日本の公的機関でのトウモロコシ (Zea mays L.)
育種のためのゲノムワイドセレクションの予備研究
として行われた RAD-Seq
法により発見された親自殖系統群内のゲノム全体に
わたる分子多型

Genomewide Molecular Polymorphisms among Maize (*Zea mays* L.) Inbred Lines Found from Restriction-Associated DNA Tag Sequencing (RAD-Seq) Analysis as a Preliminary Study on 'Genomewide Selection' for Breeding by Japanese Public Sectors

Hiroyuki TAMAKI, Shohei MITSUHASHI, Hiroshi KUDOH[1], Atsushi J. NAGANO[1,2,a] and Masaki YASUGI[1,b]

Forage Crop Research Division,
NARO Institute of Livestock and Grassland Science, Nasushiobara, 329-2793 Japan

[1] Kyoto University, Otsu, 520-2113 Japan

[2] Japan Science and Technology Agency (JST) PRESTO, Kawaguchi, 332-0012 Japan

Abstract

For accelerating yield improvement in the maize breeding programs operated by the Japanese public sectors, the authors have been interested in introducing genomewide selection (GwS), which requires the arrangement of at least 90 molecular markers in a pair of inbred lines (inbreds) having an identical genetic background. The purpose of this study was to evaluate the potential of 'restriction-associated DNA sequencing' analysis (RAD-Seq) as a genotyping tool in the maize GwS. RAD-Seq is a molecular technique with restriction enzymes and a high-speed sequencing system, whose advantages have been reported on the high reproductivity as well as on the low genotyping cost. Molecular polymorphisms among 34 inbreds from three genetic background groups were surveyed with RAD-Seq, from which 14384 polymorphic loci were found distributed genomewide. The dendrogram drawn based on the polymorphisms well accords not only with those drawn in the previous studies but also with the pedigrees. The results have also shown that 373 to 1106 polymorphic loci exist in each of most pairs of inbreds in an identical genetic background. Therefore the authors have concluded that RAD-Seq can be a powerful genotyping tool in the future maize GwS.

Key words: genomewide selection, inbred line, maize breeding, molecular polymorphism, RAD-Seq

Introduction

There are concerns in Japan over the difficulties in ensuring long-term food supply, for which the Japanese government has adopted a policy to increase food self-sufficiency. A primary plan for this goal is to raise domestic feed crop production from 4.00×10^6 metric tons on the total digestible nutrition (TDN) basis (in 2012) to 5.27×10^6 in 2020[8]. The maize (*Zea mays* L.) breeding teams of Japanese public sectors, including NARO Institute of Livestock and Grassland Science, are now expected to support this policy by releasing high-yield varieties for whole-crop silage use highly adapted to Japanese climates.

As the cost of molecular genotyping has rapidly

declined in recent years (Yan *et al.*[14]), it has become feasible especially for the maize breeding teams in Japanese public sectors to adopt molecular breeding techniques where markers should be arranged over the whole genome[12]. ʻGenomewide selection (GwS)ʼ is one of such techniques; its details and concept are explained in Bernardo and Yu[3] and Meuwissen *et al.*[9]. The focus of GwS is on accumulating favorable genes in many minor quantitative trait loci (QTLs) whereby the yield is deemed controlled[15]. In addition, GwS can be started from a biparental population, i.e. with molecular-marker information on a small number of inbred lines (hereinafter referred to as inbreds). Therefore, introducing GwS to Japanese maize breeding programs is quite valuable to accelerate yield improvement.

The purpose of this study was to evaluate the potential of ʻrestriction- associated DNA sequencingʼ analysis (hereafter referred to as RAD-Seq) as the genotyping tool of the maize GwS. The details of RAD-Seq are described in Baird *et al.*[1]. Briefly, it is a molecular technique with restriction enzymes and such a sequencing system as HiSeq™ of Illumina™, providing sequence information near the restriction sites of multiple samples with high reproductivity[10]. As numerous restriction sites are thought distributed over the whole maize genome, RAD-Seq is expected to detect a remarkable number of genomewide molecular polymorphisms among hundreds of the tested samples at once. Our previous computer-simulation study[12] has revealed that molecular markers should be placed at the intervals of 20cM or shorter for the success of GwS. Based on the reports in the website (http://www.maizegdb.org/complete_map?id=1203637) that the entire linkage-map length is about 1808cM, it means that at least 90 (=1808/20) molecular markers should be arranged for this purpose. The numbers of polymorphic molecular loci should greatly exceed 90 in most cases because the locations of such loci are thought biased over the genome. Therefore the point at which the authors were most interested has been how many reliable molecular polymorphisms were detected from RAD-Seq between a potential pair of inbreds in an identical genetic group whose F_2 progeny would be genotyped. In addition, the advantages of RAD-Seq have also been discussed from its economic aspects; Nagano[10] have reported that it costs only JPY350000 for a RAD-Seq operation for 200 samples, as small as JPY1750 per sample. Therefore the authors have

