Rapeseed (Brassica napus L.; genome AACC, 2n = 38) is one of the most economically important oilseed crops worldwide. It is thought to have originated from multiple independent natural hybridization events between B. rapa (AA, 2n = 18) and B. oleracea (CC, 2n = 18) (U 1935). The oldest description of rapeseed cultivation was found in the Indian literature of 2000 to 1500 BCE (Röbbelen et al. 1989). Nowadays, over 450 million tonnes of seeds are harvested worldwide every year, which accounts for about 20% of world grain production (Carré and Pouzet 2014). In 2012, the EU led the worldwide rapeseed production (~20 Mt), followed by Canada (15 Mt), China (12 Mt), and India (6 Mt) (ISTA Mielke GmbH 2012). In Japan, since the Meiji period (~150 years ago), many crops, including rapeseed, were imported and cultivated. Modern rapeseed breeding in Japan began in the 1930s; selective breeding of overseas rapeseed cultivars and crossing between rapeseed and B. rapa were implemented to breed cold-resistant and early maturing cultivars suitable for the Japanese climate and cultivation methods (Matsuo 1954, Shiga 1970, Yamamori 2006). During the second half of the 20th century, cross-breeding between Japanese and new European cultivars was performed to obtain single-zero (low erucic acid content) and double-zero (single-zero plus low glucosinolate content) cultivars, and many such cultivars were released (Ishida 2003).

Genetic resources are key to progress in breeding, contributing to sustainable agriculture and associated industries. More than 600 accessions, including >300 Japanese landraces and cultivars, are maintained at the Genetic Resources Center of the National Agriculture and Food Research Organization (NARO Genebank). A collection of this size is expected to contain a wide range of genetic variation and offer opportunities for trait improvement. However, screening for target traits in the whole collection is often time-consuming, laborious, and costly. To efficiently use germplasm collections in breeding, knowledge of the genetic diversity and population structure among the germplasms is vital. In addition, establishing a representative subset of the whole collection, called a core collection, will be a feasible way for management, evaluation, and utilization of genetic resources.

A core collection should represent most of the diversity

Diversity analysis of rapeseed accessions preserved in the Japanese Genebank can provide valuable information for breeding programs. In this study, 582 accessions were genotyped with 30 SSR markers covering all 19 rapeseed chromosomes. These markers amplified 311 alleles (10.37 alleles per marker; range, 3–39). The genetic diversity of Japanese accessions was lower than that of overseas accessions. Analysis of molecular variance indicated significant genetic differentiation between Japanese and overseas accessions. Small but significant differences were found among geographical groups in Japan, and genetic differentiation tended to increase with geographical distance. STRUCTURE analysis indicated the presence of two main genetic clusters in the NARO rapeseed collection. With the membership probabilities threshold, 227 accessions mostly originating from overseas were assigned to one subgroup, and 276 accessions mostly originating from Japan were assigned to the other subgroup. The remaining 79 accessions are assigned to admixed group. The core collection constructed comprises 96 accessions of diverse origin. It represents the whole collection well and thus it may be useful for rapeseed genetic research and breeding programs. The core collection improves the efficiency of management, evaluation, and utilization of genetic resources.

Key Words: Brassica napus L., Genebank, landrace, oilseed, SSR marker.
in the whole collection and should cover at least 70% of the alleles present in the whole collection (Brown 1989). Phenotypic traits have been used for constructing core collections, but they are easily affected by environmental conditions and are sometimes difficult to measure accurately. Recently, many core collections have been established by using DNA marker data, because DNA markers can be easily obtained and are highly stable and polymorphic (Tanguchi et al. 2014, Xu et al. 2016, Zhao et al. 2016). In rapeseed, several studies revealed genetic diversity of different collections using molecular markers such as RFLP (Song and Osborn 1992), RAPD (Ma et al. 2000), SRAP (Riaz et al. 2001), AFLP (Seyis et al. 2003), and SSR (Hasan et al. 2006). SSR markers are superior because they are able to detect multiple alleles per locus, are codominant at a single locus, and are relatively evenly distributed across genomes (Zalapa et al. 2012); they have been widely used in recent diversity analyses of Brassica species (El-Esawi et al. 2016, Guo et al. 2016, Yao et al. 2012).

Over the last five decades, effort devoted in Japan to rapeseed breeding has resulted in considerable progress. Despite a large number of germplasms in NARO Genebank, only a few studies analyzed the small number of either Japanese or overseas accessions in this collection (Diers and Osborn 1994, Gyawali et al. 2013, Hasan et al. 2006, Ma et al. 2000). The number of rapeseed germplasms in the NARO Genebank has increased owing to both past studies and exploration of genetic resources. The aims of this study were to examine the genetic diversity and population structure of the NARO rapeseed collection by using single-locus SSR markers, and to establish a core collection with good representation of the genetic diversity present in the whole collection.

**Materials and Methods**

**Plant materials**

The rapeseed collection in NARO Genebank hold 639 germplasms, the seeds are preserved in cool dry storage. We excluded accessions from our analysis if both parental lines were in the collection. The total number of accessions analyzed was 582 (Supplemental Table 1), including 305 Japanese accessions (landraces, breeding lines, and cultivars) and 277 overseas accessions from all inhabited continents except Africa.

