

1 **Chromosome-scale genome assembly of Japanese pear (*Pyrus pyrifolia*) variety**

2 **‘Nijisseiki’**

3

4 Kenta Shirasawa^{1*}, Akihiro Itai², and Sachiko Isobe¹

5 ¹Kazusa DNA Research Institute, Chiba, Japan

6 ²Kyoto Prefectural University, Kyoto, Japan

7

8 *To whom correspondence should be addressed.

9 Tel.: +81 438 52 3935

10 Fax: +81 438 52 3934

11 Email: shirasaw@kazusa.or.jp

12

13 **Abstract**

14 **Aim**

15 The Japanese pear (*P. pyrifolia*) variety 'Nijisseiki' is valued for its superior flesh
16 texture, which has led to its use as a breeding parent for most Japanese pear cultivars.

17 However, in the absence of genomic resources for Japanese pear, the parents of the
18 'Nijisseiki' cultivar remain unknown, as does the genetic basis of its favorable texture.

19 The genomes of pear and related species are complex due to ancestral whole genome
20 duplication and high heterozygosity, and long-sequencing technology was used to
21 address this.

22 **Methods and Results**

23 De novo assembly of long sequence reads covered 136× of the Japanese pear genome
24 and generated 503.9 Mb contigs consisting of 114 sequences with an N50 value of

1 7.6 Mb. Contigs were assigned to Japanese pear genetic maps to establish 17
2 chromosome-scale sequences. In total, 44,876 protein-encoding genes were predicted,
3 84.3% of which were supported by predicted genes and transcriptome data from
4 Japanese pear relatives. As expected, evidence of whole genome duplication was
5 observed, consistent with related species.

6 **Conclusion and Perspective**

7 This is the first genome sequence analysis reported for Japanese pear, and this resource
8 will support breeding programs and provide new insights into the physiology and
9 evolutionary history of Japanese pear.

11 **Keywords**

12 Chromosome-scale genome assembly; Japanese pear; Long-read sequence technology;
13 Whole genome duplication

15 **1. Introduction**

16 Pear (*Pyrus* spp.) is a genus of the Malinae subtribe of the Rosaceae that includes
17 European pear (*P. communis*), Chinese white pear (*P. × bretschneideri*), Japanese pear
18 (*P. pyrifolia*), and apple (*Malus × domestica*). The Japanese pear variety ‘Nijisseiki’ (a
19 Japanese term referring to 20th-century) was discovered in Matsudo (Chiba, Japan) in
20 1888¹. Due to its favorable flesh texture, ‘Nijisseiki’ was the leading Japanese pear
21 variety in Japan from the 1940s to the 1980s, and ‘Nijisseiki’ was widely used as a
22 breeding parent for the development of Japanese pear cultivars in Japan¹.

23 Genome information can enhance breeding programs² by facilitating understanding
24 of the genetic backgrounds of the breeding pedigrees³, and by enhancing physiology

1 and evolutionary studies. In the Malinae, genome sequence data are publicly available
 2 for crop species such as apple⁴, European pear⁵, and Chinese white pear⁶, as well as for
 3 some pear wild relatives (*P. betuleafolia* and (*P. ussuriensis* × *communis*) × spp.) that
 4 are used for root stocks^{7,8}. Genome analysis of the Malinae is complex due to ancestral
 5 whole genome duplication⁹ and high heterozygosity in their genomes resulting from
 6 allogamy and self-incompatibility. To simplify genomic analysis, doubled-haploid lines
 7 were developed to reduce the genome complexities in materials used for genome
 8 sequencing^{5,6}. However, neither doubled-haploid lines nor genome sequence data are
 9 available for Japanese pear, despite the publication of transcriptome data¹⁰ and genetic
 10 maps^{11,12}.

11 Long-read sequencing has several advantages over short-read technologies¹³. Long
 12 reads span repetitive sequence stretches in genomes, extending sequence contiguity and
 13 facilitating assembly. Furthermore, long reads allow haplotype phases of highly
 14 heterozygous genome sequences to be determined. In this study, long-read sequencing
 15 was used to produce a highly contiguous genome sequence assembly of the Japanese
 16 pear variety ‘Nijisseiki’. This Japanese pear genome will enhance our understanding of
 17 genetics, genomics, and breeding in Japanese pear.

