Manipulation of biosynthesis of aliphatic glucosinolates by gene replacement and RNAi gene silencing

Zhixia Niu[1], Muqiang Gao[2], Peter B. E. McVetty[1], Fouad Daayf[1], Hongyan Zhu[2], Carlos F. Quiros[3], Genyi Li[1]

[1.] *Department of Plant Science, University of Manitoba, Winnipeg, MB R3T2N2 Canada*
[2.] *Department of Plant and Soil Science, University of Kentucky, KY 40546-0236, USA*
[3.] *Department of Plant Sciences, University of California, Davis, CA 95616, USA Email: umniuzx@cc.umanitoba.ca*

Abstract

Glucosinolates are a large group of plant secondary metabolites found mainly in the order Capparales. Manipulation of biosynthesis of aliphatic glucosinolates was attempted using gene replacement and RNAi gene silencing. Experiments were conducted to replace the functional Alk allele in *Brassica rapa* with the non functional Alk allele from *Brassica oleracea* (broccoli) to develop glucoraphanin-rich *B. rapa*. Gene introgression lines but no gene replacement lines were obtained. Four MAM-like genes involved in glucosinolate biosyntheses, MAM1(At5g23010), MAML(At5g23020), MAML-3 (At1g18500) and MAML-4 (At1g74040) were silenced by RNAi techniques in *Arabidopsis* and the resulting glucosinolate profiles were studied. All four MAM-like genes, MAM1, MAML, MAML-3 and MAML-4 were shown to be involved in the aliphatic glucosinolate elongation pathway.

Key words: Gene replacement, MAM-like gene, RNAi, Alk gene, *Arabidopsis thaliana, Brassica rapa*

Introduction

Glucosinolates are a large group of plant secondary metabolites found mainly in the order Capparales, which includes a large number of economically important *Brassica* crops and the model plant Arabidopsis. Glucosinolates have a thioglucose moiety, a sulfonated oxime, and a side chain derived from aliphatic, aromatic, or indole amino acids (Mikkelsen *et al*. 2000). Myrosinase catalyses hydrolysis of the glucosinolates when tissues are damaged and when moisture is present. This produces three biologically active products, isothiocyanates, thiocyanates and nitriles (Bones and Rossiter, 1996), some of which are beneficial and others detrimental. For example they could exhibit goitrogenic or antithyroid activity, plant defense effects or antitumorogenic activity in mammals (Halkier et al., 2006). For example, the glucoraphanin, the main component of glucosinolates in broccoli hydrolyses into sulforaphane, an isothyocianate that has strong antitumorogenic effects.

The Alk gene is involved in the glucosinolate modification pathway. In *Brassica oleracea* (broccoli), the Alk gene is non-functional so broccoli accumulates glucoraphanin. On the other hand *B. rapa* has a functional Alk allele, resulting in the desaturation of of glucoraphanin into gluconapin and hydroxylation of this compound into progoitrin, an antinutrient causing goiter. Therefore, manipulation of Alk gene expression may be an effective way to develop glucoraphanin rich vegetables. We report the introgression of the non-functional Alk allele from broccoli to *B. rapa* to replace the functional allele by wide hybridization. The resulting hybrids were backcrossed and marker assistant selection was applied to develop glucoraphanin rich *B. rapa* using non-GMO techniques.

There are three steps involved in the glucosinolate biosynthesis pathway, the elongation of the amino acid carbon chain, the formation of the basic glucosinolate skeleton, and further side chain modification (Halkier et al., 2006). Aliphatic glucosinolates are derived from methionine. Genetic studies in *B. oleracea* and Arabidopsis have provided information on aliphatic glucosinolate biosynthesis pathways. In the Arabidopsis ecotype Columbia, there are two genes MAM1(At5g23010) and MAML(At5g23020) located on chromosome 5 involved in the glucosinolate elongation pathway, where MAM1 controls the ratio of three carbon and four carbon glucosinolates and MAML controls the long carbon chain synthesis (6C, 7C and 8C) (Field, 2004). Both genes have very high similarity to two other genes, MAML-3 (At1g18500) and MAML-4 (At1g74040) which are involved in the leucine biosynthesis pathway (Kroymann, 2001). Heterologous expression of all the MAM-like genes showed that only MAML-3 is involved in the leucine biosynthesis pathway while MAML-4 maybe also be involved in the glucosinolates pathway. Plants over expressing the MAML gene contained two novel amino acids (Field et al. 2004). The literature suggests that all these MAM-like genes maybe involved in both pathways depending on the plant condition and both pathways may have crosstalk. Knocking out a single MAM-like gene cannot provide an overall view of the function of these genes because of gene compensation effects. RNA inference (RNAi) is a new technique for gene function analysis, especially for multiple gene family functional analysis (Hammond et al., 2000). RNAi was used to silence more than one MAM-like gene in *Arabidopsis* simultaneously to analyze their function.