**Genotyping**

Seeds were germinated on moistened filter paper in 9-cm Petri dishes and then transplanted into a 72-cell tray filled with granular culture soil (Nippi-Engei-Baido-igou soil, Nihon Hiryo Co., Ltd., Tokyo, Japan). Seedlings were grown in a climate chamber (25/22°C 12 h/12 h) until the first true leaves had fully expanded. Genomic DNA was extracted from the first true leaf of each seedling by using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer’s protocol with some minor modifications. A total of 502 Brassica SSR markers (Supplemental Table 2; AAFC Consortium 2016, Cheng et al. 2009, Hatakeyama et al. 2010, Iniguez-Luy et al. 2009, Kim et al. 2009, Li et al. 2011, 2013, Lowe et al. 2002, Piquemal et al. 2005, Suwabe et al. 2002, Wang et al. 2012, Xu et al. 2010) were prescreened against 8 representative accessions that were chosen on the basis of their origins and breeding histories: OOMINATANE (Japan), MICHINOKUNATANE (Japan), ASAHINATANE (Japan), CASCADE (USA), HAMBURG 1 (Germany), WESTAR (Canada), PROTA (Germany) and RAPORA (Korea). SSR markers with clear reproducible polymorphic amplification products were then applied to all other accessions. PCR mixtures (10 μl) contained template DNA (10 ng), 1× KAPA 2G buffer A, 200 nM dNTP, 0.5 mM MgCl2, 0.1 U KAPA 2G Fast DNA polymerase (KAPA Biosystems Inc., Woburn, MA, USA), 2 pmol reverse primer, and 0.5 pmol forward primer. The forward primers were 5′-labeled with the fluorescent dyes 6-FAM, VIC, NED, or PET (Shimizu and Yano 2011). PCR was performed in a C1000 Thermal Cycler (Bio-Rad Laboratories, Inc., Hercules, CA, USA) as follows: initial denaturation at 94°C for 3 min; 30 cycles of 94°C for 20 s, 54°C for 30 s, and 62°C for 30 s; 3 cycles of 94°C for 20 s, 49°C for 10 s, and 72°C for 5 s; and a final extension at 72°C for 10 min. The size of the amplified fragments was estimated by using an automated DNA analyzer (model 3130xl) with a GeneScan-600LIZ size standard and GeneMarker v. 4.0 software (all Thermo Fisher Scientific, Inc., Waltham, MA, USA).

**Genetic diversity and population structure analysis**

The number of alleles, major allele frequency, number of rare alleles (frequency <5%), and polymorphism information content (PIC) were calculated for the whole collection and for each geographic group in PowerMarker v. 3.25 software (Liu and Muse 2005). Observed heterozygosity (Ho), expected heterozygosity (He), Shannon’s information index (I), pairwise F statistics (FST), and Nei’s genetic distance (Nei et al. 1983) were calculated in GenAlEx v. 6.502 software (Peakall and Smouse 2012). The components of variance between Japanese and overseas accessions, between different groups and among individuals within groups were estimated from the genetic distance matrix, as specified in the analysis of molecular variance (AMOVA) procedure using Arlequin 3.5.2.2 software (Excoffier and Lischer 2010). A nonparametric permutation procedure with 1000 permutations was used to test the significance of variance components associated with the different possible levels of genetic structure in this study (Excoffier et al. 1992).

Genotyping data for the SSR markers were analyzed by using the model-based STRUCTURE v. 2.3.4 software (Pritchard et al. 2000) to determine the most probable number of clusters (K value) and to assign rapeseed accessions to different clusters. The K value was determined by running an admixture and related frequency model with K = 1 to 20 (10 replications per K value); the burn-in period of
Analysis of genetic diversity of rapeseed genetic resources in Japan

Breeding Science Vol. 67 No. 3 BS

Results

Genetic diversity of the NARO rapeseed collection

Prescreening selected 30 of the 502 SSR markers. The selected markers amplified a total of 311 alleles in the 582 accessions, ranging from 3 (BoGMS0660 and BoEMS0049) to 39 (BrGMS0070) per marker, with a mean allelic richness of 10.37 (Table 1); 214 alleles (68.8%) with low frequency (<0.05) were regarded as rare alleles. In addition, the selected 30 SSR markers amplified only one or two alleles in each accession. Average genetic diversity indices were 0.56 for major allele frequency, 0.57 for He, and 0.52 for PIC. The average Ho value was 0.05 (range, 0.01–0.34), indicating that almost all accessions in the whole collection were highly homozygous.

The genetic diversity indices for geographical groups are summarized in Table 2. The average number of alleles per maker was 8.50 in Japanese germplasms and 9.37 in overseas germplasms. Among Japanese geographical groups, the average number of alleles per marker ranged from 2.90 (Chugoku-Shikoku area) to 5.43 (Chubu area), whereas overseas geographical groups had a wider variation, from 3.13 (Oceania) to 8.07 (Europe). The values of all other genetic diversity indices (Ho, He, I, PIC) were lower in Japanese than in overseas geographical groups. Hierarchical analysis of molecular variance revealed significant differences at all hierarchical levels (Table 3); the percentage of variation was highest among individuals within geographical groups (67.18%), followed by that between Japanese and overseas accessions (17.71%).

There was significant genetic differentiation between geographic groups in Japan, except between Kanto and Tohoku and between Kanto and Chubu. Pairwise FST ranged from 0.017 (Kanto and Chubu) to 0.192 (Hokkaido and Chugoku-Shikoku) (Table 4). The FST values between Hokkaido and western Japanese geographic groups (Kinki, Chugoku-Shikoku, and Kyushu) showed relatively high differentiation. Nei’s genetic distance among Japanese geographic groups ranged from 0.044 (Kinki and Kyushu) to 0.173 (Hokkaido and Chugoku-Shikoku). Similar to pairwise FST, the genetic distances tended to be higher between the Hokkaido group and western Japanese geographic groups.