regarded it as a highly feasible genotyping tool candidate in the maize GwS to boost breeding programs. A number of reports have been made on the genotyping results with RAD-Seq on major crops including maize (e.g. Beissinger *et al.*[2], Davey *et al.*[6]). None of them, however, have provided convincing information on its potential as the genotyping tool for GwS, especially on the following three points, (1) how many polymorphic molecular loci would be detected between a pair of inbreds in an identical genetic group (not among miscellaneous materials), (2) how many of them can be judged potential as the molecular markers for the GwS (our preliminary analyses have revealed that a remarkable number of polymorphisms seem to have come from repeated sequences within an inbred and that they cannot be adopted as markers for GwS), and (3) whether the potential polymorphic loci are distributed genomewide or not. Therefore the authors especially focus the discussion on these three points.

Materials and Methods

Plant materials and DNA preparation

Table 1 shows the list of the genotyped parental inbreds. All but three, ʻMi108ʼ, ʻKi44ʼ and ʻNa106ʼ, were also genotyped in our previous study[13] where single nucleotide polymorphisms (SNPs) were detected with ˮmaizeSNP50 beadchipˮ, a product of Illumina Inc. (San Diego, CA, U.S.) (http://www.illumina.com/products/maizesnp50_dna_analysis_kit.html). All of the inbreds were developed in the Japanese public sectors; National Agriculture and Food Research Organization (NARO) Hokkaido Agricultural Research Center (NARO/HARC), Nagano Animal Industry Experiment Station (NAIES), NARO Kyushu Okinawa Agricultural Research Center (NARO/KARC) and NARO Institute of Livestock and Grassland Science (NARO- ILGS) developed inbreds whose names start with ˮHoˮ, ˮKiˮ, ˮMiˮ and ˮNaˮ, respectively. The inbreds are classified into three groups in terms of their genetic backgrounds (they are hereafter called genetic groups); dent mainly derived from U.S. corn-belt dent (MD), semi-dent mainly developed from hybrids for summer seeding (RD) and flint mainly derived from Japanese landraces (JF). Their seeds were provided from the stocks in NARO-ILGS for breeding experiments. Only one individual per inbred was sampled for DNA preparation,

because our previous study on SNP polymorphisms[13] has revealed that more than 96% of the SNPs are identical

Table 1.　The list of 34 inbreds genotyped in this study with RAD-Seq, and the numbers of the catalogs (loci) provided from the genotyping

Group	Inbreds		Numbers of catalogs (loci)[1]		
			All	Trustable	Potential
MD[2]	Ho102	*	136173	25630	4466
	Ho104	*	113135	21938	3886
	Ho110	*	143188	28062	4724
	Ki66	*	143399	25894	4472
	Ki70	*	140445	28366	4838
	Ki74	*	108734	21455	3775
	Mi29	*	159289	27726	4808
	Mi88	*	114736	24366	4206
	Mi108		142579	29702	5052
	Na65	*	111016	21953	3771
	Na71	*	105777	20369	3425
	Na99	*	176689	30274	5156
	Na102	*	146238	31603	5262
RD[2]	Mi62	*	106205	23268	4108
	Mi71	*	138225	26417	4280
	Mi91	*	103435	20116	3412
	Mi93	*	138247	24798	4327
	Mi106	*	148407	25724	4158
JF[2]	Ho95	*	149352	26553	4535
	Ki44		234521	32703	5190
	Ki68	*	161854	28548	4778
	Ki75	*	151273	27149	4653
	Mi47	*	165855	29852	5047
	Mi103	*	169001	31770	5367
	Mi111	*	177260	35336	5592
	Na50	*	152414	26786	4365
	Na91	*	163036	26255	4136
	Na96	*	199997	30208	4953
	Na101	*	186192	26541	4622
	Na103	*	199997	37129	5707
	Na104	*	271242	37787	5866
	Na105	*[3]	169838	30514	5130
	Na106		131065	29676	4949
	N09-07	*	133226	26261	4300
"Average over the 34 inbreds"			152707	27669	4627
Total (duplicates excluded)			1280484	152658	14384