18

19 **2. Materials and methods**

20 **2.1. Plant materials and DNA extraction**

21 A single tree of Japanese pear (*P. pyrifolia*), variety ‘Nijisseiki’, which is planted at the
 22 orchard of Kyoto Prefectural University (Kyoto, Japan), was used for genome
 23 sequencing analysis. Genome DNA was extracted from the young leaves by a modified
 24 sodium dodecyl sulfate method¹⁴.

1

2 2.2.Estimation of the genome size of Japanese pear

3 A short-read sequence library was prepared using a PCR-free Swift 2S Turbo Flexible
4 DNA Library Kit (Swift Sciences, Ann Arbor, MI, USA) and converted into a DNA
5 nanoball sequencing library with an MGI Easy Universal Library Conversion Kit (MGI
6 Tech, Shenzhen, China). The library was sequenced on a DNBSEQ G400RS (MGI
7 instrument in paired-end, 101 bp mode. The obtained reads were used to estimate the
8 genome size with Jellyfish after removing low-quality bases (<10 quality value) with
9 PRINSEQ, adaptor sequences (AGATCGGAAGAGC) with fastx_clipper in
10 FASTX-Toolkit, and reads from organelle genomes^{15,16} (GenBank accession numbers:
11 AP012207 and KY563267) by read mapping with Bowtie2 on the reference sequences.

12

13 2.3.Chromosome-scale genome assembly

14 A long-read sequence library was constructed using an SMRTbell Express
15 Template Prep Kit 2.0 and sequenced on SMRT cells (1M v3 LR) in a PacBio Sequel
16 system (PacBio, Menlo Park, CA, USA). The obtained reads (≥ 15 kb) were assembled
17 with Falcon, and the two haplotype sequences, primary contigs and haplotigs, of the
18 diploid genome were resolved with Falcon-unzip. Potential sequence errors in the
19 assembled sequences were corrected with ARROW using the PacBio reads. Haplotype
20 duplications in the primary contigs were removed with Purge_Dups. Sequences from
21 the organelle genomes, which were identified by a sequence similarity search of
22 reported plastid¹⁶ (AP012207) and mitochondrial¹⁵ (KY563267) genome sequences
23 from Japanese pear using minimap2, were also deleted. The final assembly was
24 designated as PPY_r1.0. The primary contigs of PPY_r1.0 were assigned to genetic

1 maps of Japanese pear^{11,12} with ALLMAPS, in which marker sequences were aligned on
 2 the contigs with BLAST. The resultant chromosome-scale pseudomolecule sequences
 3 were named PPY_r1.0.pmol. The haplotig sequences were aligned onto PPY_r1.0.pmol
 4 with minimap2. Completeness evaluation of the assembly was performed with BUSCO
 5 (embryophyta odb9). The chromosome-scale pseudomolecule sequences
 6 (PPY_r1.0.pmol) were compared with those of apple (GDDH13, v1.1)⁴, European pear
 7 (Bartlett DH, v2.0)⁵, and Chinese white pear (Dangshan Suli, v1.1)⁶ with D-GENIES.
 8 Whole genome duplications were detected with MCScanX, with threshold values of
 9 $\geq 85\%$ sequence identity and $\leq 1e-100$ E-values, and visualized with VGSC.

10

11 2.4. Gene prediction and repetitive sequence analysis

12 Protein-encoding genes were predicted by a MAKER pipeline using two training sets
 13 and a preset of SNAP for Arabidopsis. Peptide sequences of apple (GDDH13, v1.1)⁴,
 14 European pear (Bartlett DH, v2.0)⁵, and Chinese white pear (Dangshan Suli, v1.1)⁶
 15 registered in the Genome Database for Rosaceae¹⁷, as well as transcriptome data for
 16 Japanese pear (Ppyrifolia_protein_v1.0)¹⁰, were also employed in the prediction. Short
 17 gene sequences of <300 bases, as well as genes predicted with an annotation edit
 18 distance of <0.5, were removed to facilitate selection of high-confidence genes. The
 19 genome positions of the high-confidence genes were compared with those of the peptide
 20 sequences from apple, European pear, Chinese white pear, and Japanese pear, and were
 21 aligned to PPY_r1.0.pmol in the MAKER pipeline. Functional annotation of the
 22 predicted genes was performed with Hayai-Annotation Plants.

