Article

Exploring the History of Chloroplast Capture in
Arabis Using Whole Chloroplast Genome Sequencing

Akira Kawabe * , Hiroaki Nukii and Hazuka Y. Furihata

Faculty of Life Sciences, Kyoto Sangyo University, Kyoto, Kyoto 603-8555, Japan; g1447799@cc.kyoto-su.ac.jp (H.N.); hazuka.furi@cc.kyoto-su.ac.jp (H.Y.F.)
* Correspondence: akiraka@cc.kyoto-su.ac.jp; Tel.: +81-75-705-3126

Received: 26 January 2018; Accepted: 16 February 2018; Published: 18 February 2018

Abstract: Chloroplast capture occurs when the chloroplast of one plant species is introgressed into another plant species. The phylogenies of nuclear and chloroplast markers from East Asian Arabis species are incongruent, which indicates hybrid origin and shows chloroplast capture. In the present study, the complete chloroplast genomes of A. hirsuta, A. nipponica, and A. flagellosa were sequenced in order to analyze their divergence and their relationships. The chloroplast genomes of A. nipponica and A. flagellosa were similar, which indicates chloroplast replacement. If hybridization causing chloroplast capture occurred once, divergence between recipient species would be lower than between donor species. However, the chloroplast genomes of species with possible hybrid origins, A. nipponica and A. stelleri, differ at similar levels to possible maternal donor species A. flagellosa, which suggests that multiple hybridization events have occurred in their respective histories. The mitochondrial genomes exhibited similar patterns, while A. nipponica and A. flagellosa were more similar to each other than to A. hirsuta. This suggests that the two organellar genomes were co-transferred during the hybridization history of the East Asian Arabis species.

Keywords: Arabis; chloroplast capture; Brassicaceae

1. Introduction

The genus Arabis includes about 70 species that are distributed throughout the northern hemisphere. The genus previously included many more species, but a large number of these were reclassified into other genera, including Arabidopsis, Turritis, and Boechera, Crucihimalaya, Scapiarabis, and Sinoarabis [1–6]. Because of their highly variable morphology and life histories, Arabis species have been used for ecological and evolutionary studies of morphologic and phenotypic traits [7–11]. The whole genome of Arabis alpina has been sequenced, providing genomic information for evolutionary analyses [12,13].

Molecular phylogenetic studies of Arabis species have been conducted to determine species classification and also correlation to morphological evolution of Arabis species [10,14,15]. Despite having similar morphologies, A. hirsuta from Europe, North America, and East Asia have been placed in different phylogenetic positions and are now considered distinct species. For example, East Asian A. hirsuta, which was previously classified as A. hirsuta var. nipponica, is now designated as A. nipponica [16]. Meanwhile, nuclear ITS sequences indicated that A. nipponica, A. stelleri, and A. takeshimana were closely related to European A. hirsuta. However, chloroplast trnLF sequences indicated that the species were closely related to East Asian Arabis species [14,16]. Such incongruent nuclear and organellar phylogenies have been reported from other plant species and this is generally known as “chloroplast capture” [17,18], which is a process that involves hybridization and many successive backcrosses [17]. When chloroplast capture happens, the chloroplast genome of a species is replaced by another species’ chloroplast genome. A. nipponica may have originated...
from the hybridization of *A. hirsuta* or *A. sagittata* and East Asian *Arabis* species (similar to *A. serrata*, *A. paniculata*, and *A. flagellosa*), which act as paternal and maternal parents, respectively [14,16]. However, the evolutionary history and hybridization processes of *A. nipponica* and other East Asian *Arabis* species still need to be clarified. Because these conclusions for incongruence between nuclear and chloroplast phylogenies came from analyzing a small number of short sequences, hybridized species, the divergence level, and the classification of species are somewhat ambiguous. In the present study, the whole chloroplast genomes of three *Arabis* species were sequenced in order to analyze their divergence and evolutionary history. The whole chloroplast genome sequences also provide a basis for future marker development.

2. Results

2.1. Chloroplast Genome Structure of *Arabis* Species

The structures of the whole chloroplast genomes are summarized in Table 1, which also includes previously reported *Arabis* chloroplast genomes and the chloroplast genome of the closely related species *Draba nemorosa*. The chloroplast genome structure identified in the present study is shown as a circular map (see Figure 1). The complete chloroplast genomes of the *Arabis* species had total lengths of 152,866–153,758 base pairs, which included 82,338 to 82,811 base pair long single copy (LSC) regions and 17,938 to 18,156 base pair short single copy (SSC) regions, which were separated by a pair of 26,421 to 26,933 base pair inverted repeat (IR) regions. The structure and length are conserved, and are similar to other Brassicaceae species’ chloroplast genome sequences [19–22]. The complete genomes contain 86 protein-coding genes, 37 tRNA genes, and eight rRNA genes. Of these, seven protein-coding genes, seven tRNA genes, and four rRNA genes were located in the IR regions, and were therefore duplicated. The *rps16* gene became a pseudogene in *A. flagellosa*, *A. hirsuta*, and *A. nipponica* strain Midori, which was previously reported as a related species [23]. In addition, the *rps16* sequences of *D. nemorosa*, *A. stelleri*, *A. flagellosa*, *A. hirsuta*, and *A. nipponica* shared a 10 base pair deletion in the first exon, while *A. stelleri*, *A. flagellosa*, *A. hirsuta*, and *A. nipponica* shared a 1 base pair deletion in the second exon and *D. nemorosa* lacked the second exon entirely. The *rps16* sequence of *A. alpina* also lacked part of the second exon and had mutations in the start and stop codons. Therefore, different patterns of *rps16* pseudogenization were observed in *A. alpina* and the other *Arabis* species, as was previously suggested [23]. The *A. alpina* lineage had acquired independent dysfunctional mutation(s). The patterns observed for the European *A. hirsuta* revealed that the pseudogenization of *rps16* in the other *Arabis* species might not have occurred independently but, instead, occurred before the divergence of *D. nemorosa* and other *Arabis* species after splitting from *A. alpina*.

