of downstream genes of *AS1* and *AS2* in clusters constructed by original FuzzyART The distribution of candidate downstream genes of *AS1* and *AS2*—namely, *BP/KNAT1*, *ETT/ARF3*, *KNAT6*, and *YAB5*—was investigated for the clustering of original FuzzyART, as shown in Table 4. These four genes clustered only in cluster 3, for the clustering of KB-FuzzyART; however, these genes were distributed in three clusters for original FuzzyART, because enormous clusters were constructed therein. This fact implies that original FuzzyART was not suitable for the present study.

Members of *STM* cluster The *SHOOT MERISTEMLESS* (*STM*) gene, a member of the family of class-1 *KNOX*

TABLE 2. The constructed clusters and assignment of genes

Cluster number	Gene list-1	Categories ingene list-1			Ratio of gene list-1 (%)	Cluster member	Pattern		
		1	2	3			*as1-1*	*as2-1*	*as2-1/ pAS1::AS2*
1	9	1	4	4	2	415	0	−	0
2	9	4	1	4	2	585	0	0	+
3	2	2	0	0	4	48	+	+	−
4	5	3	1	1	1	703	0	+	0
5	7	3	1	3	2	289	−	−	0
6	5	2	2	1	1	530	−	0	0
7	14	3	1	10	2	717	+	0	0
8	2	1	0	1	3	70	0	+	+
9	2	0	1	1	5	37	+	0	+
10	2	1	0	1	2	106	−	0	−
11	12	5	1	6	4	327	+	+	0
12	1	0	0	1	20	5	+	−	−
13	4	0	0	4	7	60	0	−	−
14	14	6	2	6	2	683	0	0	−
15	2	0	0	2	1	180	−	−	−
16	1	1	0	0	1	106	−	−	−
17	3	0	0	3	3	106	+	+	+
Outlier	0	0	0	0	0	98	ND	ND	ND
Over all	94	32	14	48	2	5065	ND	ND	ND

+, −, 0, ND indicate up-regulation, down-regulation, no change, and not determined, respectively. Categories in gene list-1, were cell fate determination related genes, cell division related genes, and plant hormone related genes, respectively.

TABLE 3. Members of cluster 3

Gene list-1 members	Gene symbol	Gene description	Gene list-1 members	Gene symbol	Gene description
+	BP/KNAT1	Homeobox protein knotted-1 like 1	−	COR47	Dehydrin
+	ETT/ARF3	Auxin-responsive factor	−	LTI29	Dehydrin
−	KNAT6	Homeobox transcription factor	−	LEA14	Late embryogenesis abundant protein, putative
−	YAB5	Plant-specific transcription factor YABBY family protein	−	ATGSTF11	Glutathione S-transferase, putative
−	APK	Adenylylsulfate kinase 1	−	−	Expressed protein
−	SGR5	Zinc finger (C2H2 type) family protein	−	−	Avirulence-responsive protein, putative
−	−	Protease inhibitor	−	−	Sulfotransferase family protein
−	−	Flavin-containing monooxygenase family protein	−	−	Lysine and histidine specific transporter, putative
−	−	Expressed protein	−	−	Auxin-responsive family protein
−	ERD7	Senescence/dehydration-associated protein-related	−	−	Meprin and TRAF homology domain-containing protein
−	SUR1	Aminotransferase, putative	−	−	Phenazine biosynthesis PhzC/PhzF family protein
−	LSH10	Expressed protein	−	AOP2	2-Oxoglutarate-dependent dioxygenase, putative
−	CYP79F1	Cytochrome P450 family protein	−	−	Bile acid:sodium symporter family protein
−	−	2-Oxoglutarate-dependent dioxygenase, putative	−	CYP83A1	Cytochrome P450 family protein
−	−	Expressed protein	−	−	Leucoanthocyanidin dioxygenase, putative
−	ATGSTF6	Glutathione S-transferase, putative	−	GBF6	bZIP transcription factor family protein
−	−	Ovule development protein, putative	−	AKN2	Adenylylsulfate kinase 2
−	ERD14	Dehydrin (ERD14)	−	CYP79B2	Cytochrome P450 79B2, putative
−	−	Nodulin MtN21 family protein	−	−	Aconitase C-terminal domain-containing protein
−	MT-1C	Metallothionein-like protein 1C	−	−	Lipid transfer protein, putative
−	−	Isoflavone reductase, putative	−	ATGSTF12	Glutathione S-transferase, putative
−	−	Sulfotransferase family protein	−	−	Germin-like protein, putative
−	−	Glycine-rich protein	−	−	Vesicle-associated membrane family protein
−	−	Expressed protein	−	COR78	Low-temperature-responsive protein 78

