SSR Marker Development and Genetic Identification of Pitaya (Hylocereus spp.) Collected in Okinawa Prefecture, Japan

Kenji Nashima¹*, Fumiko Hosaka², Yohei Shimajiri³, Masato Matsumura³, Kazuhiko Tarora⁴, Naoya Urasaki⁴, Moriyuki Shoda⁴, Chikako Nishitani², Yutaka Sawamura² and Toshiya Yamamoto²

¹College of Bioresource Sciences, Nihon University, Fujisawa 252-0880, Japan
²Institute of Fruit Tree and Tea Science, National Agriculture and Food Research Organization, Tsukuba 305-0852, Japan
³Okinawa Prefectural Agricultural Research Center Nago Branch, Nago 905-0012, Japan
⁴Okinawa Prefectural Agricultural Research Center, Itoman 901-0336, Japan

Pitaya (Hylocereus spp.) is a tropical fruit grown in Okinawa prefecture, Japan. In this study, we developed simple sequence repeat (SSR) markers to assess the genetic diversity of pitaya accessions collected in Okinawa prefecture. SSR markers were developed with next-generation sequencing, using the Roche 454 GS FLX+ platform. Seventy primer sets were tested for the genetic identification of pitaya, including 50 AG motif and 20 AC motif SSRs. In total, 16 SSR markers successfully amplified fragments and exhibited polymorphism among the accessions. The SSR markers revealed the genetic diversity parameters including number of alleles per locus (ranging from 6 to 12), expected heterozygosity (ranging from 0.5 to 0.85), observed heterozygosity (ranging from 0.23 to 0.93), polymorphic information content (ranging from 0.47 to 0.83), fixation index value (0.013), and out-crossing rate (0.973). Using these SSR markers, genetic identification was performed for 32 pitaya accessions collected in Okinawa prefecture. These accessions could be distinguished genetically using three SSR markers, with the exception of accessions between ‘Uruma White’/’Nago White’ and ‘Ohgimi Red 1’/‘Ohgimi Red 2’. Genetic diversity analyses suggested six groups, and grouping appeared to be based on differences in species. The obtained SSR markers and clarified genetic background information may facilitate future breeding of pitaya.

Key Words: DNA marker, dragon fruit, genetic diversity.

Introduction

Pitaya (Hylocereus spp.), also called pitahaya or dragon fruit, belongs to Hylocereus, a member of the Cactaceae family. Pitaya has high potential as an ornamental and fruit crop, as well as an industrial source of compounds (Ortiz-Hernández and Carrillo-Salazar, 2012). The main production areas of pitaya are South America, South East Asia, and Israel. The origin of pitaya is thought to be South America (Ortiz-Hernández and Carrillo-Salazar, 2012).

In Japan, 167.7 tons of pitaya fruits were produced in 22.6 ha of production area in 2015. Pitaya is mainly produced in Okinawa prefecture (85% of total production). Production and pitaya growers are increasing in Okinawa prefecture; however, a large variation in fruit quality has emerged as an issue. A possible reason for quality variation is the difference in pitaya genotype grown by each commercial grower. Most commercial growers produce indigenous pitaya, whose introduction routes are unknown. However, it was not confirmed whether such indigenous pitayas were genetically different or identical. If they are genetically different, a possible way to resolve this problem is the breeding of new elite pitaya cultivars and replacement from unknown indigenous pitaya.

A pitaya breeding program was established at the Okinawa Prefectural Agricultural Research Center (OPARC). For this breeding program OPARC collected indigenous pitayas and commercially bred seedlings across Okinawa prefecture. However, issues in the breeding process of pitaya at the OPARC emerged; col-
lected accessions had not been evaluated for genetic diversity and identity and the species were not identified. As a result, the pitaya breeding pedigrees were questionable for accuracy. Although cross breeding was performed for pitaya, it may be possible that unexpected crosses including self-pollination had occurred. Clarifying these genetic backgrounds, including genetic diversity, genetic identity, and pedigree, would enable evaluation of the breeding system at OPARC and facilitate the establishment of cross breeding strategies.

DNA marker analyses are effective to assess such genetic backgrounds. To date, several types of DNA markers have been utilized to assess the genetic diversity of pitaya, including random-amplified polymorphic DNA (RAPD) (Tel-Zur et al., 2004), inter simple sequence repeat (ISSRs) (Tao et al., 2014), amplified fragment length polymorphism (AFLP) (Pagliaccia et al., 2015), and simple sequence repeat (SSRs) (Pan et al., 2017). Among these DNA markers, SSR markers have several advantages due to high levels of polymorphism, multiple alleles, and co-dominance (Ellegren, 2004). As highly polymorphic SSR markers are necessary to assess genetic background efficiently, the development of new SSR markers with high polymorphisms is required. In recent years, next generation sequencing (NGS) analyses provided large-scale sequence information, and NGS has facilitated efficient SSR identification development (Zalapa et al., 2012).

To obtain polymorphic SSR markers, it is effective to use SSRs with larger repeat numbers because these frequently contain repeated numbers of mutations. Extracting SSRs with larger repeat numbers among the NGS sequence should be effective for developing SSR markers with many polymorphisms. In addition, such DNA markers should be useful for efficiently assessing genetic identification, parentage analyses, and genetic diversity analyses.