Materials and Methods

Plant Growth: Brassica plants were sown directly in pots grown in greenhouse under standard conditions. For screening purposes, Arabidopsis seeds were surface-sterilized and plated on half strength MS medium (Murashige and Skoog salts plus vitamins, [sigma], 3% sucrose, 0.8% agar[Fisher Chemicals], pH 5.8) with 50mg/L kanamycin and 100mg/L carbencynin

added. The seeds were stratified for 1 to 2 days in the dark before germination in a growth room. Ten to 14 days later after seeds germinated, the seedlings were transferred to small pots and grown in a growth chamber (16 hours light, 20° C / 8 hours dark, 16° C).

Interspecific crosses, backcrosses and marker assisted selection: *B. oleracea* (broccoli), cv. Luling (B12) was used as the donor parent for the non-functional Alk gene and *B. rapa*, cv. Feng-kang 70 (DH line)**,** was used as the recipient parent. First, a cross between these two species was made, followed by colchicine doubling to create resynthesizized *B. napus*. Second, the resynthesized *B. napus* was crossed with *B. rapa*, cv. RI16 (Chinese cabbage) three times, then one selfing was done to develop a series of chromosome addition lines in *B. rapa*. Third, these chromosome addition lines were used as male parents to backcross further to *B. rapa* as female parent to diminish the transmission of the additional chromosomes. Fourth, marker assisted selection (MAS) was used to select the plants carrying the non-functional Alk allele followed by selfing to obtain homozygous gene introgression or gene replacement lines.

Chromosome counting: The flower buds were fixed in a 3:1 ethanol: glacial acetic acid for 24 hrs, and then transferred to 70% ethanol. The chromosomes in pollen mother cells were counted by using the standard 1% aceto-carmine anther squash method.

Bioinformatics Analysis and Primer design: The Genome sequence and CD sequences for four MAM-like genes in *Arabidopsis thaliana* ecotype Columbia were searched through the Arabidopsis Information Resource web site - TAIR (www.arabidopsis.org) and BLAST (Basic Local Alignment Search Tool) from the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/BLAST/) was used to identify high similarity regions between the two sequences located in the same chromosomes.

Construction of RNAi constructs: MAM1, MAML, MAML-3 and MAML-4 gene sequences were isolated by PCR using genomic DNA as a template; MAM1 and MAMl-4 were also isolated by PCR using cDNA as a template. PCR products were separated on 1% agarose gels, excised and gel-purified. The purified fragments were cloned to binary vector pFGC5941or pHELLSGATE 8.

Sequencing*:* Sequencing was performed using the ABI BigDye terminator system (Applied Biosystems) according to manufacturer's instructions.

DNA extraction, RNA extraction and cDNA Synthesis: DNA was extracted from leaves of each line using a modified CTAB method. About 500mg leaf tissue samples were ground in liquid nitrogen and total RNA was extracted according to Logmann et al. (1987). Five micrograms of total RNA was used to make first strand cDNA using Superscript III (Invitrogen) in a 20 µl reaction with oligo (dT) primers according to the manufacturer's instructions and in the subsequent PCR, 0.6 µl of the products was used per 10ul of reaction mix.

