Comparison of Gene Coding Clones Content in *In vivo* and *In vitro* Methyl - Filtration Libraries of Maize (*Zea may*)

Myung Chul Lee, Rod A Wing[1], Seok-Cheol Suh* and Moo Young Eun
National Institute of Agricultural Biotechnology, RDA, Suwon 441-707, Korea
[1]Clemson University Genomics Institute, Room 100, Jordan Hall, Clemson, SC 29634, USA

Abstract - It has been hypothesized that efficient exclusion of methylated retrotransposons and repeated DNA region is one of the rapid and cost-effective approaches for comprehensive gene discovery in large genome size of maize. Three kinds of methylation-sensitive restriction enzymes, *Hap*II, *Msp*I and *Mcr*BC, were used to identify the restriction frequency of cytosine methylation sites in maize genome. Roughly 60% of total maize genomic DNA was restricted less than 500bp by *Mcr*BC, and the most of restricted small size fraction was composed retrotransposon. In order to validate the efficient construction of gene-rich shotgun library, we compare two gene-rich methyl-filtration shotgun libraries using *in vivo* and *in vitro* methyl-filtration system. The size selected DNA fraction by *Sau*3A-*Mcr*BC enzyme treated was very stable and has not appeared modification in *E.coli*, but most insert DNA size of partially digested with *Sau*3A were decrease less than 500bp by bacterial methylation-modification system. In compare of retroelements portion, A 44.6% of the sequences were retroelement in unmethyl-filtered library, and the most of them was Copia type, such as Prem, Opie and Ji. The portion of retroelement was drastically decreased to 25% and 20% by *in vivo* and *in vitro* filtration system, respectively.

Key words - Methylation, Maize, Methyl-filtration, McrBC, Retroelement

Introduction

The genome size of the grass family, such includes rice, sorghum, barely, maize, wheat and oat, is highly variable from 430 Mb in rice to 16,000 Mb in wheat (Bennet and Leitch, 1995). This variation in genome size is largely dependent on differences in an amount of repetitive and one of the most abundant repetitive elements is retrotransposons DNA (Flavell and Smith, 1976).

The maize genome, which has about 2,500Mb haploid genome size, is comprised of 60∼80% repetitive elements (SanMiguel and Bennetzen, 1998). Most repetitive elements are interspersed repeats that vary in copy number from about tens to ten thousands per haploid nucleus (SanMiguel *et al.*, 1996; Tikhonov *et al.*, 1996; Llaca and Messing 1998). A large fraction of maize repetitive DNA was primarily composed of retrotransposons and derivatives, both of which are highly methylated at cytosine in the 5' -CG-3' and 5' -CNG-3' . In contrast, the small portion of hypomethylated CpG islands is the coding region about actual functional genes, and intermixed with large stretches of repetitive or junk DNA. The gene-rich regions of maize present in very long compositionally homologous DNA segments, which represent 10∼20% of the genome. (Antequera and Bird, 1988; Springer, 1992; Bennezen *et al.*, 1994; Carels *et al.*, 1995).

Recently, several studies focused on finding a way to analyze about unmethylated genomic DNA fragments that might prove helpful in obtaining genomic coverage of gene-rich chromosomal region by remove hypermethyled repetitive region. Fu and Dooner (2000) constructed maize BAC (bacterial artificial chromosome) library by complete digestion with *Not*I, which is one of the methylation-sensitive and rare-cutting enzymes. Even though the constructed BAC library contained the two-thirds of a maize genome, the BAC library could represent highly significant gene-enriched genomic region, especially in the bronze (*bz*) allele region. Foreign methylated DNA modification system of *E.coli* was also used to construct the maize genomic shotgun library (Rabinowicz *et al.*, 1999; Balke *et al.*, 2000). Large portion of maize retroelements was highly excluded by methylation modification of *E.coli*. These results from random sequencing of maize genomic shotgun library suggested that methyl-filtration system using *E.coli* is an efficient approach to identify the numerous unique genes from complex of large genome size plants.

In the methylation restriction by *E.coli* host strain, named *in vivo*

*Corresponding author. E-mail : scsuh@rda.go.kr

methyl-filtration system, however, it was revealed that most of the insertion sizes were seriously decreased to less than 500bp from our preliminary study. Furthermore, frequency of methylation sites modification was different by the choice of the bacterial host strains, consequently conferring partial exclusion of repetitive elements. To overcome this restriction, we tested a novel method using methyl sensitive restriction enzyme, named *in vitro* methyl-filtration system, and here we reported the results.