In this study, we developed SSR markers and selected reliable SSRs, which produced no multiple, competitive or null amplification, to be used for genetic identification, pedigree confirmation by parentage analysis, and genetic diversity analysis of pitaya collected or bred in Okinawa prefecture, Japan.

Materials and Methods

Plant material and DNA extraction

For SSR marker genotyping, 32 Hylocereus accessions, shown in Table 1, were used, including one registered cultivar, 12 breeding lines bred by OPARC, nine commercial breeding lines, seven indigenous accessions collected in Okinawa prefecture, and three of unknown origin. The indigenous accessions were collected from commercial growers and renamed according to collected city and flesh color. All accessions were maintained at the OPARC Nago branch (OPARC-Nago). Pedigree, fruit skin color, and fruit flesh color, are shown in Table 1.

For DNA extraction, tender stems of each accession were sampled. DNA extraction was performed using a DNeasy Plant Mini Kit (Qiagen, Hilden, Germany).

SSR marker development

Genomic DNA collected from pitaya accession ‘No. 5’, which was predicted as a hybrid of H. undatus, was sheared by nebulization (600–900 bp in length) and amplified by emulsion polymerase chain reaction (PCR). Two runs of single read pyrosequencing were performed using a Genome Sequencer 454 GS FLX+ (Roche Diagnostics, Basel, Switzerland). The obtained sequence data were deposited in the DDBJ Sequence Read Archive (DRA) with accession nos. DRR213735 and DRR213736. Sequence reads were trimmed to remove adapter sequences, low-quality regions, and short-read sequences using CLC Genomics Workbench 7.0 (Qiagen) with the following parameters: quality trimming limit: 0.01, maximum number of ambiguities: 2, and discard reads below length: 200.

For identification of di- to hexanucleotide SSRs, the MicroSAtellite (MISA) identification tool <http://pgrc.ipkgsatersleben.de/misa/> was used for trimmed sequences. Dinucleotide motifs with at least six repeats and tri- to hexanucleotide motifs with at least five repeats were considered as SSRs. SSR markers were designed from sequences including the SSRs of 10 to 20 repeats of GA/CT or CA/GT motifs and no other SSRs within 100 b. Sequences including these SSRs were applied to Batchprimer3 (You et al., 2008). Primer picking conditions included a product size range: 150–250 bp, a primer Tm: 50–70°C (optimum 55°C), and a primer size: 18–27 b (optimum 23 b). Based on sequence similarities among the PCR amplified region of each SSR primer, redundant SSR markers were detected by BLAST (Basic Local Realignment Tools; Altschul et al., 1990) with an E-value cut off of 1e-20. For SSR analysis, 50 markers of the GA/CT motif and 20 of the CA/GT motif were selected from the longest repeat and used for SSR amplification.

Genotyping on the 32 accessions was performed for 70 SSR markers. SSR-PCR amplification was performed in a 10-μL reaction mixture containing 5 μL of GoTaq Master Mix including GoTaq DNA Polymerase (Promega, Madison, WI, USA), 1 pmol FAM or HEX-labeled forward primer, 1 pmol of unraveled reverse primer, and 5 ng of template DNA. For reverse primers, the “gtttctt” nucleotide sequence was added to the 5’ end of reverse primers as pig-tailing (Brownstein et al., 1996) to enhance adenylation and facilitate accurate genotyping. DNA was amplified in 35 cycles at 94°C for 1 min, 55°C for 1 min, and 72°C for 2 min, and a final extension of 5 min at 72°C. The amplified PCR products were separated and detected in a PRISM 3130xl DNA sequencer (Applied Biosystems, Waltham, MA, USA). The sizes of the amplified fragments were scored against internal-standard DNA (400HD-ROX;
Table 1. Pitaya accessions used in this study.