Population structure of the NARO rapeseed collection

The STRUCTURE analysis suggested K = 2 as the real value of K even though a secondary peak of ΔK exists at K = 4 (Supplemental Fig. 1), indicating there are two genetic clusters in the NARO rapeseed collection (Fig. 1). With the membership probabilities (Q) threshold of 0.80, 227 accessions, mostly originating from overseas, were assigned to subgroup 1, and 276 accessions, mostly originating from Japan were assigned to subgroup 2 (Table 5). The remaining 79 accessions are assigned to admixed group, i.e., Q < 0.8. The result of PCoA was similar to that of STRUCTURE analysis (Fig. 2). The first and second factors of the PCoA explained approximately 19.1% and 7.3% of the variation in the genetic distance matrix, respectively. A bi-dimensional PCoA scatter plot indicated differentiation between Japanese and overseas accessions (Fig. 2).

Construction of a core collection

Four different core collections were developed by different methods. Sample size and genetic diversity indices of each core collection are shown in Table 6. The random sampling strategy and CoreHunter selected the largest number of accessions (116; 19.93% of the whole collection),
whereas the CoreFinder method selected the smallest number (96; 16.49%). The core collections constructed by the two M strategy methods (PowerCore and CoreFinder) had the highest allele retention ratio (100%), indicating that they had all alleles observed in the whole collection, whereas approximately 30% of alleles were lost in the two other core collections. Other indices (except Ho) were higher in the core collections constructed by the two M strategy methods

| Marker name | Linkage group | Number of alleles | Number of rare alleles | Major allele frequency | Ho  | He  | PIC |
|-------------|--------------|------------------|------------------------|------------------------|-----|-----|-----|
| BrGMS4028a  | A1           | 9                | 5                      | 0.46                   | 0.34| 0.71| 0.67|
| BrGMS4031a  | A1           | 5                | 3                      | 0.86                   | 0.04| 0.24| 0.22|
| BRAS084b    | A1           | 22               | 17                     | 0.30                   | 0.10| 0.83| 0.81|
| BrGMS1411a  | A2           | 7                | 4                      | 0.58                   | 0.05| 0.59| 0.55|
| BrGMS0667a  | A2           | 15               | 8                      | 0.51                   | 0.06| 0.71| 0.69|
| BrGMS2498a  | A3           | 4                | 2                      | 0.52                   | 0.04| 0.52| 0.40|
| n2025a      | A4           | 10               | 5                      | 0.44                   | 0.04| 0.70| 0.65|
| BrGMS2252a  | A5           | 4                | 2                      | 0.82                   | 0.03| 0.32| 0.29|
| BrGMS0070a  | A5           | 39               | 36                     | 0.28                   | 0.06| 0.87| 0.86|
| BnEMS0753d  | A6           | 7                | 4                      | 0.63                   | 0.02| 0.52| 0.46|
| BrGMS3750a  | A6           | 6                | 2                      | 0.60                   | 0.02| 0.57| 0.53|
| BrGMS3872a  | A7           | 8                | 6                      | 0.60                   | 0.01| 0.51| 0.41|
| BnEMS0620d  | A7           | 17               | 14                     | 0.54                   | 0.08| 0.64| 0.60|
| BrGMS0742a  | A8           | 11               | 8                      | 0.61                   | 0.03| 0.55| 0.50|
| BrGMS0281a  | A9           | 8                | 4                      | 0.43                   | 0.06| 0.71| 0.66|
| BrGMS0557a  | A10          | 5                | 2                      | 0.61                   | 0.06| 0.55| 0.49|
| BrGMS3688a  | A10          | 6                | 2                      | 0.41                   | 0.06| 0.71| 0.66|
| BrGMS0086a  | A10          | 12               | 7                      | 0.41                   | 0.06| 0.76| 0.73|
| BnEMS2271c  | C1           | 9                | 6                      | 0.48                   | 0.02| 0.59| 0.50|
| BoGMS2016f  | C2           | 15               | 9                      | 0.30                   | 0.03| 0.84| 0.82|
| BoEMS0016e  | C2           | 9                | 6                      | 0.50                   | 0.05| 0.65| 0.60|
| BoGMS0660d  | C2           | 3                | 1                      | 0.61                   | 0.01| 0.48| 0.37|
| BrGMS0289a  | C3           | 21               | 18                     | 0.44                   | 0.03| 0.62| 0.55|
| BrGMS347f   | C4           | 10               | 8                      | 0.44                   | 0.04| 0.62| 0.55|
| BoGMS0037f  | C5           | 9                | 7                      | 0.85                   | 0.01| 0.26| 0.23|
| BoGMS1909f  | C6           | 13               | 10                     | 0.78                   | 0.03| 0.37| 0.36|
| BrGMS0353s  | C6           | 8                | 5                      | 0.54                   | 0.04| 0.62| 0.57|
| BoEMS0049g  | C7           | 3                | 1                      | 0.53                   | 0.01| 0.50| 0.38|
| BnEMS0336e  | C8           | 4                | 1                      | 0.81                   | 0.17| 0.33| 0.31|
| BoGMS0525g  | C9           | 12               | 11                     | 0.89                   | 0.07| 0.21| 0.21|
| Average     |              | 10.37            | 7.13                   | 0.56                   | 0.05| 0.57| 0.52|

Ho, observed heterozygosity; He, expected heterozygosity; PIC, polymorphism information content.