Materials and Methods

High-molecular-weight (HMW) DNA isolation

High-molecular-weight (HMW) DNA was isolated from young leaves of maize (*Zea may)* inbred line B73 that grown at green house conditions in order to reduce the chloroplast DNA contamination. HMW DNA plugs were prepared by embedding nuclei into 0.5% of low-melting agarose and stored in 0.5M EDTA at 4℃ with modified of NIB buffer (De Scenzo and Wise, 1996; Yang *et al.,* 1997). Modified NIB buffer component was 10mM Tris, 10mM EDTA, 100mM KCl, 500mM sucrose, 4mM spermidine, 2% polyvinylpolypyrrolidion, 1 mM spermine and 0.1% β-mercaptoethanol at pH 9.4-9.5. The HMW DNA plugs were washed two times in Tris-EDTA (TE) with 1mM phenylmethylsulfonyl fluoride (PMSF) and three times in distilled H_2O before using.

Analysis of methyl cytosine content

About 5μg HMW DNA was separately digested with three kinds of methylation-sensitive restriction enzymes, *Hap*II, *Msp*I (Promega Co.) and McrBC (New England BioLab). DNA plugs were equilibrated twice with each restriction enzyme buffer for 1hr at 4 ℃. Twenty units of each enzyme were added and the enzyme was allowed to diffuse into the plug for 1 hr on ice. Complete digestion was achieved by incubating at 37℃ for 16 hrs.

Samller than 500 bp DNA fractions after *Mcr*BC restriction were purified from the low melting point agarose gel by QuiaGen Cleaning Kit (QuiaGen Co.), treated with T4 DNA polymerase (Promega Co.) to repair fragment ends. To analyze the retroelement distribution in this fraction, end repaired fragments were ligated pGem11Zf(-) (Promega Co.) vector and transformed into *E.coli* strain DH10H (Invitrogen Co.) according to manufacture' s instructions.

Construction of maize shotgun genomic libraries

About 30μg of HMW DNA was partially digested with serially diluted *Sau3*AI enzyme (0, 0.5, 1.0, 2.5, 5.0 units) by incubation at 37℃ for 10 minutes. The digested DNA was electrophoresed on the 1% agarose gel and size selected in the size range of 0.5-1, 1-1.5, 1.5-2 and 2-3 kb. Each DNA fractions were excised and ligated with pGEM-11Zf(-). One hundred ng of ligation mixture was electrophoresed into different *E.coli* strains, DH10B (*mcr*A, *mcr*BC) and DH5α (*Mcr*A, *Mcr*BC) of ElectroMax electro-competent cells (Invitrogen Co.) and XL1-Blue (*Mcr*A, *Mcr*BC) (Stratagene Co.), to construct the normal and *in vivo* methyl-filtered maize shotgun genomic library.

About 50μg of HMW DNA were digested with 50units of *Mcr*BC by incubation at 37℃ for overnight. DNA plugs of *Mcr*BC digestion were washed twice with TE buffer and saturated *Sau3*AI buffer at 4 ℃ for 2hr. Serially diluted *Sau3*AI (0, 0.5, 1.0, 2.5, 5.0 units) were used to additional digest the high molecular weight (> 8,000 bp) DNA remaining after *Mcr*BC restriction. According to the same methods as described above, the double enzyme digested DNA was used to construct *in vivo* methyl-filtered shotgun genomic library.

DNA sequencing and analysis

Transformants were randomly selected from each size-selected library for sequencing, and plasmid DNA was purified by AutoGen 740 plasmid isolation system (Autogene Co.). DNA sequencing was performed according to manufacture' s instructions for the Big Dye terminator chemistry (Perkin Elmer) with T7 or SP6 primer, and analyzed by ABI 3700 automated sequencer. The sequence traces were base called by Phred, and vector sequences were masked by Cross-Match software package (Ewing and Green, 1998). Analysis of sequence homology search was performed using BLASTN and BLASTX at the GenBank. For further identification of unidentified clones was defined comparing with TIGR maize gene index (http://www.tigr.org/tdb/tgi.shtml). Potential exon sequences were predicted using GenScan software at the web site (http://genes. mit.edu/GENSCAN. html). Transposable elements were identified by compare with GenBank database. Additional transposable elements were detected the TFASTA searches of STC database using known MITEs (Miniature Inverted-repeat Transposable Element) (Pearson *et al.,* 1997).

