

Article

Remarkable Divergence of the Sex-Linked Region between Two Wild Spinach Progenitors, *Spinacia turkestanica* and *Spinacia tetrandra*

Hongbing She, Zhaosheng Xu, Helong Zhang, Jian Wu, Xiaowu Wang , Zhiyuan Liu * and Wei Qian * 

Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing 100081, China; shehongbing_sp@163.com (H.S.); xuzhaosheng@caas.cn (Z.X.); zhanghelong@caas.cn (H.Z.); wujian@caas.cn (J.W.); wangxiaowu@caas.cn (X.W.)

* Correspondence: liuzhiyuan01@caas.cn (Z.L.); qianwei@caas.cn (W.Q.); Tel.: +86-010-62194559 (W.Q.)

Simple Summary: There remain substantial gaps in our understanding of the sex-linked region (SLR) in two wild spinach progenitors, *Spinacia turkestanica* and *S. tetrandra*, although SLR in cultivated spinach is well understood. Using 63 *Spinacia* accessions comprising 32 *S. oleracea*, 19 *S. turkestanica* and 12 *S. tetrandra*, we found that *S. oleracea* shared a similar SLR with *S. turkestanica*, while the SLR was remarkably divergent between *S. oleracea*/*S. turkestanica* and *S. tetrandra*. Additionally, the SLR increasingly expanded via accumulating repetitive sequences and was more conserved than the pseudoautosomal region (PAR) during the evolution of *S. tetrandra* to *S. oleracea*. The results obtained in this study provide a broader understanding of the evolution of SLR in *Spinacia* species.

Abstract: The sex-linked region (SLR) plays an important role in determining the sex of a plant. The SLR of the Y chromosome, composed of a 14.1-Mb inversion and a 10-Mb Y-duplication region (YDR), was deciphered in *Spinacia oleracea* previously. However, our understanding of the SLR in its wild relatives, *S. turkestanica* and *S. tetrandra*, remains limited. In this study, we used 63 resequencing data from the three *Spinacia* species to infer the evolution of the SLR among the *Spinacia* species. In the SLR, all the cultivated spinach and *S. turkestanica* accessions were clustered into two distinct categories with both sexes, while the *S. tetrandra* accessions of both sexes were grouped. This suggests that *S. oleracea* shared a similar SLR with *S. turkestanica*, but not with *S. tetrandra*, which was further confirmed based on the population structure and principal component analysis. Furthermore, we identified 3910 fully sex-linked SNPs in *S. oleracea* and 92.82% of them were available in *S. turkestanica*, while none of the SNPs were adopted in *S. tetrandra*. Genome coverage in males and females supported the hypothesis that the YDR increasingly expanded during its evolution. Otherwise, we identified 13 sex-linked transposable element insertion polymorphisms within the inversion in both *S. oleracea* and *S. turkestanica*, demonstrating that the transposable element insertions might have occurred before the recombination suppression event of the inversion. The SLR was conserved compared with the pseudoautosomal region given that the genetic hitchhiking process occurred in the SLR during its evolution. Our findings will significantly advance our understanding of the characteristics and evolution of the SLR in *Spinacia* species.

Keywords: Y chromosome; sex-linked region (SLR); evolution; wild spinach progenitors



Citation: She, H.; Xu, Z.; Zhang, H.; Wu, J.; Wang, X.; Liu, Z.; Qian, W. Remarkable Divergence of the Sex-Linked Region between Two Wild Spinach Progenitors, *Spinacia turkestanica* and *Spinacia tetrandra*. *Biology* **2022**, *11*, 1138. <https://doi.org/10.3390/biology11081138>

Academic Editor: Valeria Terzi

Received: 13 June 2022

Accepted: 20 July 2022

Published: 29 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The origin and evolution of dioecy has been one of the most attractive topics for researchers [1]. Dioecy is rare in flowering plants, constituting approximately 6% of angiosperm species [2]. Sex chromosomes that determine the genders of individuals independently originated multiple times from autosomes [3]. Theoretical studies have predicted that recombination suppression is an indispensable event that occurs during the evolution of sex chromosomes, especially in the sex-linked region (SLR) [4,5]. The initial SLR might

be small, but it increasingly expands by accumulating abundant repetitive sequences owing to non-recombination [6]. Recently, many SLRs have been identified [7–11], providing a broader understanding of the characteristics and the evolution of sex chromosomes.

Spinach (*Spinacia oleracea* L.) is an important vegetable crop in the Amaranthaceae family and has been cultivated for more than 2000 years [12]. Spinach is a dioecious species with an XY sex-determining system and is an ideal vegetable crop for investigating sex determination and sex chromosome evolution [13]. Quite a few sex-linked markers [13–16] and male-determining bacterial artificial chromosomes (BAC) [17] have been reported in spinach, which provides a basis for further identifying SLRs. Additionally, Yu et al. (2021) identified a non-recombining region with 39 bin markers co-segregating with sex, which are located at 45.2 cM of LG1, by constructing two high-density genetic maps. The region contains accumulated abundant repetitive sequences, and its X counterpart is estimated to be approximately 18.4 Mb [18]. Previously, we identified the X and Y haplotypes based on homozygous female (XX) and male (YY) genotypes in *S. oleracea*. Furthermore, we detected a 24.1-Mb SLR on the Y chromosome, corresponding to a 13-Mb inversion on the X chromosome, and both of them exhibited low recombination. Specifically, the 13-Mb inversion located on the X chromosome ranged from 85.8 to 98.8, corresponding to 86.5–95.8 Mb (termed IV2-1) and 105.9–110.6 Mb (termed IV2-2) on the Y chromosome. Therefore, the SLR on the Y chromosome consisted of a 14.1-Mb inversion and a 10-Mb Y-duplication region (YDR) that was a male-specific region, ranging 95.8–105.9 Mb, on the Y chromosome [19]. Recently, Cai et al. (2021) reported a chromosome-scale reference genome of monoecious spinach and identified an SLR on the X chromosome, which was consistent with our previous finding [20]. Until now, no sex-determining gene has been reported in spinach. A microRNA, *sol-miR2550n*, was recently identified and defined as a male-promoting factor [21].

There are two wild spinach species, *S. turkestanica* and *S. tetrandra*, which are defined as spinach ancestors and are expected to be a potential genetic resource for spinach breeding programs [22,23]. Previous investigations have demonstrated that *S. turkestanica* is more similar to the cultivated *S. oleracea* than *S. tetrandra*, based on nuclear and chloroplast genome levels [24,25]. Similarly to cultivated spinach, the two wild relatives are dioecious [22]. However, our understanding of the sex chromosomes and SLR in the wild relatives of spinach remains limited.

To explore the evolution of the SLR in *Spinacia* species, we analyzed 63 *Spinacia* whole-genome sequence data, comprising 32 *S. oleracea*, 19 *S. turkestanica*, and 12 *S. tetrandra*. By combining them with the haplotype-resolved Y chromosome in *S. oleracea*, we assessed the divergence level of the Y chromosome, structure, and evolution of the SLR between cultivated spinach and its two wild progenitors.