23 Repetitive sequences in the PPY_r1.0.pmol assembly were identified with
 24 RepeatMasker, using repeat sequences registered in Repbase and a de novo repeat

1 library built with RepeatModeler. The repeat elements were classified into nine types:
 2 short interspersed nuclear elements (SINEs), long interspersed nuclear elements
 3 (LINEs), long terminal repeat (LTR) elements, DNA elements, Small RNA, satellites,
 4 simple repeats, low complexity repeats, and unclassified, in accordance with
 5 RepeatMasker.

6 The software tools used for the data analyses are shown in Supplementary Table
 7 S1.

8

9 **3. Results and data description**

10 **3.1. Assembly of the genome of a Japanese pear, 'Nijisseiki'**

11 In total, 76.9 Gb short-read data were used for estimation of the genome size of
 12 Japanese pear. *K*-mer distribution analysis indicated that the 'Nijisseiki' genome was
 13 highly heterozygous, and that the estimated haploid genome size was 537.1 Mb
 14 (Figure 1).

15 Long-read sequencing analysis with four SMRT cells generated 85.8 Gb data (4.5
 16 million reads with an N50 of 25.5 kb). Of these, reads ≥ 15 kb in length (73.3 Gb,
 17 136.4 $\times$ genome coverage) were selected and assembled into 750.0 Mb contigs (495
 18 sequences with an N50 of 5.9 Mb). The contig sequences were resolved into primary
 19 contigs and haplotigs, and polished to correct potential errors. The sizes of the resultant
 20 sequences were 745.8 Mb for the primary contigs and 113.3 Mb for the haplotigs. As
 21 this primary contig size was longer than the estimated haploid size, we doubted the
 22 existence of duplicate sequences. Hence, 257 haplotype duplications in the primary
 23 contigs (241.6 Mb) identified by a read depth distribution pattern of Purge_Dups
 24 (Supplementary Figure S1) were excluded. In addition, 10 sequences (283.5 kb), which

showed significant similarity with previously reported organelle genomes, were removed. The final assembly, named PPY_r1.0, consisted of 503.9 Mb primary contigs (including 114 sequences with an N50 length of 7.6 Mb) and 353.4 Mb haplotigs (including 822 sequences with an N50 length of 1.6 Mb) (Table 1). A BUSCO analysis of the primary contigs indicated that 58.3% and 39.7% of sequences were single-copy and duplicated complete BUSCOs, respectively (Table 2), supporting the genome-wide duplication state in the Japanese pear genome.

3.2. Chromosome-scale assembly construction and comparative genomics

Two previously established genetic maps for Japanese pear^{11,12} were used to establish pseudomolecule sequences. Totals of 609 and 2,388 SNP marker sequences were collected from the genetic maps of Li et al.¹¹ and Terakami et al.¹², respectively, and used to identify 587 and 2,306 marker positions on the primary contigs, respectively. Thus, 89 contig sequences (481.4 Mb in total) were assigned to the 17 linkage groups with the marker positions as anchors (Table 3), while the remaining 25 contigs (22.4 Mb) were not included. The haplotig sequences were aligned on the pseudomolecule sequences and covered 332.5 Mb in length. In total, 1,719,266 allelic sequence variants were detected between the two haplotypes.

The structure of the pseudomolecule sequences was compared with those of apple, European pear, and Chinese white pear to reveal one-to-one synteny relationships at chromosome level between the species (Figure 2). Furthermore, possible homologous chromosome pairs derived from genome-wide duplication and intra-chromosomal duplications were detected in all chromosomes (Figure 3, Table 4).

1 3.3. Gene prediction and repetitive sequence analysis

2 Initial predictions suggested 85,013 likely genes in the genome assembly. This gene
3 number was resolved to 44,876 (Table 1) after removal of 29,607 low-confidence genes
4 with an annotation edit distance of <0.5, 10,527 short genes of <300 bases in length,
5 and three genes in redundant positions. A BUSCO analysis of the 44,876 remaining
6 genes indicated that 98.4% of the sequences were complete BUSCOs (Table 2). The
7 44,876 sequences (56.0 Mb in total length) were therefore concluded to be
8 high-confidence Japanese pear genes.

9 The genome positions of 37,835 (84.3%) of the 44,876 genes overlapped those of
10 transcripts of Japanese pear and/or peptide sequences predicted from the genomes of
11 apple, European pear, and Chinese white pear (Table 5). Furthermore, 6,610 (14.7%),
12 11,664 (26.0%), and 8,168 (18.2%) sequences were assigned to Gene Ontology slim
13 terms in the biological process, cellular component, and molecular function categories,
14 respectively, and 1,967 genes had enzyme commission numbers (Supplementary Table
15 S2).