| No. | Accession name     | Type            | Origin            | Pedigree                          | Flesh color | Skin color | Predicted ploidy | Predicted species            | Pedigree suggested by SSR marker                                             |
|-----|--------------------|-----------------|-------------------|-----------------------------------|-------------|------------|------------------|--------------------------------|--------------------------------------------------------------------------------|
| 1   | Chiyuramiyarabi    | Cultivar        | Private breeding  | Sibling of Ohzato White           | White       | Red        | Diploid          | hybrid of *H. undatus*          |                                                                              |
| 2   | Impact Ruby        | Breeding line   | OPARC             | Ohgimi White × Spineless Red      | Pink        | Red        | Diploid          | hybrid of *Hylocereus* species | Ohgimi White × Spineless Red                                                  |
| 3   | Nago 1             | Breeding line   | OPARC             | Polycross progeny of No. 5        | White       | Red        | Diploid          | hybrid of *H. undatus*          | Offspring of No. 5, paternal parent unknown (discrepancy at TsuHU035)       |
| 4   | Nago 2             | Breeding line   | OPARC             | Polycross progeny of No. 5        | Pink        | Red        | Diploid          | hybrid of *H. undatus*          | Offspring of No. 5, paternal parent unknown                                  |
| 5   | No. 5              | Breeding line   | OPARC             | Unknown                          | Red         | Red        | Diploid          | hybrid of *H. undatus*          | Candidate parent: Nago Red, Ohgimi Red 1, Ohgimi White 2                     |
| 6   | No. 22             | Breeding line   | OPARC             | Unknown                          | White       | Red        | Diploid          | hybrid of *H. undatus*          | Nago White × Ohgimi White 2                                                  |
| 7   | 1-08               | Breeding line   | OPARC             | Polycross progeny of No. 5        | Purple      | Red        | Diploid          | hybrid of *H. undatus*          | Offspring of No. 5, paternal parent unknown                                 |
| 8   | 1-21               | Breeding line   | OPARC             | Polycross progeny of No. 5        | White       | Pink       | Diploid          | hybrid of *H. undatus*          | Offspring of No. 5, paternal parent unknown                                 |
| 9   | 2-27               | Breeding line   | OPARC             | Nago Red × No. 22                | Purple      | Red        | Diploid          | hybrid of *H. undatus*          | Nago Red × No. 22                                                           |
| 10  | 2-29               | Breeding line   | OPARC             | Nago Red × No. 22                | Red         | Red        | Diploid          | hybrid of *H. undatus*          | Nago Red × No. 22                                                           |
| 11  | 5-19               | Breeding line   | OPARC             | Polycross progeny of No. 5        | Purple      | Pink       | Diploid          | hybrid of *H. undatus*          | Offspring of No. 5, paternal parent unknown                                 |
| 12  | 5-21               | Breeding line   | OPARC             | Polycross progeny of No. 5        | Purple      | Red        | Diploid          | hybrid of *H. undatus*          | No. 5 × Ohgimi Red 1                                                        |
| 13  | 6-15               | Breeding line   | OPARC             | Ohgimi White × Spineless Red      | Red         | Pink       | Diploid          | hybrid of *Hylocereus* species  | Ohgimi White × Spineless Red (discrepancy at TsuHU049)                        |
| 14  | Desert King Seedling | Breeding line | Private breeding  | Offspring of Desert King          | Purple      | Orange     | Triploid         | hybrid of *H. megaranthus*      |                                                                              |
| 15  | Golden Seedling 1  | Breeding line   | Private breeding  | Offspring of yellow skin pitaya   | Red         | Red        | Diploid          | Unknown *Hylocereus* species    |                                                                              |
| 16  | Golden Seedling 2  | Breeding line   | Private breeding  | Sibling of Golden Seedling 1      | Dark red    | Red        | Diploid          | Unknown *Hylocereus* species    |                                                                              |
| 17  | Golden Seedling 3  | Breeding line   | Private breeding  | Sibling of Golden Seedling 1      | Dark purple | Red        | Diploid          | Unknown *Hylocereus* species    |                                                                              |
| 18  | Golden Seedling 4  | Breeding line   | Private breeding  | Sibling of Golden Seedling 1      | Dark red    | Red        | Diploid          | Unknown *Hylocereus* species    |                                                                              |
| 19  | Golden Seedling 5  | Breeding line   | Private breeding  | Sibling of Golden Seedling 1      | Dark purple | Red        | Diploid          | Unknown *Hylocereus* species    |                                                                              |
| 20  | Golden Seedling 6  | Breeding line   | Private breeding  | Sibling of Golden Seedling 1      | Dark purple | Red        | Diploid          | Unknown *Hylocereus* species    |                                                                              |
| 21  | Golden Seedling 7  | Breeding line   | Private breeding  | Sibling of Golden Seedling 1      | Dark purple | Red        | Diploid          | Unknown *Hylocereus* species    |                                                                              |
| 22  | Ohzato White       | Breeding line   | Private breeding  | Sibling of Chiyuramiyarabi        | White       | Red        | Diploid          | hybrid of *H. undatus*          |                                                                              |
| 23  | Nago Red           | Indigenous      | Nago city         | Unknown                          | Red         | Red        | Diploid          | hybrid of *H. undatus*          | Identical to Uruma White                                                    |
| 24  | Nago White         | Indigenous      | Nago city         | Unknown                          | White       | Red        | Diploid          | hybrid of *H. undatus*          | Identical to Ohgimi White                                                   |
| 25  | Ohgimi Red 1       | Indigenous      | Ohgimi village    | Unknown                          | Red         | Pink       | Diploid          | hybrid of *H. undatus*          | Identical to Ohgimi Red 2                                                   |
| 26  | Ohgimi Red 2       | Indigenous      | Ohgimi village    | Unknown                          | Red         | Pink       | Diploid          | hybrid of *H. undatus*          | Identical to Ohgimi Red 1                                                   |
| 27  | Ohgimi White 1     | Indigenous      | Ohgimi village    | Unknown                          | Light pink  | Pink       | Diploid          | hybrid of *H. undatus*          |                                                                              |
| 28  | Ohgimi White 2     | Indigenous      | Ohgimi village    | Unknown                          | White       | Pink       | Diploid          | hybrid of *H. undatus*          |                                                                              |
| 29  | Uruma White        | Indigenous      | Uruma city        | Unknown                          | White       | Red        | Diploid          | hybrid of *H. undatus*          | Identical to Nago White                                                     |
| 30  | Spineless Red      | Unknown         | Unknown           | Unknown                          | Red         | Red        | Diploid          | Unknown *Hylocereus* species    |                                                                              |
| 31  | Spineless Yellow   | Unknown         | Unknown           | Unknown                          | Red         | Red        | Diploid          | Unknown *Hylocereus* species    |                                                                              |
| 32  | Yellow Pitaya      | Unknown         | Unknown           | Unknown                          | White       | Yellow     | Tetraploid       | *H. megaranthus*               |                                                                              |
Applied Biosystems) using GeneMapper (Applied Biosystems). SSR markers with amplification of target, no multiple amplifications, no obvious competitive amplification, and zero or one accession with null amplification, were selected as relevant SSR markers. The scheme for SSR marker development is shown in Supplemental Figure 1.