| Species       | Strain | Nucleotide Length (bp) | GC Contents (%) | NCBI #          | Reference         |
|---------------|--------|------------------------|-----------------|-----------------|-------------------|
|               |        | Entire | LSC | SSC | IR | Entire | LSC | SSC | IR |
| *Draba nemorosa* | JO21   | 153289 | 82457 | 18126 | 26353 | 36.47 | 34.27 | 29.3 | 42.39 | AP009373 (NC009272) |
| *Arabis alpina* | 152866 | 82338 | 17938 | 26933 | 36.45 | 34.21 | 29.31 | 42.39 | HP934132 (NC023367) [25] |
| *Arabis hirsuta* Brno | 153758 | 82710 | 18156 | 26446 | 36.4 | 34.15 | 29.16 | 42.41 | LC361350 | this study |
| *Arabis flagellosa* Kifune | 153673 | 82775 | 18052 | 26423 | 36.4 | 34.13 | 29.22 | 42.41 | LC361351 | this study |
| *Arabis stelleri* | 153683 | 82807 | 18030 | 26423 | 36.39 | 34.11 | 29.22 | 42.42 | KY126841 | [23] |
| *Arabis nipponica* JO23 | 153689 | 82811 | 18036 | 26421 | 36.4 | 34.1 | 29.31 | 42.42 | AP009369 (NC009268) |
| *Arabis nipponica* Midori | 153668 | 82772 | 18052 | 26422 | 36.39 | 34.1 | 29.24 | 42.42 | LC361349 | this study |
Table 1. Summary of chloroplast genome structure in Arabis species.

| Species       | Strain   | Nucleotide Length (bp) | GC Contents (%) | NCBI #         | Reference |
|---------------|----------|------------------------|-----------------|----------------|-----------|
| Draba nemorosa| JO21     | 153,289                | 82457           | 18126          | 26,353    | 36.47  | 34.27 | 29.3 | 42.39 | AP009373 (NC009272) |
| Arabis alpina |          | 152,866                | 82338           | 17938          | 26,933    | 36.45  | 34.21 | 29.31| 42.39 | HF934132 (NC023367) [25] |
| Arabis hirsuta| Brno     | 153,758                | 82710           | 18156          | 26,446    | 36.4   | 34.15 | 29.16| 42.41 | LC361350 this study |
| Arabis flagellosa| Kifune | 153,673                | 82,775          | 18052          | 26,423    | 36.4   | 34.13 | 29.22| 42.41 | LC361351 this study |
| Arabis stelleri|          | 153,683                | 82807           | 18030          | 26,423    | 36.39  | 34.11 | 29.22| 42.42 | KY126841 [23] |
| Arabis nipponica| JO23 | 153,689                | 82,811          | 18036          | 26,421    | 36.4   | 34.1  | 29.31| 42.42 | AP009369 (NC009268) |
| Arabis nipponica| Midori | 153,668                | 82,772          | 18052          | 26,422    | 36.39  | 34.1  | 29.24| 42.42 | LC361349 this study |

Figure 1. Chloroplast genome structure of Arabis species. Genes shown outside the map circles are transcribed clockwise, while those drawn inside are transcribed counterclockwise. Genes from different functional groups are color-coded according to the key at the top right. The positions of long single copy (LSC), short single copy (SSC), and two inverted repeat (IR: IRA and IRB) regions are shown in the inner circles.

2.2. Chloroplast Genome Divergence

Phylogenetic trees were generated by using whole chloroplast genome sequences and concatenated coding sequence (CDS) regions (see Figure 2). The inclusion of other Brassicaceae members revealed that D. nemorosa should be placed within Arabis, as previously reported [24]. In both trees, the two A. nipponica strains were grouped with A. flagellosa and A. stelleri. Although several nodes were supported by high bootstrap probabilities, the nearly identical sequences of the four East Asian Arabis species made them indistinguishable.

The divergence among the Arabis chloroplast genomes was shown using a VISTA plot (see Figure 3) and this was summarized in Table 2. The genome sequences of the two Japanese A. nipponica strains differed by only 55 nucleotide substitutions (0.036% per site), while those of A. hirsuta and A. nipponica differed by about 3500 sites (2.4% per site). The chloroplast genomes of A. nipponica and the other two East Asian Arabis species were also very similar (~100 nucleotide differences, <0.1% per site). Additionally, the 35 CDS regions, 29 tRNA genes, and four rRNA genes of the four East Asian Arabis species were identical, with three, 27, and four, respectively, also found to be identical in A. hirsuta. The levels of divergence between the East Asian Arabis species were similar to previously
reported levels of variation within the local *A. alpina* population, in which 130 SNPs were identified among 24 individuals (Waterson’s $\theta = 0.02\%$) [25]. If the hybridization event had facilitated chloroplast capture, the divergence between the *A. stelleri* and *A. nipponica* chloroplast genomes should have been less than their divergence from *A. flagellosa*. However, the divergence between the potential hybrid-origin species (*A. stelleri* and *A. nipponica*: 0.068 to 0.085) was similar to their divergence from *A. flagellosa* (0.056 to 0.086). Although the level of divergence was too low to make reliable comparisons, it is possible that *A. stelleri* and *A. nipponica* originated from independent hybridization events or the introgression process may still be ongoing.