16 Repetitive sequences occupied a total of 280.3 Mb (55.6%) of the PPY_r1.0.pmol
17 (503.8 Mb) assembly. The repeats consisted of nine major types in varying proportions
18 (Table 6). The dominant repeat types in the pseudomolecule sequences were LTR
19 retroelements (172.9 Mb) followed by DNA transposons (59.4 Mb). Repeat sequences
20 that were unavailable in public databases totaled 25.0 Mb.

21

22 4. Conclusion and future perspectives

23 Here, we present the first reported genome sequence analysis of Japanese pear,
24 complementing the existing available genome sequences of other members of the *Pyrus*

1 genus⁵⁻⁸. Long-sequencing technology allowed the assembly of 114 contig sequences
 2 (N50 of 7.6 Mb) spanning 503.9 Mb, 89.6% of the estimated size of the genome (Table
 3 1). Subsequently, 481.4 Mb (95.5% of the assembled sequences) were successfully
 4 assigned to the pear chromosomes to establish pseudomolecule sequences,
 5 PPY_r1.0.pmol. This pseudomolecule assembly is the longest among the reported *Pyrus*
 6 spp.⁵⁻⁸, and PPY_r1.0.pmol can thus act as a new standard for *Pyrus* genomics. Genome
 7 structures and whole genome duplication events were conserved among pears and
 8 apples (Figure 2, Figure 3), despite their divergent genome sizes^{2,4-8}. This Japanese pear
 9 genome assembly contributes to our understanding of the evolutionary history of pear
 10 species and the wider Malinae.

11 The ‘Nijisseiki’ variety of Japanese pear was used for the genome analysis.
 12 Although the parents of the ‘Nijisseiki’ variety are unknown, its favorable flesh texture¹
 13 led to its use as a popular breeding parent for Japanese pear cultivars. The genome data
 14 generated in this study will facilitate the discovery of the likely parents of the ‘Nijisseiki’
 15 variety and will allow the ‘Nijisseiki’ haplotypes to be traced in progeny cultivars. As
 16 the superior flesh texture was unique to ‘Nijisseiki’ at the point of its discovery, and this
 17 texture was passed to progeny cultivars¹, it should be possible to discover the genomic
 18 region responsible for the texture trait through pedigree analysis. This approach was
 19 used successfully to find associations between haplotypes and phenotypes for five traits
 20 in apple cultivars bred in Japan, for which the popular ‘Fuji’ variety was the founder³.

21 In summary, the Japanese pear genome was sequenced using long-read technology,
 22 and sequences were successfully assigned to the chromosomes. Genome structure was
 23 conserved among the members of the Malinae, and whole genome duplication was
 24 apparent in the Japanese pear genome. The Japanese pear genome will provide new

1 insights into the physiology and evolutionary history of Rosaceae and will facilitate
2 Japanese pear breeding programs.

3

4 **Acknowledgements:** We thank Y. Kishida, C. Minami, H. Tsuruoka, and A. Watanabe
5 (Kazusa DNA Research Institute) for their technical assistance.

6

7 **Data availability:** Sequence reads are available from the Sequence Read Archive
8 (DRA) of the DNA Data Bank of Japan (DDBJ) under accession number DRA011164.

9 The DDBJ accession numbers of assembled sequences are

10 BNSU01000001-BNSU01000114.

11

12 **Funding:** This work was supported by the Kazusa DNA Research Institute Foundation.

13

14 **Conflict of interest:** None declared.

15

16 **Supporting information**

17 **Supplementary Table S1.** Program tools used for genome assembly and gene
18 prediction.

19 **Supplementary Table S2.** Functional annotations of predicted genes.

20 **Supplementary Figure S1.** Distribution of read depth on the primary genome
21 assemblies.

