

The jojoba genome reveals wide divergence of the sex chromosomes in a dioecious plant

Othman Al-Dossary^{1,2,†}, Bader Alsubaie^{1,2,†}, Ardashir Kharabian-Masouleh^{1,†}, Ibrahim Al-Mssallem^{2,*}, Agnelo Furtado¹ and Robert J. Henry^{1,3,*} 

¹Queensland Alliance for Agriculture and Food Innovation, University of Queensland, Brisbane 4072, Australia,

²College of Agriculture and Food Sciences, King Faisal University, Al Hofuf 36362, Saudi Arabia, and

³ARC Centre of Excellence for Plant Success in Nature and Agriculture, University of Queensland, Brisbane 4072, Australia

Received 2 August 2021; accepted 20 September 2021; published online 27 September 2021.

*For correspondence (e-mail imssallem@gmail.com; robert.henry@uq.edu.au).

[†]These authors contributed equally to this work.

SUMMARY

Most flowering plants are hermaphrodites, but around 6% of species are dioecious, having separate male and female plants. Sex chromosomes and some sex-specific genes have been reported in plants, but the genome sequences have not been compared. We now report the genome sequence of male and female jojoba (*Simmondsia chinensis*) plants, revealing a very large difference in the sex chromosomes. The male genome assembly was 832 Mb and the female 822 Mb. This was explained by the large size differences in the Y chromosome (37.6 Mb) compared with the X chromosome (26.9 Mb). Relative to the X chromosome, the Y chromosome had two large insertions each of more than 5 Mb containing more than 400 genes. Many of the genes in the chromosome-specific regions were novel. These male-specific regions included many flowering-related and stress response genes. Smaller insertions found only in the X chromosome totalled 877 kb. The wide divergence of the sex chromosomes suggests a long period of adaptation to diverging sex-specific roles. Male and female plants may have evolved to accommodate factors such as differing reproductive resource allocation requirements under the stress of the desert environment in which the plants are found. The sex-determining regions accumulate genes beneficial to each sex. This has required the evolution of many more novel sex-specific genes than has been reported for other organisms. This suggests that dioecious plants provide a novel source of genes for manipulation of reproductive performance and environmental adaptation in crops.

Keywords: sex chromosomes, sexual dimorphism, dioecious plants, jojoba, *Simmondsia chinensis*.

INTRODUCTION

Sex separation found in dioecious plants has evolved in a small portion of plant species for which this must offer some selective advantage (Leite Montalvão *et al.*, 2020). Many different plant families have dioecious species, found in diverse environments. Male and female dioecious plants may differ in their needs for environmental resources due to their distinctive reproductive roles (Thomas and LaFrankie, 1993). Natural selection for reproductive success has played a crucial role in sexual differentiation (Kumar *et al.*, 2016).

Sex differentiation and sex-specific chromosomes have been extensively studied due to their significant role in plant evolutionary biology. The genetic basis of sexual differentiation in dioecious species has been found to be linked to nuclear genes on sex chromosomes (Baránková *et al.*, 2020). Sex chromosomes in dioecious plant are often indistinguishable, and only a few of them show clear variation (Baránková *et al.*, 2020). Many sex chromosomes have been reported in

plants (Charlesworth, 2016; Smith and Smith, 1947) (Charlesworth, 2016; Smith and Smith, 1947) but have not been characterised in the same way as they have in animals (Negrutu *et al.*, 2001). In plant species, sex chromosome models broadly differ at the genus or species level (Charlesworth, 2013). A recent report found two sex-specific genes on a sex chromosome in poplar (*Populus trichocarpa*) (Xue *et al.*, 2020) and many other differences have been reported (Almeida *et al.*, 2020; Atanassov *et al.*, 2001; Massonnet *et al.*, 2020) at the genome level. However, the genome sequences of male and female plants have not yet been reported.