![Figure 2](image_url)

**Figure 2.** Chloroplast genome-based phylogenetic trees of *Arabis* species. The neighbor-joining trees were constructed using both (A) whole chloroplast genomes and (B) synonymous divergence from concatenated CDS. Numbers beside the nodes indicate bootstrap probabilities (%). Scale bars are shown at the bottom left of each tree.

**Table 2.** Divergence between species.

| Compared Species       | # of Differences | Divergence (%: Ks with JC Correction) |
|------------------------|------------------|---------------------------------------|
| *Draba nemorosa* vs.   | 4475             | 2.976                                 |
| *Arabis alpina*        |                  |                                       |
| *Draba nemorosa* vs.   | 4219             | 2.801                                 |
| *Arabis hirsuta*       |                  |                                       |
| *Draba nemorosa* vs.   | 4262             | 2.765                                 |
| *Arabis flagellosa*    |                  |                                       |
| *Draba nemorosa* vs.   | 4171             | 2.771                                 |
| *Arabis stelleri*      |                  |                                       |
| *Draba nemorosa* vs.   | 4150             | 2.757                                 |
| *Arabis nipponica* JO23|                  |                                       |
| *Draba nemorosa* vs.   | 4131             | 2.745                                 |
| *Arabis nipponica* (Midori) vs. | 4131 | 2.745 |
| *Arabis alpina* vs.    | 3566             | 2.366                                 |
| *Arabis hirsuta*       |                  |                                       |
| *Arabis alpina* vs.    | 3571             | 2.371                                 |
| *Arabis flagellosa*    |                  |                                       |
| *Arabis alpina* vs.    | 3565             | 2.366                                 |
| *Arabis stelleri*      |                  |                                       |
| *Arabis alpina* vs.    | 3564             | 2.366                                 |
| *Arabis nipponica* JO23vs. | 3547 | 2.355 |
| *Arabis alpina* vs.    | 3547             | 2.355                                 |
| *Arabis flagellosa*    |                  |                                       |
| *Arabis alpina* vs.    | 1245             | 0.815                                 |
| *Arabis hirsuta*       |                  |                                       |
| *Arabis hirsuta* vs.   | 1253             | 0.82                                  |
| *Arabis stelleri*      |                  |                                       |
| *Arabis hirsuta* vs.   | 1234             | 0.808                                 |
| *Arabis nipponica* JO23vs. | 1234 | 0.808 |
| *Arabis hirsuta* vs.   | 1214             | 0.795                                 |
| *Arabis nipponica* JO23vs. | 1214 | 0.795 |
| *Arabis flagellosa* vs. | 132             | 0.086                                 |
| *Arabis stelleri* vs.  | 111              | 0.072                                 |
| *Arabis flagellosa* vs. | 86              | 0.056                                 |
| *Arabis stelleri* vs.  | 130              | 0.085                                 |
| *Arabis nipponica* JO23vs. | 104 | 0.068 |
| *Arabis stelleri* vs.  | 104              | 0.068                                 |
| *Arabis nipponica* JO23vs. | 55  | 0.036 |
should have been less than their divergence from *A. flagellosa*. However, the divergence between the potential hybrid-origin species (*A. stelleri* and *A. nipponica*: 0.068 to 0.085) was similar to their divergence from *A. flagellosa* (0.056 to 0.086). Although the level of divergence was too low to make reliable comparisons, it is possible that *A. stelleri* and *A. nipponica* originated from independent hybridization events or the introgression process may still be ongoing.

### 2.3. Distribution of Simple Sequence Repeats in the Chloroplast Genomes

Because the extremely low divergence among the East Asian *Arabis* species made it difficult to resolve their evolutionary relationships, other highly variable markers were needed. Therefore, simple sequence repeat (SSR) regions throughout the chloroplast genome were assessed for their ability to provide high-resolution species definition. A total of 74 mono-nucleotide, 22 di-nucleotide, and two tri-nucleotide repeat regions of ≥10 base pairs in length were identified (see Table 3). However, these repeat regions were still unable to completely resolve the relationships of the East Asian *Arabis* species. Fifty of the 98 SSRs exhibited no variation among the East Asian *Arabis* species, while only 29 SSRs exhibited species-specific variation, including nine in *A. flagellosa*, 15 in *A. stelleri*, four in *A. nipponica* strain JO23, and one in *A. nipponica* strain Midori. Five of the SSRs were shared by the two *A. nipponica* strains, which suggests that they were also species-specific. Although the two *A. nipponica* strains were similar to each other, *A. flagellosa*, *A. stelleri*, and *A. nipponica* differ to a similar degree in terms of variable SSRs, which suggests that the occurrence of chloroplast capture would be independent or still ongoing. This was suggested by the patterns of nucleotide substitutions.

**Figure 3.** Alignment of the seven chloroplast genomes. VISTA-based identity plots of chloroplast genomes from six *Arabis* species and *Draba nemorosa* are compared to *A. nipponica* strain Midori. Arrows above the alignment indicate genes and their orientation. The names of genes ≥500 bp in length are also shown. A 70% identity cut-off was used for making the plots, and the Y-axis represents percent identity (50–100%), while the X-axis represents the location in the chloroplast genome. The blue and pink regions indicate genes and conserved noncoding sequences, respectively.