Detailed marker information is available in a Xu et al. (2010), b Piquemal et al. (2005), c AAFC Consortium (2016), d Wang et al. (2012), e Cheng et al. (2009), f Li et al. (2011), g Li et al. (2013).

Table 2. Genetic diversity indices for the NARO rapeseed collection and variations among geographic groups

| Geographical group | Number of accessions | Average number of alleles | Major allele frequency | Ho  | He  | I     | PIC   |
|--------------------|----------------------|--------------------------|------------------------|-----|-----|-------|-------|
| Whole Collection   | 582                  | 10.37                    | 0.56                   | 0.05| 0.57| 1.18  | 0.52  |
| Japan              | 305                  | 8.50                     | 0.67                   | 0.04| 0.44| 0.90  | 0.41  |
| Hokkaido           | 9                    | 3.30                     | 0.58                   | 0.03| 0.52| 0.91  | 0.47  |
| Tohoku             | 69                   | 5.27                     | 0.65                   | 0.06| 0.46| 0.88  | 0.44  |
| Kanto              | 33                   | 4.77                     | 0.68                   | 0.06| 0.42| 0.82  | 0.41  |
| Chubu              | 40                   | 5.43                     | 0.65                   | 0.04| 0.46| 0.90  | 0.44  |
| Kinki              | 72                   | 5.17                     | 0.73                   | 0.03| 0.37| 0.73  | 0.37  |
| Chugoku-Shikoku    | 11                   | 2.90                     | 0.70                   | 0.07| 0.37| 0.65  | 0.36  |
| Kyushu             | 65                   | 4.63                     | 0.72                   | 0.03| 0.37| 0.71  | 0.36  |
| Unknown            | 6                    | –                        | –                      | –   | –   | –     | –     |
| Overseas           | 277                  | 9.37                     | 0.50                   | 0.06| 0.59| 1.22  | 0.58  |
| Europe             | 202                  | 8.07                     | 0.53                   | 0.06| 0.56| 1.14  | 0.54  |
| Asia               | 30                   | 5.37                     | 0.60                   | 0.07| 0.49| 0.97  | 0.49  |
| Oceania            | 8                    | 3.13                     | 0.61                   | 0.05| 0.48| 0.84  | 0.45  |
| America            | 27                   | 4.40                     | 0.63                   | 0.05| 0.47| 0.88  | 0.45  |
| Unknown            | 10                   | –                        | –                      | –   | –   | –     | –     |

Ho, observed heterozygosity; He, expected heterozygosity; I, Shannon’s information index; PIC, polymorphism information content.
Analysis of genetic diversity of rapeseed genetic resources in Japan

Breeding Science
Vol. 67 No. 3 BS

Table 3. Hierarchical analysis of molecular variance in the NARO rapeseed collection

| Source of variation                  | d.f. | Percentage of variation (%) |
|--------------------------------------|------|----------------------------|
| Between Japanese and overseas accessions | 1    | 17.71 ***                  |
| Among geographical groups           | 9    | 7.01 ***                   |
| Among individuals                    | 555  | 67.18 ***                  |
| Within individuals                   | 566  | 8.10 **                    |

***,**: significant at the 0.1% and 1% levels, respectively, for 1000 permutations.

Table 4. Pairwise $F_{ST}$ (below diagonal) and Nei’s genetic distance (above diagonal) among geographic groups in Japan

|                  | Hokkaido | Tohoku | Kanto | Chubu | Kinki | Chugoku-Shikoku | Kyushu |
|------------------|----------|--------|-------|-------|------|-----------------|--------|
| Hokkaido         |          | 0.105  | 0.125 | 0.110 | 0.148| 0.173           | 0.143  |
| Tohoku           | 0.059*   |        | 0.055 | 0.056 | 0.063| 0.091           | 0.068  |
| Kanto            | 0.102**  | 0.019  |       | 0.053 | 0.048| 0.099           | 0.058  |
| Chubu            | 0.122**  | 0.031**| 0.017 |       | 0.056| 0.091           | 0.049  |
| Kinki            | 0.101**  | 0.039**| 0.047**|      | 0.058**|                | 0.114  |
| Chugoku-Shikoku  | 0.192**  | 0.084**| 0.067**| 0.076*| 0.088**|            0.055**| 0.075  |
| Kyushu           | 0.138**  | 0.034**| 0.063**| 0.092**|      |                |        |

**: significant at the 1% level; *: significant at the 5% level.

Fig. 1. Population structure of 582 rapeseed accessions determined (at K = 2) by using 30 single-locus SSR markers.

Discussion

We evaluated the genetic diversity of the NARO rapeseed collection by using single-locus SSR markers. In allopolyploid plants such as rapeseed, SSR primers usually amplify several alleles from multiple loci, which makes it difficult to assign these alleles to individual loci. Li et al. (2013) found a set of 230 single-locus rapeseed SSR markers. However, they used only 6 inbred lines and noted that not all of the markers might be true single-locus markers. In this study, we applied a total of 502 SSR markers, including the 230 markers of Li et al. (2013), to 582 rapeseed accessions, and found that 30 SSR markers (Table 1), distributed across all linkage groups, always amplified single alleles with high polymorphism. This finding indicates that these 30 SSR markers used in this study are useful for research on rapeseed genetic resources.