Genetic diversity analysis

The number of alleles ($N_a$), expected heterozygosity ($H_e$), observed heterozygosity ($H_o$), and polymorphic information content (PIC) at single-locus SSR markers in the tested pitaya accessions were calculated using Marker Toolkit (Fujii et al., 2008). For calculation of $H_e$ and $H_o$, polyplloid accessions ('Yellow Pitaya' and ‘Desert King Seedling’) were excluded. $H_e$ was calculated from the allele frequencies using an unbiased formula $1 - \sum p_i^2$ ($1 \leq i \leq m$), where $m$ is the number of alleles at the target locus and $p_i$ is the allele frequency of the $i^{th}$ allele at the target locus. $H_o$ was calculated as the ratio of the heterozygous genotypes scored at each locus. Fixation index ($F_{is}$) was calculated as $1-H_o/H_e$, and out-crossing rate was calculated using the formula $(1-F_{is})/(1+F_{is})$ (Weir and Cockerham, 1984). MinimalMarker (Fujii et al., 2013) was used to identify minimal marker subsets for distinguishing the 32 accessions. Parent–offspring relationships were tested by comparison of the SSR alleles in each accession; data were analyzed in MARCO (Fujii et al., 2008).

A tree of the 30 accessions, without polyploid accessions ('Yellow Pitaya' and ‘Desert King Seedling’), was constructed using the unweighted pair-group method with the arithmetic mean (UPGMA) based on similarities between genotypes estimated by Dice’s coefficient: $Dc = 2n_x/n_x + n_y$, where $n_x$ and $n_y$ represent the number of putative SSR alleles for materials $X$ and $Y$, respectively, and $n_{xy}$ represents the number of putative SSR alleles shared between $X$ and $Y$. NTSYS-pc v. 2.2 (Rohlf, 1998) was used for the analysis. Principal components analysis (PCA) was performed for SSR alleles of 32 accessions using the “rda” function in the vegan package in R 4.0.2 (Oksanen et al., 2019). Population genetic structure was analyzed by Bayesian clustering approach using STRUCTURE V2.3.4 (Pritchard et al., 2000) assuming an admixture model with five independent runs per $K$ Value ($K1–10$) with a burn-in period of 100,000 iterations and 10,000 Markov chain Monte Carlo (MCMC). The Delta $K$ method was used to predict optimal cluster number (Evanno et al., 2005).

Results

SSR marker development

We obtained sequence data for identifying SSRs and developing SSR markers. The obtained sequence data included 1,189,684 reads with an average length of 515 bp and a total base of 0.6 Gb. After sequence trimming, 941,924 reads with an average length of 426 bp and a total base of 0.4 Gb were obtained, approximately 0.2-fold smaller than the estimated size of the pitaya genome, 2 Gb. Among NGS sequences, 17,893, 12,470, 776, 186, and 216 for di-, tri-, tetra-, penta-, and hexanucleotide SSRs, respectively, were detected (Table 2). Most of the detected SSRs were di- and tri-nucleotide motifs. The AG/CT motif was dominant, while the GC/CG was scarce among detected SSRs.

For the development of SSR markers, 10 to 20 repeats of the di-nucleotide motifs were the target SSRs. Sequences including SSRs with AG/CT or AC/GT motif and no other SSRs within 100 b were searched, of which 1826 SSRs were found. Nine hundred and ten SSR primer pairs were designed by BatchPrimer3 and