* The inbreds also genotyped in the previous SNP study by the authors[13].

1) All loci: provided in the genotyping process by Stacks. Trustable loci: depth are seven or more, no plural alleles and identified cordinates. Potential loci: polymorphic trustable loci. (See Materials and Methods for more details).

2) MD, RD and JF are the dent genetic group from U.S. corn-belt, the semi-dent genetic group from hybrids for summer seeding, and the flint genetic group from Japanese landraces, respectively (See Mateirlas and Methods for more details).

3) Genotyped as 'N10-03' in our previous SNP study[13].

between two individuals within an inbred.

　　　　The DNA preparation was made in the following two stages. The first, from cutting young leaves from the seedlings to obtaining the boiled stocks, was completely equivalent to our previous genotyping study[13]. In short, after the leaf of each material was frozen and milled, ʻPrepMan™ Ultra Reagentʼ of Life Technologies Corporation (Carlsbad, CA, U.S.) was added to the milled leaf, and then the mixture in a plastic tube was kept in boiling water for extraction. After cooled down, the tube was centrifuged to obtain the supernatant containing the purified genomic DNA. The second stage, the authors adopted for the first time for DNA purification, followed the protocol of ʻDNeasy Plant Mini Kit™ʼ of Qiagen (Valencia, CA, U.S.) where the 100 μl of the supernatant having been collected in the former stage was subjected as the wet starting material. After the preparation, The DNA concentration was measured with ʻQubit™ 2.0 Fluorometer (with dsDNA HS assay kit)ʼ of Life Technologies, to adjust to 10 ng μl^{-1} for the subsequent sequencing.

Sequencing and genotyping

　　　　In the procedures for sequencing with ʻHiSeq 2000ʼ, a product of Illumina, the authors followed the protocol of Peterson *et al*.[11] with the following several modifications. (1) Genomic DNA of 175 maize materials including the 34 inbreds shown in Table 1 was digested with two restriction endonucleases, *Nde*I and *Bgl*II. (2) Two kinds of Y-shaped adopters (whose sequences are shown in Supplements as TruSeq_*Nde*I_adaptors 1 and 2, and TruSeq_*Bgl*II_adaptors 1 and 2, respectively) were ligated with T4 DNA ligase to the digested ends. Each adopter was arranged to ligate only to the sticky end of each restriction enzyme. (3) Primers (whose sequences are shown in Supplements as TruSeq_Univ_primer 2 and TruSeq_IP001_xxxxxx) were added for the subsequent polymerase chain reaction (PCR). The 5'-end primers also carry index sequences for material identification. (4) PCR was performed with KOD-Plus-Neo (TOYOBO) (94 °C for 2min, 20 cycles of 98°C for 10 sec, 65°C for 30 sec, 68°C for 30 sec), whereby selectively amplified were only the fragments that have both digested ends of the two restriction enzymes. (5) Approximately 270 – 300 base pair (bp) fragments were selected with E-Gel™ size select (Life Technologies) after all of the amplified DNA from each

material was mixed into a stock. And (6) 50 bp sequences of the *Bgl*II-digested side in the DNA fragments were read by HiSeq2000 (Illumina).