We identified 311 alleles in the whole collection, with an average of 10.37 alleles per locus and a PIC value of 0.52. Several studies calculated genetic diversity indices of diverse rapeseed collections on the basis of SSR markers. The average numbers of alleles per locus were 4.31 in 96 European rapeseed genotypes (Hasan et al. 2006), 3.4 in 192 inbred lines of various origins (Xiao et al. 2012), and 7.3 in 169 worldwide rapeseed lines (Gyawali et al. 2013). The PIC value in the 192 lines analyzed by Xiao et al. (2012) was 0.37. The differences in allelic richness between the NARO collection and other rapeseed collections are due partly to the differences in the SSR markers and the number of accessions used in this study. Nevertheless, the higher values of the genetic diversity indices in this study indicate
that the NARO rapeseed collection has moderately high genetic diversity. In contrast, the PIC value of the NARO rapeseed collection was lower than those reported in other *Brassica* species: 0.57 or 0.64 in *B. oleracea* (El-Esawi et al. 2016, Louarn et al. 2007) and 0.90 in *B. juncea* (Yao et al. 2012); this would reflect a narrower genetic base in rape-seed than in other *Brassica* species.

Rapeseed originated and was first domesticated in Europe, and then spread throughout the world. In Japan, rape-seed cultivation is considered to have spread after the Meiji period (from 1868 to 1912) (Matsuo 1954); since then, many landraces have been developed throughout the country. Modern rapeseed breeding in Japan began in the 1930s; government-led large-scale breeding projects initially used a limited number of accessions (Matsuo 1954), which may have restricted the genetic diversity of Japanese accessions. As expected, we found a significant genetic differentiation between Japanese and overseas accessions (Fig. 2, Tables 3, 4), and comparison of the genetic indices indicated that genetic diversity was lower in Japanese accessions than in overseas accessions (Table 2). However, the values of genetic diversity indices were not very different between

### Table 5. Geographic origin of rapeseed accessions assigned by STRUCTURE to two subgroups

| Geographical group | Number of accessions | Subgroup 1 | Subgroup 2 | Admixed |
|--------------------|----------------------|------------|------------|---------|
| Whole Collection   | 582                  | 227        | 276        | 79      |
| Japan              | 305                  | 19         | 240        | 46      |
| Hokkaido           | 9                    | 4          | 3          | 2       |
| Tohoku             | 69                   | 5          | 46         | 18      |
| Kanto              | 33                   | 2          | 27         | 4       |
| Chubu              | 40                   | 4          | 27         | 9       |
| Kinki              | 72                   | 0          | 66         | 6       |
| Chugoku-Shikoku    | 11                   | 1          | 9          | 1       |
| Kyushu             | 65                   | 1          | 60         | 4       |
| Unknown            | 6                    | 2          | 2          | 2       |
| Overseas           | 277                  | 208        | 36         | 33      |
| Europe             | 202                  | 171        | 9          | 22      |
| Asia               | 30                   | 3          | 22         | 5       |
| Oceania            | 8                    | 5          | 0          | 3       |
| America            | 27                   | 25         | 1          | 1       |
| Unknown            | 10                   | 4          | 4          | 2       |

* Accessions are considered belonging to either of two subgroup when membership probabilities of ≥0.8.

### Table 6. Genetic diversity indices of four core collections constructed by different methods

| Method              | Number of accessions | Number of alleles | Allele retention ratio | Number of effective alleles | Ho     | He     | I       | PIC    |
|---------------------|----------------------|-------------------|------------------------|----------------------------|--------|--------|---------|--------|
| PowerCore           | 103                  | 10.37             | 100.0                  | 3.77                       | 0.05   | 0.64   | 1.41    | 0.59   |
| CoreFinder          | 96                   | 10.37             | 100.0                  | 3.57                       | 0.06   | 0.62   | 1.39    | 0.58   |
| CoreHunter          | 116                  | 7.50              | 72.3                   | 3.17                       | 0.06   | 0.60   | 1.23    | 0.55   |
| Random sampling     | 116                  | 7.20*             | 69.5                   | 2.89                       | 0.07   | 0.57   | 1.17    | 0.52   |

* Significant at the 5% level between core collection and whole collection.

Ho, observed heterozygosity; He, expected heterozygosity; I, Shannon’s information index; PIC, polymorphism information content.

Maximization strategy methods.
Japanese and overseas accessions. Over the last five decades, repeated introduction of overseas germplasms and subsequent breeding may have increased genetic diversity.

In general, differences in breeding materials and targets among different breeders and regions may lead to genetic differentiation. In Japan, we found small but significant differences among geographical groups (Table 4). Genetic differentiation tended to increase with increasing geographic distance. For example, we found the highest $F_{ST}$ values and Nei's genetic distances between the Hokkaido group (northern Japan) and three groups in western Japan. The genetic differentiation might be caused not only by the diversity of germplasms used as breeding materials, but also by their adaptation to local environments. The genetic diversity indices PIC and I were higher in eastern than in western Japan. The genetic diversity of eastern Japanese accessions may have been further increased by recent active rapeseed breeding there.