22 A cutoff value at read depth of >117 to delete duplicated sequences is indicated by a
23 vertical dotted line.

24

References

1. Saito, T. 2016, Advances in Japanese pear breeding in Japan. *Breed Sci*, **66**, 46-59.
2. Yamamoto, T. and Terakami, S. 2016, Genomics of pear and other Rosaceae fruit trees. *Breed Sci*, **66**, 148-159.
3. Kuniyhis, M., Moriya, S., Abe, K., et al. 2016, Genomic dissection of a 'Fuji' apple cultivar: re-sequencing, SNP marker development, definition of haplotypes, and QTL detection. *Breed Sci*, **66**, 499-515.
4. Daccord, N., Celton, J. M., Linsmith, G., et al. 2017, High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. *Nat Genet*, **49**, 1099-1106.
5. Linsmith, G., Rombauts, S., Montanari, S., et al. 2019, Pseudo-chromosome-length genome assembly of a double haploid "Bartlett" pear (*Pyrus communis* L.). *Gigascience*, **8**.
6. Xue, H., Wang, S., Yao, J. L., et al. 2018, Chromosome level high-density integrated genetic maps improve the *Pyrus bretschneideri* 'DangshanSuli' v1.0 genome. *BMC Genomics*, **19**, 833.
7. Dong, X., Wang, Z., Tian, L., et al. 2020, De novo assembly of a wild pear (*Pyrus betuleafolia*) genome. *Plant Biotechnol J*, **18**, 581-595.
8. Ou, C., Wang, F., Wang, J., et al. 2019, A de novo genome assembly of the dwarfing pear rootstock Zhongai 1. *Sci Data*, **6**, 281.
9. Velasco, R., Zharkikh, A., Affourtit, J., et al. 2010, The genome of the domesticated apple (*Malus x domestica* Borkh.). *Nat Genet*, **42**, 833-839.
10. Koshimizu, S., Nakamura, Y., Nishitani, C., et al. 2019, TRANSNAP: a web database providing comprehensive information on Japanese pear transcriptome. *Sci Rep*, **9**, 18922.
11. Li, L., Deng, C. H., Knabel, M., et al. 2017, Integrated high-density consensus genetic map of *Pyrus* and anchoring of the 'Bartlett' v1.0 (*Pyrus communis*) genome. *DNA Res*, **24**, 289-301.
12. Terakami, S., Nishitani, C., Kuniyhis, M., et al. 2014, Transcriptome-based single nucleotide polymorphism markers for genome mapping in Japanese pear (*Pyrus pyrifolia* Nakai). *Tree Genetics & Genomes*, **10**, 853-863.
13. Michael, T. P. and VanBuren, R. 2020, Building near-complete plant genomes. *Curr Opin Plant Biol*, **54**, 26-33.
14. Teramoto, S., Kano-Murakami, Y., Hori, M. and Kamiyama, K. 1994, DNA Finger-Printing' to Distinguish Cultivar and Parental Relation of Japanese Pear.

1 *Journal of the Japanese Society for Horticultural Science*, **63**, 17-21.

2 15. Chung, H. Y., Won, S. Y., Kang, S.-H., Sohn, S.-H. and Kim, J. S. 2017, The

3 complete mitochondrial genome of Wonwhang (*Pyrus pyrifolia*). *Mitochondrial*

4 *DNA Part B*, **2**, 902-903.

5 16. Terakami, S., Matsumura, Y., Kurita, K., et al. 2012, Complete sequence of the

6 chloroplast genome from pear (*Pyrus pyrifolia*): genome structure and

7 comparative analysis. *Tree Genetics & Genomes*, **8**, 841-854.

8 17. Jung, S., Lee, T., Cheng, C. H., et al. 2019, 15 years of GDR: New data and

9 functionality in the Genome Database for Rosaceae. *Nucleic Acids Res*, **47**,

10 D1137-D1145.

11

1 **Figure legends**

2 **Figure 1** Genome size estimation for Japanese pear ‘Nijisseiki’ with the distribution of
3 the number of distinct k-mers ($k = 17$) with the given multiplicity values.

4 **Figure 2** Comparative analysis of sequence and structure similarities of the Japanese
5 pear genome.

6 Genome sequence and structure similarities of Japanese pear (*P. pyrifolia*) with apple
7 (*M. × domestica*), European pear (*P. communis*), and Chinese white pear (*P.*
8 *bretschneideri*).

9 **Figure 3** Whole genome duplication of the Japanese pear genome.