2. Results

2.1. Population Structure of *S. oleracea*, *S. turkestanica*, and *S. tetrandra* in the Sex-Linked Region

To infer the evolution of the SLR among the three *Spinacia* species, we combined 29 publicly available data [20,26], and 34 newly sequenced data of *Spinacia* accessions, altogether making a set of 63 individuals, comprising 32 *S. oleracea* (12 males and 20 females), 19 *S. turkestanica* (10 males and 9 females), and 12 *S. tetrandra* (5 males and 7 females) (Table S1). The 63 representative *Spinacia* accessions came from 26 countries where all the sequenced spinach accessions existed (Figure 1a).

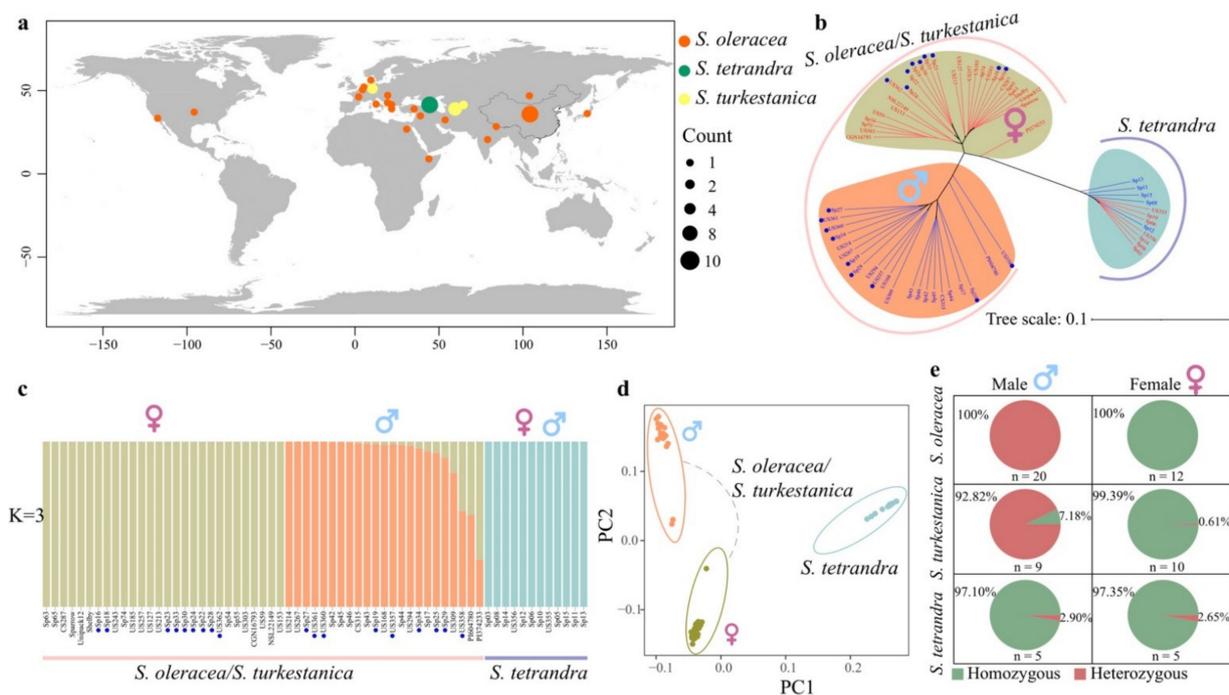


Figure 1. Population structure of the sex-linked region in *Spinacia*. (a) Geographic distribution of the 32 *Spinacia oleracea*, 19 *S. turkestanica*, and 12 *S. tetrandra* used in the study. (b) Phylogenetic tree based on 106,730 SNPs in the SLR on the Y chromosome. The accessions with red and blue colors represent females and males, respectively. The sample marked with a blue dot indicates *S. turkestanica*. (c) Population structure of the SLR in *Spinacia*. Each vertical bar represents an accession. The *S. turkestanica* accessions are marked with blue dots. (d) PCA of the SLR in the three *Spinacia*. (e) The distribution of 3910 fully sex-linked SNPs was identified using 32 *S. oleracea* in the two wild relatives.

Based on the 63 *Spinacia* accessions, we identified 1,193,638 high-quality SNPs on the Y chromosome [19], of which 106,730 SNPs appeared in the SLR, ranging from 86.58 to 110.68 Mb, and the remaining SNPs existed in the pseudoautosomal region (PAR), which excludes the SLR on the Y chromosome. Based on these SNPs within the PAR, we constructed a neighbor-joining tree of *S. oleracea* and its two wild relatives, which was consistent with the previous finding that *S. turkestanica* shares a closer genetic relationship with *S. oleracea* (Figure S1) [24]. However, a phylogenetic tree of the three *Spinacia* species based on SNPs within the SLR showed that the *S. oleracea* and *S. turkestanica* accessions were clustered into two distinct categories with males and females, while all the females and males were grouped in *S. tetrandra* (Figure 1b). Our results strongly indicate that *S. oleracea* and *S. turkestanica* have similar SLRs. The population structure and principal component analysis (PCA) of the three *Spinacia* species in the SLR support this perspective (Figure 1c,d and Figure S2). Furthermore, we identified 3910 fully sex-linked SNPs that were heterozygous genotypes in all the males and homozygous genotypes in all the females in *S. oleracea* (Table S2). *S. oleracea* shared a substantial proportion ($\geq 92.82\%$) of the fully sex-linked SNPs with *S. turkestanica*, while few SNPs *S. tetrandra* (Figure 1e and Figure S3) strongly reflected the differences in the SLRs between *S. tetrandra* and *S. oleracea/S. turkestanica*.

2.2. Evolution of the Y-Duplication Region among the *S. oleracea* and Two Wild Progenitors

YDR, a male-specific region on the Y chromosome, has been reported in many dioecious plants, and it increasingly expands due to recombination suppression [4,6]. Recently, we identified a 10-Mb YDR in *S. oleracea* [19]. The read coverage ratios of the females and males were determined to evaluate the divergence of the YDR among *S. oleracea*, *S. turkestanica*, and *S. tetrandra*. Our results showed that both *S. oleracea* and *S. turkestanica*

exhibited the YDR (Figure 2a,b). In contrast, the read coverage ratios of the females and males did not show significant differences between the PAR/IV2 and YDR in *S. tetrandra*, suggesting a smaller or absent YDR in *S. tetrandra* than those in *S. oleracea* and *S. turkestanica* (Figure 2c). Thus, to deeply survey whether *S. tetrandra* accessions have YDR, we amplified the complete genome sequence of *YY_141140.1*, located at YDR of *S. oleracea* [19], in both sexes of the *S. oleracea*, *S. turkestanica*, and *S. tetrandra* accessions. The result showed that the presence of *YY_141140.1* is restricted to males in all *Spinacia* species tested, further demonstrating that *S. tetrandra* harbored a smaller male-specific region (Figures 2d,e, S8 and S9 and Table S3) and it might have expanded stepwise during the evolution of *S. tetrandra* to *S. turkestanica*, and *S. oleracea*.