| Table 2. Number of simple sequence repeats (SSRs) in trimmed sequences. |
|-----------------------------|---------------------|
| **SSR motif**               | **Number of SSRs**  |
| Dinucleotide repeats (more than 6 repeats) |                     |
| AG/CT                       | 9621 (53.7%)        |
| AT/AT                       | 4427 (24.7%)        |
| AC/GT                       | 3773 (21.0%)        |
| CG/CN                       | 72 (0.4%)           |
| Total                       | 17,893              |
| Trinucleotide repeats (more than 5 repeats) |                     |
| AAT/ATT                     | 3744 (30.0%)        |
| AAC/GTT                     | 2754 (22.1%)        |
| ACG/ATG                     | 2329 (18.7%)        |
| AAG/CTT                     | 2010 (16.1%)        |
| Others (<5% of each)        | 1633 (13.1%)        |
| Total                       | 12,470              |
| Tetranucleotide repeats (more than 5 repeats) |                     |
| ACAT/ATGT                   | 306 (39.4%)         |
| AAAT/ATTTT                  | 244 (31.4%)         |
| AGAT/ATCT                   | 76 (9.8%)           |
| AATT/AATT                   | 48 (6.2%)           |
| Others (<5% of each)        | 102 (13.1%)         |
| Total                       | 776                 |
| Pentanucleotide repeats (more than 5 repeats) |                     |
| AAGAG/CTCTT                 | 30 (16.1%)          |
| AAAAT/ATTTTT                | 24 (12.9%)          |
| AAAAC/GTTTT                 | 23 (12.3%)          |
| ACCCC/GGGGT                 | 19 (10.2%)          |
| AAAAG/CTTTT                 | 13 (6.9%)           |
| ACCGG/CCGGT                 | 11 (5.9%)           |
| AAAC/AGTTT                  | 10 (5.4%)           |
| Others (<5% of each)        | 56 (30.1%)          |
| Total                       | 186                 |
| Hexanucleotide repeats (more than 5 repeats) |                     |
| AAGATC/ACTTTG               | 28 (13.0%)          |
| AGATAT/ATACTT               | 16 (7.4%)           |
| AAGAGG/CCCTCTT              | 15 (6.9%)           |
| ACCCAT/ATGGGT               | 12 (5.6%)           |
| Others (<5% of each)        | 145 (67.1%)         |
| Total                       | 216                 |
the resulting 623 non-redundant SSRs were extracted by BLAST search. Fifty AG/CT motif and 20 AC/GT motif SSRs were PCR-evaluated. As a result, 23 SSRs could not be used to obtain genotype information due to a complete lack of target amplification in all examined accessions (no amplification), or multiple amplification or no polymorphism among examined accessions. The remaining 47 SSRs showed polymorphism across the examined accessions. Of these, 31 were excluded because miss-genotyping was expected by null or competitive alleles. Finally, 16 markers were identified as polymorphic SSR markers with low capability of miss-genotyping. The 16 relevant SSR markers included 11 of the AG/CT motif and five AC/GT motif SSR markers for genetic identification of pitaya (Table 3).

SSR marker characteristics and genetic identification of pitaya

The 16 SSR marker characteristics and the established genotypes of the tested 32 accessions are listed in Table 3 and Supplemental Table 1. Na ranged from 6 to 12 found in TsuHU035 and TsuHU038, respectively, with an average value of 7.93. The value of He ranged from 0.5 to, with an average value of 0.75. The value of Ho ranged from 0.23 to 0.93, with an average value of 0.74. PIC ranged from 0.47 to 0.83, with an average value of 0.72. The value of $F_{st}$ was 0.013 and outcrossing rate was calculated as 0.973. Among the tested accessions, two identical pairs were found (‘Ohgimi Red 1’/‘Ohgimi Red 2’ and ‘Nago White’/‘Uruma White’ were identical to each other) as shown in Supplemental Table 1. The remaining pitaya accessions

| Marker name | Primer sequence (5′-3′) | Product size (min–max) | Repeat motif in ‘No. 5’ | Na | He | Ho | PIC | DDBJ accession |
|-------------|-------------------------|------------------------|------------------------|----|----|----|-----|----------------|
| TsuHU031    | F:ATTACTCAAGCAATCTGCAA R:GTTTTGGTAGTGATGGTCATTTCACCG | 169–200 (AG)16 | 9 | 0.78 | 0.80 | 0.76 | LC534945 |
| TsuHU035    | F:CAATCATCTTCCCTCTTAT R:GTTGCTCTGTAGAGGAGTTCGAAG | 201–219 (GA)16 | 6 | 0.75 | 0.80 | 0.70 | LC534946 |
| TsuHU038    | F:CGATACACACACAAAGTCTCT R:GTTTTCTTTGCTCTACTATTCTCAGTGCA | 161–203 (GA)12 | 8.5 | 0.93 | 0.83 | LC534947 |
| TsuHU040    | F:GTATTTGGCAGAAATGAACCA R:GTTTTCTTGGCAGCTAAGTTTGT | 160–196 (GA)16 | 7 | 0.70 | 0.80 | 0.65 | LC534948 |
| TsuHU044    | F:GCCACCTAAAGAAGTAAACCA R:GTTTTCTCTCCTCTCCTCCTCTG | 194–217 (AG)15 | 8 | 0.80 | 0.83 | 0.78 | LC534949 |
| TsuHU045    | F:AATTCCTTGCAGAAAGTAT R:GTTTTCTTTATGAAATGAACATGGAGACC | 205–229 (GA)15 | 9 | 0.72 | 0.70 | 0.69 | LC534950 |
| TsuHU046    | F:GAAATCTCTTGGGCACTT R:GTTTTCTTGGGCTCTCTCTCCTCTG | 184–210 (AG)15 | 8 | 0.75 | 0.80 | 0.71 | LC534951 |
| TsuHU047    | F:AGGCATTGACTAATGCGAAT R:GTTTTCTTCACTCAGGTCACCTCTT | 186–202 (AG)15 | 8 | 0.83 | 0.67 | 0.81 | LC534952 |
| TsuHU048    | F:CAAAAGGCGAGGATAGATACCA R:GTTTTCTTTGCCACCAAGTATAGCA | 188–210 (AG)15 | 8 | 0.79 | 0.90 | 0.76 | LC534953 |
| TsuHU049    | F:CCACATGTTTCAAGAGAAG R:GTTTTCTCTACCTGACCAAACGCTAT | 192–213 (AG)15 | 7 | 0.50 | 0.23 | 0.47 | LC534954 |
| TsuHU050    | F:AAAGCCCAGTGTGTAGAAG R:GTTTTCTTAGGATAAAAATGGAATG | 211–222 (AG)15 | 7 | 0.72 | 0.77 | 0.68 | LC534955 |
| TsuHU296    | F:TAGGAATCGCTAGAGACTTAA R:GTTTTCTTATCACTCAGCAGT | 191–230 (CA)15 | 8 | 0.83 | 0.57 | 0.71 | LC534956 |
| TsuHU300    | F:GTTCGCTGTCGCAATGAGTTT R:GTTTTCTTGCAAATATCTGTCTTAA | 194–217 (TG)18 | 7 | 0.79 | 0.83 | 0.77 | LC534957 |
| TsuHU301    | F:CCTGATGAAATTGGTACCTG R:GTTTTCTGATACGTATGATGAGATC | 183–244 (AC)17 | 8 | 0.82 | 0.83 | 0.80 | LC534958 |
| TsuHU304    | F:ATTTCGCTACAGGACATCTCT G:GTTTTCTTTTCATTTGAGCTGTCCTT | 181–204 (TG)17 | 8 | 0.60 | 0.70 | 0.68 | LC534959 |
| TsuHU312    | F:GAATGACGATTAGAGGAGAAT R:GTTTTCTTATTAAAGTGTTCCTCTT | 207–232 (CA)15 | 8 | 0.61 | 0.70 | 0.58 | LC534960 |