　　　　To remove reads containing low-quality bases and adapter sequences, raw sequence reads were preprocessed with ‘Trimmomatic’ ver. 0.32 (Bolger *et al.*[5]), where the followings were the adopted conditions; ILLUMINACLIP: TruSeq3- SE.fa:2:30:10, LEADING: 19, TRAILING: 19, LIDINGWINDOW: 30:20, AVGQUAL: 20, MINLEN: 51. Sequence reads were then processed with Stacks ver. 1.20 (http://creskolab.uoregon.edu/stacks/), with the reference to the entire genomic sequence information of ‘B73’, which is a maize inbred recognized worldwide as the most standard, shown in the website (http://www.ebi.ac.uk/ena/data/view/ GCA_000005005.5). The default setting options were adopted in this process, namely, ‘m’ and ‘n’ values were set to 1 and 0, respectively, in running the program ‘ref_map.pl’. The data having been provided from Stacks were at first sorted in terms of ‘catalogs’ each of which carries information of its standard 50 bp sequence in addition to its coordinate (location in the genome with referring the entire genomic sequence information), i.e. the information of each catalog was connected with that of each inbred on depth (the number of read copies) and allele(s) (hereafter a

‘catalog’ will be called a locus because, in the maize GwS and other molecular breeding, a catalog will be regarded and handled as a molecular locus). Subsequently, loci of each inbred were filtered out if their depths were less than seven (because such small numbers of repetitions were not thought trustable enough in determining the relevant genotypes) or if plural alleles were judged to exist within the inbred (because such plural alleles within an inbred might be the reflection of repeated sequences from plural genome regions). The loci not filtered out were defined in this study as ‘trustable loci’. Then each trustable locus was surveyed on polymorphisms (any kinds of different sequence information including insertion or deletion) over the 34 inbreds, and those having any polymorphism(s) were defined as ‘potential loci’. The following analyses and discussion were made based only on the potential loci.

Dendrogram drawing

　　　　Based on the polymorphisms on the potential loci, the authors drew a dendrogram of the 34 inbreds for the verification of the reliability of the obtained data by comparing with those drawn in the previous studies on simple sequence repeat (SSR) (Enoki *et al.*[7]) and on SNP genotyping[13], and with the pedigree of each inbred. The

Table 2. The numbers of trustable and potential loci (found among the 34 inbreds), and the sizes and lengths of each chromosome

	Chromosome numbers										Total
	1	2	3	4	5	6	7	8	9	10	
Trustable loci[1]	22188 (.145)	16649 (.109)	16673 (.109)	19381 (.127)	16455 (.108)	12869 (.084)	13159 (.086)	12545 (.082)	12022 (.079)	10717 (.070)	152658
Potential loci[2]	1986 (.138)	1611 (.112)	1647 (.115)	1744 (.121)	1504 (.105)	1206 (.084)	1260 (.088)	1208 (.084)	1146 (.080)	1072 (.075)	14384
Size (in Mbp)	301 (.146)	238 (.116)	232 (.113)	242 (.118)	218 (.106)	169 (.082)	177 (.086)	175 (.085)	157 (.076)	150 (.073)	2059
Length (in cM)	286 (.158)	183 (.101)	209 (.116)	189 (.105)	173 (.096)	138 (.076)	158 (.087)	172 (.095)	164 (.091)	136 (.075)	1808

* The value in each parenthesis indicates the ratio of the relevant chromosome toward the total.
　The size and length of each chromosome are quoted from the websites
　http://www.ebi.ac.uk/ena/data/view/GCA_000005005.5 and http://www.maizegdb.org/complete_map?id=1203637, respectively.

[1] A molecular locus is judged trustable if it fulfill the following three conditions in at least one of the 34 inbred lines shown in Table 1, i.e.,
　(1) its depth is seven or more,
　(2) no plural alleles exist within the relevant inbred lines, and
　(3) its cordinates have been identified.

[2] A trustable locus is judged potential if polymorphisms have been found among the inbred lines.

See Materials and Methods for more details of trustable and potential loci.

rules for the drawing were mostly equivalent to those in our previous SNP genotyping study[13], where the calculation of genetic distance (GD) of each pair have been made on the ratio of polymorphic potential loci toward all shared ones, and where the drawing was made in the unweighted pair group method with arithmetic mean (UPGMA). The exceptional difference in the drawing rules between this study and the previous one was that the polymorphism judgement was made only on whether the sequences of the two relevant inbreds were equivalent or not, i.e. the number of polymorphic nucleotide(s) per locus was disregarded.