Results

Estimation of methylcytosine amount in maize

We estimated relative amount of cytosine methylation in maize

by comparing the digestion frequency by *Msp*I, *Hap*II and *Mcr*BC. *Msp*I and *Hap*II are isoschisomeres with recognition site 5' -CCGG-3' , which differ in sensitivity to internal cytosine residue' s methylation status. *Msp* I cleave the only outer cytosine methylation at 5' -CCGG-3' sites, whereas *Hap* II is sensitive both of cytosine methylation. To opposite recognition sites of consensus double strands of *Msp*I, *Hap*II, *Mcr*BC is a novel type I nucleotide-dependent restriction enzyme, and specifically cleaves either 5-methylcytosine or 4-methylcytosine of single or double strands DNA (Sutherland *et al.,* 1992). DNA cleavage by *Mcr*BC requires at least two $R^{m}C$ sites that are optimally separated by 55-103 bp, but poor cleavage efficiency as far apart as 2kb, so that the dual recognition sites are $R^{m}C$ ($N_{55\text{-}103}$) $R^{m}C$. Fig. 1 depicts a typical restriction band pattern obtained after the digestion of HMW maize nuclear DNA with three different restriction enzymes. In total nuclear maize DNA, more *Msp*I sites were found than *Hpa*II sites, but rarer than *Mcr*BC sites. Roughly, *Mcr*BC digested 60～70% of total maize HMW DNA into less than 500 bp. To analysis of the repetitive element distribution in smaller than 500 bp fraction after *Mcr*BC digestion, we constructed the library with smaller than 500 bp fraction, which was appeared like main band, and analyzed the sequence about randomly selected 96 clones. The results were appeared that all clones of this fraction were repetitive elements (data not shown).

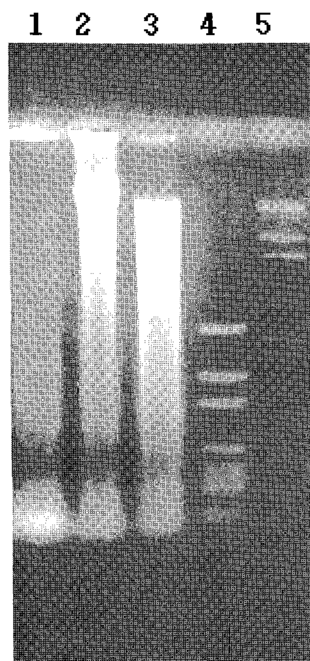

Fig. 1. Genomic band patterns of maize by methylsensitive restriction enzymes. 1: *Mcr*BC, 2: *Hna* Ⅱ, 3: *Mso* Ⅰ, 4, 5: DNA Markers.

Analysis of stability of inserted DNA in shotgun genomic library

To compare the modification of maize DNA in *E.coli,* we transformed two ligation mixtures, *Sau3*A1and *Mcr*BC-*Sau3*A1 restriction fragment, into *E.coli* DH5*α*. The *E.coli* DH5*α* strain has methylated foreign DNA modification genes, *Mcr*A, *Mcr*BC and *Mrr*. The *Mcr*A and *Mcr*BC gene represent *Hpa*II and *Mcr*BC restriction enzyme, respectably. Fig. 2 shows the band patterns of insert size that are randomly selected clones from each fraction libraries. Insert sizes of most clones from *in vivo* methyl-filtered libraries were less than 500 bp. Furthermore, as the insert size of maize genome DNA is bigger, the ratio of generation smaller than 500 bp DNA was increased from 12% to 83%. In contrast, less than 1% clones were reduced size in *in vitro* methyl-filtered libraries. These results seems that *Mcr*BC restriction sites were covered most of cytosine methylation in maize genomic DNA.