* Na: number of alleles.  
* He: expected heterozygosity.  
* Ho: observed heterozygosity.  
* PIC: polymorphic information content.
could be differentiated with two of the three markers as follows; TsuHU038, TsuHU044, and TsuHU047; and TsuHU047, TsuHU300, and TsuHU304.

One or two fragments of SSR markers were observed in most accessions, amplification of three or four alleles was observed in ‘Desert King Seedling’ and ‘Yellow Pitaya’ (Supplemental Table 1). For ‘Yellow Pitaya’, four alleles were observed by SSR TsuHU038. For the ‘Desert King Seedling’, three alleles were observed by nine SSRs. Therefore, it is possible that the ‘Yellow Pitaya’ was tetraploid and the ‘Desert King Seedling’ was triploid.

According to parentage analysis, both parents were predicted and compared with pedigree according to the SSR marker genotypes (Table 1; Supplemental Table 2). There were six accessions including ‘Impact Ruby’, ‘No. 22’, ‘2-27’, ‘2-29’, and ‘6-15’ without SSR marker discrepancy. For ‘6-15’, both parents were suggested with one marker discrepancy, except for ‘6-15’ having one marker discrepancy, five accessions including ‘Nago 1’, ‘1-08’, ‘1-21’, ‘5-19’, and ‘Nago 2’ had one parent suggested with one marker discrepancy in ‘Nago 2’. ‘No. 5’ had three possible parents. Suggested parentages were coincident with pedigree when accepting a one marker discrepancy.

Genetic diversity analysis
For genetic diversity analysis, an UPGMA tree and PCA plot based on SSR marker genotypes were constructed (Figs. 1 and 2). PCA showed that the three first principal components explained 46.4% of total variability. The contributions of PC1, PC2, and PC3 were 18.52%, 15.52%, and 12.37%, respectively. According to the UPGMA tree and PCA plot, the examined accessions could be classified into six groups. Group I included 19 accessions that had white to purple flesh with pink to red skin. According to the UPGMA tree, Group I was divided into three subgroups. Subgroup I-1 included identical accessions ‘Uruma White’ and ‘Nago White’. Subgroup I-2 included ‘No. 22’ and its progenies ‘2-27’ and ‘2-29’. Subgroup I-3 included ‘No. 5’ and its polycross progenies (‘1-08’, ‘1-21’, ‘5-19’, ‘5-21’, ‘Nago 1’, and ‘Nago 2’). Group II included ‘Spineless Red’. Group III included ‘6-15’ and ‘Impact Ruby’, which were hybrids between ‘Ohgimi White 2’ in Group I, and ‘Spineless Red’ in Group II. Group IV
included all the ‘Golden Seedling’ accessions obtained from imported pitaya fruit. Group V included ‘Spineless Yellow’. Although Group II and V were relatively close on PCA plots, they were not grouped in the same group in the UPGMA tree (Figure 2). Thus, groups II and V were treated as different groups. Group VI included ‘Yellow Pitaya’ and ‘Desert King Seedling’, both of which had yellowish fruit skin. According to SSR marker genotypes, these two accessions may exhibit polyploidy. Classification of groups I to V was supported by STRUCTURE analysis (Supplemental Fig. 2). Cluster number of $K = 4$ was presumed as optimal based on delta $K$ method. Other genetic structures, $K = 3$ to $K = 6$ are also displayed in Supplemental Figure 2. Clustering based on $K = 5$ and $K = 6$ was almost in accordance with the UPGMA tree result. Three subgroups (I-1, I-2, and I-3) which were also suggested by the UPGMA tree were similarly divided in a STRUCTURE bar plot.