Results and Discussion

As shown in Table 1, 1.28×10^6 catalogs (substantially equivalent to loci) were provided in the genotyping process by Stacks on the 34 inbreds. About 12% of them, 1.53×10^5, were judged trustable in at least one inbred, and 14384 trustable loci were judged potential, namely had any kinds of polymorphisms among the inbreds. These trustable and potential loci were found distributed genomewide without unevenness (Table 2). Figure 1 is a dendrogram drawn based on the calculated GDs. The GD between two inbreds ranged from 0.049 (=135/2747, between 'Ki66' and 'Ki74') to 0.435 (=1317/3026, between 'Na99' and 'Na103') with the average of 0.344. The dendrogram well accords with the previous studies and the pedigrees. For example, (1) it shows that molecular polymorphisms clearly divide the three genetic groups MD, RD and JF, as reported in Enoki *et al.*[7]. (2) It accords with the pedigrees that 'Ho102' has been developed from a three-way cross of ('Na7' × 'Mi29') × 'Mi29', and that 'Mi111' and 'N09-07' have been developed from the same F_2 line. And, (3) this dendrogram shows what have been found in our previous SNP study[13], i.e., that 'Ki66', 'Ki70' and 'Ki74' have extremely close genetic backgrounds to each other, that 'Ho95' is quite distant from any other inbreds belonging to the same JF group, and that the GDs between 'Na50' and 'Na101' and between 'Mi47' and 'Na103' are relatively close though they have been developed independently. (Hereafter the five pairs, 'Ho102' and 'Mi29', 'Mi111' and 'N09-07', 'Ki66' and 'Ki70', 'Ki66' and 'Ki74', and 'Ki70' and 'Ki74', are excluded from the subsequent discussion on the potential of RAD-Seq for GwS, because GwS for such a pair of inbreds
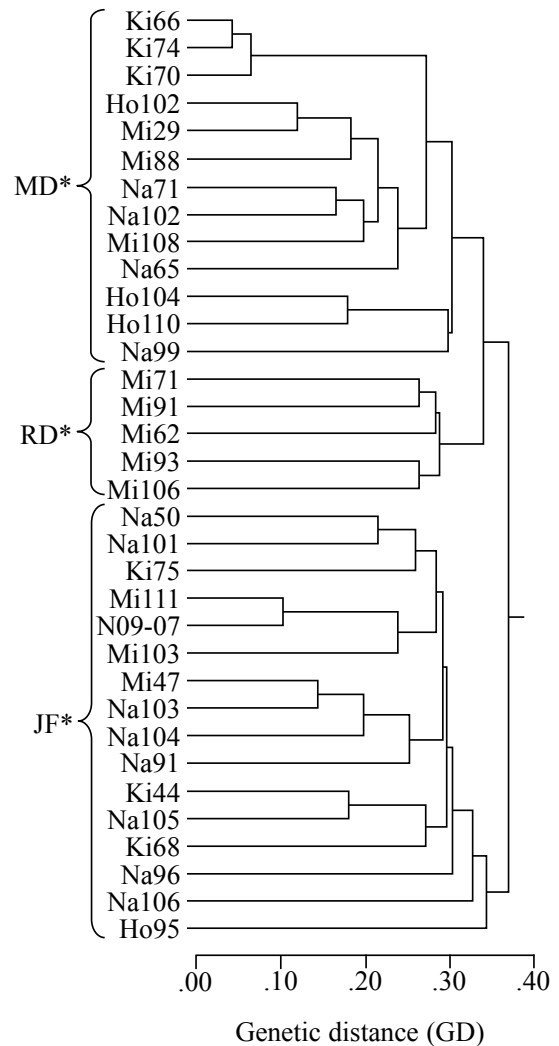

Fig. 1. The unweighted pair group method with arithmetic mean (UPGMA) dendrogram of genetic relationships among 34 inbred lines calculated on the basis of genetic distances of the 4627 trustable polymorphic loci obtained.

* MD, RD and JF are the dent genetic group from U.S. corn-belt, the semi-dent genetic group from hybrids for summer seeding, and the flint genetic group derived from Japanese landraces, respectively (See Materials and Methods for more details).

having extremely close genetic backgrounds to each other is thought impractical.) Table 3 shows the numbers of polymorphic potential loci found from pairs in identical genetic groups. The numbers ranged from 373 (between 'Mi88' and 'Na71') to 1106 (between 'Na103' and 'Na106') unless the pairs have extremely close genetic backgrounds to each other. Figure 2 is a mixed graph to describe how many potential loci are shared by how many inbred lines. It shows the tendency that each of most loci are shared only by a small number of inbreds; more than 75% (10871) of the

Table 3. The number of polymorphic potential loci found in each pair of inbreds belonging to the same genetic groups.