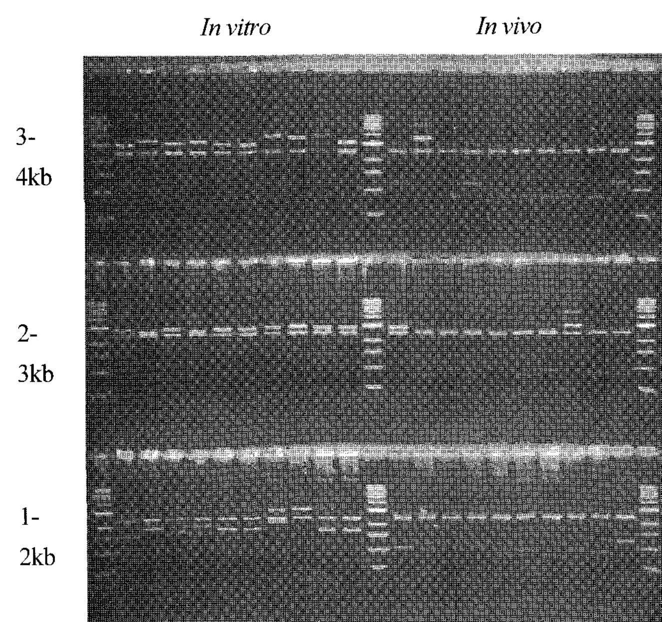

Fig. 2. Comparison of insertion size by different methyl-filtration system. *In vitro* system is that *Sau3*A1and *Mcr*BC-*Sau3*A1 restriction fragment were transformed into *E.coli* DH10B (*mcrA, mcrBC, mrr*). *In vivo* is that *Sau3*A1 restriction fragment were transformed into *E.coli* DH5*α* (*McrA, McrBC, Mrr*).

Repetitive sequence analysis

We sequenced 1,434 clones and have been submitted to GenBank (BH 254961 to BH 256395) by the randomly selected clones from normal and methyl-filtered shotgun maize library. The 307 clones were generated from normal genomic libraries. The 313 and 314 clones were generated from *in vivo* methly-filtration

libraries using *E.coli* strain DH5α and XL1-Blue. The 488 clones from *in vitro* methyl-filtered libraries were sequenced.

The BLAST results and sequence assembly were used to estimate the abundance of different repeats in the maize shotgun libraries. Clones revealed to have overlapped regions or considered to be generated from the same retroelement, were classified into the same retroelement. Comparisons of the number of retroelement presents on individual libraries indicated that repetitive sequences are less enriched in the *in vitro* methyl-filtration library than *in vivo* methyl-filtration libraries. Contents of repetitive sequence showed 44.7% in normal shotgun library, which account for 40.1% of retrotransposon and 4.6% of other repeats, such as MITEs, transposon and tendomly repetitive elements. Klaas and Amasino (1989) reported that methyl-sensitive DNA components of pea, barely and maize contained approximately half concentration of methyl cytosine found in the overall genome. Sequencing results of clone-by-clone from in this study were very similar with these results. The contents of retroelements in the *in vivo* methyl-filtered libraries were 23.6% (DH5α cell) and 29.5% (XL1Blue cell) respectably, whereas 19.8% in *in vitro* methyl-filtered libraries.

Meyers *et al.* (2001) reported Hick, Ji and Opie were the largest propagated retrotransposons in the maize genome and also included Prem-2 (Mascarenhas and Turcich, 1994), Zeon (Fu *et al.,* 2000) and Grande. In compare of each element, even though the most high copy elements were drastically reduced by methyl-filtration, but Prem-2 and Opie families were remained as the most redundancy clone among the retroelements in each library (Table 1). As minor retro-elements, Cinful, Mare, Milt, Stonor, Grandel, Tekai, A188 and Shadowspawn1 were detected. Overall, *in vivo* methyl-filtration library represented 19.5% (DH5α), 22.9% (XL1Blue) of T1/copia type retrotransposons and 1.9%(DH5α), 3.2% (XL1Blue) of Ty3/Gypsy type retrotransposons. In contrast, only 8% of Ty1/copia type retrotransposons were represented in the *in vitro* methyl filtration library, however, 7.6% of Ty3/gypsy type retrotransposons were detected. However, *in vitro* methyl-filtered libraries, the proportion of Huck retro-element (4.2%) was some increased compare with *in vivo* filtered and non-filtered libraries.

