Please cite this article in press as: Yan et al., Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato, Molecular Plant (2024), https://doi.org/10.1016/j.molp.2023.12.019

Molecular Plant
Research article

CellPress
Partner Journal

Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato

Mengxiao Yan[1,12], Ming Li[2,3,12], Yunze Wang[1,4], Xinyi Wang[1,4], M-Hossein Moeinzadeh[5], Dora G. Quispe-Huamanquispe[6], Weijuan Fan[1], Yijie Fang[1,4], Yuqin Wang[1,4], Haozhen Nie[1], Zhangying Wang[7], Aiko Tanaka[8], Bettina Heider[9], Jan F. Kreuze[9,*], Godelieve Gheysen[6,*], Hongxia Wang[1,10,*], Martin Vingron[5,*], Ralph Bock[11,*] and Jun Yang[1,10,*]

[1]Shanghai Key Laboratory of Plant Functional Genomics and Resources, Shanghai Chenshan Botanical Garden, Shanghai 201602, China

[2]College of Life Sciences, Chongqing Normal University, Chongqing 401331, China

[3]Biotechnology and Nuclear Technology Research Institute, Sichuan Academy of Agricultural Sciences, Chengdu 610061, China

[4]College of Life Sciences, Shanghai Normal University, Shanghai 200234, China

[5]Department of Computational Molecular Biology, Max Planck Institute for Molecular Genetics, Ihnestraße 63–73, 14195 Berlin, Germany

[6]Department of Biotechnology, Ghent University, 9000 Ghent, Belgium

[7]Guangdong Provincial Key Laboratory of Crops Genetics and Improvement, Crop Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou 510640, China

[8]Graduate School of Bioagricultural Sciences, Nagoya University, Chikusa, Nagoya 464-8601, Japan

[9]International Potato Center (CIP), Lima, Peru

[10]CAS Center for Excellence of Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai 200233, China

[11]Max-Planck-Institut für Molekulare Pflanzenphysiologie, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany

[12]These authors contributed equally to this article.

***Correspondence: Jan F. Kreuze (j.kreuze@cgiar.org), Godelieve Gheysen (godelieve.gheysen@ugent.be), Hongxia Wang (hxwang@cemps.ac.cn), Martin Vingron (vingron@molgen.mpg.de), Ralph Bock (rbock@mpimp-golm.mpg.de), Jun Yang (jyang03@cemps.ac.cn)**

https://doi.org/10.1016/j.molp.2023.12.019

ABSTRACT

The hexaploid sweetpotato (*Ipomoea batatas*) is one of the most important root crops worldwide. However, its genetic origin remains controversial, and its domestication history remains unknown. In this study, we used a range of genetic evidence and a newly developed haplotype-based phylogenetic analysis to identify two probable progenitors of sweetpotato. The diploid progenitor was likely closely related to *Ipomoea aequatoriensis* and contributed the B_1 subgenome, *Ib*T-DNA2, and the lineage 1 type of chloroplast genome to sweetpotato. The tetraploid progenitor of sweetpotato was most likely *I. batatas* 4x, which donated the B_2 subgenome, *Ib*T-DNA1, and the lineage 2 type of chloroplast genome. Sweetpotato most likely originated from reciprocal crosses between the diploid and tetraploid progenitors, followed by a subsequent whole-genome duplication. In addition, we detected biased gene exchanges between the subgenomes; the rate of B_1 to B_2 subgenome conversions was nearly three times higher than that of B_2 to B_1 subgenome conversions. Our analyses revealed that genes involved in storage root formation, maintenance of genome stability, biotic resistance, sugar transport, and potassium uptake were selected during the speciation and domestication of sweetpotato. This study sheds light on the evolution of sweetpotato and paves the way for improvement of this crop.

Key words: sweetpotato, origin, gene conversion, domestication, *Ib*T-DNA, HPA

Yan M., Li M., Wang Y., Wang X., Moeinzadeh M.-H., Quispe-Huamanquispe D.G., Fan W., Fang Y., Wang Y., Nie H., Wang Z., Tanaka A., Heider B., Kreuze J.F., Gheysen G., Wang H., Vingron M., Bock R., and Yang J. (2024). Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato. Mol. Plant. **17**, 1–20.

Published by the Molecular Plant Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and CEMPS, CAS.

Please cite this article in press as: Yan et al., Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato, Molecular Plant (2024), https://doi.org/10.1016/j.molp.2023.12.019

INTRODUCTION

Global food security in the face of the growing world population is one of the key challenges of this century. Concurrently, crop production is threatened by global climate change (Godfray et al., 2010; Wheeler and Von Braun, 2013; Prosekov and Ivanova, 2018). Development of climate change–resilient crops with maximum net production is necessary to ensure food security (Dhankher and Foyer, 2018; Tian et al., 2021). As a nutritious crop with high production and adaptability to diverse environments, sweetpotato (*Ipomoea batatas*, 2n = 6x = 90) has the potential to address issues of food and nutrition security (Kwak, 2019). Sweetpotato is an important staple crop worldwide, with an annual production of ~89 million tons, and an important source of dietary calories, proteins, vitamins, and minerals (Padmaja, 2009; Food and Agriculture Organization, 2019). This important root crop has a critical role in food security, especially in developing countries (Food and Agriculture Organization, 2019), and orange-fleshed sweetpotato plays a crucial part in combating vitamin A deficiency in Africa (Kurabachew, 2015). Sweetpotato can be highly productive in the global hot and dry environment caused by climate change (Nedunchezhiyan and Ray, 2010). Continued research on sweetpotato breeding will therefore contribute to food and nutrition security.

Understanding the origin and domestication history of crops is vital for breeding and genetic engineering efforts. It is also key to evaluating strategies for genetic resource conservation that involve wild relatives. Identifying the wild/progenitor species of a given crop is often very important for breeding because wild/progenitor species typically have useful genetic traits, such as high resistance and high yield (Hajjar and Hodgkin, 2007; Kole, 2011). Through hybridization or gene editing, scientists and breeders can introduce genetic traits from the progenitor species into the crop to create better varieties (Wallace et al., 2018). For example, artificial cultivars with higher yield have been created by crossing two progenitor species of wheat, markedly increasing the lost genetic diversity of the wheat D subgenome (Das et al., 2016; Jafarzadeh et al., 2016). For decades, significant time and effort have been spent in attempts to generate artificial hexaploids from diploid and tetraploid wild relatives of sweetpotato (Nishiyama et al., 1975). However, because the progenitor species of sweetpotato are still unclear, these experiments have not achieved promising results. Revealing the origin and domestication history of sweetpotato is vital for supporting further studies of its biology, genetics, and genetic engineering and for utilization and conservation of its wild relatives.

Because of its highly heterozygous and complex hexaploid genome (Yang et al., 2017; Wu et al., 2018), the origin of cultivated sweetpotato has been extensively debated. Three polyploidization scenarios have been proposed: the autopolyploid hypothesis, the segmental allopolyploid hypothesis, and the allopolyploidy hypotheses. The autopolyploid hypothesis suggests that sweetpotato has an autopolyploid origin, with *Ipomoea trifida* as the single progenitor. This hypothesis has gained support from phylogenetic (Roullier et al., 2013; Muñoz-Rodríguez et al., 2018), genetic linkage (Ukoskit and Thompson, 1997; Kriegner et al., 2003; Cervantes-Flores et al., 2008; Zhao et al., 2013; Mollinari et al., 2020), and cytogenetic analyses (Shiotani, 1988; Shiotani and Kawase, 1989). The segmental allopolyploid hypothesis, based on an analysis of 811 conserved single-copy genes (Gao et al., 2020), proposes that sweetpotato is a segmental allopolyploid that carries three partially differentiated subgenomes originally from *I. trifida*, *Ipomoea tenuissima*, and *Ipomoea littoralis*. The allopolyploidy hypotheses are diverse and less consistent. Based on cytogenetic analysis, Nishiyama (1971) suggested that sweetpotato originated from *I. trifida* 3x, which is a hybrid between *Ipomoea* × *leucantha* and *I. littoralis*. Austin (1988) suggested that the cultivated sweetpotato was derived from a hybridization event between *I. trifida* and *Ipomoea triloba* on the basis of morphological data. Gao et al. (2011), using *Waxy* (*Wx*) intron sequence variation, suggested that sweetpotato arose via hybridization between *I. tenuissima* and *I. littoralis*. However, both cytogenetic and recent genomic analyses suggest that sweetpotato ($B_1B_1B_2B_2B_2B_2$) is composed of two subgenomes and arose from a cross between a diploid and a tetraploid progenitor (Shiotani and Kawase, 1987; Yang et al., 2017). *I. trifida* has been proposed as the diploid progenitor, whereas the tetraploid progenitor has remained a subject of debate (Yan et al., 2022). On the basis of a phylogenetic analysis of homologous haplotypes, Yan et al. (2021) suggested that the tetraploid progenitor of sweetpotato was *I. batatas* 4x. Subsequently, Muñoz-Rodríguez et al. (2022) identified *Ipomoea aequatoriensis* as the tetraploid progenitor of sweetpotato on the basis of morphological and phylogenetic analyses. However, these hypotheses fail to provide a reasonable explanation for the origins of the subgenomes, the chloroplast genome, and the *Ib* transfer DNAs (*Ib*T-DNAs) and for the genetic pattern conflict between the nuclear and chloroplast genomes of sweetpotato. The genetic origin and domestication history of sweetpotato thus remain unclear.

In this study, we used comparative studies of *Ib*T-DNA insertions and nuclear, and chloroplast genome variations, as well as a newly developed haplotype-based phylogenetic analysis (HPA), to identify the possible progenitors of sweetpotato. We clarified the two progenitors of sweetpotato and determined their contributions to sweetpotato germplasm, including nuclear subgenomes, chloroplast genomes, and *Ib*T-DNA insertions. We also identified biased gene conversion events between the subgenomes on the basis of homoelogous haplotypes. We obtained new insights into the role played by selection during the domestication of cultivated sweetpotato, and identified useful candidate genes for future breeding and genetic engineering efforts and evolutionary studies. Our results provide valuable insights into the evolution and domestication of sweetpotato, paving the way for genetic improvement of this important crop.

RESULTS

Phylogeny and population structure of sweetpotato and its wild relatives

To investigate the phylogenetic relationships among sweetpotato and its wild relatives, we analyzed 23 sweetpotato cultivars/landraces and all putative genomic donors of sweetpotato, representing a wide range of taxonomic groups, geographic distributions, and ploidy levels (Figure 1A; Supplemental Table 1). Among the

Please cite this article in press as: Yan et al., Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato, Molecular Plant (2024), https://doi.org/10.1016/j.molp.2023.12.019

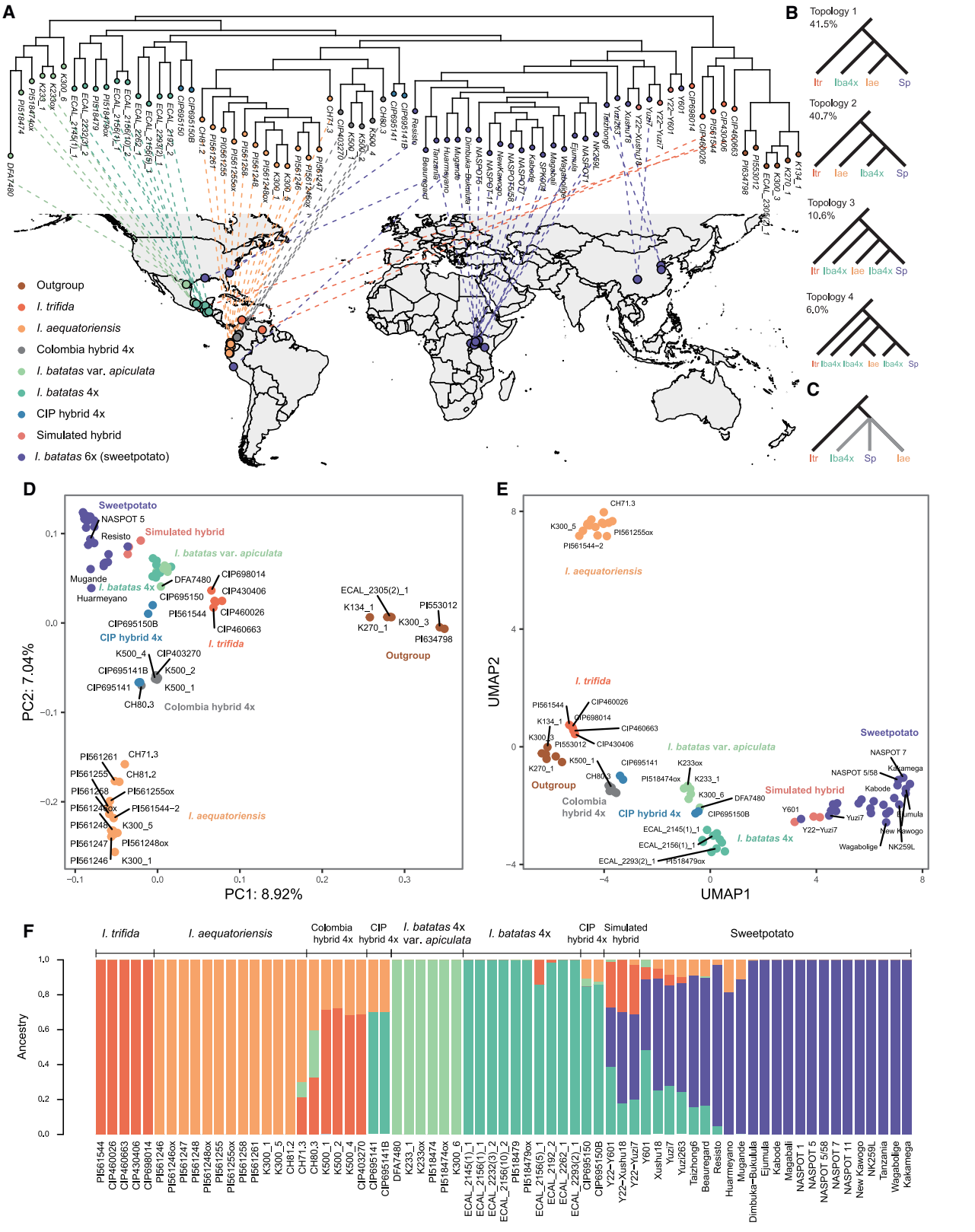

Figure 1. Phylogeny and population structure from sweetpotato and its wild relatives.

(A) Phylogenetic tree constructed using the genome-wide genetic variants inferred by the maximum likelihood (ML) method. All nodes are 100% supported by bootstrap values. The groups are color coded, and the colors in **(D)–(F)** are consistent. Dashed lines link the phylogenetic position on the tree

(legend continued on next page)

Please cite this article in press as: Yan et al., Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato, Molecular Plant (2024), https://doi.org/10.1016/j.molp.2023.12.019

diploid relatives, we included five accessions of *I. trifida*, which is proposed as the most likely diploid progenitor of sweetpotato, as well as two wild relatives, *I. triloba* and *I. tenuissima*. We also sampled the 43 wild tetraploid individuals that have previously been suggested as the possible tetraploid progenitor of sweetpotato, including *Ipomoea tiliacea*, *I. aequatoriensis*, *I. batatas* var. *apiculata*, *Ipomoea tabascana*, *I. batatas* 4x (also named *I. trifida* 4x in some studies), and potential hybrid *Ipomoea* accessions. Detailed information, including ploidy, original location, and source, can be found in Supplemental Table 1. The phenotypes of the wild relatives are shown in Supplemental Figures 1–4.