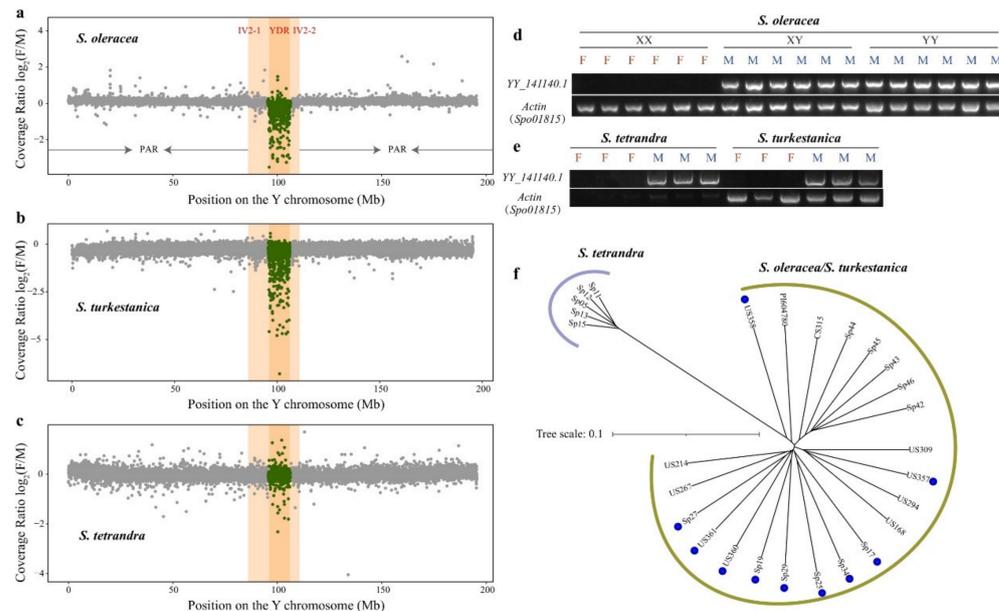


Figure 2. Characteristics of the Y-duplication region in cultivated spinach and two wild progenitors. (a) The gray dots show the sex-mapped read coverage ratio $\text{Log}_2(F/M)$ for 20 females and 12 males of *S. oleracea*. (b) The gray dots show the sex-mapped read coverage ratio $\text{Log}_2(F/M)$ for nine females and ten males of *S. turkestanica*. (c) The gray dots show the sex-mapped read coverage ratio $\text{Log}_2(F/M)$ for five females and five males of *S. tetrandra*. The read coverage per 20 kb bin was counted. YDR: Y-duplication region; IV: inversion; PAR: pseudoautosomal region. Complete male-specific conservation of *YY_141140.1* within the YDR in the genomes of (d) *S. oleracea* and (e) its two wild relatives, *S. turkestanica* and *S. tetrandra*. F: female; M: male. *Spo01815* is an actin gene in spinach. (f) Phylogenetic tree of 12 males of *S. oleracea*, 10 males of *S. turkestanica*, and 5 males of *S. tetrandra* based on the SNPs within the Y-duplication region. The accessions with blue dots are *S. turkestanica*.

A total of 16,366 high-quality SNPs were identified within the YDR-based 12 males of *S. oleracea*, 10 males of *S. turkestanica*, and 5 males of *S. tetrandra*. A neighbor-joining tree of these males in the YDR showed a remarkable divergence between *S. tetrandra* and *S. turkestanica*/*S. oleracea*. In contrast, both *S. oleracea* and *S. turkestanica* clustered, suggesting a recent divergence (Figure 2f). YDR is referred to as a vital region because many sex-determining genes are found within the region [27–29]. Among the 16,366 SNPs within the YDR, only 1.53% (251) were exonic SNPs and 2.84% (465) were intronic SNPs because of the YDR with abundant repetitive sequences (92.32%) and low gene-density (49 genes) (Figure S4) [19]. Twenty-eight out of 49 genes exhibited exonic SNPs, comprising 86 synonymous, 155 nonsynonymous, and 10 stopgains, and 30 genes bore intronic SNPs. Altogether, 37 out of 49 genes exhibited SNPs, suggesting that the remaining 12 genes might be more conservative (Tables S4 and S5 and Figure S5).

2.3. Landscapes of Transposable Element Insertion Polymorphisms in *S. oleracea* Sex Chromosomes

To deeply survey the landscape of transposable element (TE) insertion polymorphisms (TIPs) in spinach sex chromosomes, where the SLR is more likely to accumulate TEs [4,30], we identified 3616 TIPs, comprising 2053 TE insertions on the X chromosome (termed X-TIP), and 1563 TE insertions on the Y chromosome (termed Y-TIP) (Figure 3a; Tables S6 and S7). A total of 1571 (43.44%) TIPs were long terminal repeats (LTR)/Copia, which is a major LTR type in spinach [19,20] (Figure S6). For the SLR, we identified 313 TIPs, comprising 160 X-TIPs and 153 Y-TIPs, and 137 (43.76%) TIPs were LTR/Copia, which is consistent with the TIPs on the whole sex chromosomes (Figure 3b). Additionally, the SLR shared substantially higher TIP densities than the PAR (Student’s *t*-test, $p = 0.002$), while there was no significant difference between the X-TIP and Y-TIP densities in the SLR (Student’s *t*-test, $p = 0.631$). The Y-TIPs (median length = 2776 bp) appeared longer than the X-TIPs (median length = 898 bp) in the SLR (Student’s *t*-test, $p = 0.0002$) (Figure 3c).