The similarity analysis of gene related sequences

To identify the proportion of genic region sequences, we

Table 1. Distribution of retroelements percent in different shotgun libraries

Type	Name	unfiltered	*In vivo* filtered		*In vitro* filtered
		DH10B	DH5α	XL1-Blue	McrBC
Copia	Prem-1	2.9	1.9	0.6	0.2
	Prem-2	15.3	9.9	13.1	2.9
	Opie-1	0.7	0	0.3	0
	Opie-2	9.4	6.7	7.0	4.3
	Opie-3	0.3	0	0	0.2
	Ji	2.3	1	1.6	0.4
	Sto	0.7	0	0.3	0
	Stl-19	0.3	0	0	0
Gypsy	Cinful	2.3	0.3	1.9	0.4
	Huck	0.7	0	0.6	4.5
	MARE	0.3	0	0	0
	Tekay	0.7	0.6	0.3	0.2
	Milt3	0.3	0	0	0
	Grande	1.6	0.3	0.3	1.8
	Romani-Zm1	0.3	0.6	0	0
Other reotransposon		2.0	0.6	1.0	2.9
MITE		1.6	0	0	0
Etc		2.9	1.6	2.5	1.4
Total		44.6%	23.5%	29.5%	19.8%

compared the sequences with GenBank, and additionally compared with the TIGR maize gene index that represent EST database of maize. We classified the sequence composition of libraries based on the probability cutoff value (E value) of at least 9.9 10-5 for Blast X search. For EST search are least 9.9 10-10 E values were used. For the analysis of unknown sequences, we used GENSCAN (Burge *et al.*, 1997) program based on maize sequence data set and 1.0 sub-optimal exon cutoff value. The analysis of protein sequence similarity indicated that all methyl-filtration libraries had more clones containing protein sequences than the non-filtered libraries. *In vivo* methyl-filtered libraries showed 12% (DH5α), 10.5% (XL1Blue) and 8.6% *in vitro* methyl-filtration of the protein sequences composition, compare to 5% of unfiltered library (Fig. 3). However, the proportion of maize EST matched sequence was similar between methyl-filtered and non-filtered libraries. The EST matched sequence composition of *in vitro* methyl-filtration library (16.8%) was slightly higher than unfiltered library (12.7%). For further survey of gene related sequences, we applied no homology sequences to GENSCAN program for identification of potential exon sequences. All methyl-filtration libraries contained more potential exon sequences than 4.9% of unfiltered library. For instance, about 9% showed *in vivo* methyl-filtration, and 12.3% *in vitro* methyl-filtration.

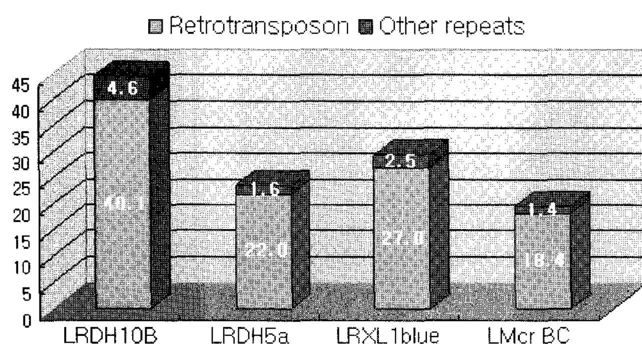

Fig. 3. Comparison of retroelemants in each shot-gun genomic library

The GC content analysis

The GC content means the proposition of total guanine and cytosine base composition against adenine and thymine; it is highly correlated with gene density, repeat element and DNA methylation (Mouchiroud *et al.*, 1991; Duret *et al.*, 1995; Jabbari and Bernardi, 1998). Using bouyant density and Tm, average GC content was already reported 49.5% for the maize genome (Hake and Walbot, 1980). Meyers *et al.* (2001) reported 47% of GC content from maize genomic sequence analysis. To detect the variation of GC content between the different methyl-filtration libraries and non-filtered libraries, we investigated the average of GC composition from sequences, containing homology with ESTs, proteins and no homology sequences. In *in vivo* methyl-filtration library, approximately 10% of GC content was increased compare to non-filtered libraries. The GC content of other XL1Blue *in vivo* methyl-filtration library also slightly increased from 2.8% to 7.6% than unfiltered library. The assay of EST sequence GC content revealed that both *in vivo* and *in vitro* methyl-filtration libraries contained slightly more GC content than the non-filtered library. In contrast, only in root DH5α and XL1Blue methyl-filtration libraries, the ESTs of GC content were decreased compare to non-filtered root library. The GC content of non-filtered libraries, however, showed more than about 50%.

Discussion

The direct approach of whole genome sequencing of maize required the large amount of time and budget. Recently some of studies suggested the efficient methods to characterize the maize genome structure (Rabinowicz *et al.*, 1999; Meyers *et al.*, 2001). Most of these approaches were largely dependent on maize genomic composition. The genic regions of maize intermixed with large portion of hyper ethylated repetitive DNA and removal of hyper ethylated repetitive DNA enriches for transcription active regions.