10 Genome regions showing sequence similarities are connected with lines.

11

12

1 **Table 1** Statistics of the primary contig sequences of Japanese pear ‘Nijisseiki’

	PPY_r1.0 (primary contigs)	PPY_r1.0. haplotigs (haplotig contigs)
Total contig size (bp)	503,888,133	353,451,501
Number of contigs	114	822
Contig N50 length (bp)	7,676,629	1,608,651
Longest contig size (bp)	28,225,311	7,932,431
Gap (%)	0	0
#Predicted genes	44,876	Not available

2

1 **Table 2** Completeness evaluation of genome assembly and predicted genes

	Genome (PPY_r1.0)	Predicted Genes
Complete BUSCOs	98.00%	98.40%
Single-copy BUSCOs	58.30%	58.60%
Duplicated BUSCOs	39.70%	39.80%
Fragmented BUSCOs	0.50%	0.90%
Missing BUSCOs	1.50%	0.70%

2

1 **Table 3** Statistics of the Japanese pear 'Nijisseiki' pseudomolecule sequences,

2 PPY_r1.0.pmol

Chromosome	#Contigs	%	Contig size (bp)	%	#Genes	%
Chr01	3	2.6%	23,195,497	4.6%	2,121	4.7%
Chr02	7	6.1%	31,777,058	6.3%	2,828	6.3%
Chr03	6	5.3%	31,637,635	6.3%	2,765	6.2%
Chr04	5	4.4%	19,501,643	3.9%	1,625	3.6%
Chr05	3	2.6%	35,298,035	7.0%	3,320	7.4%
Chr06	6	5.3%	22,071,348	4.4%	2,016	4.5%
Chr07	8	7.0%	36,191,613	7.2%	3,009	6.7%
Chr08	6	5.3%	26,500,184	5.3%	2,508	5.6%
Chr09	3	2.6%	24,999,441	5.0%	2,399	5.3%
Chr10	2	1.8%	21,149,732	4.2%	2,159	4.8%
Chr11	7	6.1%	35,209,706	7.0%	2,788	6.2%
Chr12	5	4.4%	24,643,494	4.9%	2,227	5.0%
Chr13	6	5.3%	29,570,600	5.9%	2,662	5.9%
Chr14	8	7.0%	22,573,349	4.5%	2,119	4.7%
Chr15	4	3.5%	39,634,418	7.9%	3,697	8.2%
Chr16	5	4.4%	25,480,041	5.1%	2,187	4.9%
Chr17	5	4.4%	31,990,191	6.3%	2,635	5.9%
Subtotal (Chr01-Chr17)	89	78.1%	481,423,985	95.5%	43,065	96.0%
Unassigned contigs	25	21.9%	22,747,703	4.5%	1,811	4.0%
Total	114	100.0%	504,171,688	100.0%	44,876	100.0%

3

1 **Table 4** Chromosome pairs showing sequence and structure homology in the Japanese

2 pear genome

Chromosome	Homologous chromosome ^a
1	7, 15
2	2 , 7, 15, 17
3	3 , 11
4	12
5	5 , 10
6	14, 16
7	1, 2, 7
8	8 , 15
9	17
10	5
11	3
12	4, 12 , 14
13	13 , 16
14	6, 12
15	1, 2, 8
16	6, 13
17	2, 9

3 ^aBold indicates intra-chromosome duplications.

4

1 **Table 5** Number of *P. pyrifolia* genes supported by other species of the Rosaceae

Species supporting genes ^a	Number of genes	%
Pb, Pc, Pp, and Md	17,502	39.0%
Pb, Pc, and Pp	735	1.6%
Pb, Pc, and Md	4,603	10.3%
Pb, Pp, and Md	1,737	3.9%
Pc, Pp, and Md	1,832	4.1%
Pb and Pc	699	1.6%
Pb and Pp	437	1.0%
Pb and Md	1,639	3.7%
Pc and Pp	850	1.9%
Pc and Md	1,350	3.0%
Pp and Md	576	1.3%
Pb	882	2.0%
Pc	2,010	4.5%
Pp	1,580	3.5%
Md	1,403	3.1%
None	7,041	15.7%
Total	44,876	100.0%

2 ^aPb: Chinese white pear (Dangshan Suli, v1.1)⁶; Pc: European pear (Bartlett DH, v2.0)⁵,