GSL extraction and HPLC analysis: Glucosinolates were extracted and purified using sephadex/sulfatase protocols previously described (Kliebenstein et al., 2001). Forty to fifty microliters of the glucosinolate extract was run on a RP18 column (Lichrocart, Fisher, Canada) on a Hewlett-Packard 1100 series HPLC. Compounds were detected at 229 nm and separated using the programs previously described   (Kliebenstein et al., 2001). The identity of HPLC peaks was based on a comparison of retention time and UV absorption spectrum as determined on a diode-array detector with those of standards. Results are given as µmol per g fresh weight tissue calculated from HPLC peak areas using response factors computed for pure de-sulfo glucosinolate standards at A229 nm (Brown et al 2003). Each line was run in triplicate.

Results

ALK Gene introgression from *B. oleracea**:* Three backcross cycles with marker assisted selection were performed after the *B. rapa* chromosome addition lines were obtained. Two lines (14-7-2 and S1-10-5) carried both the functional allele Alk[+] and the non functional allele Alk- (Figure 1A). Chromosome counting indicated that there was no extra chromosome in these lines (Figure 1B). Selfing or test crossing these two lines with *B. rapa* showed that the functional allele Alk+ was always present and that the non functional allele Alk- segregated 1 to 1 in the test cross and 3 to 1 in selfed progeny. Both lines displayed monogenic   Mendelian   inheritance for the non functional allele Alk [-]. Thus, the non functional allele Alk[-] from *B. oleracea* had been introgressed in the *B. rapa* genome. However, no change in   glucosinolate profile was observed   in the progeny in relation to *B. rapa* (data not shown).
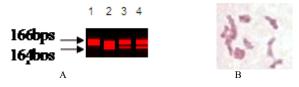

Figure 1: A) Molecular markers for the alk+ and alk- alleles (1: alk+ allele; 2: alk- allele; 3: 14-7-2; 4: S1-10-5); B) chromosome squash for line 14-7-2 showing normal chromosome number

Phenotypes of the RNAi transformed lines: Six RNAi constructs for MAM-like gene in *Arabidopsis thaliana (Columbia)* were assembled: four from genome sequence (1: at1g18500, 2: at1g74040, 5: at5g23010 and 6: at5g23020) and two from cDNA sequence (1B:at1g74040 and 5A: at5g23010). For each RNAi construct, two clones were chosen to transform *Arabidopsis thaliana (Columbia)* through the floral dip method (Clough and Bent 1998). After screening the seeds on half strength MS with 50mg/l kanamycin, the survivors had three different phenotypes: a) normal type, b) short and weak with normal leaves and c) big crinkled leaves with wearied siliques (data not shown).

RT-PCR: RT-PCR was performed for the 1B RNAi construct knockout line. All four MAM-like gene RNA expression levels were tested simultaneously (data not shown). Results showed that the 1B construct decreased the expression of three MAM-like genes (MAML, MAML-3 and MAML-4) simultaneously but did not affect the MAM1gene.

Analysis of Glucosinolates profiles of transformed lines: The glucosinolate profiles were analyzed by HPLC (Table 2). Results showed that the glucosinolate profiles changed significantly in the RNAi knockout lines. The total content of aliphatic glucosinolate and the composition of the aliphatic glucosinolates changed. In 1B (23), 3C glucosinolates decreased about 8 times and 4C glucosinolates decreased around 12 times compared to the wild type while the 3C/4C ratio was similar to the wild type. In 5A (37), 3C glucosinolates did not change significantly, but 4C glucosinolates decreased more than 138 times. In both lines, long chain glucosinolates decreased significantly.

Discussion

In the interspecific crosses, when alien chromosomes existed, tri- or quadrivalents formed between homoeologs depending on sequence similarity. The Alk gene in *B. rapa* and *B. oleracea* has high sequence similarity (only two base deletion differences). Theoretically, gene replacement should happen at very low rate, however, no replacement lines for the Alk gene were found. This suggests that there could be more than one functional Alk locus in *B. rapa* and to replace all of them simultaneously would be very difficult.