In this study, we have investigated several of *in vivo* and *in vitro* methyl filtration systems to remove the hyper ethylated repetitive DNA. However, we discovered some restriction for the generation of gene-enriched DNA sequences from *in vivo* methyl-filtration system. It was revealed that variation of insert DNA size is very serious, such as production of the clones that the cloning of unmethylated large DNA sequences more than 2kb was difficult and it caused many of less than 200bp insert DNA clones from *in vivo* methyl-filtration system. Although the method of *in vivo* methyl-filtration system has the effects of removing repetitive DNA sequences, the instability of insert DNA is not suitable for the cloning method of gene-enriched DNA sequences, especially library construction. The phenomenon of unstable insert DNA such as low transformation efficiency and deleted insert DNA was found in *mcr*BC genotype of bacterial strains when methylated plasmid fragments were transformed (Williamsson *et al.*, 1993). Recently, it was reported that the stability of repetitive DNA depends on *dam*

(DNA adenine methyltransferase) *E. coli* system (Troester *et al.*, 2000). According to this paper, however, repetitive DNA deletion mechanism might be related with methylation restriction system not dam methylation system. Because DH10B host strain whose mcr A, BC and mrr methylation system. Because DH10B host strain whose mcr A, BC and mrr methylation restriction systems are eliminated and not mutated dam/dcm, presented intact insert DNA. On the other hands, in the DH5α and XL1 Blue host strains that has mcrA, mcrBC, mrr and dam/dcm genes, revealed many of less than 200bp insert DNA.

We have shown that *in vitro* methyl-filtration system had same effect of enriching protein, EST, and potential exon sequences as *in vivo* methyl-filtration system (Rabinowicz *et al.*, 1999; Meyer *et al.*, 2001). The proportions of genic clones were increased approximately two fold in the methyl-filtered versus the non-filtered libraries. In vitro methyl-filtered library contained about 2～3% of lower percent of protein sequences, but about 4～5% of higher percent of EST and potential exon sequences compare to the *in vivo* methyl-filtered libraries. These results suggest that more than 500bp of Mcr BC restriction enzyme completely digested DNA were unmethylated genic DNA fragments.

The average protein GC content of non-filtered libraries was 44%. While we obtained the average protein GC content of 55% (DH5α), 49.6% (XL1Blue) from *in vivo* filtered leaf and root library, and 52% from *in vitro* filtered library. The GC content of maize genic region was 52% from previous reported (Meyers *et al.*, 2001); those results indicate that methyl-filtration system facilitate cloning of GC rich unmethylated genic region compare to the no methyl-filtration system. Interestingly, the difference of EST GC content between the non-filtered leaf and root libraries was 1.8%, however, the leaf *in vivo* methyl-filtered libraries contained more 7.1% (DH5α) and 9.2% (XL1Blue) EST GC content than the root *in vivo* methyl-filtered libraries. It might be related with differential methylation status from different tissues (Banks and Fedoroff, 1989; Lund *et al.*, 1995; Rossi *et al.*, 1997; Walker, 1998). The root genome might be enriched GC methylated DNA sequences compare to GC unmethylated DNA. We found that repetitive elements have been successfully restricted in the *in vivo* methyl-filtration system, in agreement with previous studies (Rabinowicz *et al.*, 1999; Meyers *et al.*, 2001). The *in vitro* methyl-filtration system also showed the decreased proportion of repetitive elements compare to non-filtered library. *In vivo* methylation restriction system contained about 21% of copia and 2.5% gypsy type retrotransposon family. On the other hand, in the *in vitro* methyl-filtration system, just 8% of copia type retrotransposon, however, 7.6% of gypsy type retrotransposons were facilitated. That different composition pattern seems that derived from *Sau3*A1 partial digested genomic library since non-filtered library were made up 32% of copia and 6.2% of gypsy families of LTR retrotransposons. Moreover, 15.3% of Prem-2 and 10.4% of Opie were the most abundant retrotransposons from non-filtered library. In contrast, Mayers *et al.* (2001) reported that maize genome composed by 19% of copia and 23% of gypsy type retrotransposon, including 10.7% of Huck, 9.4% of Ji and 7.1% of Opie retrotransposons as a most ubiquitous retrotransposons. The different constitution of copia and gypsy type retrotransposons might be caused by the cloning methods that by nebulization and *Sau3*A1 partial digestion.