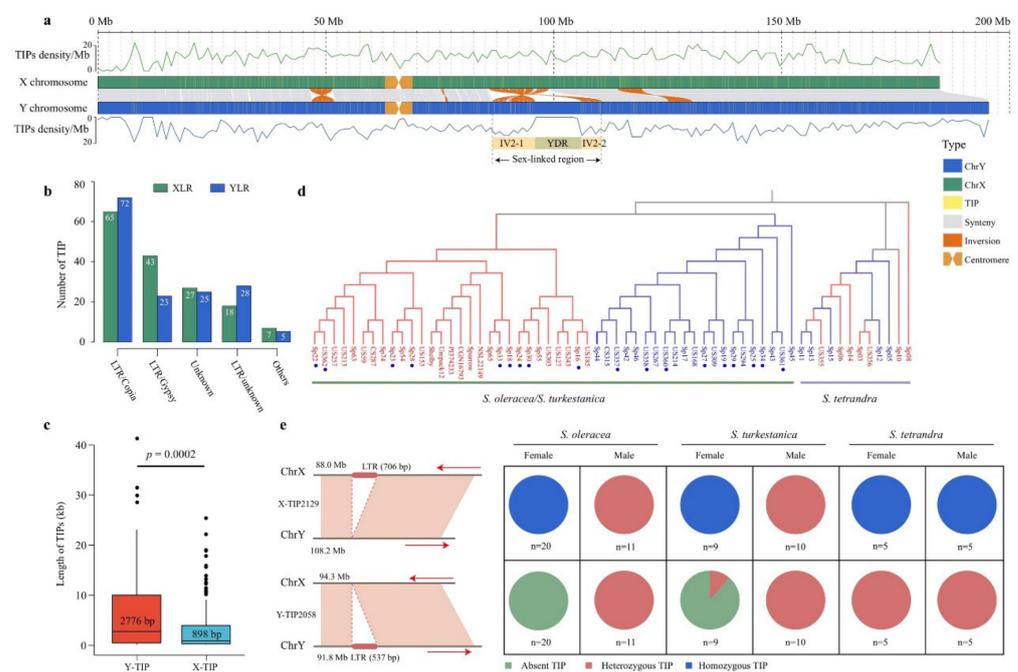


Figure 3. Identification of transposable element insertion polymorphisms (TIPs) in the sex-linked region. (a) Distribution of TIPs on the sex chromosomes in spinach. The top and bottom lines indicate TIP densities per Mb on the X and Y chromosomes, respectively. YDR: Y-duplication region; IV: inversion region. (b) Number of TIPs per TE family in the sex-linked region. XLR: X-linked region; YLR: Y-linked region. (c) Boxplot of the length of the TIPs detected in the sex-linked region. Y-TIP represents TIP on the Y-linked region. X-TIP represents TIP on the X-linked region. Significant difference analysis was performed using Student’s *t*-test. (d) Phylogenetic tree of 62 *Spinacia* accessions based on TIPs in the SLR. The accessions with red and blue colors represent females and males, respectively. The accession marked with a blue dot indicates *S. turkestanica*. (e) An example of TIP on the XLR and YLR in 62 *Spinacia* accessions. The right arrow represents the plus strand, while the left arrow represents the minus strand.

The genotypes of 63 *Spinacia* accessions based on TIPs on the sex chromosomes was determined, as described by Cai et al. (2022) [31]. One sample (PI604780) was removed because of the high missing rate. At least 1776 TIPs on the sex chromosomes were genotyped using the 62 *Spinacia* accessions, and 180 out of the 1776 TIPs were located in the SLR. Similar to the phylogenetic tree constructed using SNPs in the SLR, the phylogenetic tree of the 62 *Spinacia* accessions based on TIPs in the SLR also showed that samples from *S. oleracea* and *S. turkestanica* were grouped into two distinct clades with females and males, whereas accessions from *S. tetrandra* were not (Figure 3d). Furthermore, we identified

13 fully sex-linked TIPs in *S. oleracea* and *S. turkestanica* (Figure S7). For example, X-TIP2129 and Y-TIP2058 shared a heterozygous genotype in all the males because the males are heterogametic (XY), and X-TIP2129 shared a homozygous genotype in the homogametic (XX) females, while Y-TIP2058 was absent in almost all females in *S. oleracea* and *S. turkestanica* (Figure 3e). However, the 13 sex-linked TIPs could not cosegregate with the sexes in *S. tetrandra*, suggesting recombination suppression in the SLR might have occurred after the 13 TE insertion, thus forming a conserved TIP in *S. oleracea* and *S. turkestanica*.

2.4. Patterns of the Y Chromosome Divergence between *S. oleracea* and Its Two Wild Progenitors

All the *S. oleracea* and *S. turkestanica* accessions clustered based on variants in the SLR, while they separately clustered in the PAR (Figures 1b and S1). *S. tetrandra* did not cluster with *S. oleracea* and *S. turkestanica* in the PAR or SLR, suggesting a high genetic divergence between *S. tetrandra* and *S. oleracea*/*S. turkestanica*. To quantify the differences in the Y chromosome among the three *Spinacia* species, the population differentiation index (F_{ST}) value between the *S. oleracea*, *S. turkestanica*, and *S. tetrandra* accessions were estimated. As shown in Figure 4a, F_{ST} for the Y chromosome was low between *S. oleracea* and *S. turkestanica* (mean F_{ST} = 0.033) and strikingly high for *S. tetrandra* (mean F_{ST} = 0.273) (Table S8).

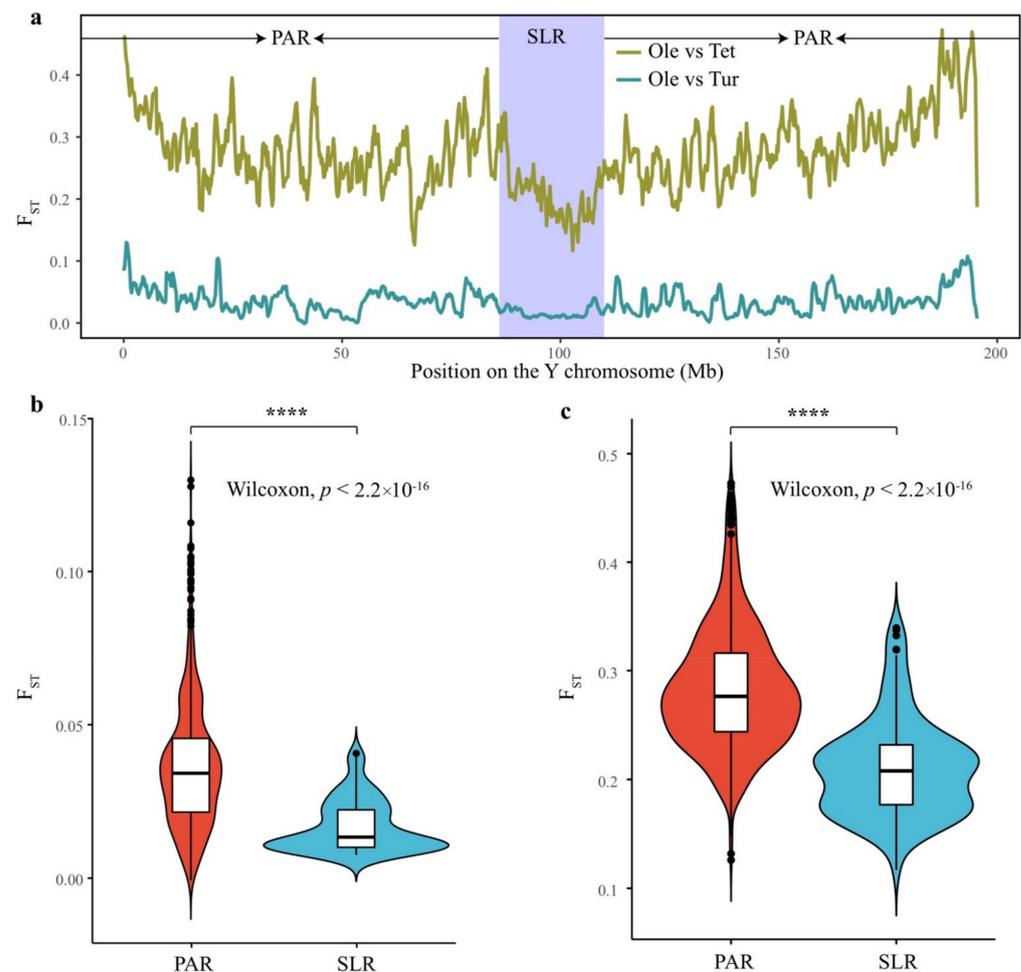


Figure 4. Pairwise F_{ST} between *S. oleracea* and its two wild progenitors on the Y chromosome. (a) The F_{ST} values between *S. oleracea* and its two wild relatives. Comparison of the F_{ST} values between the PAR and SLR in (b) *S. oleracea* vs. *S. turkestanica* and (c) *S. oleracea* vs. *S. tetrandra*. The Wilcoxon test was used for comparison between the two regions. PAR: pseudoautosomal region; SLR: sex-linked region. Ole: *S. oleracea*; Tet: *S. tetrandra*; Tur: *S. turkestanica*. The symbol **** indicates significant difference at the $p < 2.2 \times 10^{-16}$.