Our STRUCTURE analysis indicated the presence of two main genetic clusters in the NARO rapeseed collection and demonstrated differentiation between Japanese and overseas accessions: with a membership probability threshold of 0.80, most of the overseas accessions were assigned to the subgroup 1, and the Japanese accessions were assigned to the subgroup 2 (Table 5). Yet the presence of several accessions in each subgroup is not consistent with their geographical origins (Table 5), as similarly observed in previous studies (Chen et al. 2008, Hu et al. 2007, Li et al. 2012, Xiao et al. 2012). All of the Canadian cultivars, known by the trade name ‘Canola’ and bred in the 1970s or later, were classified into the subgroup 1. Many accessions from Asian countries other than Japan belonged to the subgroup 2. The principal coordinate analysis (PCoA) further confirmed the STRUCTURE results (Fig. 2). These results indicate that Asian rapeseed may have genetically differentiated from that in other regions of the world, and that Japanese, Chinese, and Korean landraces and cultivars may be closely genetically related (Fig. 1). Some studies have revealed a wide diversity of Chinese and Australian rapeseed accessions (Chen et al. 2008, Guo et al. 2016, Wang et al. 2009). The NARO collection holds only 7 accessions from China and none from Australia. Obtaining accessions from these countries would further expand the diversity of the collection.

The most important factor in core collection construction is sample selection strategy, and many strategies based on stratified sampling and clustering methods are available (Zhang et al. 2011). Because different strategies result in different core collections (Thachuk et al. 2009, Wang et al. 2007), we used four different methods. As the M strategy is based on maximizing the number of alleles, it can automatically generate a sampling ratio on the basis of the genetic diversity of the species; this strategy has been widely used in recent years (Belaj et al. 2012, Liu et al. 2015, Zhang et al. 2011). In this study, both PowerCore and CoreFinder, which are based on the M strategy, constructed core collections that retained 100% of alleles, and in this respect were superior to the other two methods used. The genetic diversity indices tended to be nearly same in both the PowerCore collection and the CoreFinder collection (Table 6). The number of accessions are fewer in the CoreFinder collection (Table 6), so we finally recommend it as the core collection of the NARO rapeseed collection. The core collection comprised 57 overseas accessions (20.6% of total) and 39 Japanese accessions (12.8%). The higher percentage of overseas accessions reflects their high allelic richness. However, the fact that approximately 40% of accessions in the core collection originated in Japan indicates their relatively high genetic diversity.

In conclusion, our study revealed the genetic structure of the NARO rapeseed collection and genetic relationships among accessions on the basis of single-locus SSR markers. Some accessions were genetically identical or closely related to each other. This information is important for decreasing redundancy in the collection, thereby reducing the management cost and avoiding unnecessary distribution of such accessions to breeders. Further integration of the data from other collections maintained in different countries will make it possible to exploit and preserve the whole rapeseed gene pool and to retain the largest number of allelic variants for genes controlling the most important agronomic traits in the NARO collection. Our candidate core collection can be used in further research such as genome-wide association studies to identify genomic regions controlling important agronomic traits. We did not analyze the phenotypic data recorded in the NARO rapeseed collection because they have been obtained by different investigators and methods under various environmental conditions. Therefore, it will be necessary to evaluate the phenotypes of all the accessions under the same environmental conditions, and integrate the data on genotypic and phenotypic diversity. This will make the core collection, which comprises accessions that are genetically diverse at both genotypic and phenotypic levels, available to breeders for enhancing the genetic potential of this crop.

**Acknowledgements**

We thank NARO Genebank for providing rapeseed seeds used in this study. This work was partly supported by a grant from the Ministry of Agriculture, Forestry and Fisheries of Japan (Genomics-based Technology for Agricultural Improvement, GRA-102).