### Discussion

In this study, we developed 16 SSR markers with 154 alleles at an average PIC of 0.72. The SSR markers developed in this study suggested high polymorphism of DNA markers. Using these markers, genetic identification was successfully performed using three SSR markers with the exceptions of ‘Uruma White’/‘Nago White’ and ‘Ohgimi Red 1’/‘Ohgimi Red 2’.

As the developed SSR markers were highly polymorphic across examined accessions, they could be applied to confirm pedigree and predict unknown parents. Regarding confirmation of the pedigree in breeding lines bred by OPARC for 10 accessions without unknown pedigree accessions as in ‘No. 5’ and ‘No. 22’, pedigree was confirmed with less than one marker discrepancy in the SSR marker genotype. For accessions with a one marker discrepancy, there were possibilities of unexpected crosses, seedling confusion, and erroneous genotyping caused by allele competition or stutter bands. Assessing additional SSR markers would clarify the cause of discrepancy in the two accessions. Although a one marker discrepancy was observed in the two accessions, the breeding pedigrees were mostly accurate. This suggested that the crossbreeding technique in OPARC was appropriate.

For prediction of the unknown parent of an accession, we successfully predicted both parents in two accessions. Two OPARC breeding lines, ‘No. 5’ and ‘No. 22’ had unknown parents, and one parent was unknown in six accessions, the polycross progeny of ‘No. 5’. According to SSR marker genotypes, both parents were successfully predicted in ‘No. 22’ and ‘5-21’. In contrast, the parent combination was not suggested in ‘No. 5’ and polycross progenies of ‘No. 5’ including ‘Nago 1’, ‘Nago 2’, ‘1-08’, ‘1-21’, and ‘5-19’. For a polycross, pollen from various accessions are mixed and used for pollination. OPARC possesses pitaya accessions, other than the 32 accessions used in this study, including other breeding lines. Such accessions were suggested as the paternal parent of these accessions. For ‘No. 5’, a combination of parents without discrepancy was not suggested. This suggested that at least one parent was not included in the examined accessions. In addition, one parent could not be specified in ‘No. 5’ because three accessions had no SSR marker discrepancies as parents. To identify the parent in ‘No. 5’, additional SSR markers should be applied for parentage analysis. Including SSR markers described in Pan et al. (2017) for accessions used in this study or developing new SSR markers from NGS data obtained in this study would facilitate further parentage analyses.

The accessions examined in this study were not species-clarified. Based on genetic diversity analyses, examined accessions were genetically classified into six groups (I–VI). These groupings appeared to reflect species differences. White fleshed accessions in Group I showed typical characteristics of H. undatus, which have white flesh and a good taste (Weiss et al., 1994). Group I also included red fleshed accessions. This suggested that hybrids between H. undatus and other species were also included in Group I. For Group VI, H. megalanthus and its hybrid accessions were suggested. H. megalanthus has tasty medium-sized yellow skinned fruit (Nerd et al., 2002; Weiss et al., 1994). In addition, while most Hylocereus species have diploid (2n = 22) chromosomes, H. megalanthus has tetraploid (2n = 44) chromosomes and the hybrid between diploid Hylocereus species produce triploid hybrids (Lichtenzveig et al., 2000; Tel-Zur et al., 2003). Group IV showed a darker fruit color, which is a typical characteristic for H. costaricensis or H. monacanthus (Weiss et al., 1994). For Groups II and V, it was difficult to judge the species based on fruit characteristics. Group III was suggested as a hybrid of Groups I and II. In this study, we could not include accessions identified as belonging to specific species (e.g., H. undatus, H. costaricensis, H. megalanthus, and H. monacanthus), and out-group accessions belonging to the Cactaceae family (e.g., Epiphyllum species and Opuntia species). Adding these accessions for genetic diversity analyses may provide more reliable species prediction.

For breeding of pitaya cultivars in OPARC, fruit with good taste (approximately more than 13% brix) and large size (approximately more than 350 g) are required as the main target traits. Group I accessions were used to obtain such cultivars in the OPARC. In addition, various fruit colors, including yellow skin or dark flesh color, are expected. Accessions in Groups IV and VI could be considered as breeding parents to obtain dark red color and yellow skin cultivars. To obtain good taste and large fruit with such fruit color cultivars, a cross between Group I and Groups IV or VI could be considered. Although an embryo rescue technique is required, interspecific crossed progenies could be obtained in
Hylocereus (Cisneros and Tel-Zur, 2010). Including interspecific cross, cultivars with target traits are expected to be obtained by future breeding. Therefore, we concluded that the tested accessions were sufficiently diverse for future breeding. Although self-incompatibility was reported for *H. monacanthus* (Weiss et al., 1994) and *H. costaricensis* (Lichtenzevig et al., 2000), a weak self-incompatibility reaction was observed in *H. undatus* and *H. megalanthus* (Lichtenzevig et al., 2000). In this study, breeding lines included no self progenies according to SSR marker genotypes. In addition, the examined accessions showed high heterozygosity as the Ho of SSR markers were high among the majority of examined SSR markers. The obtained out-crossing rate suggested an almost complete out-crossing nature. These results may reflect the self-incompatibility of pitaya examined in this study.