The SLR exhibited a significant lower F_{ST} value than the PAR for both *S. oleracea* vs. *S. turkestanica* and *S. oleracea* vs. *S. tetrandra* (Figure 4b,c). This finding indicates that the SLR was a more conserved region than the PAR. The genetic hitchhiking process that occurs in a suppressed recombination region could result in low within-population diversity in the Y-linked region [32]. Previously, we confirmed that recombination in the SLR on both the X and Y chromosomes was suppressed [19]. Thus, the low F_{ST} values in the SLR (mean F_{ST} = 0.061 and 0.209 for *S. oleracea* vs. *S. turkestanica* and *S. tetrandra*, respectively) were likely to be a property of the Y chromosome.

3. Discussion

A comparison of cultivated spinach and its wild relatives is an available approach for studying the evolution of sex chromosomes and their dynamic structure [32]. In this study, we first inferred that *S. turkestanica* exhibited an SLR similar to that of *S. oleracea*, while *S. tetrandra* did not. Sharing almost fully sex-linked SNPs/TIPs, a size similar to that of the YDR, and a low F_{ST} value in the SLR between *S. oleracea* and *S. turkestanica* strongly suggests a very recent divergence of the SLR (Figures 1e, 3e and 4). However, a substantially diverged SLR was observed between *S. tetrandra* and *S. oleracea* due to a long enough evolutionary time [24,25]. Specifically, the fully sex-linked variants of the IV2 region for *S. oleracea*/*S. turkestanica* were not available in *S. tetrandra*, indicating that the IV2 region did not exist in *S. tetrandra*. Assembling the *S. tetrandra* genome will confirm this in the future.

Based on the plant chromosome evolution process, the sex-determining locus occurred first, followed by recombination suppression close to the sex-determining locus, resulting in a stepwise expanded SLR [5,6]. The wild relatives of the *S. tetrandra* possess smaller YDRs than those of *S. oleracea* and *S. turkestanica* (Figure 2c,d), and a male-specific gene *YY_141140.1* within the YDR was identified among the three *Spinacia* species (Figure 2d,e). Thus, the conserved male-specific gene within the YDR might be referred to as a potential sex-determining gene in spinach. As described in kiwifruit (*Actinidia chinensis*), the sex-determining genes, *SyGl* and *FrBy*, are specifically present in a wide variety of male *Actinidia* species [27,29].

Recombination suppression could facilitate plant sex chromosome evolution, prevent neuter individuals (both male and female sterility) from existing in the population and maintain the stable characteristics of dioecious plants [4,6,33]. In our understanding, *S. tetrandra* is dioecious and has no neuter individuals [22]; thus, the initial recombination close to the sex-determining locus might be suppressed in *S. tetrandra*. The SLR is expected to accumulate abundant repetitive sequences due to suppressed recombination [34]. Here, we also found significantly higher TIPs in the IV2 region than in the PAR (Figure 3a). Importantly, 13 fully sex-linked TIPs within the IV2 regions of *S. oleracea* and *S. turkestanica* were obtained (Figure 3e), suggesting that the conserved TIPs occurred before the recombination suppression event in the IV2 region. Furthermore, we speculated that the nonrecombination regions in the IV2 region originated from the 14.1-Mb large-scale inversion as chromosomal rearrangements (particularly inversion), which could result in recombination suppression, such as in papaya [35], humans [36], and ostrich [37]. Our findings show that excluding the 13 fully sex-linked TIPs, the remaining TE insertions on the sex chromosomes may have originated from the inversion.

During the evolution of *S. tetrandra* to cultivated spinach, the SLR was more conserved than the PAR (Figure 4). Furthermore, the phylogenetic relationship of the *Spinacia* species in the SLR did not correlate with the geographic regions (Figure 1a,b) although a small number of accessions were collected, further indicating a conserved SLR. The SLR, however, exhibited significant divergence between *S. tetrandra* and *S. oleracea*/*S. turkestanica*; thus, we hypothesized that there might be other *Spinacia* species between *S. tetrandra* and *S. turkestanica* that we did not collect. Additionally, given the limited sequencing technology in the study, we could not decipher the detailed evolution process of the SLR in the *Spinacia* species, including whether the TEs in *S. oleracea* increased compared with those in

S. turkestanica and how the YDR was formed during evolution. High-quality genomes of the two wild progenitors will contribute to an extensive study of these in the future.

4. Conclusions

In the present study, we used 63 resequencing data from the three *Spinacia* species (cultivated and its two wild progenitors) to survey the evolution of the sex-linked region (SLR). Phylogenetic tree, population structure, and PCA analysis indicated that *S. oleracea* and *S. turkestanica* have similar SLR, while remarkably divergent SLR were found between *S. oleracea*/*S. turkestanica* and *S. tetrandra*. Additionally, *S. tetrandra* harbored a smaller male-specific region than that in *S. oleracea* and *S. turkestanica*. The SLR was conserved compared with the PAR during the evolution of sex chromosomes in *Spinacia* species. Further, we found 13 sex-linked TIPs that were inserted before the recombination suppression event on the sex chromosomes. Altogether, our results provide a basis to deeply investigate the evolution of sex chromosomes in the *Spinacia* species.

5. Materials and Methods

5.1. Plants Materials

A total of 34 individual *Spinacia* species, comprising 5 females and 5 males of *S. tetrandra*; 8 females and 6 males of *S. turkestanica*; and 5 females and 5 males of *S. oleracea*, were used for resequencing in this study. Among them, *S. tetrandra* and *S. turkestanica* accessions from the U.S Department of Agriculture (<https://www.usda.gov>; accessed on 23 May 2017) were introduced. The remaining *S. oleracea* were cultivated by the Spinach Breeding Group at the Institute of Vegetables and Flowers (IVF) of the Chinese Academy of Agricultural Sciences (CAAS). All the above plants were planted in the field at the IVF, CAAS, in spring 2018.