2.3. Distribution of Simple Sequence Repeats in the Chloroplast Genomes

Because the extremely low divergence among the East Asian *Arabis* species made it difficult to resolve their evolutionary relationships, other highly variable markers were needed. Therefore, simple sequence repeat (SSR) regions throughout the chloroplast genome were assessed for their ability to provide high-resolution species definition. A total of 74 mono-nucleotide, 22 di-nucleotide, and two tri-nucleotide repeat regions of ≥10 base pairs in length were identified (see Table 3). However, these repeat regions were still unable to completely resolve the relationships of the East Asian *Arabis* species. Fifty of the 98 SSRs exhibited no variation among the East Asian *Arabis* species, while only 29 SSRs exhibited species-specific variation, including nine in *A. flagellosa*, 15 in *A. stelleri*, four in *A. nipponica* strain JO23, and one in *A. nipponica* strain Midori. Five of the SSRs were shared by the two *A. nipponica* strains, which suggests that they were also species-specific. Although the two *A. nipponica* strains were similar to each other, *A. flagellosa*, *A. stelleri*, and *A. nipponica* differ to a similar degree in terms of variable SSRs, which suggests that the occurrence of chloroplast capture would be independent or still ongoing. This was suggested by the patterns of nucleotide substitutions.
Table 3. Simple sequence repeats (SSRs) in *Arabis* chloroplast genome.

| Position in *Arabis* (Midori) Genome | UNIT       | A. nipponica | A. stelleri | A. flagellosa | A. hirsuta | A. alpina |
|--------------------------------------|------------|--------------|-------------|---------------|------------|-----------|
| from to                             | Midori     | JO23         | from to     | Midori       | JO23       | from to   | Midori | JO23 | from to   | Midori | JO23 | from to   | Midori | JO23 | from to   | Midori | JO23 | from to   | Midori | JO23 |
| 287 to 318                         | AT         | 16           | 15          | 15            | 12         | 13 with 2 mutations | 29 bp with several mutations |
| 1922 to 1932                        | A          | 11           | 11          | 9             | 12         | 11         | 9        |
| 3029 to 3038                        | T          | 10           | 9           | 10            | 9          | 7          | 7        |
| 4258 to 4270                        | T          | 13           | 18          | 18            | 17         | 13         | 13       |
| 7713 to 7727                        | T          | 15           | 15          | 15            | 15         | 12         | 11       |
| 7729 to 7738                        | A          | 10           | 10          | 10            | 9          | 10         | 10       |
| 8203 to 8216                        | TA         | 7            | 6           | 7             | 7          | 6          | 6        |
| 8273 to 8282                        | TA         | 5            | 5           | 5             | 5          | 5          | 5        |
| 8289 to 8302                        | TA         | 7            | 7           | 6             | 7          | 8          | 6        |
| 8321 to 8330                        | TA         | 5            | 5           | 4             | 5          | deletion  |          |
| 9677 to 9690                        | T          | 14           | 14          | T4GT10        | 15         | 14         | 14       |
| 11,660 to 11,669                    | A          | 10           | 10          | 9             | 10         | 7          | 7        |
| 12,406 to 12,414                    | T          | 9            | 9           | 10            | 10         | T3AT6     | T3AT6    |
| 13,010 to 13,018                     | T          | 9            | 9           | 10            | 9          | T7AT2     | T7AT2    |
| 15,777 to 15,786                    | T          | 10           | 10          | 10            | 10         | 10         | 10       |
| 17,261 to 17,281                    | T          | 11           | 11          | 11            | 11         | 12         | 9        |
| 22,549 to 22,558                    | TA         | 5            | 5           | 5             | 5          | 5          | 5        |
| 30,351 to 30,355                    | AT         | 5            | 5           | 5             | 5          | 5          | 5        |
| 35,538 to 35,555                    | AT         | 9            | 9           | 9             | 9          | 3          | deletion |
| 45,018 to 45,022                    | T          | 15           | 13          | 11            | 13         | 18nt      | 10nt     |
| 46,652 to 46,658                    | T          | 15           | 15          | 12            | 12         | 12         | 12       |
| 48,987 to 48,995                    | T          | 16           | 16          | 16            | 16         | 15         | 15       |
| 53,088 to 53,097                    | T          | 10           | 10          | 10            | 10         | 2nt shorter | 2nt shorter |
| 54,805 to 54,812                    | TA         | 5            | 5           | 5             | 5          | 19nt      | deletion |
| 59,933 to 59,942                    | TA         | 9            | 9           | 9             | 9          | 11         | 11       |
| 63,472 to 63,478                    | AT         | 5            | 5           | 5             | 5          | 5          | 5        |
| 64,283 to 64,290                    | TA         | 7            | 7           | 6             | 7          | 8          | 6        |
| 65,636 to 65,645                    | T          | 13           | 13          | 13            | 13         | 8          | C2TGC7   |
| 66,253 to 66,262                    | AT         | 5            | 5           | 5             | 5          | 5          | 5        |
| 66,851 to 66,864                    | A          | 14           | 14          | 14            | 14         | 17         | 12       |
| 68,965 to 69,977                    | T          | 13           | 13          | 13            | 13         | 11         | 11       |
| 69,965 to 69,975                    | T          | 11           | 11          | 12            | 11         | 11         | 8        |
| 75,328 to 75,340                    | A          | 13           | 14          | 14            | 13         | 19         | 14       |
### Table 3. Cont.