We constructed phylogenies with coalescence-based and concatenation approaches based on 6,326 447 whole-genome variations, and these topologies were consistent (Figure 1A and Supplemental Figures 5 and 6). In these phylogenetic analyses, the basal clade was formed by diploid *I. trifida* and outgroup species, including diploid *I. triloba*, *I. tenuissima*, and tetraploid *I. tiliacea*. The *I. batatas* 4x lineages, including the *I. batatas* 4x and *I. tabascana* clade (hereafter referred to as *I. batatas* 4x) and the *I. batatas* var. *apiculata* clade, resided at the base of a large lineage composed of sweetpotato cultivars and a monophyletic tetraploid lineage. The monophyletic tetraploid lineage consisted of two monophyletic clades, including tetraploid *I. aequatoriensis* from Ecuador and tetraploid hybrids from Colombia (Colombia hybrid 4x). Sweetpotato cultivars/landraces formed a sister monophyletic lineage to the monophyletic tetraploid lineage. To infer reticulate phylogenetic relationships between sweetpotato and its two closest wild relatives (*I. aequatoriensis* and *I. batatas* 4x), we defined 5-kb non-overlapping windows across the concatenated variation matrix for tree construction. We then counted the topologies of the resulting 1092 trees. The four most frequent topologies (topology 1 to topology 4) accounted for 41.5% to 6.0% of the trees (Figure 1B), indicating significant conflict among the windows. Topology 1, accounting for 41.5% of the trees, was consistent with the species trees derived from coalescence-based and concatenation methods (Figure 1B). However, the second most prevalent topology (topology 2), accounted for 40.7% of the trees, only marginally less than topology 1. Topology 2 showed that sweetpotato and *I. batatas* 4x were sister clades, with *I. aequatoriensis* positioned at the base of their clades (Figure 1B). The poorly resolved topologies (Figure 1C) highlighted reticulation history during sweetpotato formation and suggested that *I. aequatoriensis* and *I. batatas* 4x were likely involved in the allopolyploid reticulation.

Both principal-component analysis (PCA) and uniform manifold approximation and projection (UMAP) analyses clustered all accessions into six major groups: outgroup, *I. trifida*, *I. aequatoriensis*, Colombia hybrid 4x, *I. batatas* 4x group (including *I. batatas* 4x, *I. batatas* var. *apiculata*, and *I. tabascana*), and sweetpotato (Figure 1D and 1E; Supplemental Figure 7A, 7B, 7D, and 7E). These results are consistent with the phylogenetic clades of sweetpotato and its wild relatives. Admixture-based analysis achieved the lowest cross-validation error when K = 3 (Supplemental Figure 8B) and delineated three populations: 1) *I. trifida* and the *I. batatas* 4x group, 2) *I. aequatoriensis*, and 3) sweetpotato (Supplemental Figure 8A). When K = 2, sweetpotato merged into the population of *I. trifida* and the *I. batatas* 4x group, whereas *I. aequatoriensis* remained isolated (Supplemental Figure 8A). These results demonstrate a close relationship among *I. trifida*, the *I. batatas* 4x group, and sweetpotato while indicating a relatively distant relationship between *I. aequatoriensis* and sweetpotato. However, when K = 5, admixture-based analysis revealed a fine population structure that was consistent with the phylogenetic, PCA, and UMAP analyses (Figure 1F and Supplemental Figure 8A). The detailed classification of all samples is provided in Supplemental Table 1.

There are two speculations regarding the relationship between the *I. batatas* 4x group and sweetpotato. The first considers the *I. batatas* 4x group to be the tetraploid progenitor of sweetpotato (Yan et al., 2021), whereas the second treats the *I. batatas* 4x group as hybrid offspring of crosses between sweetpotato and *I. trifida* (Muñoz-Rodríguez et al., 2022). Therefore, it is necessary to establish an effective standard to distinguish the tetraploid progenitor and hybrid offspring using real hybrids or simulated data (Yan et al., 2022). We simulated 3 tetraploid hybrids of *I. trifida* and sweetpotato by randomly sampling reads from the closest known accession of *I. trifida* (CIP698014) related to sweetpotato, as well as from three sweetpotato cultivars, at a ratio of 1:3. We then integrated these sampled reads for phylogenetic and clustering analyses. All analyses revealed that the simulated tetraploid hybrids consistently grouped within the sweetpotato clade or cluster in all analyses and were distinctly separated from other wild tetraploid relatives, including the *I. batatas* 4x group (Figure 1A, 1D, and 1E). This suggests that the *I. batatas* 4x group is unlikely to comprise hybrids derived from *I. trifida* and sweetpotato.

Some wild hybrids (Colombia hybrid 4x) and artificial hybrids (International Potato Center [CIP] hybrid 4x) were identified or confirmed by PCA, UMAP, and admixture-based analysis. Colombia hybrid 4x lay in the middle of *I. trifida*, the *I. batatas* 4x group, and *I. aequatoriensis* (Figure 1D and 1E), and the population structure also supports the scenario in which Colombia hybrid 4x is likely a hybrid originating from *I. trifida*, the *I. batatas* 4x group, or *I. aequatoriensis* (Figure 1F). CIP

with the geographic location on the map for each accession. Accessions lacking geographic data are not linked to the map. The same accessions from Muñoz-Rodríguez et al. (2022) are labeled as "ox" at the end of the accession name.

(B) The four most common topologies across windows. The value in the top left corner is the percentage of all 5-kb windows that recover that topology. Itr, *I. trifida*; Iba4x, the *I. batatas* 4x group; Iae, *I. aequatoriensis*; Sp, sweetpotato.

(C) Consensus cladogram showing the poorly resolved relationship of sweetpotato and its two closest wild relatives.

(D) Principal-component analysis (PCA) of sweetpotato and its relatives. The proportions of variance explained by principal component 1 (PC1) and PC2 are illustrated on the axes.

(E) Uniform manifold approximation and projection (UMAP) using the first three PCs.

(F) Population structure analysis of sweetpotato and its close wild relatives for K = 5.

Please cite this article in press as: Yan et al., Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato, Molecular Plant (2024), https://doi.org/10.1016/j.molp.2023.12.019

hybrid 4x is an artificial hybrid between tetraploid relatives that were crossed and recorded by the CIP. Individuals from the *I. batatas* 4x group and *I. aequatoriensis* are involved in the pedigrees of CIP hybrid 4x accessions, and their genetic backgrounds are described in the Supplementary Note. The admixture-based analysis confirmed their hybrid origins; individuals within CIP hybrid 4x exhibited a genetic blend of both the *I. batatas* 4x group and *I. aequatoriensis* (Figure 1F).

Horizontally transferred *Ib*T-DNAs reveal two progenitors of sweetpotato

Given that the genomes of nearly all sweetpotato cultivars/landraces contain horizontally transferred *Ib*T-DNA1 and/or *Ib*T-DNA2 sequences from *Agrobacterium* spp. (Kyndt et al., 2015; Quispe-Huamanquispe et al., 2019), it is very likely that these *Ib*T-DNAs were inherited from the progenitors of sweetpotato. Therefore, *Ib*T-DNAs may serve as natural genetic markers for tracing the origins of sweetpotato (Quispe-Huamanquispe et al., 2019). In this study, *I. tenuissima* was the only diploid species that contained *Ib*T-DNA1, but its *Ib*T-DNA1 sequence diverged significantly from that of sweetpotato (Figure 2A; Supplemental Table 1). Among the tetraploid relatives, six accessions of the *I. batatas* 4x group and three hybrid accessions (CIP695141, CIP695150B, and CIP403270) possessed *Ib*T-DNA1 (Figure 2A; Supplemental Table 1). Among these, the non-hybrid wild relative with sweetpotato-like *Ib*T-DNA1 sequences was found only in the *I. batatas* 4x group (Figure 2A). Consequently, the progenitor that passed on *Ib*T-DNA1 to sweetpotato is likely to be found among the *I. batatas* 4x group. Interestingly, *Ib*T-DNA1 sequences of several accessions from the *I. batatas* 4x group were partially covered with sequencing reads (Figure 2A), suggesting that these accessions gradually lost the *Ib*T-DNA1 sequence. This explains why the other accessions in the *I. batatas* 4x group exhibit a closer affinity to sweetpotato but lack the *Ib*T-DNA1 insertion.

I. trifida has been reported previously as the diploid progenitor of sweetpotato. After screening 37 accessions of *I. trifida*, we identified six positive accessions for *Ib*T-DNA2 (Supplemental Table 1; Supplemental Figure 9). Nevertheless, *Ib*T-DNA2 sequences from all six positive accessions formed a sister lineage to those of sweetpotato and its tetraploid wild relatives (Figure 2B and Supplemental Figure 9), indicating that these sequences in *I. trifida* are genetically distant from those of sweetpotato. Among the tetraploid wild relatives, nearly all accessions of *I. aequatoriensis*, along with three artificial hybrid accessions (CIP695141, CIP695141B, and CIP695150B), contain the *Ib*T-DNA2 insertion (Figure 2B; Supplemental Table 1) and align within the same lineage as sweetpotato (Figure 2B; Supplemental Figure 9). Therefore, *I. aequatoriensis*, as the only non-hybrid wild relative with sweetpotato-like *Ib*T-DNA2, is likely related to the progenitor that passed on *Ib*T-DNA2 to sweetpotato.

Subgenome origins revealed by HPA

Relationships between sweetpotato and its potential progenitors are informative for determining which subgenome was contributed by each progenitor. Considering the dosage effect, the progenitor that contributed four copies of the B_2 subgenome (tetraploid progenitor) is genetically closer to sweetpotato than the other progenitor that contributed two copies of the B_1 subgenome (diploid progenitor). The *I. batatas* 4x group shows a closer relationship to sweetpotato than *I. aequatoriensis* in the PCA (specifically PC1 vs. PC2), UMAP plots (Figure 1D and 1E), and genome-wide nucleotide diversity analysis (Supplemental Figures 10 and 11). Nevertheless, these analyses regard both the hexaploid sweetpotato and its tetraploid relatives as diploids, thereby artificially reducing the allelic variations among polyploids.

To determine the precise relationship between sweetpotato and its progenitors, we developed an HPA pipeline that uses homoeologous haplotypes from polyploids for high-throughput phylogenetic analyses (Supplemental Figure 12). We initially mapped the sequence reads of sweetpotato cultivars and tetraploid wild relatives to the sweetpotato genome and independently reconstructed haplotypes based on the genomic variations (Supplemental Figure 12A). Taking into account both phylogenetic lineages and geographic locations, we selected three representative cultivars (Figure 1A and Supplemental Figures 5 and 6): Huameyano, NK259L, and Yuzi7. We identified 439 555–760 769 haplotype blocks in the three sweetpotato cultivars (Supplemental Table 2; Supplemental Figure 13B) and 380 895–1 007 206 haplotype blocks in the 38 tetraploid accessions (Supplemental Table 3; Supplemental Figure 13A). We recorded the start and end position for each haplotype block to the genome. Next, we extracted the syntenic haplotype blocks shared between each sweetpotato cultivar and each tetraploid accession. By comparing the start and end positions, we extracted overlapping regions of haplotype blocks between sweetpotato and the tetraploid accession as syntenic haplotype blocks (Supplemental Figure 12B). Through this process, we identified 606 246–1 154 274 syntenic haplotype blocks (Supplemental Table 4; Supplemental Figure 14). Finally, we removed 1) redundant syntenic haplotype blocks that had overlapping regions with other blocks and 2) blocks that consisted of short sequences (less than 20 bp). As a result, we extracted 412 632–866 522 syntenic haplotype blocks, comprising 28.2%–41.7% of the sweetpotato genome (Supplemental Table 5; Supplemental Figure 15).

We used the syntenic haplotype blocks between each sweetpotato cultivar and each tetraploid accession for independent phylogenetic reconstructions (Supplemental Figure 12C). The phylogenetic trees were generated by two methods: unweighted pair-group method with arithmetic mean (UPGMA) and maximum likelihood (ML). We calculated the monophyletic ratios, the Nsp–Nwr distances, and thenucleotide diversity (PI) indices to measure the relationship between the investigated tetraploid accession and the representative hexaploid sweetpotato (Supplemental Figure 12D). Both the monophyletic ratio and the Nsp–Nwr distance are calculated on the basis of tree topology. The monophyletic ratio is defined as the proportion of trees in which sweetpotato haplotypes form a monophyletic clade (Supplemental Figure 12D). The Nsp–Nwr distance is defined as the tree branch length between the most recent common ancestor (MCRA) node of sweetpotato haplotypes (Nsp) and the MCRA node of the tetraploid accession (Nwr) (Supplemental Figure 12D). When either sweetpotato or the wild relative is not

Please cite this article in press as: Yan et al., Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato, Molecular Plant (2024), https://doi.org/10.1016/j.molp.2023.12.019

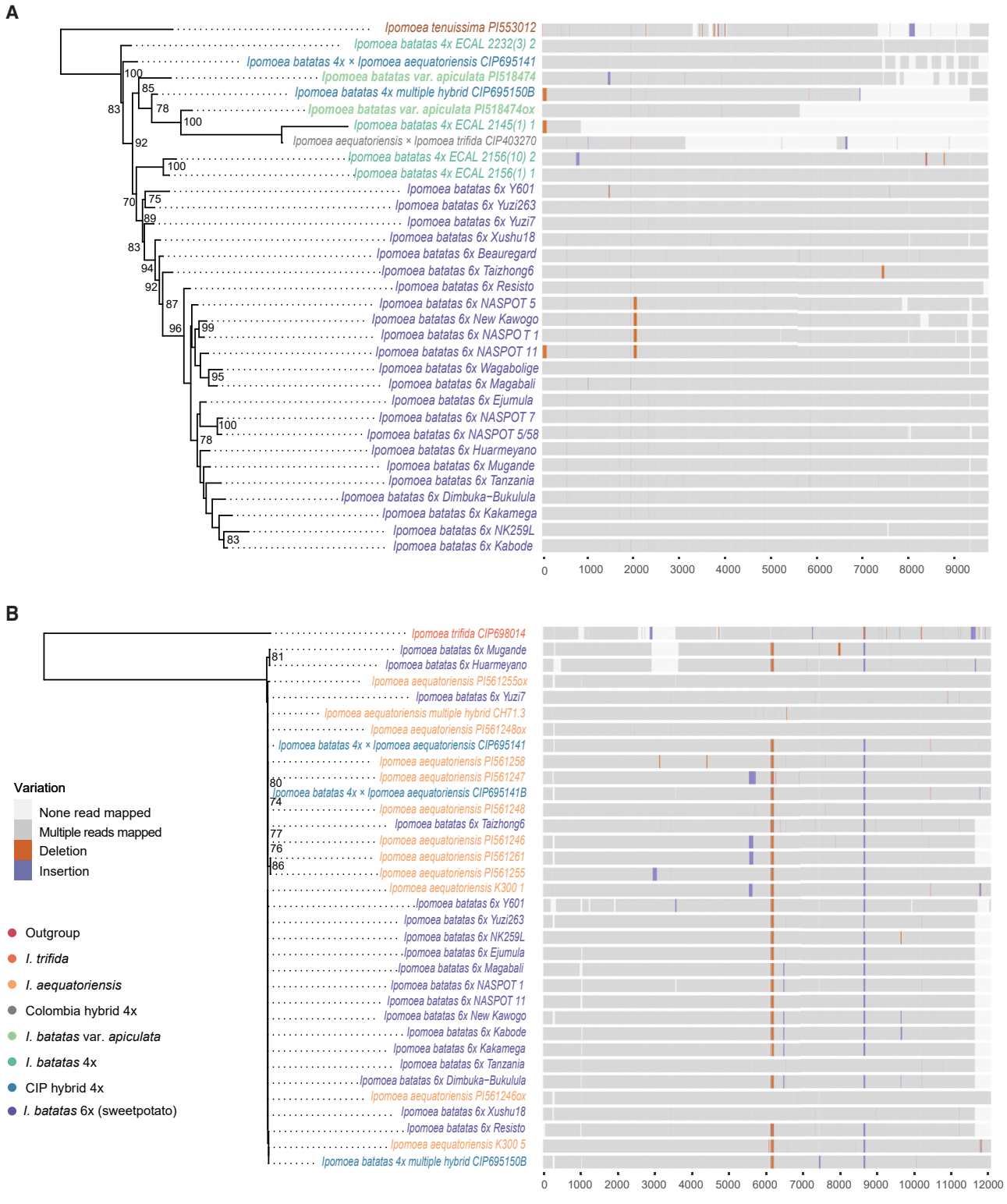

Figure 2. Phylogeny and structure of *Ib*T-DNAs from sweetpotato and its wild relatives.