3 Pp: Japanese pear (Ppyrifolia_protein_v1.0)¹⁰; and Md: apple (GDDH13, v1.1)⁴.

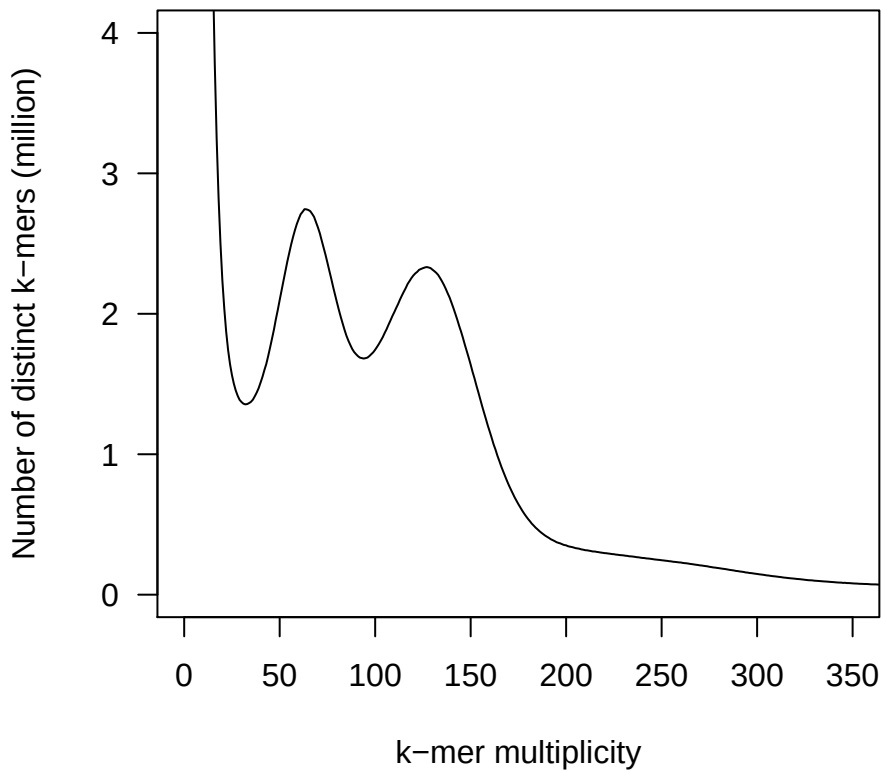
4

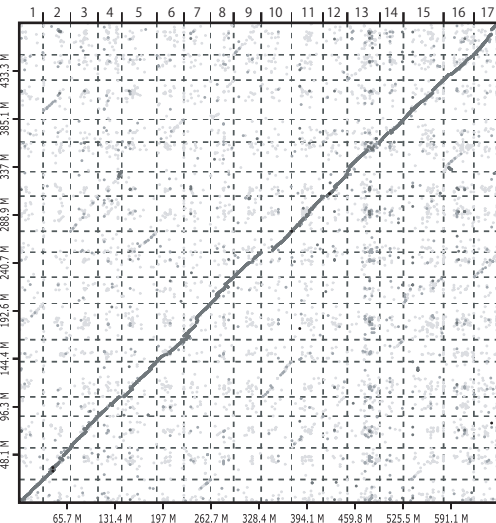
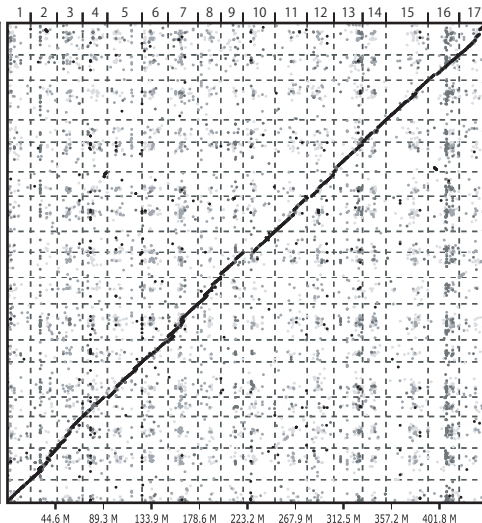
1 **Table 6** Repetitive sequences in the Japanese pear 'Nijisseiki' sequences, PPY_r1.0

Repeat type	Number of elements	Length occupied (bp)	% ^a
SINEs	27,012	3,844,706	0.80%
LINEs	21,048	7,222,875	1.40%
LTR elements	162,422	172,911,968	34.30%
DNA transposons	246,048	59,478,780	11.80%
Unclassified	131,375	25,065,317	5.00%
Small RNA	27,106	4,153,084	0.80%
Satellites	648	109,271	0.00%
Simple repeats	121,341	4,868,197	1.00%
Low complexity	20,350	974,726	0.20%

2 ^aPercentage of sequence length of PPY_r1.0 (503,888,133 bp).

3



Malus x domestica*Pyrus communis**Pyrus x bretschneideri*