Since two pitaya pairs, ‘Uruma White’/’Nago White’ and ‘Ohgimi Red 1’/’Ohgimi Red 2’, appeared identical by SSR analysis, there were 5 different indigenous pitaya genotypes. Therefore, it was confirmed that commercial growers cultivated and traded genetically different pitayas with the same trading name. Quality variation in traded pitaya is an issue in Okinawa prefecture, and genetic variation in cultivated accessions was suggested to be an important cause. Breeding of new cultivars and supplying them to commercial farmers would facilitate advances in the quality of pitaya grown in Okinawa prefecture.

**Literature Cited**

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215: 403–410.

Brownstein, M. J., J. D. Carpten and J. R. Smith. 1996. Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. BioTechniques 20: 1004–1010.

Cisneros, A. and N. Tel-Zur. 2010. Embryo rescue and plant regeneration following interspecific crosses in the genus *Hylocereus* (Cactaceae). Euphytica 174: 73–82.

Ellegren, H. 2004. Microsatellites: simple sequences with complex evolution. Nat. Rev. Genet. 5: 435–445.

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.

Fujii, H., T. Ogata, T. Shimada, T. Endo, H. Iketani, T. Shimizu, T. Yamamoto and M. Omura. 2013. Minimal marker: an algorithm and computer program for the identification of minimal sets of discriminating DNA markers for efficient variety identification. J. Bioinform. Comput. Biol. 11: 1250022. DOI: 10.1142/S0219720012500229.

Fujii, H., H. Yamashita, F. Hosaka, S. Terakami and T. Yamamoto. 2010. Development of a software MARCO to presume the parental-child relationship using the result of DNA marker typing. Hort. Res. (Japan) 9 (Suppl. 1): 34 (In Japanese).

Fuji, H., H. Yamashita, T. Shimada, T. Endo, T. Shimizu and T. Yamamoto. 2008. MarkerToolKit: an analysis program for data sets consist of DNA marker types obtained from various varieties. DNA Polymorph. 16: 103–107 (In Japanese).

Lichtenzevig, J., S. Abbo, A. Nerd, N. Tel-Zur and Y. Mizrahi. 2000. Cytology and mating systems in the climbing cacti *Hylocereus* and *Selenicereus*. Am. J. Bot. 87: 1058–1065.

Nerd, A., N. Tel-Zur and Y. Mizrahi. 2002. Fruits of vine and columnar Cacti. p. 185–197. In: P. S. Novel (ed.). Cacti: biology and uses. University of California Press, California.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O’Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs and H. Wagner. 2019. vegan: Community Ecology Package. R package version 2.5-6. Available at <https://cran.r-project.org/web/packages/vegan/index.html>.

Ortiz-Hernández, Y. D. and J. A. Carrillo-Salazar. 2012. Pitahaya (*Hylocereus* spp.): a short review. Comun. Sci. 3: 220–237.

Pagliaccia, D., G. Vidalakis, G. Douthan, R. Lobo and G. Taniakizi. 2015. Genetic characterization of pitahaya accessions based on amplified fragment length polymorphism analysis. HortScience 50: 332–336.

Pan, L., J. Fu, R. Zhang, Y. Qin, F. Lu, L. Jia, Q. Hu, C. Liu, L. Huang and G. Liang. 2017. Genetic diversity among germplasms of pitaya based on SSR markers. Sci. Hortic. 225: 171–176.

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

Rohlf, F. J. 1998. NTSYS, numerical taxonomy and multivariate analysis system, Ver. 2.01. Exeter Publishing, Ltd., Setauket, New York.

Tao, J., G. Qiao, X. P. Wen, G. L. Gao, T. Liu, Z. J. Peng, Y. Q. Cai, N. Chen, F. X. Yan and B. X. Zhang. 2014. Characterization of genetic relationship of dragon fruit accessions (*Hylocereus* spp.) by morphological traits and ISSR markers. Sci. Hortic. 170: 82–88.

Tel-Zur, N., S. Abbo, D. Bar-Zvi and Y. Mizrahi. 2003. Chromosome doubling in vine Cacti hybrids. J. Hered. 94: 329–333.

Tel-Zur, N., S. Abbo, D. Bar-Zvi and Y. Mizrahi. 2004. Clone identification and genetic relationship among vine cacti from the genera *Hylocereus* and *Selenicereus* based on RAPD analysis. Sci. Hortic. 100: 279–289.

Weir, B. S. and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38: 1358–1370.

Weiss, J., A. Nerd and Y. Mizrahi. 1994. Flowering behavior and pollination requirements in climbing cacti with fruit crop potential. HortScience 29: 1487–1492.

You, F. M., N. Huo, Y. Q. Gu, M. Luo, Y. Ma, D. Hane, G. R. Lazo, J. Dvorak and O. D. Anderson. 2008. BatchPrimer3: a high throughput web application for PCR and sequencing primer design. BMC Bioinform. 9: 253. DOI: 10.1186/1471-2105-9-253.

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.