**Literature Cited**

AAFC Consortium (2016) http://brassica.agr.gc.ca/index_e.shtml
Belaj, A., M. del Carmen Dominguez-Garcia, S.G. Atienza, N.M. Urdiroz, R. De la Rosa, Z. Satovic, A. Martin, A. Kilian, J. Trujillo, V. Valpuesta et al. (2012) Developing a core collection of olive (*Olea europaea* L.) based on molecular markers (DaRts, SSRs, SNPs) and agronomic traits. Tree Genet. Genomes 8: 365–378.
Brown, A.H.D. (1989) Core collections: a practical approach to genetic
resources management. Genome 31: 818–824.
Carré, P. and A. Pouzet (2014) Rapeseed market, worldwide and in Europe. OCIL 21: D102.
Chen, S., M.N. Nelson, K. Ghamkhar, T. Fu and W.A. Cowling (2008) Divergent patterns of allelic diversity from similar origins: the case of oilseed rape (Brassica napus L.) in China and Australia. Genome 51: 1–10.
Cheng, X., J. Xu, S. Xia, J. Gu, Y. Yang, J. Fu, X.J. Qian, S.C. Zhang, J.S. Wu and K. Liu (2009) Development and genetic mapping of microsatellite markers from genome survey sequences in Brassica napus. Theor. Appl. Genet. 118: 1121–1131.
Cipriani, G., A. Spadotto, I. Jurman, G. Di Gaspero, M. Crespìan, S. Meneghetti, E. Frare, R. Vignani, M. Cresti, M. Morgante et al. (2010) The SSR-based molecular profile of 1005 grapevine (Vitis vinifera L.) accessions uncovers new synonymy and parentages, and reveals a large admixture amongst varieties of different geographic origin. Theor. Appl. Genet. 121: 1569–1585.
De Beukelaer, H., P. Smýkal, G.F. Davenport and V. Fack (2012) Core Hunter II: fast core subset selection based on multiple genetic diversity measures using Mixed Replica search. BMC Bioinformatics 13: 312.
Diers, B.W. and T.C. Osborn (1994) Genetic diversity of oilseed Brassica napus germplasm based on restriction fragment length polymorphisms. Theor. Appl. Genet. 88: 662–668.
Earl, D.A. (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4: 359–361.
El-Esawi, M.A., K. Germaine, P. Bourke and R. Malone (2016) Genetic diversity and population structure of Brassica oleracea germplasm in Ireland using SSR markers. C. R. Biol. 339: 133–140.
Evanno, G., S. Regnaut and J. Goudet (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14: 2611–2620.
Excoffier, L., P.E. Smouse and I.M. Quattro (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479–491.
Excoffier, L. and H.E. Lischer (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10: 564–567.
Guo, S., X. Zhang, Y. Niu, L. Wang and Y. Liu (2016) Genetic diversity of resynthesized Brassica napus lines from SW China assessed by main agronomic traits and SSR markers in comparison with common B. napus lines. Euphytica 207: 95–108.
Gyawali, S., D.D. Hegedus, A.P. Isobel, J. Poon, E. Higgins, K. Horner, D. Bekkaoui, C. Coutu and L. Buchwaldt (2013) Genetic diversity and population structure in a world collection of Brassica napus L. gene pool using SSR markers. Genet. Resour. Crop Evol. 60: 1717–1725.
Hasan, M., F. Seyis, A.G. Badani, J. Pons-Kühnemann, W. Friedt, W. Lühs and R.J. Snowdon (2006) Analysis of genetic diversity in the Brassica napus L. gene pool using SSR markers. Genet. Resour. Crop Evol. 53: 793–802.
Hatakeyama, K., A. Horisaki, S. Niikura, Y. Narusaka, H. Abe, H. Yoshiaki, M. Ishida, H. Fukuo and S. Matsumoto (2010) Mapping of quantitative trait loci for high level of self-incompatibility in Brassica rapa L. Genome 53: 257–265.
Hu, J., J. Zhu and H.M. Xu (2000) Methods of constructing core collections by stepwise clustering with three sampling strategies based on the genotypic values of crops. Theor. Appl. Genet. 101: 264–268.
Hu, S., C. Yu, H. Zhao, G. Sun, S. Zhao, M. Vyvadilova and V. Kucera (2007) Genetic diversity of Brassica napus L. germplasm from China and Europe assessed by some agronomically important characters. Euphytica 154: 9–16.
Iniguez-Luy, F.L., L. Lukens, M.W. Farnham, R.M. Amasino and T.C. Osborn (2009) Development of public immortal mapping populations, molecular markers and linkage maps for rapid cycling Brassica rapa and B. oleracea. Theor. Appl. Genet. 120: 31–43.
Ishida, M. (2003) Rapeseed production, utilization and breeding program in Japan. J. Agric. Sci. 58: 205–210.
ISTA Mielke GmbH (2012) Oil production. ISTA Mielke GmbH, Hamburg, p. 15.
Kim, H., S.R. Choi, J. Bae, C.P. Hong, S.Y. Lee, M.J. Hossain, D.V. Nguyen, M. Jin, B-S. Park, J-W. Bang et al. (2009) Sequenced BAC anchored reference genetic map that reconciles the ten individual chromosomes of Brassica rapa. BMC genomics 10: 432.
Kim, K.W., H.K. Chung, G.T. Cho, K.H. Ma, D. Chandrabalan, J.G. Gwag, T.S. Kim, E.G. Cho and Y.J. Park (2007) PowerCore: a program applying the advanced M strategy with a heuristic search for establishing core sets. Bioinformatics 23: 2155–2162.
Li, H., X. Chen, Y. Yang, J. Xu, J. Gu, J. Fu, X.J. Qian, S.C. Zhang, J.S. Wu and K.D. Liu (2011) Development and genetic mapping of microsatellite markers from whole genome shotgun sequences in Brassica oleracea. Mol. Breed. 28: 585–596.
Li, H., M. Younas, X. Wang, X. Li, L. Chen, B. Zhao, X. Chen, J.S. Xu, F. Hou, B.H. Hong et al. (2013) Development of a core set of single-locus SSR markers for allotetraploid rapeseed (Brassica napus L.). Theor. Appl. Genet. 126: 937–947.
Li, W., W. Jiang, M. Vyvadilova and M. Stamm (2012) Genetic diversity of rapeseed accessions from different geographic locations revealed by expressed sequence tag-simple sequence repeat and random amplified polymorphic DNA markers. Crop Sci. 52: 201–210.
Liu, J., Z.R. Wang, C. Li, Y.B. Bian and Y. Xiao (2015) Evaluating genetic diversity and constructing core collections of Chinese Lentinula edodes cultivars using ISSR and SRAP markers. J. Basic Microbiol. 55: 749–760.
Liu, K. and S.V. Muse (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21: 2128–2129.
Louarn, S., A.M. Torp, I.B. Holme, S.B. Andersen and B.D. Jensen (2007) Database derived microsatellite markers (SSRs) for cultivar differentiation in Brassica oleracea. Genet. Resour. Crop Evol. 54: 1717–1725.
Lowe, A.J., A.E. Jones, A.F. Raybould, M. Trick, C.L. Moule and K.J. Edwards (2002) Transferability and genome specificity of a new set of microsatellite primers among Brassica species of the U triangle. Mol. Ecol. Notes 2: 7–11.
Ma, C., Y. Kimura, H. Fujimoto, T. Sakai, J. Imamura and T. Fu (2000) Genetic diversity of Chinese and Japanese rapeseed (Brassica napus L.) varieties detected by RAPD markers. Breed. Sci. 50: 257–265.
Matsuo, T. (1954) Detailed exposition of breeding. Yokendo, Tokyo, pp. 319–323.
Nei, M., F. Tajima and Y. Tateno (1983) Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. J. Mol. Evol. 19: 153–170.
Peakall, R. and P.E. Smouse (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28: 2537–2539.
Piquemal, J., E. Cinquin, F. Couton, C. Rondeau, E. Seignoret, I. Doucet, D. Perret, M.-J. Villeger, P. Vincourt and P. Blanchard (2005) Construction of an oilseed rape (Brassica napus L.) genetic map with SSR markers. Theor. Appl. Genet. 111: 1514–1523.