(A) MD*

	Na102	Na99	Na71	Na65	Mi108	Mi88	Mi29	Ki74	Ki70	Ki66	Ho110	Ho104
Ho102	738	763	463	597	754	404	*389*[2]	641	697	706	808	652
Ho104	772	758	536	540	834	642	764	630	749	699	469	
Ho110	962	841	631	695	1004	761	915	732	907	838		
Ki66	765	886	576	656	855	672	770	*135*[1]	*231*[1]			
Ki70	742	878	530	669	868	694	779	*196*[1]				
Ki74	666	755	500	568	749	602	717					
Mi29	674	794	509	525	646	602						
Mi88	658	757	**373**[3]	577	662							
Mi108	654	920	460	530								
Na65	603	732	460									
Na71	408	617										
Na99	894											

(B) RD*

	Mi106	Mi93	Mi91	Mi71
Mi62	601	709	556	640
Mi71	625	668	508	
Mi91	551	644		
Mi93	636			

(C) JF*

	N09-07	Na106	Na105	Na104	Na103	Na101	Na96	Na91	Na50	Mi111	Mi103	Mi47	Ki75	Ki68	Ki44
Ho95	799	940	956	1040	1051	992	935	829	865	978	982	1008	867	903	919
Ki44	766	862	590	933	966	847	968	758	758	919	1001	782	787	737	
Ki68	795	901	832	858	828	840	786	707	690	983	930	781	764		
Ki75	668	888	878	870	779	806	840	714	608	759	920	700			
Mi47	687	1024	953	751	539	990	864	597	842	937	973				
Mi103	697	1038	1025	919	978	960	920	757	768	833					
Mi111	*341*[2]	1031	998	992	1006	942	981	814	630						
Na50	622	810	787	778	803	587	663	665							
Na91	663	788	715	752	625	719	756								
Na96	790	968	971	962	932	950									
Na101	856	746	746	970	1050										
Na103	771	**1106**[3]	1061	751											
Na104	763	1052	1074												
Na105	894	785													
Na106	861														

* MD, RD and JF are the dent genetic group from U.S. corn-belt, the semi-dent genetic group from hybrids for summer seeding, and the flint genetic group from Japanese landraces, respectively (see Materials and Methods for more details).

[1] Because 'Ki66', 'Ki70' and 'Ki74' are judged genetically extremely close to each other, the relevant three pairs (shown in the italic font with underlines) have been excluded from discussion on the potential of GwS (see Materials and Methods for more details).

[2] The pairs of 'Ho102' and 'Mi29' and 'Mi111' and 'N09-07' (shown in the italic font with underlines) have been excluded from discussion on the potential of GwS because of their extremely close pedigrees (see Materials and Methods for more details).

[3] The pairs of 'Na103' and 'Na106' and 'Mi88' and 'Na71' (shown in the bold font with underlines) have been found to have the most and least potential loci among all pairs shown above but the five mentioned above.

14384 potential loci are shared by less than a half (17) of the all 34 inbreds. Table 1 also suggests this tendency, showing that each inbred holds only 3412 to 5866 potential loci.