We suggest that *in vitro* methyl-filtration system is a more efficient method for the library construction of stable, gene-enriched genomic DNA than *in vivo* methyl-filtration system. Also, this cloning system could be useful for supplement mean of genome-wide analysis of large genome organisms with EST cloning system, because methyl-filtration method is cheap and getting more information about genic legion.

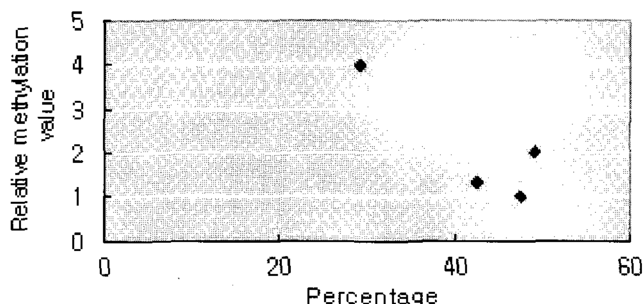

Fig. 4. Relationship between G^{m}C content region and gene related sequence size of DNA.

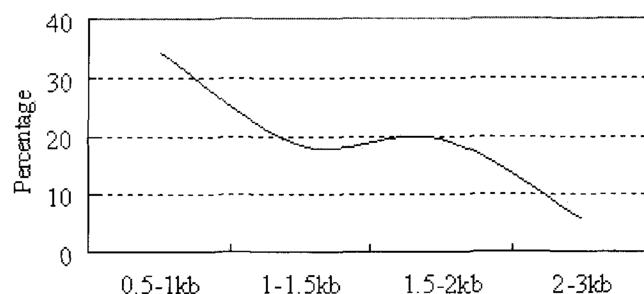

Fig. 5. Estimation of GC methylation frequency in different size of retrotransposon.

Acknowledgements

This study was supported by National Institute Agricultural Science. Project No is 02-2-11-6-1.

Literature Cited

Antequera, F. and A. P. Bird. 1988. Unmethylated CpG islands associated with genes in higher plant DNA. EMBO J. 7: 2295-2299.

Banks, J. A. and N. Fedoroff. 1989. Patterns of developmental and heritable change in methylation of the *suppressor-mutator* transposable element. Dev. Genet 10: 425-437.

Bennetzen, J. L., K. Chrick, P. S. Springer, W. E. Brown and P. SanMiguel. 1994. Active maize genes are unmodified and flanked by diverse classes of modified, highly repetitive DNA. Genome 37: 565-576.

Bennetzen, J. L., K. Schrick, P. S. Springer, W. E. Brown and P. SanMiguel. 1994. Active maize genes are unmodified and flanked by diverse classes of modified, highly repetitive DNA. Genome 37: 565-576.

Burge, C. and S. Karlin. 1997. Prediction of complete gene structures in human genomic DNA. J. Mol. Biol. 268: 78-94.

Carels, N., A. Barakat and G. Bernardi. 1995. The Gene Distribution of the Maize Genome. PNAS 92: 11057-11060.

Constancia, M., B. Pickard, G. Kelsey and W. Reik. 1998. Imprinting Mechanisms. Genome Res. 8: 881-900.

Cost, G. J. and J. D. Boeke. 1998. Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. Biochemistry 37: 18018-93.

De Scenzo, R. A. and R. P. Wise. 1996. Variation in the ratio of physical to genetic distance in intervals adjacent to the Mla locus on barley chromosome 1H. Mol. & Gen. Genet 251: 472-482.

Duret, L., D. Mouchiroud and C. Gautier. 1995. Statistical analysis of vertebrate sequences reveals that long genes are scarce in GC rich isochores. J. Mol. Evol. 40: 308-317.

Fu, H. and H. K. Dooner. 2000. A gene-enbriched BAC library for cloning large allele-specific fragments from maize:isolation of a 240-kb contig of the bronze region. Genome Res. 10: 866-873.

Hake, S. and V. Walbot. 1980. The genome of Zea mays, its organization and homology to related grasses. Chromosoma 79: 251-270.

Jabbri, K. and G. Bernardi. 1998. CpG doublets, CpG islands and Alu repeat elements in long human DNA sequences from different isochores families. Gene 224: 123-128.