| Position in A. nipponica (Midori) Genome | UNIT | A. nipponica | A. stelleri | A. flagellosa | A. hirsuta | A. alpina |
|----------------------------------------|------|--------------|-------------|--------------|-----------|-----------|
| 76,614                                 | T    | 13           | 13          | 13           | 13        | 13        |
| 78,154                                 | TTG  | 5            | 3           | 3            | 4         | 2         |
| 80,484                                 | A    | 10           | 11          | 10           | 10        | 9         |
| 81,019                                 | T    | 17           | 17          | 17           | 17        | 17        |
| 81,178                                 | T    | 14           | 14          | 14           | 14        | 18        |
| 82,568                                 | A    | 11           | 10          | 9            | 10        | 9         |
| 83,492                                 | TA   | 5            | 5           | 5            | 5         | 4         |
| 83,506                                 | TA   | 5            | 5           | 5            | 5         | 4         |
| 85,147                                 | A    | 11           | 10          | 11           | 11        | 10        |
| 93,329                                 | AT   | 5            | 5           | 5            | 5         | 7         |
| 93,333                                 | T    | 12           | 11          | 13           | 13        | T2(AT)4T7 |
| 93,335                                 | TA   | 5            | 5           | 5            | T2TGTGA   |
| 93,335                                 | AT   | 5            | 5           | 5            | 5         | 4         |
| 93,335                                 | T    | 8            | 8           | 10           | 8         | 7         |
| 93,335                                 | A    | A7CA2        | A7CA2       | 10           | A7CA2     | A7CA2     |
| 107,287                                | T    | 5            | 5           | 5            | 5         | 7         |
| 107,287                                | A    | 10           | 10          | 10           | 10        | 10        |
| 107,287                                | T    | 12           | 11          | 13           | 13        | T2(AT)4T7 |
| 111,490                                | TA   | 5            | 5           | 5            | 5         | 4         |
| 111,490                                | AT   | 5            | 5           | 5            | 5         | 4         |
| 111,665                                | T    | 8            | 8           | 10           | 8         | 7         |
| 111,665                                | A    | A7CA2        | A7CA2       | 10           | A7CA2     | A7CA2     |
| 112,472                                | A    | 10           | 10          | 10           | 10        | 10        |
| 116,636                                | T    | 10           | 9           | 10           | 11        | T7AT3     |
| 123,173                                | T    | 12           | 12          | 12           | 12        | 12        |
| 123,285                                | T    | 10           | 10          | 10           | 10        | 10        |
| 123,884                                | T    | 10           | 10          | 10           | 10        | 10        |
| 123,975                                | A    | 13           | 13          | 13           | 13        | 13        |
| 124,356                                | TA   | 5            | 5           | 5            | 5         | 4         |
| 124,874                                | TA   | 13           | 13          | 13           | 13        | 13        |
| 125,029                                | A    | 13           | 13          | 13           | 13        | 13        |
| 126,052                                | T    | 15           | 15          | 15           | 15        | 15        |
| 126,087                                | T    | 11           | 11          | 11           | 11        | 11        |
| 126,117                                | A    | 12           | 12          | 12           | 12        | 12        |
| 126,952                                | T    | 11           | 11          | 11           | 11        | T8CT2     |
| 127,241                                | A    | 12           | 12          | 12           | 12        | 12        |

#### 2.4. Mitochondrial Genome Analysis

Chloroplast capture could have originated from hybridization events that also affected other cytoplasmic genomes. Due to this, variation in the mitochondrial genome sequences was analyzed. Mapping next-generation sequencing (NGS) reads to the *Eruca vesicaria* mitochondrial genome revealed that 29 sites with five or more mapped reads varied among the *A. nipponica* strain Midori, *A. flagellosa*, and *A. hirsuta* (see Table 4). Twenty-eight of the sites were conserved among *A. nipponica* and *A. flagellosa*. One site was specific to *A. nipponica* and provided 100% support for the relationship between *A. nipponica* and *A. flagellosa*. Even though reliability decreased, 123 of 125 sites with two or more reads (98.4%) also supported the similarity of the *A. nipponica* and *A. flagellosa* mitochondrial genomes. These findings suggest that the hybridization history of the species affects both the chloroplast and the mitochondrial genomes similarly.

### Table 4. Nucleotide variation in the mitochondrial genome of *Arabis* species.

| Number of Mapped Reads | 5 and More | 4 and More | 3 and More | 2 and More |
|------------------------|------------|------------|------------|------------|
| Number of variable sites Total | 29 | 46 | 74 | 129 |
| Specific to | | | | |
| *A. nipponica* | 1 | 1 | 4 | 12 |
| *A. flagellosa* | 0 | 0 | 0 | 3 |
| *A. hirsuta* | 14 | 25 | 35 | 62 |
| Shared with | | | | |
| *A. flagellosa* and *A. nipponica* | 14 | 19 | 31 | 46 |
| *A. nipponica* and *A. hirsuta* | 0 | 0 | 1 | 1 |
| *A. flagellosa* and *A. hirsuta* | 0 | 0 | 1 | 1 |
| other type | 0 | 1 | 2 | 4 |
3. Discussion

Chloroplast capture results in the incongruence of chloroplast and nuclear phylogenies, which has been reported in many plant taxa and is considered common among plants [17,18,26–37]. Furthermore, it is possible that the introgression of chloroplast genomes occurs more frequently than that of nuclear genomes as a result of uniparental inheritance, lack of recombination, and low selective constraint [38–40]. Chloroplast capture could occur by using several factors including sampling error, convergence, evolutionary rate heterogeneity, wrong lineage sorting, and hybridization/introgression [17]. Introgression-induced chloroplast capture occurred through hybridization between distant but compatible species, which was followed by backcrossing with pollen donor species [41,42].