    The authors have concluded from these results that RAD-Seq can be a powerful genotyping tool in the future maize GwS for the following two reasons. One is that the dendrogram of this study well accords to those of the previous ones (Figure 1), meaning that the genotyping information is trustable enough as a whole. The other is that the RAD-Seq operation of this study provided sufficient numbers not only of trustable molecular loci on each material (3412 to 5866) distributed genomewide (Tables 1 and 2), but also of polymorphic potential loci found in each of most pairs of inbreds in identical genetic groups (more than 373 unless the pair of inbreds have extremely close genetic backgrounds to each other, Table 3). Considering the
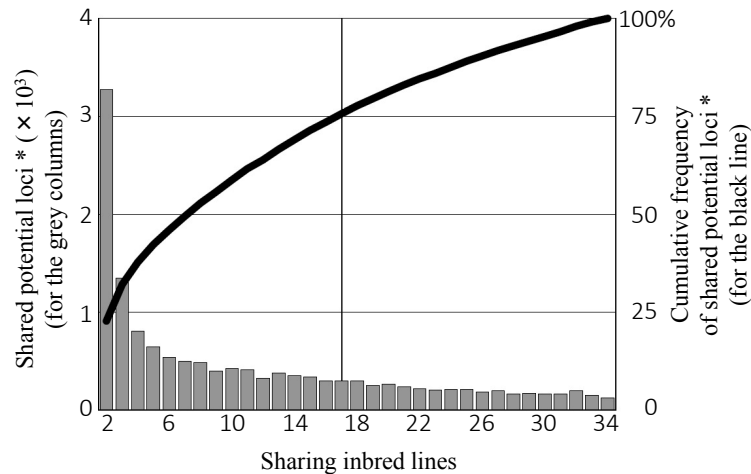
Fig. 2. A mixed graph to describe how many potential loci* are shared by how many inbred
lines (by the grey columns with the left scale), and the cumulative frequency toward
the total 14384 potential loci (by the black line with the right scale).

* A molecular locus is defined potential if it fulfills the following three conditions in at least
  two of the 34 inbred lines shown in Table 1,
  (1) its depth (the number of copies) is seven or more,
  (2) plural alleles do not exist within the relevant inbred line, and
  (3) polymorphism exists between/among the inbred lines.
  See Materials and Methods for more details.

findings from our previous computer-simulation study[12] that the GwS requires at least 90 molecular markers arranged genomewide, RAD-Seq can be judged to provide sufficient numbers of polymorphic potential loci, and to be a highly potential genotyping tool for the maize GwS. The previous studies on RAD-Seq handling miscellaneous materials (e.g. Chattopadhyay *et al*.[4]) have also reported the tendency shown in Figure 2 that each of most loci are shared only by a small number of entries. It will be an annoying problem in genotyping materials having been derived from a wide genetic background, but not in the maize GwS the authors consider to adopt, because the genotyping will be made in this GwS not on such miscellaneous materials but on an F_2 population derived from a pair of inbreds. Therefore RAD-Seq can be regarded as a feasible genotyping tool for the GwS also in terms of this point.

References

1) Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A. and Johnson, E.A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers, PLoS ONE, 3(10), e3376.

2) Beissinger, T.M., Hirsch, C.N., Sekhon, R.S., Foerster, J.M., Johnson, J.M., Muttoni, G., Vaillancourt, B, Buell, C.R., Kaeppler, S.M. and de Leon, N. (2013). Marker density and read-depth for genotyping populations using genotyping-by- sequencing, Genetics, 193, 1073–1081.

3) Bernardo, R. and Yu, J. (2007). Prospects for genomewide selection for quantitative traits in maize, Crop Sci., 47, 1082–1090.

4) Chattopadhyay, B., Garg, K.M. and Ramakrishnan, U. (2014). Effect of diversity and missing data on genetic assignment with RAD-Seq markers, BMC Res. Notes, 7, 841.

5) Bolger, A.M., Lohse, M. and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data, Bioinformatics, btu170.

6) Davey, J.W., Hohenlohe, P.A., Etter, P.D., Boone, J.Q., Catchen, J.M. and Blaxter, M.L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing, Nat. Rev. Genet., 12, 499–510.

7) Enoki, H., Sato, H., and Koinuma, K. (2002). SSR analysis of genetic diversity among maize inbred lines adapted to cold regions of Japan, Theor. Appl. Genet., 104, 1270–1277.

8) MAFF (2014). Fiscal Year 2013 Annual Report on Food, Agriculture and Rural Areas in Japan -Summary, 1-36, Ministry of Agriculture, Forestry and Fisheries of Japan.

9) Meuwissen, T.H.E., Hayes, B.J. and Goddard, M.E. (2001). Prediction of total genetic value using genome-wide dense marker maps, Genetics, 157, 1819–1829.

10) Nagano, A.J. (2013). Genomewide SNP analyses with the cost starting from JPY3000 per sample, Center for ecological research news Kyoto University, 119, 5 (in Japanese).

11) Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S. and Hoekstra, H.E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species, PLoS ONE 7(5), e37135.