Jablonka, E., R. Goiten, M. Marcus and H. Cedar. 1985. DNA hypomethylation causes an increase in Dnase I sensitivity and advanse in the timing of replication of the entire X chromosome. Chromosoma 93: 152-156.

Llaca, V. and J. Messing. 1998. Amplicons of maize zein genes are conserved within genic but expanded and constructed in intergenic region. Plant J. 15: 211-220.

Lund, G., J. Messing and A. Viotti. 1995. Endosperm-specific demethylation and activation of specific alleles of alpha-tubulin genes of *Zea mays L.* Mol. Gen. Genet 246: 716-722.

Mayers, B. C., S. V. Tingey and M. Morgante. 2001. Abundance, distribution, and Transcriptional activaty of repetitive elements in the maize genome. Genome Research 11: 1660-1676

Mouchiroud, D., G. D' onofrio, B. Aissanni, G. Macaya and C. Gautier. 1997. The distribution of genes in the human genome. Gene 100: 181-187.

Pearson, W. R., T. Wood, Z. Zhang and W. Miller. 1997. Comparison of DNA sequences with protein sequences. Genomics 46: 24-36.

Rossi, V., M. Motto and L. Pellegrini. 1997. Analysis of the methylation pattern of the maize opaque-2 (O2) promotor and in vitro binding studies indicate that the O2 B-Zip protein and other endosperm factors can bind to methylated target sequences. J. Biol. Chem. 272: 13758-13765.

Rabinowicz, P. D., K. Schutz, N. Dedhia, C. Yordam, L. D. Parnell, L. Stein, W. R. McCombie and R. A. Martienssen. 1999. Differential methylation of genes and retrotransposons facilitates shotgun sequencing of the maize genome. Nat. Genet 23: 305-308.

Razin, A. 1998. CpG methylation, chromatin structure and gene silencing a three-way connection. EMBO J. 17: 4905-4908.

Razin, A. and H. Ceder. 1991. DNA methylation and gene expression. Microbiol. Rev. 55: 451-458.

SanMiguel, P. and J. L. Bennetzen. 1998. Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. Ann. Bot. 82: 37-44.

SanMiguel, P., A. Tikhonov, Y. K. Jin, N. Motchoulskaia, D. Zakharov, B. A. Melake, P. S. Springer, K. J. Edwards, Z. Avramova and J. L. Bennetzen. 1996. Nested retrotransposons in the intergenic regions of the maize. Science 274: 765-768

SanMiguel, P., B. S. Gaut, A. Tikhonov, Y. Nakajima and J. L.

Bennetzen. 1998. The paleontology og intergene retrotransposons of maize. Nature Genet 20: 43-45.

Springer, P. S. 1992. Genomic organization of Zea mays and its close relatives. PhD. Dissertation (West Lafayette, IN:Purdue University).

Sutherland, E, L. Coe and E. A. Raleigh. 1992. McrBC cleavage on distance between recognition endonuclease. J. Mol. Biol. 225: 327-348.

Tikhonov, A. P., P. J. SanMiguel, Y. Nakajima, N. M. Gorensten, J. L. Bennetzen and Z. Avramova. 1999. Colinearity and its exceptions in orthologous adh regions of maize and sorghum. Pro. Natl. Sci. USA 96: 7409-7414.

Walker, E. L. 1998. Paramutation of the r1 locus of maize is associated with increased cytosine methylation. Genetics 148: 1973-1981.

Williamson, M. D., J. P. Doherty and D. M. Woodcock. 1993. Modified-cytosine restriction-system-induced recombinant cloning artifacts in *Escherichia coli.* Gene. 124: 37-44.

Xiong, Y. and T. H. Eickbush. 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J. 9: 3353-62.

Yang, D., A. Parco, S. Nandi, P. Subudhi, Y. Zhu, G. Wang and N. Huang. 1997. Construction of a bacterial artificial chromosome (BAC) library and identification of overlapping BAC clones with chromosome 4-specific RFLP markers in rice. Theor. Appl. Genet. 95: 1147-1154.

Yu, Y. 2000. Development and application of genomics tools for analysis of grass genome. PhD. Dissertation (Clemson, IN: Clemosn University).

Zhang H.B., X. Zhao, X. Ding, A. H. Paterson and R. A.Wing. 1995. Preparation of megabase-size DNA from plant nuclei. Plant J. 7: 175-184.

(Received 30 June 2007 ; Accepted 20 July 2007)