(A) ML tree of *Ib*T-DNA1 based on variants from positive accessions. Nodes supported by bootstrap values greater than 70% are noted. Structural diagrams of *Ib*T-DNA1 are shown on the right. Regions with no mapped reads are likely to be deletions and are colored in light gray. Indels are also color coded.

(B) ML tree of *Ib*T-DNA2 based on sequence variants from positive accessions. Nodes with bootstrap values greater than 70% are noted. Structural diagrams of *Ib*T-DNA2 in positive accessions are shown on the right. Regions with no mapped reads are likely to be deletions and are colored in light gray. Indels are also color coded.

Please cite this article in press as: Yan et al., Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato, Molecular Plant (2024), https://doi.org/10.1016/j.molp.2023.12.019

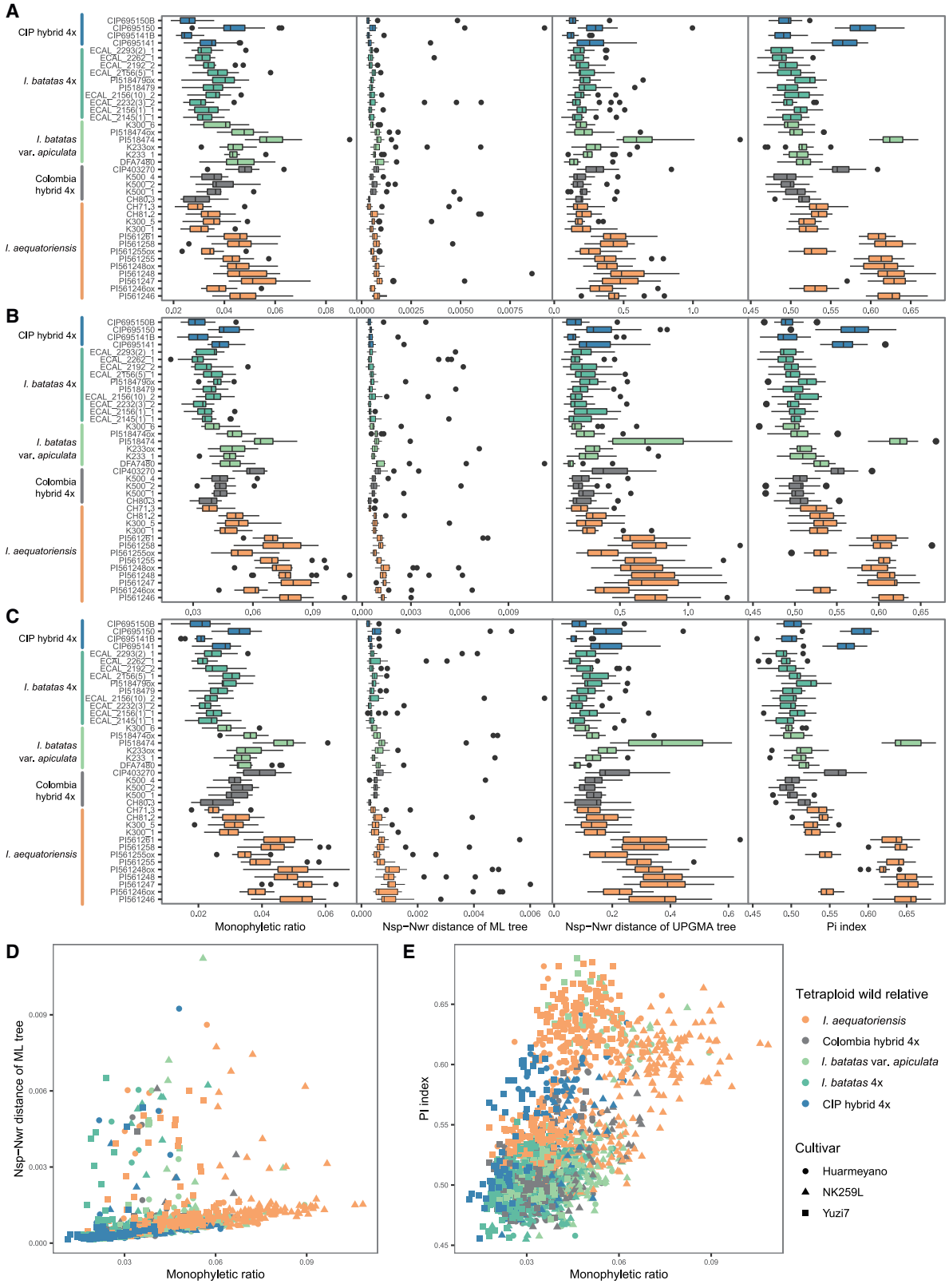

Figure 3. Relationships between sweetpotato cultivars and tetraploid accessions as revealed by haplotype-based phylogenetic analysis (HPA).

(A–C) Boxplots of the monophyletic ratio, the Nsp–Nwr distance based on two tree-building methods, and the PI index of 15 chromosomes among 38 tetraploid accessions: **(A)** results for cultivar Huarmeyano, **(B)** results for cultivar NK259L, and **(C)** results for cultivar Yuzi7. Monophyletic ratio, the

(legend continued on next page)

Please cite this article in press as: Yan et al., Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato, Molecular Plant (2024), https://doi.org/10.1016/j.molp.2023.12.019

monophyletic, the Nsp–Nwr distance becomes zero. If the wild relative is sufficiently close to sweetpotato, then neither sweetpotato nor its wild relative is likely to be monophyletic, resulting in a low monophyletic ratio and a short Nsp–Nwr distance. While, the PI index is a coefficient that quantifies the difference in nucleotide diversity between the haplotypes from sweetpotato and the tetraploid accession (Supplemental Figure 12D). If the wild relative shares a close genetic affinity with sweetpotato, the difference in nucleotide diversities between sweetpotato and the wild relative is small, leading to a low PI index. To enhance accuracy, we limited our analysis to trees that consistently received the same monophyletic judgment from both tree-building methods, using them to compute the monophyletic ratio, Nsp–Nwr distance, and PI index. Of all syntenic haplotype blocks examined, the 6:4 dataset (composed of six haplotypes of sweetpotato and four haplotypes of tetraploid accessions) produced the most robust results, because the results of the 6:4 dataset were consistent across all three indices using the three sweetpotato cultivars (Figure 3A–3C; Supplemental Figures 16–30).

HPA provided a better resolution and relatively consistent results in resolving the relationship between sweetpotato and its tetraploid relatives. The strong correlations between monophyletic ratio and Nsp–Nwr distance, as well as between monophyletic ratio and PI index (Figure 3D and 3E), demonstrated that the three indices revealed a consistent relationship. All three HPA indices showed that, among the non-hybrid tetraploid relatives, the *I. batatas* 4x group contained the closest relatives of sweetpotato, and *I. aequatoriensis* was the most distant tetraploid relative (Figures 3A–3C). HPA also indicated that *I. batatas* 4x was more closely related to sweetpotato than *I. batatas* var. *apiculata* (Figure 3A–3C). Combined with the evidence from *Ib*T-DNA1, these results indicate that *I. batatas* 4x is most likely the tetraploid progenitor that contributed the four B_2 subgenomes to sweetpotato. As revealed by *Ib*T-DNA2 evidence, *I. aequatoriensis* is closely related to another progenitor species. Because *I. aequatoriensis* has a relatively distant relationship to sweetpotato compared with *I. batatas* 4x, it is most likely related to the diploid progenitor, which contributed the two B_1 subgenomes to sweetpotato. Because *I. aequatoriensis* is an autotetraploid species with low heterozygosity (Muñoz-Rodríguez et al., 2022), it probably formed through a whole-genome duplication of its diploid form. Therefore, its ancient diploid form is likely the diploid progenitor of sweetpotato.

The CIP hybrid 4x exhibited the closest genetic affinity to sweetpotato (Figure 3A–3C). Samples of CIP hybrid 4x are artificial hybrids derived from tetraploid relatives, incorporating the *I. batatas* 4x group and *I. aequatoriensis* in their pedigrees (Supplemental Note). Their close genetic relationship to sweetpotato suggests that these hybridizations likely mimicked the natural origin of sweetpotato, except for the differing ploidy levels.

Chloroplast genome analyses confirm the identification of the two sweetpotato progenitors

Consistent with previous reports (Muñoz-Rodríguez et al., 2018, 2022), the chloroplast haplotypes of sweetpotato were divided into two lineages, lineage 1 and lineage 2 (Figure 4; Supplemental Figure 31), and the chloroplast haplotypes of *I. trifida* were distantly related to the two lineages of sweetpotato. However, *I. batatas* 4x was nested in lineage 2 of sweetpotato, and the closest individuals were five accessions of *I. batatas* 4x (ECAL_2156(1)_1, ECAL_2156(10)_2, ECAL_2192_2, ECAL_2262_1, and ECAL_2293(2)_1) (Figure 4; Supplemental Figure 31). Other accessions of *I. batatas* 4x, *I. batatas* var. *apiculata*, and Colombia hybrid 4x also belong to lineage 2, but they showed a relatively distant relationship to sweetpotato haplotypes (Figure 4; Supplemental Figure 31). The closest relationship was observed between the chloroplast genomes of *I. batatas* 4x and sweetpotato, providing further evidence that *I. batatas* 4x is the probable progenitor of sweetpotato. In addition, *I. aequatoriensis* was the only non-hybrid species nested within lineage 1 of sweetpotato haplotypes (Figure 4; Supplemental Figure 31), which suggests that *I. aequatoriensis* resembles the progenitor of sweetpotato that contributed the lineage 1 type of chloroplast genome to sweetpotato. Consequently, it is highly probable that the two chloroplast haplotype lineages of sweetpotato are directly inherited from its two progenitors. It is also probable that the two progenitors crossed reciprocally and thus passed on the two chloroplast genome lineages and the identical nuclear genome conformation to sweetpotato (Figure 5A). The haplotypes of *I. trifida* exhibited greater genetic distance from sweetpotato than did the two progenitors (Figure 4; Supplemental Figure 31), indicating that extant *I. trifida* may not be the diploid progenitor of sweetpotato.

Gene conversion between sweetpotato subgenomes

Gene conversion in polyploids refers to sequence exchanges between homoelogous genes from different subgenomes in which one progenitor allele overwrites another (Wang and Paterson, 2011; Cenci et al., 2012; Chen et al., 2016). The sweetpotato genome comprises two B_1 and four B_2 subgenomes ($B_1B_1B_2B_2B_2B_2$). Subgenomes B_1B_1 originated from the diploid progenitor, whereas subgenomes $B_2B_2B_2B_2$ were contributed by the tetraploid progenitor (Shiotani and Kawase, 1987). In the absence of any conversion events, each syntenic haplotype block between sweetpotato and *I. batatas* 4x should have two copies of the B_1 subgenome from sweetpotato, four copies of the B_2 subgenome from sweetpotato, and four copies of the B_2 subgenome from *I. batatas* 4x (Figure 5C). If genes converted between the B_1 and B_2 subgenomes, both the copy numbers of the subgenomes and the tree topology are expected to deviate from the standard 2:8 ratio between B_1 and B_2 in hexaploid sweetpotato and *I. batatas* 4x (Figure 5D and 5E). We therefore identified gene conversion events by examining tree topology, in the Supplemental Note. To detect potential gene conversion

proportion of trees in which sweetpotato haplotypes form a monophyletic clade; Nsp–Nwr distances, the tree branch length between the most recent common ancestor (MCRA) node of sweetpotato haplotypes (Nsp) and the MCRA node of the tetraploid accession (Nwr); PI index, a coefficient that calculates the difference between haplotype nucleotide diversity of sweetpotato and that of the tetraploid accession.

(D) Dot plot between the monophyletic ratio and the Nsp–Nwr distances of the ML tree.

(E) Dot plot between the monophyletic ratio and the PI index.

Please cite this article in press as: Yan et al., Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato, Molecular Plant (2024), https://doi.org/10.1016/j.molp.2023.12.019

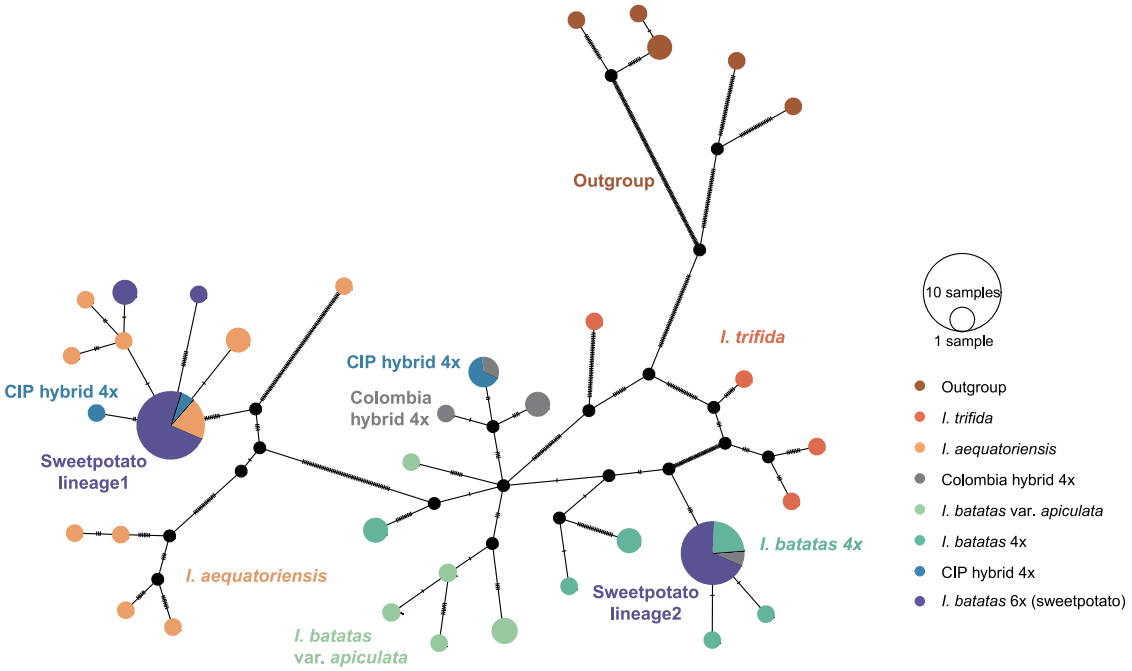

Figure 4. Phylogenetic network of chloroplast genomes.
The phylogenetic network of chloroplast genomes from sweetpotato and its wild relatives was inferred using the TSC network. Circle size is proportional to the frequency of the corresponding haplotype across all populations. The number of short lines between two haplotypes represents mutational steps. Filled black circles indicate either unsampled haplotypes or extinct ancestral haplotypes.

events, we only analyzed syntenic haplotype blocks located in gene regions, which included six haplotypes of sweetpotato and four haplotypes of *I. batatas* 4x. For the representative sweetpotato cultivars, we selected one cultivar from each lineage in the sweetpotato phylogeny (Figure 1A and Supplemental Figures 5 and 6): Huameyano, NASPOT5/58, NK259L, Y601, and Yuzi7. The closest *I. batatas* 4x accession (ECAL_2262_1), which resembles the tetraploid progenitor, was selected as a reference. The analysis pipeline is illustrated in Supplemental Figure 32.