12) Tamaki, H., Sato, H., Kikawada, T. and Mitsuhashi, S. (2012) Preliminary study on ˹genomewide selection˼ for maize (*Zea mays* L.) breeding in Japanese public sectors: Estimated selection response in the first selfing cycle of single-cross progeny, Grassl. Sci., 58, 20–27.

13) Tamaki, H., Matsumoto, T., Mitsuhashi, S., Okumura, N., Kikawada, T. and Sato, H. (2014). A study on ˹genomewide selection˼ for maize (*Zea mays* L.) breeding in Japanese public sectors: single nucleotide polymorphisms observed among parental inbred lines, Bull. NARO Inst. Livest. Grassl. Sci., 14, 1–11.

14) Yan, J., Yang, X., Shah, T., Sa´nchez-Villeda, H., Li, J., Warburton, M., Zhou, Y., Crouch, J.H. and Xu, Y. (2010). High-throughput SNP genotyping with the GoldenGate assay in maize, Mol. Breed. 25, 445–451.

15) Yu, J. and Buckler, E.S. (2006). Genetic association mapping and genome organization of maize, Curr. Opin. Biotech., 17, 155–160.

Supplements

Oligo-nucleotide sequences

TruSeq_NdeI_adaptor 1
/5Phos/T*A*GAGATCGGAAGAGCACACGTCTGAACTC CAGTC*A*C

TruSeq_NdeI_adaptor 2
G*T*CAAGTTTCACAGCTCTTCCGATC*T*C

TruSeq_BglII_adaptor 1
A*A*TGATACGGCGACCACCGAGATCTACACTCTTTC CCTACACGACGCTCTT*C*C

TruSeq_BglII_adaptor 2
G*A*TCGGAAGAGCTGTGCAGA*C*T

TruSeq_Univ_primer 2
AATGATACGGCGACCACCGAGATCTACACTCTTTCCC TACACGA

TruSeq_IP001_xxxxxx
CAAGCAGAAGACGGCATACGAGATXXXXXXGTGACTG GAGTTCAGACGTGT

˺ * ˼ = phosphorothioate bond

XXXXXX = 6 mer index sequence

日本の公的機関でのトウモロコシ（*Zea mays* L.）育種のためのゲノムワイドセレクションの予備研究として行われた RAD-Seq 法により発見された親自殖系統群内のゲノム全体にわたる分子多型

玉置宏之・三ツ橋昇平・工藤洋[1]・永野惇[1,2,a]・八杉公基[1,b]

農研機構畜産草地研究所　草地管理研究領域, 那須塩原市, 329-2793
[1] 京都大学, 大津市, 520-2113
[2] 科学技術振興機構さきがけ研究員, 川口市, 332-0012

摘　　要

　著者らは, 日本の公的機関が行っているサイレージ用トウモロコシ育種における収量性改良の効率化のためにゲノムワイドセレクション (GwS) の導入を検討しているが, GwS では同一の遺伝背景を持つ一対の親自殖系統間においてゲノム全体にわたり 90 個以上の分子マーカーを配置する必要がある。本研究の目的は, GwS におけるジェノタイピング手法としての RAD-Seq 法の有用性を評価することであった。制限酵素と次世代シークエンサーを用いてゲノム全体のジェノタイピングを行う RAD-Seq 法の長所は高い再現性と少ない解析費用である。日本の公的機関が育成した 3 群の遺伝背景に由来する 34 親自殖系統の分子多型を RAD-Seq 法により調査した結果, ゲノム全体にわたる 14384 の多型が発見された。これらの多型を基に描かれた樹形図は, 既往の研究および各親自殖系統の育成記録とよく一致した。また同一の遺伝背景を持つ一対の親自殖系統の間に発見された多型の数は, 両者が極端に近縁でない限り 373 個から 1106 個であった。以上のことから著者らは, RAD-Seq 法が将来の GwS におけるジェノタイピング手法として高い潜在性を持っていると結論した。

キーワード： ゲノムワイドセレクション, 自殖系統, トウモロコシ育種, 分子多型, RAD-Seq 法

[a] 現　龍谷大学, 大津市, 520-2194
[b] 現　基礎生物学研究所, 岡崎市, 444-8787