In recent years, it has been reported that MAM1 and MAML genes are mainly involved in the glucosinolate side chain elongation pathway. MAM1 controls 3C to 4C variation in Arabidopsis, and MAML controls long chain elongation (Field et al. 2004). Results showed that knocking out MAML-3 and MAML-4 genes but leaving a functional MAM1gene also reduced the 3C and 4C glucosinolate content. This agrees with the recent research results which showed that expression of a *Brassica* isopropylmalate synthesis gene in *Arabidopsis* perturbs both glucosinolate and amino acid metabolism (Field, 2006). In our experiments, three MAM-like genes were silenced simultaneously using one RNAi construct. This could be very useful for gene function analysis in polyploidy species.

Table 2 Glucosinolates content (umol/g fresh wt) in RNAi knockout lines and wild type Arabidopsis

	GSL name	Response Factor	Col-0	5a(37)	1b(23)
1	3MSOP	1.2	1.67	1.17	0.22
2	4OHB	1.4	4.39		0.32
3	4MSOB	0.9	10.48	0.11	0.93
4	6MSOH	1	0.13		
7	4OHI3M	1	0.39	0.28	0.15
8	7MSOH	1	0.71		
9	4MTB	0.9	0.40		
10	8MSOO	1.1	3.82		0.13
11	I3M	0.3	3.11	6.40	3.00
12	4MOI3M	0.3	1.78	0.69	0.70
13	1MOI3M	0.2	0.38	0.17	0.09
14	7MTH	1	1.73	0.20	
15	8MTO	1	0.40	0.26	0.24
sum	3C		1.67	1.17	0.22
sum	4C		15.27	0.11	1.26
sum	7C		2.44	0.20	
sum	8C		4.22	0.26	0.38
sum	indol		5.66	7.54	3.94
Ratio	3C/4C		0.11	10.64	0.17
Total			29.26	9.28	5.8

References

Bones A.M., and Rossiter J.T. 1996. The myrosinase-glucosinolate system, its organization and biochemistry. Plant Physiology **97**:184-208.

Brown P. D., Tokuhisa J. G, Reichelt M and Gershenzon J. 2003. Variation of glucosinolate accumulation among different organs and developmental stages of *Arabidopsis thaliana*. Phytochemistry **62**: 471-481.

Clough S.J., Bent A.F. 1998. Floral dip: a simplified method for Agrobacterium mediated transformation of *Arabidopsis thaliana*. Plant J **16**: 735–743.

Field B., Cardon G., Traka M., Botterman J., Vancanneyt G., and Mithen R. 2004. Glucosinolate and Amino Acid Biosynthesis in Arabidopsis. Plant Physiology **135**:828-839.

Field B., Furniss C., Wilkinson A., and Mithen R. 2006. Expression of a *Brassica* isopropylmalate synthesis gene in Arabidopsis perturbs both glucosinolate and amino acid metabolism. Plant Molecular Biology **60**:717:727.

Halkier B.A. and Gershenzon J. 2006. Biology and Biochemistry of Glucosinolates. Annual Review of Plant Biology **507**: 303-333.

Hammond S. M., Bernstein, E., Beach D., and Hannon G. J. 2000. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* **404**: 293-6.

Kliebenstein D.J., Kroymann J., Brawn P., Figuth A., Pedersen D., Gershenzon J., and Mitchell-Olds T. 2001. Genetic contorol of natural variation in *Arabidopsis* glucosinolate accumulation. Plant Physiology **126**:811-825.

Kroymann J., Textor S., Tokuhisa J.G., Falk K.L., Bartram S., Gershenzon J., and Mitchell-Olds T. 2001. A gene controlling variation in Arabidopsis glucosinolate composition is part of the methionine chain elongation pathway. Plant Physiol **127**: 1077–1088.

Logemann, J., Schell J. and Willmitzer L. 1987. Improved method for the isolation of RNA from plant tissues. Analytical Biochemistry **163**: 16-20

Mikkelsen M.D., Hansen C.H., Wittstock U., and Halkier B.A. 2000. Cytochrome P450 CYP79B2 from *Arabidopsis* catalyzes the conversion of tryptophan to indole-3 acetaldoxime, a precursor of indole glucosinolates and indole-3-acetic acid. Journal of Biological Chemistry **275**: 33712-33717.