East Asian Arabis species have previously been reported to show evidence of chloroplast capture [14,16]. More specifically, detailed phylogenetic analyses of nuclear and chloroplast marker genes has suggested that A. nipponica, A. stelleri, and A. takedimana originated from the hybridization of A. hirsuta (or A. sagittata) and East Asian Arabis species (close to A. serrata, A. paniculata, and A. flagellosa), which act as paternal and maternal parents, respectively [14,16]. In the present study, comparing the whole chloroplast genomes of four plants from three East Asian Arabis species (two A. nipponica, one each of A. stelleri, and A. flagellosa) revealed genome-wide similarities that indicated chloroplast capture by A. nipponica and A. stelleri. The study also compared the species’ partial mitochondrial genomes, which indicated a closer relationship between A. nipponica and A. flagellosa than between the former and European A. hirsuta. This suggested that A. nipponica also has a history of mitochondrial capture. This is not surprising, because hybridization and backcrossing could have similar effects on both organellar genomes. Also, cyto-nuclear incompatibility caused by a mitochondrial genome could lead cytoplasmic replacement to exhibit chloroplast capture [17,41,42]. The pattern of variation in the mitochondrial genomes suggested that both the chloroplast and mitochondrial genomes were co-transmitted during the evolutionary history of East Asian Arabis species. Future research should focus on the process of chloroplast (organellar) capture. Simple backcrossing could show the mechanisms of cytoplasm replacement and could produce results in as few as a hundred generations under certain conditions [42].

In the present study, the divergence between the genomes of hybrid-origin species and putative pollen-donor species was similar to the divergence observed within species, which suggests that the hybridization event was relatively recent. Nuclear genome markers are needed to estimate the proportion of parental genome fragments in the current nuclear genome of A. nipponica.

4. Materials and Methods

4.1. Plant Materials

Arabis nipponica (A. hirsuta var. nipponica, sampled from Midori, Gifu Prefecture, Japan), A. flagellosa (sampled from Kifune, Kyoto Prefecture, Japan), and A. hirsuta (strain Brno from Ulm Botanical Garden, Germany) were used in the present study.

4.2. DNA Isolation, NGS Sequencing, and Genome Assembly

Chloroplasts were isolated from A. hirsuta and A. nipponica as described in Okegawa and Motohashi [43]. DNA was isolated from the chloroplasts using the DNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA), while the total DNA was isolated from leaves of A. flagellosa. NGS libraries were constructed using the Nextera DNA Sample Preparation Kit (Illumina, San Diego, CA, USA) and sequenced as single-ended reads using the NextSeq500 platform (Illumina). About 2 Gb (1.4 Gb, 12 M clean reads) of sequences were obtained for A. flagellosa (43 Mb mapped reads, 282.69× coverage). Additionally, 400 Mb (300 Mb, 2.5 M clean reads) were obtained for both A. hirsuta (64 Mb mapped reads, 417.17× coverage) and A. nipponica (72 Mb mapped reads, 455.87× coverage). The generated reads were assembled using velvet 1.2.10 [44] and assembled into complete chloroplast genomes by mapping to previously published whole chloroplast genome sequences. Sequence gaps were
resolved using Sanger sequencing. Genes were annotated using DOGMA [45] and BLAST. The newly constructed chloroplast genomes were deposited in the DDBJ database under the accession numbers LC361349-51. Finally, the circular chloroplast genome maps were drawn using OGDRAW [46].

4.3. Molecular Evolutionary Analyses

The whole chloroplast genome sequences of *A. nipponica* (strain JO23: AP009369), *A. stelleri* (KY126841) [23], *A. alpina* (HF934132) [25], and *D. nemorosa* (strain JO21: AP009373) in the GenBank were also used. Whole chloroplast sequences were aligned in order to construct neighbor-joining trees with Jukes and Cantor distances. The sequences of 77 known functional genes were linked in a series after excluding initiation and stop codons and were then used for phylogenetic analyses along with sequences from the related clade species *Brassica oleracea* (KR233156) [47], *B. rapa* (DQ231548), *Eutrema salsugineum* (KR584659) [48], *Raphanus sativus* (KJ716483) [49], *Scherenkiella parvula* (KT222186) [48], *Sinapis arvensis* (KU050690), and *Thlaspi arvense* (KX886351) [21] using *A. thaliana* (AP000423) [50] as an outgroup. The synonymous divergence of the concatenated CDS was estimated using the Nei and Gojobori method. All phylogenetic analyses were performed using MEGA 7.0 [51]. Levels of divergence throughout the chloroplast genome were visualized using mVISTA [52] with Shuffle-LAGAN alignment [53].

4.4. Mapping NGS Reads to Mitochondrial Genome Sequences

Because the chloroplast isolation method used in the present study did not completely exclude mitochondria, about 1% of the sequence reads were derived from mitochondrial genomes. Although this proportion is too low to be useful for assembling whole mitochondrial genomes, the reads were nevertheless mapped to the mitochondrial genome of *Eruca vesicaria* (KF442616) [54] in order to measure mitochondrial genome divergence. Regions with at least five mapped reads were used for the analysis.

Acknowledgments: This study was supported in part by JSPS KAKENHI Grant Number 17K19361 and Grants-in-Aid from MEXT-Supported Program for the Strategic Research Foundation at Private Universities (S1511023) to Akira Kawabe.