To collect as many spinach accessions as possible from different countries, 22 *Spinacia* accessions, comprising 15 *S. oleracea*, 5 *S. turkestanica*, and 2 *S. tetrandra* accessions, were downloaded from a previous investigation [20]. Meanwhile, seven *S. oleracea* accessions were also obtained [26]. A total of 63 *Spinacia* accessions from 26 countries were used in this study. The detailed information is summarized in Supplementary Table S1.

5.2. DNA Extraction and Whole-Genome Resequencing

Fresh leaves from each individual were collected and stored in liquid nitrogen. Genomic DNA was extracted using the DNeasy plant mini kit (Qiagen, Frankfurt, Germany). Paired-end (PE) Illumina libraries with an average insert size of 300 bp were constructed using the Illumina Genomic DNA Sample Preparation kit according to the manufacturer's instructions (Illumina, San Diego, CA, USA). Then sequencing was performed using a HiSeq 2500 instrument (Illumina, San Diego, CA, USA) to generate 150-bp PE reads, representing an average of 10× coverage.

5.3. Read Mapping and Variant Calling

Raw PE reads from 63 *Spinacia* plants were processed to remove adapters and low-quality sequences using fastp (v0.20.0) with the parameter '-q 20' [38]. Cleaned PE reads were aligned to the Sp_YY_v1 genome [19] using Burrows–Wheeler Aligner (v0.7.17) [39] with default parameters. Variants were identified using BCFtools (v1.8) [40] with the parameters '-q 20 -Q 30 -C 50' and filtered using VCFtools (v0.1.16) [40] with the parameters '-maf 0.05, -mac 4 -minQ 30 -max-missing 0.9'. The annotation of the SNPs was performed using annovar [41]. Based on the variants of the 63 *Spinacia* accessions, the genders of the accessions collected from others were determined using a fully sex-linked marker, D4.3, used in spinach [16]. Reads coverage were calculated using BedTools (v2.26.0) with parameter settings 'coverage-bga' [42].

5.4. Amplification of Gene within the YDR among the Three *Spinacia* Species

To determine whether YY_141140.1 is male-specific in the *Spinacia* species, we amplified the genes in 18 *S. oleracea* (from inbred line 10S15 [19]), six *S. turkestanica* (Sp16,

Sp18, Sp22, Sp25, Sp27, and Sp29), and six *S. tetrandra* accessions (Sp03, Sp06, Sp08, Sp11, Sp12, and Sp13). The primers of *YY_141140.1* and actin were designed using the online software Primer 3 (v0.4.0; <https://bioinfo.ut.ee/primer3-0.4.0/>; accessed on 3 March 2022). The polymerase chain reaction (PCR) was performed in a total reaction volume of 10 μ L containing 5 μ L of 2xTaq Master Mix (CoWin Biosciences, Taizhou, China), 0.25 μ L of forward and reverse primers each, respectively, 3.6 μ L of ddH₂O and 1 μ L of DNA/cDNA. The reaction was performed on a Veriti 96 Well Thermal Cycler (Applied Biosystems, Foster City, CA, USA) under the following conditions: 5 min at 94 °C, followed by 35 cycles of 30 s at 94 °C, 30 s at 60 °C, and 40 s at 72 °C. The result was assessed through electrophoresis with 1.5% agarose gels.

5.5. Population Analysis

A neighbor-joining phylogeny was constructed based on the *P* distance matrix calculated using VCF2Dis (v1.4.3) (<https://github.com/BGI-shenzhen/VCF2Dis>; accessed on 23 March 2022). PCA was performed using PLINK (v1.90b4.6; <http://pngu.mgh.harvard.edu/purcell/plink/>; accessed on 23 March 2022). The population differentiation indexes (F_{ST}) across the Y chromosome were calculated in each 1000-kb window with a step size of 200 kb using VCFtools (v0.1.16) [40]. The population structure of the SLR was analyzed with the cluster number *K* ranging from 1 to 7 using ADMIXTURE (v 1.3.0) [43]. Each *K* was run 10 times, and the correct *K* value was chosen using ADMIXTURE's cross-validation procedure.

5.6. Identification of Transposable Element Insertion Polymorphisms (TIPs)

To identify TIPs between the X and Y chromosomes in spinach, we first aligned the X chromosome to the Y chromosome from the Sp_XX_v1 and Sp_YY_v1 assemblies [19] using Mummer (v4.0.0rc1) with the parameter settings “-g 1000-c 90-l 40” [44]. The alignment block was then further filtered and one-to-one alignment was identified using delta-filter with the parameter setting “-l-i 90”. Then, the potential insertion and deletion (≥ 50 bp) relative to the Y chromosome were identified using show-diff in Mummer (v4.0.0rc1). Furthermore, the insertion and deletion sequences overlapped with the gap region in the respective sex chromosomes, and sequences with feature type “BRK” were ruled out. Meanwhile, the insertion and deletion in the syntenic regions were used in this study. Finally, the sequences of the insertion and deletion were mapped to the *S. oleracea* TE library [19] using BLASTN (v2.9.0). When the identification and coverage of the alignment was greater than 80%, then the insertion or deletion was defined as a TE insertion on the X (X-TIP) and Y (Y-TIP) chromosomes, respectively.

5.7. Genotyping of TIPs Using 63 *Spinacia* Accessions

We estimated the genotypes of each *Spinacia* accession as described by Cai et al. (2022) [31]. Specifically, we first defined each TE insertion sequence and their 1 kb upstream and downstream as the TE reference. Then, each clean PE read of the 63 *Spinacia* accessions was mapped to the TE reference using BWA-MEM (0.7.17-r1188) with parameter settings “-T 20-Y”. Next, based on the alignment position of the TE insertion, the genotype of each accession was obtained. Finally, the genotypes of TIPs with minor allele frequency (MAF) ≤ 0.05 and missing ≥ 0.08 were filtered.

5.8. Identification of Fully Sex-Linked SNPs/TIPs

For the SNPs on the Y chromosome, we defined one SNP as a fully sex-linked SNP when the genotype of the SNP was “0/1” in all males and “1/1” in all females. “0” indicated that the allele was consistent with the reference (the Y chromosome), while “1” indicated that the allele was consistent with the alternative. For the X-TIPs, one TIP was regarded as a fully sex-linked TIP when all the females appeared as homologous X-TIP, while all of the males appeared as heterozygous X-TIP. In contrast, one Y-TIP was regarded as a fully

sex-linked TIP when the TIP was absent in all the females while sharing a heterozygous Y-TIP in all the males.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biology11081138/s1>, Figure S1: Phylogenetic tree of 63 *Spinacia* accessions based on SNPs within the pseudoautosomal region (PAR); Figure S2: The cross-validation (CV) error for different *K* values in admixture analysis; Figure S3: Fully sex-linked SNPs in *S. oleracea*, *S. turkestanica*, and *S. tetrandra*; Figure S4: Annotation of SNPs within the Y-duplication region; Figure S5: Venn diagram of genes with exonic or intronic SNPs in the Y-duplication region; Figure S6: Number of detected TIPs per TE family; Figure S7: Phylogenetic tree based on 13 conserved TIPs in 62 *Spinacia* accessions; Figures S8 and S9: Original images of full amplification of *YY_141140.1*; Table S1: List of 63 *Spinacia* accessions; Table S2: Fully sex-linked SNPs in *S. oleracea*; Table S3: Primer sequences used in this study; Table S4: Exonic SNPs within the Y-duplication region; Table S5: Intronic SNPs within the Y-duplication region; Table S6: TE insertion on the ChrY chromosome; Table S7: TE insertion on the X chromosome; Table S8: F_{ST} value between *S. oleracea* and the two wild progenitors on the Y chromosome.