We ultimately obtained 13 535–27 867 homologous haplotype blocks of sweetpotato cultivars and the closest *I. batatas* 4x accession, which we then used to identify gene conversion events between subgenomes (Supplemental Table 6). Using the 5 sweetpotato cultivars as references, 47.1%–48.3% of the gene regions in sweetpotato exhibited signs of subgenome conversion (Figure 5B; Supplemental Table 6). Our findings indicated that gene conversions from B_1 to B_2 subgenomes (38.1%–39.3%) occurred much more frequently than those from B_2 to B_1 (8.9%–9.6%) (Figure 5B; Supplemental Table 6). This result aligns with expectations, because gene conversion is known to be a copy number–dependent process (Khakhlova and Bock, 2006).

Genomic signatures of selective sweeps in sweetpotato

We compared the genetic diversity between sweetpotato and its two progenitors by estimating the genome-wide nucleotide diversity (π) across a sweetpotato population consisting of 23 cultivars and landraces as well as the populations of its two progenitors (*I. aequatoriensis* and the *I. batatas* 4x group) in sliding windows. Sweetpotato exhibited higher genome-wide nucleotide diversity than its diploid progenitor, a trend evident across all chromosomes (Supplemental Figure 10). The average nucleotide diversity of sweetpotato was also higher than that of the diploid progenitor ($\pi_{sweetpotato}$ = 0.0035, $\pi_{diploid\ progenitor}$ = 0.0026). However, the genome-wide nucleotide diversity of sweetpotato was remarkably similar to that of its tetraploid progenitor across all chromosomes (Supplemental Figure 11), with only a slightly higher average nucleotide diversity (Supplemental Figure 10; $\pi_{tetraploid\ progenitor}$ = 0.0034).

We used three metrics to detect potential signatures of selection during sweetpotato domestication: π ratio ($\pi_{wild\ relative}/\pi_{sweetpotato}$), population differentiation (F_{ST}), and cross-population composite likelihood ratio (XP-CLR). We calculated the three metrics using 100-kb sliding windows with 10-kb steps. Regions ranking in the top 1% for π ratio, F_{ST}, and XP-CLR scores were identified as selective sweep regions. In total, we identified 466 potential selective sweep regions between sweetpotato and its two progenitor populations. These regions of potential selective sweeps were probably linked to natural selection and domestication.

The selective sweep regions between sweetpotato and its diploid progenitor encompass 1014, 1020, and 2921 genes, as identified by the three respective metrics (Supplemental Figure 33A; Supplemental Table 7). These genes are primarily involved in various biological pathways, including cell surface receptor signaling, symbiotic interaction, regulation of amino acid transport, glutathione peroxidase activity, triterpenoid biosynthetic process, ethylene-activated signaling pathway, mitotic cell cycle, plant-type hypersensitive response, enzyme

Please cite this article in press as: Yan et al., Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato, Molecular Plant (2024), https://doi.org/10.1016/j.molp.2023.12.019

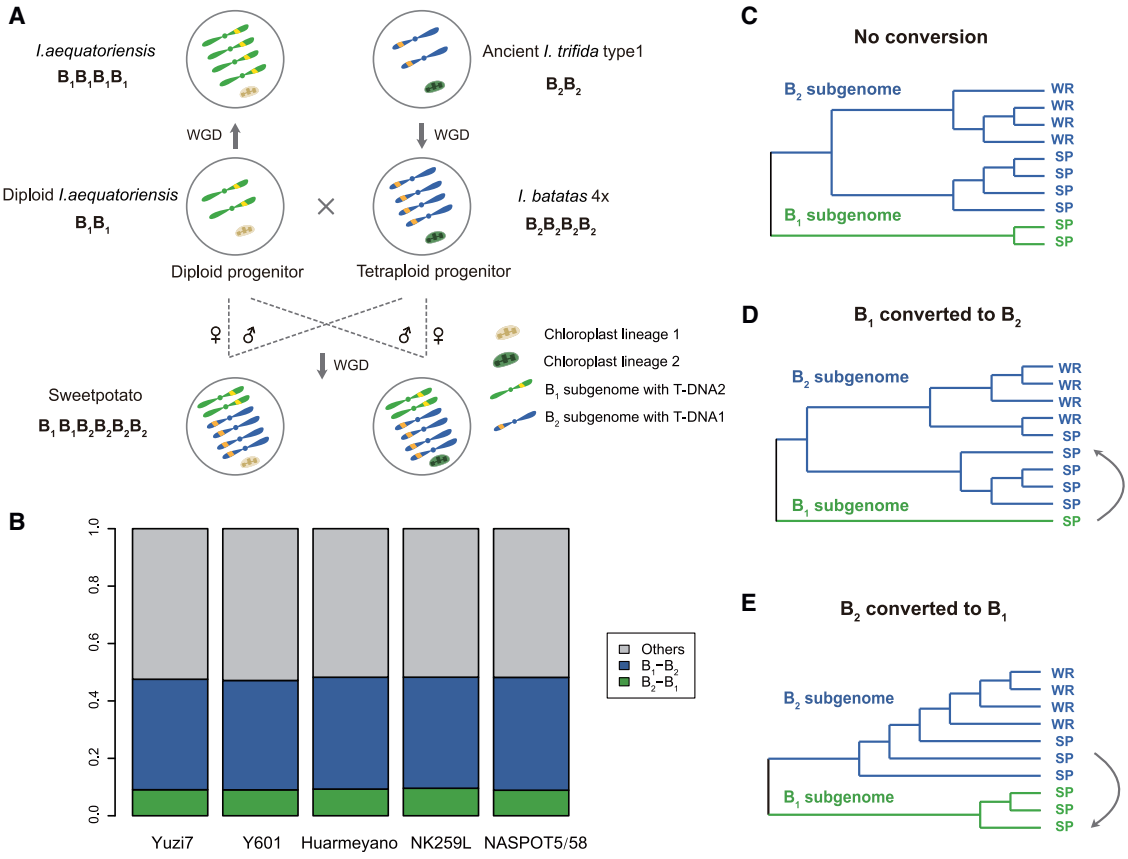

Figure 5. Hypothesized origin of sweetpotato and gene conversions between subgenomes.

(A) Hypothesized origin of sweetpotato. The diploid progenitor (likely the diploid form of *I. aequatoriensis*) contributed the B_1 subgenome, *Ib*T-DNA2, and the lineage 1 type of chloroplast genome to sweetpotato. The tetraploid progenitor of sweetpotato, identified as *I. batatas* 4x, probably originated from genome duplication in ancient *I. trifida*. *I. batatas* 4x contributed the B_2 subgenome, *Ib*T-DNA1, and the lineage 2 type of chloroplast genome. Sweetpotato is derived from reciprocal crosses between the diploid and tetraploid progenitors, followed by a subsequent whole-genome duplication (WGD). **(B)** Gene conversion ratios in five hexaploid sweetpotato cultivars/landraces using the closest known non-hybrid accession (ECAL_2262_1) as a reference to resemble the tetraploid progenitor. B_1–B_2, gene conversion events from the B_1 to the B_2 subgenome; B_2–B_1, conversion events from the B_2 to the B_1 subgenome; Others, other scenarios, including no conversion and unresolved scenarios. **(C–E)** Examples of tree topologies under the scenarios of no conversion **(C)**, B_1 to B_2 gene conversion **(D)**, and B_2 to B_1 gene conversion **(E)**. The B_1 subgenome is shown in green and the B_2 subgenome in blue. SP, sweetpotato. WR, wild relative (tetraploid progenitor).

inhibitor activity, etc. (Supplemental Table 8). Selection of candidate domestication genes was based on their ranking in the top 1% of least two metrics and on consideration of their biological functions and expression levels (Figure 6A–6C; Supplemental Table 9). Among these candidates, genes encoding two basic helix loop helix (*bHLH*) transcription factors, heat shock cognate protein 70-1 (*HSP70-1*), and auxin response factor 2A-like (*ARF2*) are known to be involved in initiation and/or development of storage roots (Cao et al., 2012; Ravi et al., 2014, 2017). Genes encoding two lateral organ boundaries (LOB) domain-containing proteins (*LBD1* and *LBD33*), U-box domain-containing protein 13-like (*U-box13*), cysteine-rich RECEPTOR-like protein kinase 2 (*CRK2*), 14-3-3 protein (*14-3-3*), NAC (NAM, ATAF1,2, CUC2) transcription factor 56-like (*NAC056*), and transmembrane protein 18 (*Tmem18*) are functionally related to root development in *Arabidopsis* and other plants (Berckmans et al., 2011; Mayfield et al., 2012; Dou et al., 2016; Hunter et al., 2019; Yamauchi et al., 2019; Kim et al., 2021; Xu et al., 2022). Aside from *LBD1* and *LBD33*, which were predominantly expressed in the stem, all other genes were highly expressed in the root (Supplemental Figure 34A). Their expression also tended to increase with root development, suggesting that they may play significant roles in storage root development in sweetpotato. Potassium is crucial for yield of sweetpotato storage roots because it influences photosynthesis, translocation of nutrients, and initiation and thickening of storage roots (George et al., 2002). We identified a gene, integrin-linked protein kinase 1-like (*ILK1*), that promotes potassium uptake and is highly expressed at the early stage of storage root development (Supplemental Figure 34A). We identified two well-known plant defense genes, specifically *N*, which encodes tobacco mosaic virus (TMV) resistance protein N-like, and *CAMTA*, which encodes a calmodulin-binding transcription activator protein. In addition, we detected sporamin B, a major storage protein in sweetpotato storage roots that plays significant roles in stress tolerance (Senthilkumar and Yeh, 2012). Maintaining genomic stability is particularly challenging in polyploids owing to the complex meiotic behavior of chromosome sets and the complexity of recombination events (Comai, 2005; Hollister, 2015). We

Please cite this article in press as: Yan et al., Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato, Molecular Plant (2024), https://doi.org/10.1016/j.molp.2023.12.019

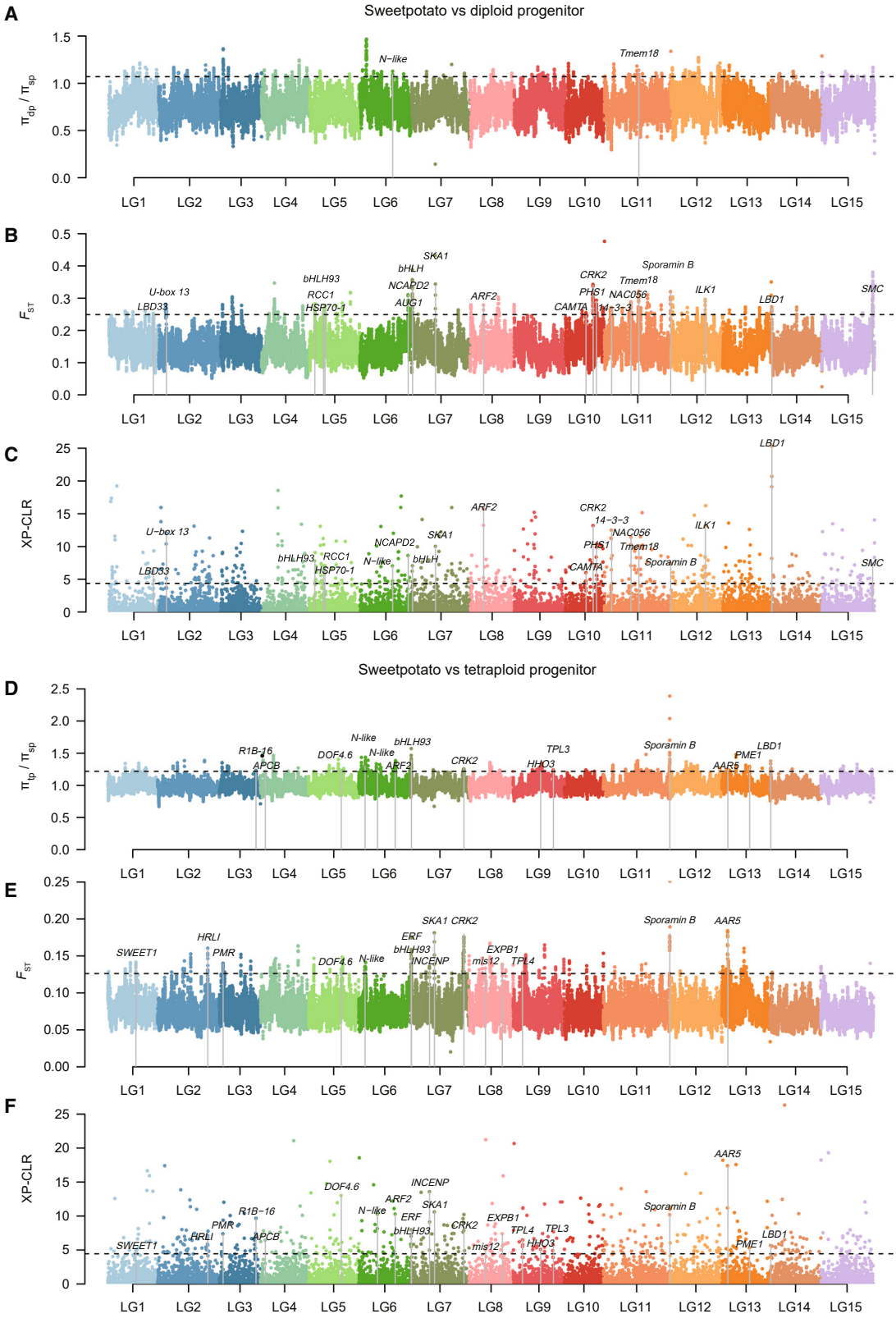

Figure 6. Profiling of selective sweeps during sweetpotato speciation and domestication.

(A–C) Selective sweep regions identified between sweetpotato and the tetraploid form of the diploid progenitor (*I. aequatoriensis*). Selective sweep regions identified by π ratio ($\pi_{\text{diploid progenitor}}/\pi_{\text{sweetpotato}}$) **(A)**, population differentiation (F_{ST}) **(B)**, and cross-population composite likelihood ratio

(legend continued on next page)

Please cite this article in press as: Yan et al., Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato, Molecular Plant (2024), https://doi.org/10.1016/j.molp.2023.12.019

identified six genes essential for accurate chromosome segregation and maintenance of genomic stability: regulator of chromosome condensation family protein (*RCC1*), condensin complex subunit 1 (*NCAPD2*), AUGMIN subunit 1 (*AUG1*), spindle and kinetochore-associated protein 1 (*SKA1*), poor homologous synapsis 1 (*PHS1*), and structural maintenance of chromosomes protein (*SMC*).