Pritchard, J.K., M. Stephens and P. Donnelly (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Riaz, A., G. Li, Z. Quresh, M.S. Swati and C.F. Quiros (2001) Genetic diversity of oilseed Brassica napus inbred lines based on sequence-related amplified polymorphism and its relation to hybrid performance. Plant Breed. 120: 411–415.

Röbbelen, S., R.K. Downey and A. Ashri (1989) Oilcrops of the world: their breeding and utilization. McGraw-Hill Publishing Company, New York, p. 10.

Seys, F., R.J. Snowdon, W. Luhs and W. Friedt (2003) Molecular characterization of novel resynthesized rapeseed (Brassica napus) lines and analysis of their genetic diversity in comparison with spring rapeseed cultivars. Plant Breed. 122: 473–478.

Shiga, T. (1970) Rape breeding by interspecific crossing between Brassica napus and Brassica campestris in Japan. Jpn. Agric. Res. Quart. 5: 5–10.

Shimizu, T. and K. Yano (2011) A post-labeling method for multiplexed and multicolored genotyping analysis of SSR, indel and SNP markers in single tube with bar-coded split tag (BStag). BMC Res. Notes 4: 161.

Song, K. and T.C. Osborn (1992) Polyphyletic origins of Brassica napus: new evidence based on organelle and nuclear RFLP analyses. Genome 35: 992–1001.

Suwabe, K., H. Iketani, T. Nunome, T. Kage and M. Hirai (2002) Isolation and characterization of microsatellites in Brassica rapa L. Theor. Appl. Genet. 104: 1092–1098.

Taniguchi, F., K. Kimura, T. Saba, A. Ogino, S. Yamaguchi and J. Tanaka (2014) Worldwide core collections of tea (Camellia sinensis) based on SSR markers. Tree Genet. Genomes 10: 1555–1565.

Thachuk, C., J. Crossa, J. Franco, S. Dreisigacker, M. Warburton and G.F. Davenport (2009) Core Hunter: an algorithm for sampling genetic resources based on multiple genetic measures. BMC Bioinformatics 10: 245.

U, N. (1935) Genome analysis in Brassica with special reference to the experimental formation of B. napus and peculiar mode of fertilization. Japan. J. Bot. 7: 389–452.

Wang, F., X. Wang, X. Chen, Y. Xiao, H. Li, S. Zhang, J. Xu, J. Fu, L. Huang, C. Liu et al. (2012) Abundance, marker development and genetic mapping of microsatellites from unigenes in Brassica napus. Mol. Breeding 30: 731–744.

Wang, J., S. Kaur, N.O.I. Cogan, M.P. Dobrowolski, P.A. Salisbury, W.A. Burton, R. Baillie, M. Hand, C. Hopkins, J.W. Forster et al. (2009) Assessment of genetic diversity in Australian canola (Brassica napus L.) cultivars using SSR markers. Crop Pasture Sci. 60: 1193–1201.

Wang, J.C., J. Hu, H.M. Xu and S. Zhang (2007) A strategy on constructing core collections by least distance stepwise sampling. Theor. Appl. Genet. 115: 1–8.

Xiao, Y., D. Cai, W. Yang, W. Ye, M. Younas, J. Wu and K. Liu (2012) Genetic structure and linkage disequilibrium pattern of a rapeseed (Brassica napus L.) association mapping panel revealed by microsatellites. Theor. Appl. Genet. 125: 437–447.

Xu, J., X. Qian, X. Wang, R. Li, X. Cheng, Y. Yang, J. Fu, S. Zhang, G.J. King, J. Wu et al. (2010) Construction of an integrated genetic linkage map for the A genome of Brassica napus using SSR markers derived from sequenced BACs in B. rapa. BMC Genomics 11: 594.

Xu, Y., C. Chen, D. Ji, K. Xu, X. Xie and C. Xie (2016) Developing a core collection of Pyropia haitanensis using simple sequence repeat markers. Aquaculture 452: 351–356.

Yamamori, M. (2006) Present status of rapeseed breeding with emphasis on fatty acid composition. Oleoscience 6: 189–194.

Yao, Q.L., F.B. Chen, P. Fang, G.F. Zhou, Y.H. Fan and Z.R. Zhang (2012) Genetic diversity of Chinese vegetable mustard (Brassica juncea Coss) landraces based on SSR data. Biochem. Syst. Ecol. 45: 41–48.

Zalapa, J.E., H. Cuevas, H. Zhu, S. Steffan, D. Senalik, E. Zeldin, B. McCown, R. Harbut and P. Simon (2012) Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. Am. J. Bot. 99: 193–208.

Zhao, J., Y. Tong, T. Ge and J. Ge (2016) Genetic diversity estimation and core collection construction of Sinojackia huangmeiensis based on novel microsatellite markers. Biochem. Syst. Ecol. 64: 74–80.