genes, is required for the development of the shoot apical meristem, as well as for the maintenance of stem-cell identity throughout the life of the plant. Transcripts of the *STM* gene accumulated in *as1* leaves, although the relative levels were lower than those of the *BP* gene (4). Levels of *STM* transcripts in *as2* leaves sometimes accumulate at lower

TABLE 4. Distribution of downstream genes of *AS1* and *AS2* in clusters constructed by original FuzzyART

Gene symbol	Gene description	Cluster no. (KB-FuzzyART)	Cluster no. (Original FuzzyART)
BP/KNAT1	Homeobox protein knotted-1 like 1	3	30
ETT/ARF3	Auxin-responsive factor	3	141
KNAT6	Homeobox transcription factor	3	141
YAB5	Plant-specific transcription factor YABBY family protein	3	207

levels (4). As shown in Table 2, the pattern of gene expression for cluster 11 was up-regulation and up-regulation for *as1-1* and *as2-1*, respectively; however, cluster 11 included genes that had not been down-regulated (or had been weakly down-regulated) in *as2-1/pAS1::AS2*. Cluster 11 included 12 gene list-1 members and 327 genes from among the 5065 genes (Table 2). Since cluster 11 included *KAN2*, which is an abaxial determinant gene and was shown to be a candidate of the downstream genes of *AS1* and *AS2* (13), as well as *STM*, which is a member of the class 1 *KNOX* gene, it is possible that cluster 11 also includes genes regulated by *AS1* and *AS2*.

In the present study, we constructed gene list-1 and developed KB-FuzzyART, and found that our new method could extract candidate genes such as *KANT6* and *YAB5*. In addition, KB-FuzzyART required a short calculation time. Although our new method requires any knowledge, our results showed that gene list-1 included appropriate genes with regard to plant leaf development and differentiation in *A. thaliana* for the construction of a suitable variety of clusters. However, there are generally no remaining situations where a gene list for analysis cannot be constructed. Therefore, if we can obtain gene expression data from some mutants, the KB-FuzzyART can extract candidates of downstream genes from gene expression data on mutant strains for not only plants, but also bacteria, fungi, and animal cells, among others. These results suggest that the KB-FuzzyART has the high potential to function as a new method of candidate gene selection for thousands of gene or protein expression profiles obtained from DNA microarray, mass spectrometry (MS), and two-dimensional polyacrylamide gel electrophoresis (2D-PAGE).

ACKNOWLEDGMENTS

This work was supported in part by a grant-in-aid for General Scientific Research (no. 19060003, to H.T, C.M., and Y.M) from the Ministry of Education, Culture, Sports, Science, and Technology, Japan, and by a grant from the "Academic Frontier" Project for Private Universities (matching fund subsidy from MEXT18, 2005–2009).