The selective sweep regions between sweetpotato and its tetraploid progenitor contained 928, 945, and 3082 genes, as identified by the three respective metrics (Supplemental Figure 33B; Supplemental Table 10). These genes are enriched in various biological pathways, such as strictosidine synthase activity, protein autophosphorylation, endonuclease activity, terpene synthase activity, root morphogenesis, peptidase inhibitor activity, plant-type hypersensitive response, mitotic G2/M transition checkpoint, etc (Supplemental Table 11). Candidate domestication genes were identified as described above (Figure 6D–6F; Supplemental Table 12). Among these candidatesthere were five genes with significant roles in the initiation and/or development of storage roots: DNA binding with one finger (dof) zinc-finger protein DOF4.6-like (*DOF4.6*), auxin response factor 2A-like (*ARF2*), an ethylene-responsive transcription factor (*ERF*), a bHLH93-like transcription factor (*bHLH93*), and expansin-like B1 (*EXPB1*) (Tanaka et al., 2009; Noh et al., 2013; Ravi et al., 2014, 2017). Six genes known to participate in *Arabidopsis* root development were also identified, including cysteine-rich RECEPTOR-like protein kinase 2 (*CRK2*, topless-related protein-like (*TPL3* and *TPL4*), MYB-like transcription factor HHO3-like (*HHO3*), DCN1-like protein 5 (*AAR5*), and LOB domain-containing protein 1-like (*LBD1*) (Biswas et al., 2007; Espinosa-Ruiz et al., 2017; Hunter et al., 2019; Yamauchi et al., 2019; Li et al., 2021). With the exception of *LBD1*, these genes were predominantly expressed in the roots, and their expression levels increased with root development (Supplemental Figure 34B). Pectin is important for cell wall properties and storage root development (Guillemin et al., 2005; Dong et al., 2020). We identified two genes, pectinesterase 1-like (*PME1*) and probable pectate lyase 20 (*PMR*), that alter the composition of the plant cell wall through pectin modification. In addition, we identified the sugar transporter SWEET1, which is known as a bidirectional uniporter/facilitator with central roles in phloem loading of sugar for long-distance transport (Ji et al., 2022). Five plant defense genes were also identified: hypersensitive response (HR)-like lesion-inducing protein-like protein (*HRLI*), putative late blight resistance protein (*R1B-16*), aspartyl protease (*APCB*), TMV resistance protein N-like, and *sporamin B*. We identified three genes essential for maintenance of genomic stability: inner centromere protein-like (*INCENP*), spindle and kinetochore-associated protein 1 (*SKA1*), and minichromosome instability 12-like protein (*mis12*).

We detected strong selective sweep signals for *sporamin* genes between sweetpotato and its two potential progenitors using all three metrics. Two tandem repeats of *sporamin* genes with 18 and 17 genes, respectively, are present on LG11 (linkage group 11, corresponding to chromosome 11) (Figure 7A). We identified selective sweeps in the second tandem repeat of *sporamin* genes between sweetpotato and its two potential progenitors (Figures 6 and 7B). Sporamin B is a major storage protein in sweetpotato storage roots and plays significant roles in plant defense (Senthilkumar and Yeh, 2012). Downregulation of *sporamin* genes in RNAi lines led to a decrease in storage root yield (Figure 7C–7E). These results suggest that *sporamin* genes not only have roles in plant defense but also are crucial for storage root formation. These genes were likely selected during the domestication and improvement of sweetpotato.

DISCUSSION

Understanding the genetic origins of crops is crucial for breeding and genetic engineering. It is particularly important for genetic improvement and development of genetic resource conservation strategies that involve wild relatives. The genetic origin of sweetpotato has been a subject of ongoing debate (Magoon et al., 1970; Ukoskit and Thompson, 1997; Rajapakse et al., 2004; Srisuwan et al., 2006; Gao et al., 2011, 2020; Muñoz-Rodríguez et al., 2018, 2022; Yan et al., 2021), primarily owing to its highly heterozygous hexaploid genome (Yang et al., 2017; Wu et al., 2018), the limitations of analyses that use genome consensus sequences, and the limited inclusion of wild relatives with close genetic affinity. The subgenomes of sweetpotato exhibit high similarity, largely due to the close genetic relationship between its diploid and tetraploid progenitor species (Magoon et al., 1970). As a result, conventional strategies for investigating the origins of allopolyploids, as applied to crops like rapeseed (*Brassica napus*), bread wheat (*Triticum aestivum*), *Echinochloa* spp., and polyploid bamboo (*Bambusa* spp.) (An et al., 2019; Guo et al., 2019; Lu et al., 2019; Ye et al., 2020; Zhou et al., 2020), are not suitable for sweetpotato. To address these challenges, we incorporated all currently available close wild relatives of sweetpotato and used a variety of genetic markers, such as *Ib*T-DNA insertions, nuclear variations, and chloroplast genotypes. In addition, we developed an HPA pipeline that takes full advantage of homologous variation while maintaining the genuine nucleotide diversity of the polyploid species. This enabled us to accurately delineate the relationships between sweetpotato and its tetraploid relatives.

On the basis of the findings presented here, we propose a hypothesis for the origin of sweetpotato (Figure 5A). This hypothesis accounts for the origins of the two subgenomes, the *Ib*T-DNA insertions, and the two chloroplast genome lineages within cultivated sweetpotato. The diploid progenitor (closely related to *I. aequatoriensis*) contributed the B_1 subgenome, *Ib*T-DNA2, and the lineage 1 type of chloroplast genome to sweetpotato. The tetraploid progenitor of sweetpotato, identified as *I. batatas* 4x and likely derived from duplication of an ancient *I. trifida*, contributed the B_2 subgenome, *Ib*T-DNA1,

(XP-CLR) **(C)**. Dashed lines indicate regions that ranked in the top 1% for π ratio, F_{ST}, and XP-CLR values. Gray vertical bars highlight the positions of candidate domestication genes. sp, sweetpotato; dp, diploid progenitor; LG, linkage group, corresponding to chromosome.
(D–F) Selective sweep regions identified between sweetpotato and the tetraploid progenitor (the *I. batatas* 4x group). Selective sweep regions identified by π ratio ($\pi_{\text{tetraploid progenitor}}/\pi_{\text{sweetpotato}}$) **(D)**, F_{ST} **(E)**, and XP-CLR **(F)**. Dashed lines indicate regions that ranked in the top 1% for π ratio, F_{ST}, and XP-CLR values. Gray vertical bars highlight the positions of candidate domestication genes.

Please cite this article in press as: Yan et al., Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato, Molecular Plant (2024), https://doi.org/10.1016/j.molp.2023.12.019

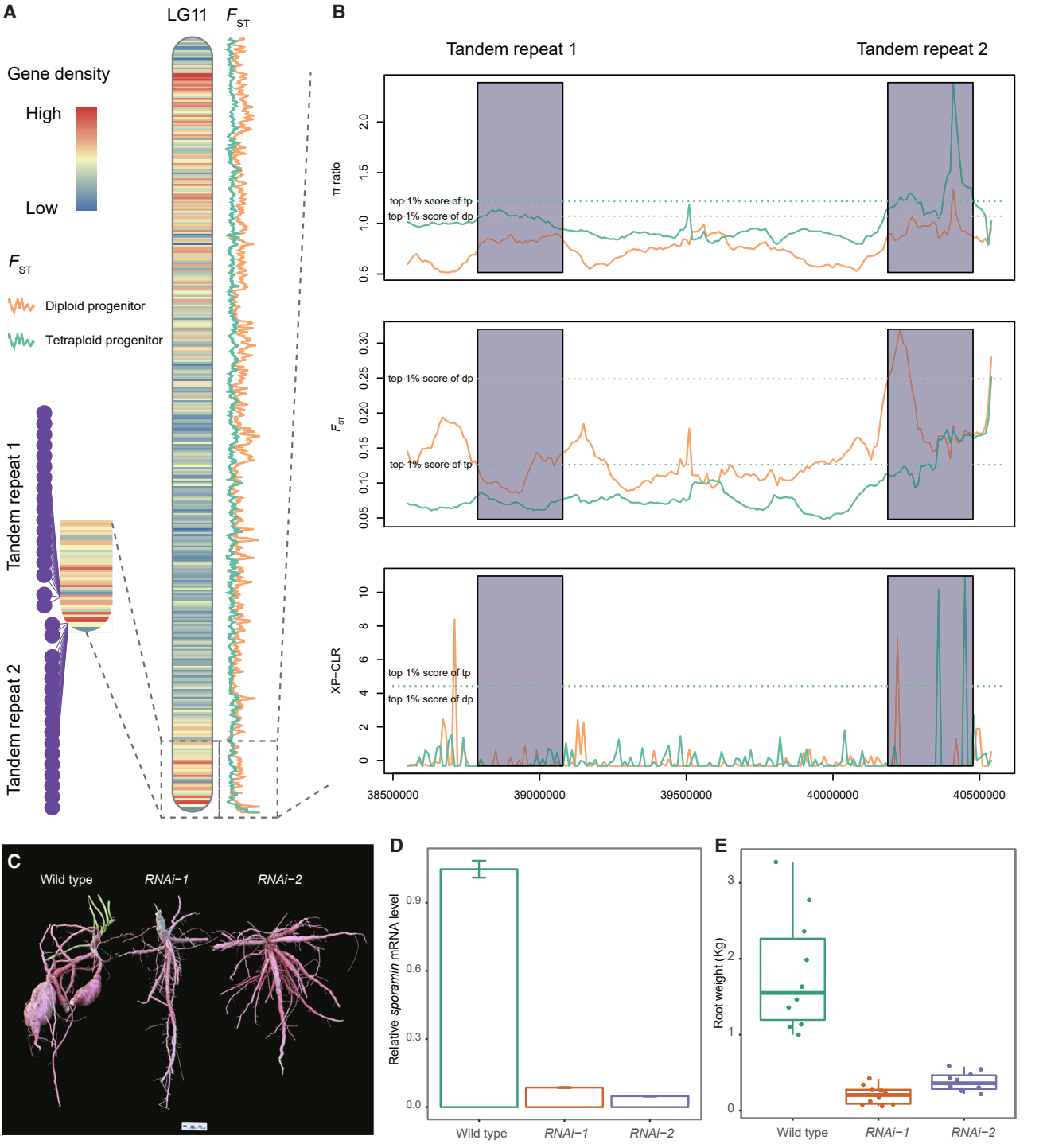

Figure 7. Selective sweeps of *sporamin* genes.

(A) Population differentiation (F_{ST}) between wild relatives and sweetpotato cultivars is illustrated on the right of the LG11 (chromosome 11) karyotype. The colors on the karyotype represent the gene density along LG11. Two tandem repeats of *sporamin* are located on one end of LG11; each gene is labeled with a purple dot and linked to its position on the chromosome.**(B)** π ratio ($\pi_{\text{wild relative}}/\pi_{\text{sweetpotato}}$), F_{ST}, and XP-CLR between sweetpotato and tetraploid populations for *sporamin*-containing regions. Dashed lines indicate the top 1% of values for the π ratio, F_{ST}, and XP-CLR values. dp, diploid progenitor; tp, tetraploid progenitor.**(C)** Root phenotypes of the wild type and two RNAi lines of the *sporamin* gene in sweetpotato.**(D)** Relative *sporamin* mRNA levels in roots (2-mm diameter) of the wild type and *sporamin* RNAi lines.**(E)** Root weights of the wild type and *sporamin* RNAi lines.

and the lineage 2 type of chloroplast genome. Hexaploid sweetpotato is probably derived from reciprocal crosses between the diploid and tetraploid progenitors, followed by a whole-genome duplication. Given that the two identified progenitors and their closest relatives are distributed in Central America, including southern Mexico, Guatemala, Ecuador, and

Please cite this article in press as: Yan et al., Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato, Molecular Plant (2024), https://doi.org/10.1016/j.molp.2023.12.019

Venezuela, sweetpotato may have originated from Central America. The fact that sweetpotato diversity is highest in Central America provides further evidence that this region is likely to be the center of origin for the crop (Huang and Sun, 2000; Zhang et al., 2000). Regrettably, the definitive diploid form of *I. aequatoriensis* has yet to be identified. Because the *Ib*T-DNA2 sequence has proven to be an effective marker for identifying the diploid progenitor of sweetpotato, it will be necessary to conduct a more extensive and comprehensive examination of sweetpotato-type *Ib*T-DNA2 in the diploid species of Central America to ultimately discover the diploid progenitor of sweetpotato. Quispe-Huamanquispe et al. (2019) demonstrated that screening a single gene (*ORF13*) coupled with phylogenetic analysis was sufficient for preliminary identification of the diploid sweetpotato progenitor.

For a long time, extant *I. trifida* has been considered either the diploid progenitor or the sole progenitor of sweetpotato because it is the closest diploid species to sweetpotato, as revealed by DNA sequence and cytogenetic evidence (Shiotani, 1988; Shiotani and Kawase, 1989; Roullier et al., 2013; Muñoz-Rodríguez et al., 2018). However, when closer wild relatives are incorporated in the analyses, all available data suggest that extant *I. trifida* is unlikely to be the diploid or sole progenitor of sweetpotato. First, extant accessions of *I. trifida* show a more distant genetic relationship to sweetpotato compared with *I. batatas* 4x and *I. aequatoriensis*, as revealed in this study and previous studies using nuclear genome variations (Yan et al., 2021; Muñoz-Rodríguez et al., 2022). Second, the chloroplast genomes of *I. trifida* form an independent lineage clearly distinct from the two existing lineages of sweetpotato. Therefore, previous hypotheses that considered extant *I. trifida* the diploid progenitor or sole progenitor failed to account for the formation of the two distinct lineages in sweetpotato. The first hypothesis suggests that asymmetric hybridization between diploid *I. trifida* and an original hexaploid sweetpotato led to chloroplast capture from *I. trifida* (Muñoz-Rodríguez et al., 2018, 2022). However, this explanation overlooks the fact that hybridization between a diploid and a hexaploid results in tetraploids rather than hexaploids in most instances (Orjeda et al., 1991). The second hypothesis suggests that the hybridization between sweetpotato and *I. trifida* produced a new allotetraploid entity that subsequently hybridized with *I. trifida* to form a new hexaploid. The newly formed hexaploid then repeatedly backcrossed with the original hexaploid *I. batatas*, thus progressively losing the *I. trifida* component of its nuclear genome while maintaining an *I. trifida*-like chloroplast (Muñoz-Rodríguez et al., 2018). This explanation appears to be unlikely because it requires two separate hybridization events to form a new allotetraploid and repeated asymmetric hybridizations in a specific direction. Therefore, the extant *I. trifida* is unlikely to have served as the diploid or sole progenitor of sweetpotato. Our study offers a more straightforward explanation for the perplexing discrepancies between the genetic structures of the nuclear and chloroplast genomes. During the formation of sweetpotato, the two progenitors crossed reciprocally, thereby passing on the two types of chloroplast genome to sweetpotato. Meanwhile, the reciprocal crossing maintained the identity of the nuclear genome in sweetpotato, because we observed that nuclear genomes of sweetpotato cultivars formed a monophyletic clade throughout their evolution and domestication. However, aside from reciprocal hybridization, asexual chloroplast capture could also account for the presence of two types of chloroplast genomes in sweetpotato. In such scenarios, chloroplast capture would occur in the absence of detectable nuclear introgression (Stegemann and Bock, 2009; Stegemann et al., 2012).

Nearly half of the gene regions in sweetpotato show signs of conversion between the subgenomes. As detected using HPA, B_1 to B_2 conversion events were approximately three times more frequent than B_2 to B_1 conversions (Figure 5B). Rampant gene conversion and conversion biases have increased genome complexity in sweetpotato and may indicate an important role of gene conversion in sweetpotato genome evolution and domestication. Subgenome-biased conversion has been reported in various allopolyploid crops, including cotton, canola, peanut, and strawberry (Paterson et al., 2012; Chalhoub et al., 2014; Chen et al., 2016; Edger et al., 2019). Nonetheless, the molecular mechanisms underlying such conversion bias remain largely unexplored. For sweetpotato, the dosage effect (of the tetraploid B_2 genome versus the diploid B_1 genome) could account for the more prevalent conversion of B_1 alleles to B_2 alleles, given that gene conversion is a copy number–dependent process (Khakhlova and Bock, 2006). However, the currently phased haplotypes remain fragmented and short (Supplemental Table 5; Supplemental Figure 35). To confirm the gene conversion scenario, extended haplotype phasing using nanopore or PacBio sequences or a fully phased genome assembly will be necessary.