Author Contributions: Funding acquisition, W.Q.; Methodology, J.W., X.W., Z.L. and W.Q.; Resources, Z.X., H.Z., Z.L. and W.Q.; Writing—original draft, H.S.; Writing—review and editing, H.S. and W.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This work was performed at the Key Laboratory of Biology and Genetic Improvement of Horticultural Crops, Ministry of Agriculture, Beijing, China, and was supported by the China Agricultural Research System (CARS-23-A-17), the Chinese Academy of Agricultural Sciences Innovation Project (CAAS-ASTIP-IVFCAAS, CAAS-ZDRW202103), Beijing Joint Research Program for Germplasm Innovation and New Variety Breeding (G20220628003).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The resequencing reads used in the study have been deposited in the Genome Warehouse in the BIG Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences, under accession number CRA004067, and are publicly accessible at <http://bigd.big.ac.cn> (accessed on 4 April 2022).

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

- Henry, I.M.; Akagi, T.; Tao, R.; Comai, L. One hundred ways to invent the sexes theoretical and observed paths to dioecy in plants. *Annu. Rev.* **2018**, *69*, 553–575. [[CrossRef](#)] [[PubMed](#)]
- Renner, S.S. The relative and absolute frequencies of angiosperm sexual systems: Dioecy, monoecy, gynodioecy, and an updated online database. *Am. J. Bot.* **2014**, *101*, 1588–1596. [[CrossRef](#)] [[PubMed](#)]
- Charlesworth, B. The evolution of chromosomal sex determination and dosage compensation. *Curr. Biol.* **1996**, *6*, 149–162. [[CrossRef](#)]
- Pannell, J.R.; Gerchen, J. Sex determination: Sterility genes out of sequence. *Curr. Biol.* **2018**, *28*, R80–R83. [[CrossRef](#)] [[PubMed](#)]
- Charlesworth, D.; Charlesworth, B.; Marais, G. Steps in the evolution of heteromorphic sex chromosomes. *Heredity* **2005**, *95*, 118–128. [[CrossRef](#)]
- Ming, R.; Bendahmane, A.; Renner, S.S. Sex chromosomes in land plants. *Annu. Rev. Plant Biol.* **2011**, *62*, 485–514. [[CrossRef](#)]
- Jia, H.; Jia, H.; Cai, Q.; Wang, Y.; Zhao, H.; Yang, W.; Wang, G.; Li, Y.; Zhan, D.; Shen, Y.; et al. The red bayberry genome and genetic basis of sex determination. *Plant Biotechnol. J.* **2019**, *17*, 397–409. [[CrossRef](#)]
- Harkess, A.; Zhou, J.; Xu, C.; Bowers, J.E.; Van der Hulst, R.; Ayyampalayam, S.; Mercati, F.; Riccardi, P.; McKain, M.R.; Kakrana, A.; et al. The asparagus genome sheds light on the origin and evolution of a young Y chromosome. *Nat. Commun.* **2017**, *8*, 1279. [[CrossRef](#)]
- Yang, W.; Wang, D.; Li, Y.; Zhang, Z.; Tong, S.; Li, M.; Zhang, X.; Zhang, L.; Ren, L.; Ma, X.; et al. A general model to explain repeated turnovers of sex determination in the Salicaceae. *Mol. Biol. Evol.* **2020**, *38*, 968–980. [[CrossRef](#)]
- Almeida, P.; Proux-Wera, E.; Churcher, A.; Soler, L.; Dainat, J.; Pucholt, P.; Nordlund, J.; Martin, T.; Ronnberg-Wastljung, A.C.; Nystedt, B.; et al. Genome assembly of the basket willow, *Salix viminalis*, reveals earliest stages of sex chromosome expansion. *BMC Biol.* **2020**, *18*, 78. [[CrossRef](#)]
- Pucholt, P.; Wright, A.E.; Conze, L.L.; Mank, J.E.; Berlin, S. Recent sex chromosome divergence despite ancient dioecy in the willow *Salix viminalis*. *Mol. Biol. Evol.* **2017**, *34*, 1991–2001. [[CrossRef](#)] [[PubMed](#)]