Author Contributions: Akira Kawabe designed the study. Hiroaki Nukii and Hazuka Y. Furihata performed the experiments. Akira Kawabe and Hazuka Y. Furihata analyzed the data. Akira Kawabe wrote the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Al-Shehbaz, I.A. Transfer of most North American species of Arabis to Boechera (Brassicaceae). *Novon* 2003, 13, 381–391. [CrossRef]
2. O’Kane, S.L.; Al-Shehbaz, I.A. Phylogenetic Position and Generic Limits of Arabidopsis (Brassicaceae) Based on Sequences of Nuclear Ribosomal DNA. *Ann. Mo. Bot. Gard.* 2003, 90, 603–612. [CrossRef]
3. Al-Shehbaz, I.A.; Beilstein, M.A.; Kellogg, E.A. Systematics and phylogeny of the Brassicaceae (Cruciferae): An overview. *Plant Syst. Evol.* 2006, 259, 89–120. [CrossRef]
4. Al-Shehbaz, I.A.; German, D.A.; Karl, R.; Ingrid, J.T.; Koch, M.A. Nomenclatural adjustments in the tribe Arabideae (Brassicaceae). *Plant Div. Evol.* 2011, 129, 71–76. [CrossRef]
5. Koch, M.A.; Karl, R.; German, D.A.; Al-Shehbaz, I.A. Systematics, taxonomy and biogeography of three new Asian genera of Brassicaceae tribe Arabideae: An ancient distribution circle around the Asian high mountains. *Taxon* 2012, 61, 955–969.
6. Kiefer, M.; Schmickl, R.; German, D.A.; Mandáková, T.; Lysak, M.A.; Al-Shehbaz, I.A.; Franzke, A.; Mummenhoff, K.; Stamatakis, A.; Koch, M.A. BrassiBase: Introduction to a novel knowledge database on Brassicaceae evolution. *Plant Cell Physiol.* 2014, 55, e3. [CrossRef] [PubMed]
7. Ansell, S.W.; Grundmann, M.; Russell, S.J.; Schneider, H.; Vogel, J.C. Genetic discontinuity, breeding-system change and population history of Arabis alpina in the Italian Peninsula and adjacent Alps. *Mol Ecol.* 2008, 17, 2245–2257. [CrossRef] [PubMed]

8. Bergonzzi, S.; Albani, M.C.; van der Loren van Themaat, E.; Nordström, K.J.; Wang, R.; Schneeberger, K.; Moerland, P.D.; Coupland, G. Mechanisms of age-dependent response to winter temperature in perennial flowering of Arabis alpina. *Science* 2013, 340, 1094–1097. [CrossRef] [PubMed]

9. Karl, R.; Koch, M.A. A world-wide perspective on crucifer speciation and evolution: Phylogenetics, biogeography and trait evolution in tribe Arabidaceae. *Ann. Bot.* 2013, 112, 983–1001. [CrossRef] [PubMed]

10. Toräng, P.; Vikström, L.; Wunder, J.; Wötzel, S.; Coupland, G.; Ågren, J. Evolution of the selfing syndrome: Anther orientation and herkogamy together determine reproductive assurance in a self-compatible plant. *Evolution* 2017, 71, 2206–2218. [CrossRef] [PubMed]

11. Heidel, A.J.; Kiefer, C.; Coupland, G.; Rose, L.E. Pinpointing genes underlying annual/perennial transitions with comparative genomics. *BMC Genom.* 2016, 17, 921. [CrossRef] [PubMed]

12. Willing, E.M.; Rawat, V.; Mandáková, T.; Maumus, F.; James, G.V.; Nordström, K.J.; Becker, C.; Warthmann, N.; Chica, C.; Szarzynska, B.; et al. Genome expansion of Arabis alpina linked with retrotransposition and reduced symmetric DNA methylation. *Nat. Plants* 2015, 1, 14023. [CrossRef] [PubMed]

13. Jiao, W.B.; Accinelli, G.G.; Hartwig, B.; Kiefer, C.; Baker, D.; Severing, E.; Willing, E.M.; Piednoel, M.; Woetzel, S.; Madrid-Herrero, E.; et al. Improving and correcting the contiguity of long-read genome assemblies of three plant species using optical mapping and chromosome conformation capture data. *Genome Res.* 2017, 27, 778–786. [CrossRef] [PubMed]

14. Koch, M.A.; Karl, R.; Kiefer, C.; Al-Shehbaz, I.A. Colonizing the American continent: Systematics of the genus Arabis in North America (Brassicaceae). *Am. J. Bot.* 2010, 97, 1040–1057. [CrossRef] [PubMed]

15. Karl, R.; Kiefer, C.; Ansell, S.W.; Koch, M.A. Systematics and evolution of Arctic-Alpine Arabis alpina (Brassicaceae) and its closest relatives in the eastern Mediterranean. *Am. J. Bot.* 2012, 99, 778–794. [CrossRef] [PubMed]

16. Karl, R.; Koch, M.A. Phylogenetic signatures of adaptation: The Arabis hirsuta species aggregate (Brassicaceae) revisited. *Perspect. Plant Ecol. Evol. Syst.* 2014, 16, 247–264. [CrossRef]

17. Rieseberg, L.H.; Soltis, D.E. Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolut. Trends Plants* 1991, 5, 65–84.

18. Soltis, D.E.; Kuzoff, R.K. Discordance between nuclear and chloroplast phylogenies in the Heuchera group (Saxifragaceae). *Evolution* 1995, 49, 727–742. [CrossRef] [PubMed]

19. Ruhfel, B.R.; Gitzendanner, M.A.; Soltis, P.S.; Soltis, D.E.; Burleigh, J.G. From algae to angiosperms-inferring the phylogeny of green plants (Viridiplantae) from 360 plastid genomes. *BMC Evol. Biol.* 2014, 14, 23. [CrossRef] [PubMed]

20. Hohmann, N.; Wolf, E.M.; Lysak, M.A.; Koch, M.A. A Time-calibrated road map of brassicaceae species radiation and evolutionary history. *Plant Cell* 2015, 27, 2770–2784. [CrossRef] [PubMed]

21. Guo, X.; Liu, J.; Hao, G.; Zhang, L.; Mao, K.; Wang, X.; Zhang, D.; Ma, T.; Hu, Q.; Al-Shehbaz, I.A.; Koch, M.A. Plastome phylogeny and early diversification of Brassicaceae. *BMC Genom.* 2017, 18, 176. [CrossRef] [PubMed]