The domestication syndrome of vegetatively propagated field crops includes the mode of reproduction, the yield of edible plant parts, the time and ease of harvest, defense adaptations, and plant architecture (Denham et al., 2020). Several of these phenotypic traits are likely associated with domestication in sweetpotato. Compared with those of its wild relatives, the most remarkable feature of sweetpotato is its large edible storage roots. The evolutionary history of the sweetpotato storage root has remained unclear. Storage roots are found in certain strains of *I. trifida* (Li et al., 2019) (Supplemental Figure 1B; Supplemental Table 1). Significant enlargement of fibrous roots has also been reported in both *I. batatas* 4x and *I. aequatoriensis* (Díaz et al., 1992; Muñoz-Rodríguez et al., 2022), although either the tetraploid accessions included in this study lack tuberous roots or the relevant data are missing (Supplemental Figures 1–3; Supplemental Table 1). The original root characteristics of primitive sweetpotato roots remain unclear because sweetpotato originated during pre-human times (Muñoz-Rodríguez et al., 2018). The wild hexaploid sweetpotato (accession Y601) produces a slightly enlarged storage root (Supplemental Figure 1F; Supplemental Table 1). All available data indicate that both the progenitor species and the wild hexaploid sweetpotato have the potential for storage root formation. However, the typical edible storage roots of modern cultivars are the result of domestication, having been selected by early hunter-gatherers, farmers, and breeders. Thus, the development of a starchy storage root has probably undergone strong selective pressure during the human-led domestication of sweetpotato (Liu, 2017), a claim supported by genomic signatures. Most selective sweeps are functionally related to

Please cite this article in press as: Yan et al., Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato, Molecular Plant (2024), https://doi.org/10.1016/j.molp.2023.12.019

root development. We identified nine candidate domestication genes (*DOF4.6*, two *ARF2s*, *ERF*, three *bHLHs*, *EXPB1*, and *HSP70-1*) that are required for initiation and development of storage roots in sweetpotato. We also identified 11 additional candidate genes (two *CRK2*, *TPL3*, *TPL4*, *HHO3*, *AAR5*, *LBD1*, *LBD33*, *U-box13*, *NAC056*, and *Tmem18*) that are known to participate in root development of *Arabidopsis* and other plants (Schiefelbein and Somerville, 1990; Pi et al., 2015; Ma and Li, 2018). These genes represent promising targets for future functional validation studies and genetic improvement of tuberous root crops.

Starch accumulation is another agronomically important trait that was under selection pressure during domestication of sweetpotato (Liu, 2017). Starch biosynthesis relies on sugars as its essential substrates. Hence, sugar transport is critical for both source–sink dynamics and starch accumulation in the storage root. Our study identified SWEET1, a sugar transporter, as a target of selection during sweetpotato domestication. SWEET1 likely functions as a bidirectional glucose transporter in sweetpotato. Furthermore, both abiotic and biotic resistance have been intensely selected during both human and natural selection (Liu, 2017), as reflected in enriched pathways within selected genomic regions. The entire sweetpotato plant is edible, attracting not only herbivorous insects but also pathogenic viruses, fungi, and bacteria (Ogawa and Komada, 1984; Jansson and Raman, 1991; Gutiérrez et al., 2003; Jang et al., 2004). Therefore, plant defense is essential for both survival and storage root yield in sweetpotato. In this regard, our study identified nine well-known plant resistance genes (*HRLI*, *R1B-16*, two *sporamin B*s, three *N-like*s, *APCB*, and *CAMTA*) that carry signatures of selective sweeps.

The discovery of the two sweetpotato progenitors will accelerate the development of better varieties involving the natural resources of the two progenitor species. In addition, as information on domestication-related genes and their genomic and subgenomic distribution continues to accumulate, new opportunities will become available to improve sweetpotato by increasing yields and developing tailor-made varieties. Clearly, a combination of applied and theoretical approaches (involving computational and systems biology–based models) will be required to meet the challenges involved (Vaughan et al., 2007). The insights into the genomics and domestication of sweetpotato gained from this study will contribute to this goal and aid future breeding and genetic engineering approaches for this important staple crop.

METHODS

Plant materials

Because previous global-scale monographic studies of *Ipomoea* have revealed the phylogenetic relationships among sweetpotato and its wild relatives (Muñoz-Rodríguez et al., 2018, 2019; Wood et al., 2020), we focused our sampling on the closest wild relatives previously identified as progenitors of sweetpotato. Seven diploid wild relatives of sweetpotato (including five accessions of *I. trifida*, one accession of *I. triloba*, and one accession of *I.* sp.), 43 tetraploid wild relatives of sweetpotato (four individuals of *I. tiliacea*, seven individuals of *I. batatas* var. *apiculata*, eight individuals of *I. batatas* 4x, two individuals of *I. tabascana*, 12 individuals of tetraploid *I. aequatoriensis*, and 10 individuals of *Ipomoea* hybrids), and 23 sweetpotato cultivars/landraces were used for phylogenetic analyses of nuclear and chloroplast genomes. Data for 12 accessions were downloaded from NCBI (Muñoz-Rodríguez et al., 2022). For sweetpotato cultivars/landraces, sequencing data from cultivars Taizhong6, Xushu18, Y601, Yuzi263, and Yuzi7 were newly generated in this study. Other data for sweetpotato cultivars/landraces were downloaded from NCBI, including cultivars Tanzania and Beauregard and 16 cultivars in the Mwanga diversity panel (Wu et al., 2018). Three tetraploid hybrids of *I. trifida* and sweetpotato were simulated by randomly sampling reads from *I. trifida* accession CIP698014 and three sweetpotato cultivars/landraces (Xushu18, Y601, and Yuzi7) at a ratio of 1:3 using seqtk (v.1.3) (Shen et al., 2016). Sampled reads were integrated between *I. trifida* and each sweetpotato cultivar. In addition, we incorporated 27 accessions of *I. trifida* with low-depth sequencing data (~1×) specifically for the *Ib*T-DNA analyses. Detailed information on the plant materials is given in Supplemental Table 1 and Supplemental Figures 1–4. The process of generating transgenic sweetpotato is described in the Supplemental Note.

Resequencing and population analysis

Variant calling

The whole-genome resequencing (WGS) paired-end reads were aligned to the reference sweetpotato genome (http://sweetpotao.com/download_genome.html) using bwa-mem v.0.7.17 (Li, 2013) and sorted using samtools v.1.10 (Li et al., 2009) with default parameters. Picard v.2.23.4 (Broad Institute, 2019) was used to label PCR duplicates on the basis of mapping coordinates. 120 369 840 genetic variants, including SNPs and insertions or deletions (indels), were detected as diploid using the Genome Analysis Toolkit v.4.1.8.1 (Poplin et al., 2017). Approximately 88% of raw variants were filtered out using VCFtools v.0.1.17 (Danecek et al., 2011) with the following parameters: –minDP 3 –minQ 30 –max-missing 0.8 –maf 0.05. SNPs were then filtered on the basis of linkage disequilibrium using PLINK v.1.90b6.24 (Purcell et al., 2007) (–indep-pairwise 200 10 0.5) and VCFtools. Finally, a total of 6 326 447 variants were selected and used in phylogenetic and population genetic diversity analyses.

Phylogenetic analysis

Vcf2phylip v.2.7 (Ortiz, 2019) was used to generate a fasta file by concatenating all SNPs from the VCF file. A phylogenetic tree of sweetpotato cultivars/landraces and wild relatives was reconstructed using IQ-TREE v.1.6.12 (Nguyen et al., 2014) with 1000 ultrafast bootstrap replicates. The nucleotide substitution model (GTR+F+I+G4) was selected by IQ-TREE. The phylogenetic tree was rooted with the diploid wild relatives as the outgroup, and all accessions were plotted onto a world map using the R package phytools v.0.7-70 (Revell, 2012). For the coalescence-based tree, we defined 1092 submatrices with 5000 non-overlapping SNPs by splitting the concatenated SNP supermatrix. We built an ML tree for each submatrix with the same methods used for the tree of concatenated SNPs. We then generated a species tree using ASTRAL v.5.7.3 (Zhang et al., 2018). The topology was determined on the basis of the main clade of the taxa, with one to two outlier individuals being ignored.

Population structure analysis

PCA was performed using PLINK v.1.90b6.24 (Purcell et al., 2007). The PCA plots were visualized using the R package ggplot2 (Wickham, 2016). UMAP was performed using the R package umap (Konopka, 2022). The binary PLINK input files were transformed from VCF (variant call format) files using PLINK. Ancestral population stratification was inferred using Admixture v.1.3.0 (Alexander et al., 2009) with ancestral population sizes K = 1–10.

Population genetic diversity

Linkage disequilibrium (LD) decay was calculated using PopLDdecay v.3.27 (Zhang et al., 2019) with default parameters. Nucleotide diversities ($\theta\pi$) were determined for two populations of tetraploid wild relatives (16 accessions of the *I. batatas* 4x group and 12 accessions of *I. aequatoriensis*) and the sweetpotato population (23 cultivars/landraces) with VCFtools v.0.1.17

Please cite this article in press as: Yan et al., Haplotype-based phylogenetic analysis and population genomics uncover the origin and domestication of sweetpotato, Molecular Plant (2024), https://doi.org/10.1016/j.molp.2023.12.019

(Danecek et al., 2011) using a 100-kb sliding window with a 10-kb step size. A composite likelihood approach (XP-CLR) was applied to scan for genome-wide selective sweeps (Chen et al., 2010) using a Python version (https://github.com/hardingnj/xpclr) with a 100-kb sliding window and a 10-kb step size. We considered each tetraploid wild relative population as the reference population and the sweetpotato population as the query population to identify potential evolution/breeding sweeps. To detect selective sweeps, we calculated the π ratios ($\pi_{\text{wild relatives}}/\pi_{\text{sweetpotato}}$) within the sliding windows. Regions with values of the π ratio, F_{ST}, or XP-CLR that ranked in the top 1% were defined as candidate domestication sweeps. Gene Ontology enrichment of domestication genes was assessed with the R package clusterProfiler (Wu et al., 2021).

HPA

We developed the HPA pipeline to investigate the relationship between each tetraploid accession and cultivated sweetpotato (Supplemental Figure 12).

Haplotyping

WGS paired-end reads from the sweetpotato cultivars and *I. batatas* 4x were mapped to the sweetpotato reference genome using bwa-mem v.0.7.17-r1188. FreeBayes v.1.3.1-17-gaa2ace8 (Garrison and Marth, 2012) was used to call variants (setting -p 6 for sweetpotato and -p 4 for the tetraploid wild relative). Ranbow v.2.0 (Moeinzadeh et al., 2020) was used for genome haplotyping.

Phylogenetic analysis

The syntenic haplotype blocks between each sweetpotato cultivar and each tetraploid accession were extracted and filtered using the HPA pipeline. Sequences within each syntenic haplotype block were aligned using MAFFT v.7.471 (Katoh and Standley, 2013). For each syntenic haplotype block, the UPGMA tree and ML tree were reconstructed independently using MEGA-CC v.10.1.8 (Kumar et al., 2012, 2018) and IQ-TREE, respectively. The monophyletic ratio, Nsp–Nwr distance, and PI index were calculated using the HPA pipeline. To enhance accuracy, only trees that had the same monophyletic judgment with two tree-building methods (i.e., trees generated based on the same syntenic block by two methods were both monophyletic or both not monophyletic) were used to calculate the monophyletic ratio, Nsp–Nwr distance, and PI index. For detailed identification procedures, refer to the Supplemental Note.

Gene conversion

The syntenic haplotype blocks located within gene regions, containing six haplotypes of sweetpotato and four haplotypes of *I. batatas* 4x, were extracted to detect gene conversion between subgenomes. Disregarding reverse gene conversion, if there wasno gene conversion in a specific syntenic haplotype block, the block is expected to have two B_1 subgenome haplotypes and four B_2 subgenome haplotypes from sweetpotato and four B_2 subgenome haplotypes from *I. batatas* 4x. The resulting phylogenetic tree is expected to form two clades, each corresponding to haplotypes of a particular subgenome. If a gene was converted between subgenomes, then the number of haplotypes and the tree topology is expected to vary. Gene conversions were identified by examining the tree topology (Supplemental Figure 32). The detailed procedures for determining gene conversion are provided in the Supplemental Note.

*Ib*T-DNA analysis

IbT-DNA detection

PCR detection of *Ib*T-DNA1 and *Ib*T-DNA2 genes was performed as described previously by Quispe-Huamanquispe et al. (2019). The WGS paired-end reads were aligned to the *Ib*T-DNA1 and *Ib*T-DNA2 reference sequences (GenBank: KM052616 and KM052617) using bwa-mem v.0.7.17 (Li, 2013). The bam files were visualized in Integrative Genomics Viewer v.2.8.2 (Thorvaldsdóttir et al., 2013) to check the presence/absence of T-DNA insertions.

Phylogenetic and structure variation

Picard v.2.23.4 (Broad Institute, 2019) was used to label PCR duplicates on the basis of the mapping coordinates of the bam files. Genetic variants, including SNPs and indels, were detected as diploid using the Genome Analysis Toolkit v.4.1.8.1 (Poplin et al., 2017). The phylogenetic analysis methods were the same as those described under Resequencing and population analysis. The PMB+F+G4 nucleotide substitution model was selected for *Ib*T-DNA1, and the TVM+F model was selected for *Ib*T-DNA2.

BEDTools genomecov v.2.25.0 (Quinlan, 2014) was used to calculate the sequencing depth of each locus and record the results as a bedgraph file. Indel variations were extracted from VCF files and incorporated into the bedgraph file. The R package ggplot2 (Wickham, 2016) was used to visualize the annotated bedgraph files, displaying a schematic of the T-DNA structure and the indels for T-DNA-positive accessions.

Chloroplast genome assembly and phylogenetic analysis

The chloroplast genomes were assembled using SPAdes v.3.14.1 (Prjibelski et al., 2020) and GetOrganelle v.1.7.5 (Jin et al., 2020) and ranged in size from 160 892 to 161 955 bp. The chloroplast genome sequences were aligned with MAFFT (Katoh and Standley, 2013) as the default, and the alignment was further refined using MUSCLE (Edgar, 2004) implemented in MEGA X (Kumar et al., 2018). Gblocks v.0.91b (Castresana, 2000) was used to remove poorly aligned positions (-b4=5 -b5=h), resulting in a final alignment length of 161 200 bp. The phylogeny was reconstructed using IQ-TREE v.1.6.12 (Nguyen et al., 2014) with 1000 ultrafast bootstrap replicates. The nucleotide substitution model (TVM+F+I) was selected by IQ-TREE. The chloroplast network was generated using the TCS Network method implemented in PopART v.1.7 (Leigh and Bryant, 2015).