12. Liu, Z.Y.; She, H.B.; Xu, Z.S.; Zhang, H.L.; Li, G.L.; Zhang, S.F.; Qian, W. Quantitative trait loci (QTL) analysis of leaf related traits in spinach (*Spinacia oleracea* L.). *BMC Plant Biol.* **2021**, *21*, 290. [[CrossRef](#)] [[PubMed](#)]
13. She, H.B.; Xu, Z.S.; Zhang, H.L.; Li, G.L.; Wu, J.; Wang, X.W.; Li, Y.; Liu, Z.Y.; Qian, W. Identification of a male-specific region (MSR) in *Spinacia oleracea*. *Hortic. Plant J.* **2021**, *7*, 341–346. [[CrossRef](#)]
14. Okazaki, Y.; Takahata, S.; Hirakawa, H.; Suzuki, Y.; Onodera, Y. Molecular evidence for recent divergence of X- and Y-linked gene pairs in *Spinacia oleracea* L. *PLoS ONE* **2019**, *14*, e0214949. [[CrossRef](#)]
15. Akamatus, T.; Suzuki, T.; Uchimiya, H. *Determination of Male or Female of Spinach by Using DNA Marker*; Sakata no tane KK: Yokohama, Japan, 1998.
16. Liu, D.; Qian, W.; Zhang, H.; Fan, G.; Xu, Z. Development and application of molecular markers linked with sex gene X/Y in spinach. *Hortic. Plant J.* **2015**, *42*, 1583–1590.
17. Kudoh, T.; Takahashi, M.; Osabe, T.; Toyoda, A.; Hirakawa, H.; Suzuki, Y.; Ohmido, N.; Onodera, Y. Molecular insights into the non-recombining nature of the spinach male-determining region. *Mol. Genet. Genom.* **2017**, *293*, 557–568. [[CrossRef](#)]
18. Yu, L.; Ma, X.K.; Deng, B.; Yue, J.J.; Ming, R. Construction of high-density genetic maps defined sex determination region of the Y chromosome in spinach. *Mol. Genet. Genom.* **2021**, *296*, 41–53. [[CrossRef](#)]
19. She, H.; Liu, Z.; Xu, Z.; Zhang, H.; Cheng, F.; Wang, X.; Qian, W. The female(XX) and male(Y) genomes provide insights into the sex determination mechanism in spinach. *bioRxiv* **2020**. [[CrossRef](#)]
20. Cai, X.F.; Sun, X.P.; Xu, C.X.; Sun, H.H.; Wang, X.L.; Ge, C.H.; Zhang, Z.H.; Wang, Q.X.; Fei, Z.J.; Jiao, C.; et al. Genomic analyses provide insights into spinach domestication and the genetic basis of agronomic traits. *Nat. Commun.* **2021**, *12*, 7246, Erratum in *Nat. Commun.* **2022**, *13*, 7246. [[CrossRef](#)]
21. Li, N.; Wang, Y.Y.; Wang, J.W.; Zhang, W.Q.; Meng, Z.W.; Wang, Y.S.; Zhang, Y.L.; Li, S.F.; Gao, W.J.; Deng, C.L. Identification of Sex Differentiation-Related microRNAs in Spinach Female and Male Flower. *Int. J. Mol. Sci.* **2022**, *23*, 4090. [[CrossRef](#)]
22. Ribera, A.; Bai, Y.; Wolters, A.; Treuren, R.V.; Kik, C.J.E. A review on the genetic resources, domestication and breeding history of spinach (*Spinacia oleracea* L.). *Euphytica* **2020**, *216*, 48. [[CrossRef](#)]
23. Treuren, R.; Groot, L.; Hisoriev, H.; Khassanov, F.; Farzaliyev, V.; Melyan, G.; Gabrielyan, I.; Soest, L.; Kik, C. Acquisition and regeneration of *Spinacia turkestanica* Iljin and *S. tetrandra* Steven ex M. Bieb. to improve a spinach gene bank collection. *Genet. Resour. Crop Evol.* **2020**, *67*, 549–559. [[CrossRef](#)]
24. She, H.B.; Liu, Z.Y.; Xu, Z.S.; Zhang, H.L.; Cheng, F.; Wu, J.; Wang, X.W.; Qian, W. Comparative chloroplast genome analyses of cultivated spinach and two wild progenitors shed light on the phylogenetic relationships and variation. *Sci. Rep.* **2022**, *12*, 856. [[CrossRef](#)] [[PubMed](#)]
25. Xu, C.; Jiao, C.; Sun, H.; Cai, X.; Wang, X.; Ge, C.; Zheng, Y.; Liu, W.; Sun, X.; Xu, Y.; et al. Draft genome of spinach and transcriptome diversity of 120 *Spinacia* accessions. *Nat. Commun.* **2017**, *8*, 15275. [[CrossRef](#)] [[PubMed](#)]
26. Hulse-Kemp, A.M.; Bostan, H.; Chen, S.Y.; Ashrafi, H.; Stoffel, K.; Sanseverino, W.; Li, L.Z.; Cheng, S.F.; Schatz, M.C.; Garvin, T.; et al. An anchored chromosome-scale genome assembly of spinach improves annotation and reveals extensive gene rearrangements in euasterids. *Plant Genome* **2021**, *14*, e20101. [[CrossRef](#)]
27. Akagi, T.; Pilkington, S.M.; Varkonyi-Gasic, E.; Henry, I.M.; Sugano, S.S.; Sonoda, M.; Firl, A.; McNeilage, M.A.; Douglas, M.J.; Wang, T.; et al. Two Y-chromosome-encoded genes determine sex in kiwifruit. *Nat. Plants* **2019**, *5*, 801–809. [[CrossRef](#)]
28. Akagi, T.; Henry, I.M.; Tao, R.; Comai, L. A Y-chromosome-encoded small RNA acts as a sex determinant in persimmons. *Science* **2014**, *346*, 646–650. [[CrossRef](#)]
29. Akagi, T.; Henry, I.M.; Ohtani, H.; Morimoto, T.; Beppu, K.; Kataoka, I.; Tao, R. A Y-encoded suppressor of feminization arose via lineage-specific duplication of a cytokinin response regulator in kiwifruit. *Plant Cell* **2018**, *30*, 780–795. [[CrossRef](#)]
30. Charlesworth, D. Plant contributions to our understanding of sex chromosome evolution. *New Phytol.* **2015**, *208*, 52–65. [[CrossRef](#)]
31. Cai, X.; Lin, R.M.; Liang, J.L.; King, G.J.; Wu, J.; Wang, X.W. Transposable element insertion: A hidden major source of domesticated phenotypic variation in *Brassica rapa*. *Plant Biotechnol. J.* **2022**, *20*, 1298–1310. [[CrossRef](#)]
32. Vanburen, R.; Zeng, F.; Chen, C.; Zhang, J.; Wai, C.M.; Han, J.; Aryal, R.; Gschwend, A.R.; Wang, J.; Na, J.K. Origin and domestication of papaya Yh chromosome. *Genome Res.* **2015**, *25*, 524–533. [[CrossRef](#)] [[PubMed](#)]
33. Rifkin, J.L.; Beaudry, F.E.; Humphries, Z.; Choudhury, B.I.; Barrett, S.C.; Wright, S.I. Widespread recombination suppression facilitates plant sex chromosome evolution. *Mol. Biol. Evol.* **2021**, *38*, 1018–1030. [[CrossRef](#)]
34. Renner, S.S.; Muller, N.A. Plant sex chromosomes defy evolutionary models of expanding recombination suppression and genetic degeneration. *Nat. Plants* **2021**, *7*, 392–402. [[CrossRef](#)] [[PubMed](#)]
35. Wang, J.; Na, J.K.; Yu, Q.; Gschwend, A.R.; Han, J.; Zeng, F.; Aryal, R.; VanBuren, R.; Murray, J.E.; Zhang, W.; et al. Sequencing papaya X and Yh chromosomes reveals molecular basis of incipient sex chromosome evolution. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 13710–13715. [[CrossRef](#)] [[PubMed](#)]
36. Cotter, D.J.; Brotman, S.M.; Sayres, M.A.W. Genetic diversity on the human X chromosome does not support a strict pseudoautosomal boundary. *Genetics* **2016**, *203*, 485–492. [[CrossRef](#)]
37. Yazdi, H.P.; Ellegren, H. A genetic map of Ostrich Z chromosome and the role of inversions in Avian sex chromosome evolution. *Genome Biol. Evol.* **2018**, *10*, 2049–2060. [[CrossRef](#)]
38. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **2018**, *34*, 884–890. [[CrossRef](#)]
39. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997.

40. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T.; et al. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [[CrossRef](#)]
41. Kai, W.; Li, M.; Hakon, H.J.N.A.R. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **2010**, *38*, e164.
42. Quinlan, A.R.; Hall, I.M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **2010**, *26*, 841–842. [[CrossRef](#)] [[PubMed](#)]
43. Alexander, D.H.; Novembre, J.; Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **2009**, *19*, 1655–1664. [[CrossRef](#)] [[PubMed](#)]
44. Kurtz, S.; Phillippy, A.; Delcher, A.L.; Smoot, M. Versatile and open software for comparing large genomes. *Genome Biol.* **2004**, *5*, R12. [[CrossRef](#)] [[PubMed](#)]