REFERENCES

1. **Tomida, S., Hanai, T., Honda, H., and Kobayashi, T.:** Analysis of expression profile using fuzzy adaptive resonance theory. Bioinformatics, **18**, 1073–1083 (2002).
2. **Takahashi, H., Tomida, S., Kobayashi, T., and Honda, H.:** Inference of common genetic network using fuzzy adaptive resonance theory associated matrix method. J. Biosci. Bioeng., **96**, 154–160 (2003).
3. **Terakura, S., Ueno, Y., Tagami, H., Kitakura, S., Machida, C., Wabiko, H., Aiba, H., Otten, L., Tsukagoshi, H., Nakamura, K., and Machida, Y.:** An oncoprotein from the plant pathogen agrobacterium has histone chaperone-like activity. Plant Cell, **19**, 2855–2865 (2007).
4. **Semiarti, E., Ueno, Y., Tsukaya, H., Iwakawa, H., Machida, C., and Machida, Y.:** The ASYMMETRIC LEAVES2 gene of *Arabidopsis thaliana* regulates formation of a symmetric lamina, establishment of venation and repression of meristem-related homeobox genes in leaves. Development, **128**, 1771–1783 (2001).
5. **Byrne, M. E., Barley, R., Curtis, M., Arroyo, J. M., Dunham, M., Hudson, A., and Martienssen, R. A.:** Asymmetric leaves1 mediates leaf patterning and stem cell function in *Arabidopsis*. Nature, **408**, 967–971 (2000).
6. **Ori, N., Eshed, Y., Chuck, G., Bowman, J. L., and Hake, S.:** Mechanisms that control *knox* gene expression in the *Arabidopsis* shoot. Development, **127**, 5523–5532 (2000).
7. **Sun, Y., Zhou, Q., Zhang, W., Fu, Y., and Huang, H.:** ASYMMETRIC LEAVES1, an *Arabidopsis* gene that is involved in the control of cell differentiation in leaves. Planta, **214**, 694–702 (2002).
8. **Xu, L., Xu, Y., Dong, A., Sun, Y., Pi, L., Xu, Y., and Huang, H.:** Novel as1 and as2 defects in leaf adaxial-abaxial polarity reveal the requirement for ASYMMETRIC LEAVES1 and 2 and ERECTA functions in specifying leaf adaxial identity. Development, **130**, 4097–4107 (2003).
9. **Zgurski, J. M., Sharma, R., Bolokoski, D. A., and Schultz, E. A.:** Asymmetric auxin response precedes asymmetric growth and differentiation of asymmetric leaf1 and asymmetric leaf2 *Arabidopsis* leaves. Plant Cell, **17**, 77–91 (2005).
10. **Ueno, Y., Ishikawa, T., Watanabe, K., Terakura, S., Iwakawa, H., Okada, K., Machida, C., and Machida, Y.:** Histone deacetylases and ASYMMETRIC LEAVES2 are involved in the establishment of polarity in leaves of *Arabidopsis*. Plant Cell, **19**, 445–457 (2007).
11. **Guo, M., Thomas, J., Collins, G., and Timmermans, M. C.:** Direct repression of KNOX loci by the ASYMMETRIC LEAVES1 complex of *Arabidopsis*. Plant Cell, **20**, 48–58 (2008).
12. **Phelps-Durr, T. L., Thomas, J., Vahab, P., and Timmermans, M. C.:** Maize rough sheath2 and its *Arabidopsis* orthologue ASYMMETRIC LEAVES1 interact with HIRA, a predicted histone chaperone, to maintain *knox* gene silencing and determinacy during organogenesis. Plant Cell, **17**, 2886–2898 (2005).
13. **Iwakawa, H., Iwasaki, M., Kojima, S., Ueno, Y., Soma, T., Tanaka, H., Semiarti, E., Machida, Y., and Machida, C.:** Expression of the ASYMMETRIC LEAVES2 gene in the adaxial domain of *Arabidopsis* leaves represses cell proliferation in this domain and is critical for the development of properly expanded leaves. Plant J., **51**, 173–184 (2007).
14. **Garcia, D., Collier, S. A., Byrne, M. E., and Martienssen, R. A.:** Specification of leaf polarity in *Arabidopsis* via the trans-acting siRNA pathway. Curr. Biol., **16**, 933–938 (2006).
15. **Kadota, K., Tominaga, D., Akiyama, Y., and Takahashi, K.:** Detecting outlying samples in microarray data: a critical assessment of the effect of outliers on sample classification. Chem-Bio Inform. J., **3**, 30–45 (2003).
16. **Li, H., Xu, L., Wang, H., Yuan, Z., Cao, X., Yang, Z., Zhang, D., Xu, Y., and Huang, H.:** The putative RNA-dependent RNA polymerase RDR6 acts synergistically with ASYMMETRIC LEAVES1 and 2 to repress BREVIPEDICELLUS and MicroRNA165/166 in *Arabidopsis* leaf development. Plant Cell, **17**, 2157–2171 (2005).