RNA sequencing (RNA-seq) analysis

Total RNA was isolated from leaf, stem, and root samples of the sweetpotato cultivar Xushu22. A total of 2 μg RNA per sample was used as input material for RNA sample preparation. Sequencing libraries were generated using the NEBNext Ultra RNA Library Prep Kit for Illumina (E7530L, New England Biolabs, USA). The RNA concentration of sequencing libraries was measured using a Qubit 3.0 fluorometer, and insert size was assessed using the Bioanalyzer 2100 system (Agilent Technologies, CA, USA). The libraries were sequenced on the Illumina NovaSeq 6000 platform to generate 150-bp paired-end reads. We also downloaded RNA-seq data from different root developmental stages of the sweetpotato cultivar Taizhong6 (He et al., 2021). The raw data were first processed with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to remove adapters and low-quality sequences. The RNA-seq reads were mapped to the reference genome using HISAT2 v.2.1.0 (Kim et al., 2019). We assembled and quantified expressed genes using StringTie v.2.1.4 (Pertea et al., 2015) to calculate the gene expression levels in each sample as fragments per kilobase of transcript per million fragments mapped. The expression matrix was extracted using the R package Ballgown (Frazee et al., 2015), and the expression heatmap was visualized using the R package pheatmap (Kolde, 2019). The method of RT–qPCR analysis is described in the Supplemental Note.

DATA AND CODE AVAILABILITY

The raw DNA sequencing data have been deposited in BIGD under accession number PRJCA004953.

The HPA pipeline and relevant instructions are available at the Github website (https://github.com/YanMengxiao/HPA). Other analysis command lines are given in the Supplemental Information.

SUPPLEMENTAL INFORMATION

Supplemental information is available at *Molecular Plant Online*.

FUNDING

This work was funded by the Ministry of Science and Technology of the People's Republic of China (2019YFD1000703 to J.Y., 2019YFD1000704-2 to M.Y., and 2019YFD1000701-2 to W.F.), the National Natural Science Foundation of China (32300207 to M.Y., 32272228 to M.L., and 31771854 to H.W.), the "1 + 9" Open Competition Project of the Sichuan Academy of Agricultural Sciences to select the best candidates (sweetpotato part of 1+9KJGG001 to M.L.), the Chongqing Normal University Foundation (23XLB033 to M.L.), the Shanghai Municipal Afforestation & City Appearance and Environmental Sanitation Administration (G222413 to M.Y., G222411 to H.W., G232405 to H.N., and G242407 to W.F.), the Science and Technology Commission of Shanghai Municipality (22JC1401300 to H.W.), the Youth Innovation Promotion Association CAS (to J.Y.) and the Bureau of Science and Technology for Development CAS (KFJ-BRP-017-42 to J.Y.).

AUTHOR CONTRIBUTIONS

M.Y. was involved in the conception of this study, developed the HPA pipeline, conducted most of the data analyses, and wrote the manuscript. M.L. provided a portion of the plant materials and sequencing data. Y.W. assembled chloroplast genomes and performed the T-DNA analysis. X.W. reconstructed the chloroplast network. M.-H.M. updated the Ranbow software. D.G.Q.-H. and H.N. performed the PCR screening for *IbT*-DNA. W.F. prepared materials for RNA sequencing and performed the PCR screening for *IbT*-DNA. Y.F. performed RT–qPCR analysis and morphological analysis of wild-type and *sporamin* RNAi lines. Y.W. was involved in assembly of chloroplast genomes of MDP. Z.W. helped to access the plants and DNA samples. A.T. provided morphology photos. B.H. tracked back the collection information for the CIP accessions. J.F.K. and G.G. discussed and contributed to the early phase of the project. H.W. revised the manuscript. M.V. discussed and contributed to the haplotyping analysis. R.B. suggested and contributed to the gene conversion analysis and revised the manuscript. J.Y. designed the study and contributed to the original concept of the project.

ACKNOWLEDGMENTS

We thank Dr. Robert Jarret from USDA-ARS for providing materials and revising the manuscript. We also thank NARO Genebank, especially Dr. Masaru Tanaka and Dr. Hiroshi Nemoto, for providing access to their materials and morphology photos. We thank Dr. Fenhong Hu and Prof. Xin Zhou for their important advice during proof-reading. No conflict of interest is declared.

Received: June 7, 2023
Revised: November 10, 2023
Accepted: December 25, 2023

REFERENCES

Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. **19**:1655–1664.

An, H., Qi, X.S., Gaynor, M.L., Hao, Y., Gebken, S.C., Mabry, M.E., McAlvay, A.C., Teakle, G.R., Conant, G.C., Barker, M.S., et al. (2019). Transcriptome and organellar sequencing highlights the complex origin and diversification of allotetraploid *Brassica napus*. Nat. Commun. **10**:2878.

Austin, D.F. (1988). The taxonomy, evolution and genetic diversity of sweet potatoes and related wild species. In Exploration, Maintenance, and Utilization of Sweetpotato Genetic Resources (Lima, Peru: International Potato Center), pp. 27–60.

Berckmans, B., Vassileva, V., Schmid, S.P.C., Maes, S., Parizot, B., Naramoto, S., Magyar, Z., Alvim Kamei, C.L., Koncz, C., Bögre, L., et al. (2011). Auxin-dependent cell cycle reactivation through transcriptional regulation of *Arabidopsis E2Fa* by lateral organ boundary proteins. Plant Cell **23**:3671–3683.

Biswas, K.K., Ooura, C., Higuchi, K., Miyazaki, Y., Van Nguyen, V., Rahman, A., Uchimiya, H., Kiyosue, T., Koshiba, T., Tanaka, A., et al. (2007). Genetic characterization of mutants resistant to the antiauxin *p*-chlorophenoxyisobutyric acid reveals that *AAR3*, a gene encoding a DCN1-like protein, regulates responses to the synthetic auxin 2, 4-dichlorophenoxyacetic acid in *Arabidopsis* roots. Plant Physiol. **145**:773–785.

Broad Institute. (2019). Picard toolkit. http://broadinstitute.github.io/picard/.

Cao, Q., Shao, H.H., Gu, Y.H., Tao, X., Huang, C.L., and Zhang, Y.Z. (2012). Cloning, characterization and expression analysis of HSP70 gene from sweet potato. Afr. J. Microbiol. Res. **6**:6325–6332.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. **17**:540–552.

Cenci, A., Combes, M.-C., and Lashermes, P. (2012). Genome evolution in diploid and tetraploid Coffea species as revealed by comparative analysis of orthologous genome segments. Plant Mol. Biol. **78**:135–145.

Cervantes-Flores, J.C., Yencho, G.C., Kriegner, A., Pecota, K.V., Faulk, M.A., Mwanga, R.O.M., and Sosinski, B.R. (2008). Development of a genetic linkage map and identification of homologous linkage groups in sweetpotato using multiple-dose AFLP markers. Mol. Breed. **21**:511–532.

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I.A.P., Tang, H., Wang, X., Chiquet, J., Belcram, H., Tong, C., Samans, B., et al. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. Science **345**:950–953.

Chen, H., Patterson, N., and Reich, D. (2010). Population differentiation as a test for selective sweeps. Genome Res. **20**:393–402.

Chen, X., Li, H., Pandey, M.K., Yang, Q., Wang, X., Garg, V., Li, H., Chi, X., Doddamani, D., Hong, Y., et al. (2016). Draft genome of the peanut A-genome progenitor (*Arachis duranensis*) provides insights into geocarpy, oil biosynthesis, and allergens. Proc. Natl. Acad. Sci. USA **113**:6785–6790.

Comai, L. (2005). The advantages and disadvantages of being polyploid. Nat. Rev. Genet. **6**:836–846.

Díaz, J., De La Puente, F., and Austin, D.F. (1992). Enlargement of fibrous roots in *Ipomoea* section Batatas (Convolvulaceae). Econ. Bot. **46**:322–329.

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. (2011). The variant call format and VCFtools. Bioinformatics **27**:2156–2158.

Das, M.K., Bai, G., Mujeeb-Kazi, A., and Rajaram, S. (2016). Genetic diversity among synthetic hexaploid wheat accessions (*Triticum aestivum*) with resistance to several fungal diseases. Genet. Resour. Crop Evol. **63**:1285–1296.

Denham, T., Barton, H., Castillo, C., Crowther, A., Dotte-Sarout, E., Florin, S.A., Pritchard, J., Barron, A., Zhang, Y., and Fuller, D.Q. (2020). The domestication syndrome in vegetatively propagated field crops. Ann. Bot. **125**:581–597.

Dhankher, O.P., and Foyer, C.H. (2018). Climate resilient crops for improving global food security and safety. Plant Cell Environ. **41**:877–884.

Dong, W., Li, L., Cao, R., Xu, S., Cheng, L., Yu, M., Lv, Z., and Lu, G. (2020). Changes in cell wall components and polysaccharide-degrading enzymes in relation to differences in texture during sweetpotato storage root growth. J. Plant Physiol. **254**, 153282.

Dou, X.Y., Yang, K.Z., Ma, Z.X., Chen, L.Q., Zhang, X.Q., Bai, J.R., and Ye, D. (2016). *AtTMEM18* plays important roles in pollen tube and vegetative growth in *Arabidopsis*. J. Integr. Plant Biol. **58**:679–692.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**:1792–1797.

Edger, P.P., Poorten, T.J., VanBuren, R., Hardigan, M.A., Colle, M., McKain, M.R., Smith, R.D., Teresi, S.J., Nelson, A.D.L., Wai, C.M., et al. (2019). Origin and evolution of the octoploid strawberry genome. Nat. Genet. **51**:541–547.

Espinosa-Ruiz, A., Martínez, C., de Lucas, M., Fàbregas, N., Bosch, N., Caño-Delgado, A.I., and Prat, S. (2017). TOPLESS mediates brassinosteroid control of shoot boundaries and root meristem development in *Arabidopsis thaliana*. Development **144**:1619–1628.

Food and Agriculture Organization. (2019). FAOSTAT Statistics Database. http://www.fao.org/faostat/.

Frazee, A.C., Pertea, G., Jaffe, A.E., Langmead, B., Salzberg, S.L., and Leek, J.T. (2015). Ballgown bridges the gap between transcriptome assembly and expression analysis. Nat. Biotechnol. **33**:243–246.

Gao, M., Soriano, S.F., Cao, Q., Yang, X., and Lu, G. (2020). Hexaploid sweetpotato (*Ipomoea batatas* (L.) Lam.) may not be a true type to either auto-or allopolyploid. PLoS One **15**, e0229624.

Gao, M., Ashu, G.M., Stewart, L., Akwe, W.A., Njiti, V., and Barnes, S. (2011). *Wx* intron variations support an allohexaploid origin of the sweetpotato [*Ipomoea batatas* (L.) Lam]. Euphytica **177**:111–133.

Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. Preprint at arXiv. https://doi.org/10.48550/arXiv.1207.3907.

George, M.S., Lu, G., and Zhou, W. (2002). Genotypic variation for potassium uptake and utilization efficiency in sweet potato (*Ipomoea batatas* L.). Field Crops Res. **77**:7–15.

Godfray, H.C.J., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Nisbett, N., Pretty, J., Robinson, S., Toulmin, C., and Whiteley, R. (2010). The future of the global food system. Philos. Trans. R. Soc. Lond. B Biol. Sci. **365**:2769–2777.

Guillemin, F., Guillon, F., Bonnin, E., Devaux, M.-F., Chevalier, T., Knox, J.P., Liners, F., and Thibault, J.-F. (2005). Distribution of pectic epitopes in cell walls of the sugar beet root. Planta **222**:355–371.

Guo, Z.H., Ma, P.F., Yang, G.Q., Hu, J.Y., Liu, Y.L., Xia, E.H., Zhong, M.C., Zhao, L., Sun, G.L., Xu, Y.X., et al. (2019). Genome sequences provide insights into the reticulate origin and unique traits of woody bamboos. Mol. Plant **12**:1353–1365.

Gutiérrez, D.L., Fuentes, S., and Salazar, L.F. (2003). Sweetpotato virus disease (SPVD): distribution, incidence, and effect on sweetpotato yield in Peru. Plant Dis. **87**:297–302.

Hajjar, R., and Hodgkin, T. (2007). The use of wild relatives in crop improvement: a survey of developments over the last 20 years. Euphytica **156**:1–13.

He, S., Wang, H., Hao, X., Wu, Y., Bian, X., Yin, M., Zhang, Y., Fan, W., Dai, H., Yuan, L., et al. (2021). Dynamic network biomarker analysis discovers *IbNAC083* in the initiation and regulation of sweet potato root tuberization. Plant J. **108**:793–813.

Hollister, J.D. (2015). Polyploidy: adaptation to the genomic environment. New Phytol. **205**:1034–1039.

Huang, J.C., and Sun, M. (2000). Genetic diversity and relationships of sweetpotato and its wild relatives in *Ipomoea* series *Batatas* (Convolvulaceae) as revealed by inter-simple sequence repeat (ISSR) and restriction analysis of chloroplast DNA. Theor. Appl. Genet. **100**:1050–1060.

Hunter, K., Kimura, S., Rokka, A., Tran, H.C., Toyota, M., Kukkonen, J.P., and Wrzaczek, M. (2019). CRK2 enhances salt tolerance by regulating callose deposition in connection with PLD α 1. Plant Physiol. **180**:2004–2021.

Jafarzadeh, J., Bonnett, D., Jannink, J.-L., Akdemir, D., Dreisigacker, S., and Sorrells, M.E. (2016). Breeding value of primary synthetic wheat genotypes for grain yield. PLoS One **11**, e0162860.

Jang, I.C., Park, S.Y., Kim, K.Y., Kwon, S.Y., Kim, J.G., and Kwak, S.S. (2004). Differential expression of 10 sweetpotato peroxidase genes in response to bacterial pathogen, *Pectobacterium chrysanthemi*. Plant Physiol. Biochem. **42**:451–455.

Jansson, R.K., and Raman, K.V. (1991). Sweet Potato Pest Management: A Global Perspective (CRC Press).

Ji, J., Yang, L., Fang, Z., Zhang, Y., Zhuang, M., Lv, H., and Wang, Y. (2022). Plant SWEET family of sugar transporters: structure, evolution and biological functions. Biomolecules **12**:205.

Jin, J.J., Yu, W.B., Yang, J.B., Song, Y., DePamphilis, C.W., Yi, T.S., and Li, D.Z. (2020). GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. **21**:241–331.

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. **30**:772–780.

Khakhlova, O., and Bock, R. (2006). Elimination of deleterious mutations in plastid genomes by gene conversion. Plant J. **46**:85–94.

Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. **37**:907–915.

Kim, D.Y., Lee, Y.J., Hong, M.J., Kim, J.H., and Seo, Y.W. (2021). Genome wide analysis of U-box E3 ubiquitin ligases in wheat (*Triticum aestivum* L.). Int. J. Mol. Sci. **22**:2699.

Kolde, R. (2019). Pheatmap: Pretty Heatmaps. https://CRAN.R-project.org/package=pheatmap.

Kole, C. (2011). Wild Crop Relatives: Genomic and Breeding Resources: Industrial Crops (Springer).

Konopka, T. (2022). Umap: Uniform Manifold Approximation and Projection. https://CRAN.R-project.org/package=umap.

Kriegner, A., Cervantes, J.C., Burg, K., Mwanga, R.O., and Zhang, D. (2003). A genetic linkage map of sweetpotato [*Ipomoea batatas* (L.) Lam.] based on AFLP markers. Mol. Breed. **11**:169–185.

Kumar, S., Stecher, G., Peterson, D., and Tamura, K. (2012). MEGA-CC: computing core of molecular evolutionary genetics analysis program for automated and iterative data analysis. Bioinformatics **28**:2685–2686.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. **35**:1547–1549.

Kurabachew, H. (2015). The role of orange fleshed sweet potato *(Ipomea batatas*) for combating vitamin A deficiency in Ethiopia: A review. Int. J. Food Sci. Nutr. Eng. **5**:141–146.