22. Mandáková, T.; Hloušková, P.; German, D.A.; Lysak, M.A. Monophyletic origin and evolution of the largest crucifer genomes. *Plant Physiol.* 2017, 174, 2062–2071. [CrossRef] [PubMed]

23. Raman, G.; Park, V.; Kwak, M.; Lee, B.; Park, S. Characterization of the complete chloroplast genome of Arabis stellari and comparisons with related species. *PLoS ONE* 2017, 12, e0183197. [CrossRef] [PubMed]

24. Jordan-Thaden, I.; Hase, I.; Al-Shehbaz, I.A.; Koch, M.A. Molecular phylogeny and systematics of the genus Draba (Brassicaceae) and identification of its most closely related genera. *Mol. Phylogenet. Evol.* 2010, 55, 524–540. [CrossRef] [PubMed]

25. Melodelima, C.; Lobréaux, S. Complete Arabis alpina chloroplast genome sequence and insight into its polymorphism. *Meta Gene* 2013, 1, 65–75. [CrossRef] [PubMed]

26. Acosta, M.C.; Premoli, A.C. Evidence of chloroplast capture in South American Nothofagus (subgenus Nothofagus, Nothofagaceae). *Mol. Phylogenet. Evol.* 2010, 54, 235–242. [CrossRef] [PubMed]

27. Dorado, O.; Rieseberg, L.H.; Arias, D.M. Chloroplast DNA introgression in southern California sunflowers. *Evolution* 1992, 46, 566–572. [CrossRef] [PubMed]
28. Fehrer, J.; Gemeinholzer, B.; Chrtek, J.; Bräutigam, S. Incongruent plastid and nuclear DNA phylogenies reveal ancient intergeneric hybridization in Pilosella hawkweeds (Hieracium, Cichorieae, Asteraceae). Mol. Phylogenet. Evol. 2007, 42, 347–361. [CrossRef] [PubMed]

29. Gurushidze, M.; Fritsch, R.M.; Blattner, F.R. Species-level phylogeny of Allium subgenus Melanocrommyum: Incomplete lineage sorting, hybridization and trnF gene duplication. Taxon 2010, 59, 829–840.

30. Liston, A.; Kadereit, J.W. Chloroplast DNA evidence for introgression and long distance dispersal in the desert annual Senecio fuscus (Asteraceae). Plant Syst. Evol. 1995, 197, 33–41. [CrossRef]

31. Mir, C.; Jarne, P.; Sarda, V.; Bonin, A.; Lumaret, R. Contrasting nuclear and cytoplasmic exchanges between phylogenetically distant oak species (Quercus suber L. and Q. ilex L.) in Southern France: Inferring crosses and dynamics. Plant Biol. 2009, 11, 213–226. [CrossRef] [PubMed]

32. Okegawa, Y.; Motohashi, K. Chloroplastic thioredoxin m functions as a major regulator of Calvin cycle enzymes during photosynthesis in vivo. Plant J. 2015, 85, 10–21. [CrossRef] [PubMed]

33. Schilling, E.E.; Panero, J.K. Phylogenetic reticulation in subtribe Helianthinae. Am. J. Bot. 1996, 83, 939–948. [CrossRef]

34. Wolfe, A.D.; Elisens, W.J. Evidence of chloroplast capture and pollen-mediated gene flow in Penstemon sect. Peltanthera (Scrophulariaceae). Syst. Bot. 1995, 20, 395–412. [CrossRef]

35. Yi, T.S.; Jin, G.H.; Wen, J. Chloroplast capture and intra- and inter-continental biogeographic diversification in the Asian–New World disjunct plant genus Osmorhiza (Apiaceae). Mol. Phylogenet. Evol. 2005, 22, 285–296. [CrossRef] [PubMed]

36. Yuan, Y.W.; Olmstead, R.G. A species-level phylogenetic study of the Verbena complex (Verbenaceae) indicates two independent intergeneric chloroplast transfers. Mol. Phylogenet. Evol. 2008, 48, 23–33. [CrossRef] [PubMed]

37. Avise, J.C. Molecular Markers, Natural History and Evolution, 2nd ed.; Sinauer: Sunderland, MA, USA, 2004.

38. Rieseberg, L.H.; Wendel, J. Introgression and its consequences in plants. In Hybrid Zones and the Evolutionary Process; Harrison, R., Ed.; Oxford University Press: New York, NY, USA, 1993; pp. 70–109.

39. Martinsen, G.D.; Whitham, T.G.; Turek, R.J.; Keim, P. Hybrid populations selectively filter gene introgression between species. Evolution 2001, 55, 1325–1335. [CrossRef] [PubMed]

40. Rieseberg, L.H. The role of hybridization in evolution: Old wine in new skins. Am. J. Bot. 1995, 82, 944–953. [CrossRef]
51. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* 2016, 33, 1870–1874. [CrossRef] [PubMed]

52. Frazer, K.A.; Pachter, L.; Poliakov, A.; Rubin, E.M.; Dubchak, I. VISTA: Computational tools for comparative genomics. *Nucleic Acids Res.* 2004, 32, W273–W279. [CrossRef] [PubMed]

53. Brudno, M.; Malde, S.; Poliakov, A.; Do, C.B.; Couronne, O.; Dubchak, I.; Batzoglou, S. Glocal Alignment: Finding rearrangements during alignment. *Bioinformatics* 2003, 19, i54–i62. [CrossRef] [PubMed]

54. Wang, Y.; Chu, P.; Yang, Q.; Chang, S.; Chen, J.; Hu, M.; Guan, R. Complete mitochondrial genome of *Eruca sativa* Mill. (Garden rocket). *PLoS ONE* 2014, 9, e105748. [CrossRef] [PubMed]

© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).