Kwak, S.S. (2019). Biotechnology of the sweetpotato: ensuring global food and nutrition security in the face of climate change. Plant Cell Rep. **38**:1361–1363.

Kyndt, T., Quispe, D., Zhai, H., Jarret, R., Ghislain, M., Liu, Q., Gheysen, G., and Kreuze, J.F. (2015). The genome of cultivated sweet potato contains *Agrobacterium* T-DNAs with expressed genes: an example of a naturally transgenic food crop. Proc. Natl. Acad. Sci. USA **112**:5844–5849.

Leigh, J.W., and Bryant, D. (2015). POPART: full-feature software for haplotype network construction. Methods Ecol. Evol. **6**:1110–1116.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at arXiv. https://doi.org/10.48550/arXiv.1303.3997.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics **25**:2078–2079.

Li, M., Yang, S., Xu, W., Pu, Z., Feng, J., Wang, Z., Zhang, C., Peng, M., Du, C., Lin, F., et al. (2019). The wild sweetpotato (*Ipomoea trifida*) genome provides insights into storage root development. BMC Plant Biol. **19**:119.

Li, Q., Zhou, L., Li, Y., Zhang, D., and Gao, Y. (2021). Plant NIGT1/HRS1/HHO transcription factors: key regulators with multiple roles in plant growth, development, and stress responses. Int. J. Mol. Sci. **22**:8685.

Liu, Q.C. (2017). Improvement for agronomically important traits by gene engineering in sweetpotato. Breed Sci. 16126.

Lu, K., Wei, L., Li, X., Wang, Y., Wu, J., Liu, M., Zhang, C., Chen, Z., Xiao, Z., Jian, H., et al. (2019). Whole-genome resequencing reveals *Brassica napus* origin and genetic loci involved in its improvement. Nat. Commun. **10**:1154.

Ma, L., and Li, G. (2018). FAR1-related sequence (FRS) and FRS-related factor (FRF) family proteins in *Arabidopsis* growth and development. Front. Plant Sci. **9**:692.

Magoon, M.L., Krishnan, R., and Vijaya Bai, K. (1970). Cytological evidence on the origin of sweet potato. Theor. Appl. Genet. **40**:360–366.

Mayfield, J.D., Paul, A.-L., and Ferl, R.J. (2012). The 14-3-3 proteins of *Arabidopsis* regulate root growth and chloroplast development as components of the photosensory system. J. Exp. Bot. **63**:3061–3070.

Moeinzadeh, M.-H., Yang, J., Muzychenko, E., Gallone, G., Heller, D., Reinert, K., Haas, S., and Vingron, M. (2020). Ranbow: A fast and accurate method for polyploid haplotype reconstruction. PLoS Comput. Biol. **16**, e1007843.

Mollinari, M., Olukolu, B.A., Pereira, G.d.S., Khan, A., Gemenet, D., Yencho, G.C., and Zeng, Z.-B. (2020). Unraveling the hexaploid sweetpotato inheritance using ultra-dense multilocus mapping. G3 (Bethesda). **10**:281–292.

Muñoz-Rodríguez, P., Wells, T., Wood, J.R., Carruthers, T., Anglin, N.L., Jarret, R.L., and Scotland, R.W. (2022). Discovery and Characterisation of Sweetpotato's Closest Tetraploid Relative. New Phytol.

Muñoz-Rodríguez, P., Carruthers, T., Wood, J.R.I., Williams, B.R.M., Weitemier, K., Kronmiller, B., Ellis, D., Anglin, N.L., Longway, L., Harris, S.A., et al. (2018). Reconciling conflicting phylogenies in the origin of sweet potato and dispersal to Polynesia. Curr. Biol. **28**:1246–1256.e12.

Muñoz-Rodríguez, P., Carruthers, T., Wood, J.R.I., Williams, B.R.M., Weitemier, K., Kronmiller, B., Goodwin, Z., Sumadijaya, A., Anglin, N.L., Filer, D., et al. (2019). A taxonomic monograph of *Ipomoea* integrated across phylogenetic scales. Nat. Plants **5**:1136–1144.

Nedunchezhiyan, M., and Ray, R.C. (2010). Sweet potato growth, development production and utilization: overview. In Sweet Potato: Post-Harvest Aspects in Food, Feed and Industry (Food Science and Technology), R.C. Ray and K.I. Tomlins, eds. (New York: Nova Science Publishers Inc.), pp. 1–26.

Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. **32**:268–274.

Nishiyama, I. (1971). Evolution and domestication of the sweet potato. Bot. Mag. **84**:377–387.

Nishiyama, I., Miyazaki, T., and Sakamoto, S. (1975). Evolutionary autoploidy in the sweet potato (*Ipomoea batatas* (L.) Lam.) and its progenitors. Euphytica **24**:197–208.

Noh, S.A., Lee, H.-S., Kim, Y.-S., Paek, K.-H., Shin, J.S., and Bae, J.M. (2013). Down-regulation of the *IbEXP1* gene enhanced storage root development in sweetpotato. J. Exp. Bot. **64**:129–142.

Ogawa, K., and Komada, H. (1984). Biological control of Fusarium wilt of sweet potato by non-pathogenic *Fusarium oxysorum*. JJP **50**:1–9.

Orjeda, G., Freyre, R., and Iwanaga, M. (1991). Use of *Ipomoea trifida* germ plasm for sweet potato improvement. 3. Development of 4x interspecific hybrids between *Ipomoea batatas* (L.) Lam. (2n=6x=90) and *I. trifida* (H.B.K) G. Don. (2n=2x=30) as storage-root initiators for wild species. Theor. Appl. Genet. **83**:159–163.

Ortiz, E.M. (2019). vcf2phylip v2.0: Convert a VCF Matrix into Several Matrix Formats for Phylogenetic Analysis.

Padmaja, G. (2009). Uses and nutritional data of sweetpotato. In The Sweetpotato, G. Loebenstein and G. Thottappilly, eds. (Dordrecht: Springer), pp. 189–234.

Paterson, A.H., Wendel, J.F., Gundlach, H., Guo, H., Jenkins, J., Jin, D., Llewellyn, D., Showmaker, K.C., Shu, S., Udall, J., et al. (2012). Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. Nature **492**:423–427.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. **33**:290–295.

Pi, L., Aichinger, E., van der Graaff, E., Llavata-Peris, C.I., Weijers, D., Hennig, L., Groot, E., and Laux, T. (2015). Organizer-derived WOX5 signal maintains root columella stem cells through chromatin-mediated repression of *CDF4* expression. Dev. Cell **33**:576–588.

Poplin, R., Ruano-Rubio, V., DePristo, M.A., Fennell, T.J., Carneiro, M.O., Van der Auwera, G.A., Kling, D.E., Gauthier, L.D., Levy-Moonshine, A., and Roazen, D. (2017). Scaling accurate genetic variant discovery to tens of thousands of samples. Preprint at bioRxiv. https://doi.org/10.1101/201178.

Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., and Korobeynikov, A. (2020). Using SPAdes de novo assembler. Curr. Protoc. Bioinfor. **70**, e102.

Prosekov, A.Y., and Ivanova, S.A. (2018). Food security: The challenge of the present. Geoforum **91**:73–77.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. **81**:559–575.

Quinlan, A.R. (2014). BEDTools: the Swiss-army tool for genome feature analysis. Curr. Protoc. Bioinfor. **47**:11.12.11–11.12.34.

Quispe-Huamanquispe, D.G., Gheysen, G., Yang, J., Jarret, R., Rossel, G., and Kreuze, J.F. (2019). The horizontal gene transfer of *Agrobacterium* T-DNAs into the series Batatas (Genus *Ipomoea*) genome is not confined to hexaploid sweetpotato. Sci. Rep. **9**:12584–12613.

Rajapakse, S., Nilmalgoda, S.D., Molnar, M., Ballard, R.E., Austin, D.F., and Bohac, J.R. (2004). Phylogenetic relationships of the sweetpotato in *Ipomoea* series Batatas (Convolvulaceae) based on nuclear β-amylase gene sequences. Mol. Phylogenet. Evol. **30**:623–632.

Ravi, V., Chakrabarti, S., Makeshkumar, T., and Saravanan, R. (2014). Molecular regulation of storage root formation and development in sweet potato. Hortic. Rev. **42**:157–208.

Ravi, V., Chakrabarti, S.K., Saravanan, R., Makeshkumar, T., and Sreekumar, J. (2017). Differential gene expression signatures of auxin response factors and auxin/indole 3-acetic acid genes in storage root as compared to non-tuber forming fibrous root of sweet potato (*Ipomoea batatas*). Indian J. Agric. Sci. **87**:512–520.

Revell, L.J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. **3**:217–223.

Roullier, C., Duputié, A., Wennekes, P., Benoit, L., Fernández Bringas, V.M., Rossel, G., Tay, D., McKey, D., and Lebot, V. (2013). Disentangling the origins of cultivated sweet potato (*Ipomoea batatas* (L.) Lam.). PLoS One **8**, e62707.

Schiefelbein, J.W., and Somerville, C. (1990). Genetic control of root hair development in *Arabidopsis thaliana*. Plant Cell **2**:235–243.

Senthilkumar, R., and Yeh, K.W. (2012). Multiple biological functions of sporamin related to stress tolerance in sweet potato (*Ipomoea batatas* Lam). Biotechnol. Adv. **30**:1309–1317.

Shen, W., Le, S., Li, Y., and Hu, F. (2016). SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q File manipulation. PLoS One **11**, e0163962.

Shiotani, I. (1988). Genomic structure and the gene flow in sweet potato and related species. In Exploration and Maintenance and Utilization of Sweet Potato Genetic Resources (Lima, Peru: International Potato Centre), pp. 61–73.

Shiotani, I., and Kawase, T. (1987). Synthetic hexaploids derived from wild species related to sweet potato. Jpn. J. Breed. **37**:367–376.

Shiotani, I., and Kawase, T. (1989). Genomic structure of the sweet potato and hexaploids in *Ipomoea trifida* (HBK) Don. Jpn. J. Breed. **39**:57–66.

Srisuwan, S., Sihachakr, D., and Siljak-Yakovlev, S. (2006). The origin and evolution of sweet potato (*Ipomoea batatas* Lam.) and its wild relatives through the cytogenetic approaches. Plant Sci. **171**:424–433.

Stegemann, S., and Bock, R. (2009). Exchange of genetic material between cells in plant tissue grafts. Science **324**:649–651.

Stegemann, S., Keuthe, M., Greiner, S., and Bock, R. (2012). Horizontal transfer of chloroplast genomes between plant species. Proc. Natl. Acad. Sci. USA **109**:2434–2438.

Tanaka, M., Takahata, Y., Nakayama, H., Nakatani, M., and Tahara, M. (2009). Altered carbohydrate metabolism in the storage roots of sweetpotato plants overexpressing the *SRF1* gene, which encodes a Dof zinc finger transcription factor. Planta **230**:737–746.

Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings Bioinf. **14**:178–192.

Tian, Z., Wang, J.W., Li, J., and Han, B. (2021). Designing future crops: challenges and strategies for sustainable agriculture. Plant J. **105**:1165–1178.

Ukoskit, K., and Thompson, P.G. (1997). Autopolyploidy versus allopolyploidy and low-density randomly amplified polymorphic DNA linkage maps of sweetpotato. J. Am. Soc. Hortic. Sci. **122**:822–828.

Vaughan, D.A., Balázs, E., and Heslop-Harrison, J.S. (2007). From crop domestication to super-domestication. Ann. Bot. **100**:893–901.

Wallace, J.G., Rodgers-Melnick, E., and Buckler, E.S. (2018). On the road to breeding 4.0: unraveling the good, the bad, and the boring of crop quantitative genomics. Annu. Rev. Genet. **52**:421–444.

Wang, X.Y., and Paterson, A.H. (2011). Gene conversion in angiosperm genomes with an emphasis on genes duplicated by polyploidization. Genes **2**:1–20.

Wheeler, T., and Von Braun, J. (2013). Climate change impacts on global food security. Science **341**:508–513.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (New York: Springer-Verlag).

Wood, J.R.I., Muñoz-Rodríguez, P., Williams, B.R.M., and Scotland, R.W. (2020). A foundation monograph of *Ipomoea* (Convolvulaceae) in the New World. PhytoKeys **143**:1–823.

Wu, S., Lau, K.H., Cao, Q., Hamilton, J.P., Sun, H., Zhou, C., Eserman, L., Gemenet, D.C., Olukolu, B.A., Wang, H., et al. (2018). Genome sequences of two diploid wild relatives of cultivated sweetpotato reveal targets for genetic improvement. Nat. Commun. **9**:4580–4612.

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation **2**, 100141.

Xu, P., Ma, W., Hu, J., and Cai, W. (2022). The nitrate-inducible NAC transcription factor *NAC056* controls nitrate assimilation and promotes lateral root growth in *Arabidopsis thaliana*. PLoS Genet. **18**, e1010090.

Yamauchi, T., Tanaka, A., Inahashi, H., Nishizawa, N.K., Tsutsumi, N., Inukai, Y., and Nakazono, M. (2019). Fine control of aerenchyma and lateral root development through AUX/IAA-and ARF-dependent auxin signaling. Proc. Natl. Acad. Sci. USA **116**:20770–20775.

Yan, M., Nie, H., Wang, Y., Wang, X., Jarret, R., Zhao, J., Wang, H., and Yang, J. (2022). Exploring and exploiting genetics and genomics for sweetpotato improvement: Status and perspectives. Plant Commun. **3**, 100332.

Yan, M., Li, M., Moeinzadeh, M.-H., Quispe-Huamanquispe, D.G., Fan, W., Nie, H., Wang, Z., Heider, B., Jarret, R., and Kreuze, J. (2021). Haplotype-based phylogenetic analysis uncovers the tetraploid progenitor of sweet potato. Res. Sq.

Yang, J., Moeinzadeh, M.H., Kuhl, H., Helmuth, J., Xiao, P., Haas, S., Liu, G., Zheng, J., Sun, Z., Fan, W., et al. (2017). Haplotype-resolved sweet potato genome traces back its hexaploidization history. Nat. Plants **3**:696–703.

Ye, C.Y., Wu, D., Mao, L., Jia, L., Qiu, J., Lao, S., Chen, M., Jiang, B., Tang, W., Peng, Q., et al. (2020). The genomes of the allohexaploid *Echinochloa crus-galli* and its progenitors provide insights into polyploidization-driven adaptation. Mol. Plant **13**:1298–1310.

Zhang, C., Rabiee, M., Sayyari, E., and Mirarab, S. (2018). ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinf. **19**:153–230.

Zhang, C., Dong, S.S., Xu, J.Y., He, W.M., and Yang, T.L. (2019). PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. Bioinformatics **35**:1786–1788.

Zhang, D., Cervantes, J., Huamán, Z., Carey, E., and Ghislain, M. (2000). Assessing genetic diversity of sweet potato (*Ipomoea batatas* (L.) Lam.) cultivars from tropical America using AFLP. Genet. Resour. Crop Evol. **47**:659–665.

Zhao, N., Yu, X., Jie, Q., Li, H., Li, H., Hu, J., Zhai, H., He, S., and Liu, Q. (2013). A genetic linkage map based on AFLP and SSR markers and mapping of QTL for dry-matter content in sweetpotato. Mol. Breed. **32**:807–820.

Zhou, Y., Zhao, X., Li, Y., Xu, J., Bi, A., Kang, L., Xu, D., Chen, H., Wang, Y., Wang, Y.G., et al. (2020). *Triticum* population sequencing provides insights into wheat adaptation. Nat. Genet. **52**:1412–1422.