There are many plants with homomorphic sex chromosomes with a small non-recombining sex-determining region for which the sequence has been reported (Charlesworth, 2016). However, here we report the chromosome-level assembly of heteromorphic sex chromosomes in a plant.

Jojoba (*Simmondsia chinensis*) is a dioecious desert shrub that inhabits the Sonoran Desert between the United

States and Mexico. This perennial woody plant is the only species in the family Simmondsiaceae (Kumar et al., 2019). Jojoba is used (Alotaibi et al., 2020; Arya and Khan, 2016; Kumar et al., 2019; Sturtevant et al., 2020) to produce a unique liquid wax ester (jojoba oil), which represents 45–55% of the jojoba female seed weight (Inoti, 2018). Jojoba oil has many important characteristics, which makes it a renewable candidate source of biodiesel production (Sandouqa and Al-Hamamre, 2019). Beside the significant interest in the oil, its cultivation has expanded in arid to semi-arid regions because of its ability to tolerate different stresses and poor soil environments (Wang et al., 2019). The morphological and physiological characteristics of jojoba play an essential role in its adaptation to various stresses as well as its sex-based variation. In a study of biochemical and physiological responses to drought, male jojoba plants expressed higher tolerance than females (Kumar et al., 2016), suggesting very different environmental responses to stress. Jojoba seed germination is male-biased, with a successful male:female germination ratio of 5:1 having been reported (Heikrujam et al., 2015a). This creates a challenge in production with plant sex not being revealed until a very late growing stage. Several efforts have been made to identify the genetic basis of sexual dimorphism using previously identified sex-related genetic markers in other plants (Heikrujam et al., 2015b; Hosseini et al., 2011). However, to understand the sex differences reference genomes for both sexes are needed. The expansion of genomic resources for jojoba will create an opportunity to address differences between the sexes and sex chromosome evolution in a dioecious plant.

Recently, the first reference genome of a jojoba plant of undefined sex was reported (Sturtevant et al., 2020). The 887-Mb genome of jojoba was comprised of 26 chromosomes including 23 490 protein-coding genes and suggested an ancient whole genome triplication with no recent duplications in the jojoba genome. The study focused on machinery for lipid synthesis and storage based on transcriptome, proteome, and lipidome data.

However, to define the sexual dimorphism in jojoba, separate reference genomes for both sexes are required. We now report the sequencing of male and female jojoba plants and analysis of the genomes to define in more detail the differences between the genome sequences of a male and female plant for the first time.

RESULTS AND DISCUSSION

Structural differences between male and female jojoba genomes

The genomes of male and female jojoba plants were sequenced using long-read sequencing (Murigneux et al., 2020) (PacBio HiFi) generating genome assemblies produced using Improved Phased Assembly (IPA) (Sharma

et al., 2021). The male assembly was significantly larger than the female assembly (male 832 Mb [N50 5.7 Mb and Benchmarking Universal Single-Copy Orthologs [BUSCO] completeness 96.9%] and female 822 Mb [N50 4.9 Mb and BUSCO completeness 97.4%]). The k-mer analysis, while supporting a larger male genome, gave a lower estimate of genome size than that provided by the HiFi assembly (Table S1), presumably due to the nature of short sequence reads and the challenge of estimating genome size for genomes with highly repetitive elements (Guo et al., 2015; Panfilio et al., 2019). The male genome had 1.53% heterozygosity and the female 1.31% (Figure S1). Proximity ligation of chromatin fragmented by a sequence-independent endonuclease (Omni-C) was used to produce a chromosome-level assembly of the male jojoba genome (Table S2). Assembly of the HiFi contigs produced pseudomolecules representing all 26 chromosomes (Table S3, Figure 1). Completeness of this assembly was 97% based upon analysis of the presence of conserved viridiplantae genes (BUSCO, Table S4). Jojoba has 52 chromosomes and some have suggested the possibility of it being a tetraploid ($n = 13$, $4n = 52$) (Tobe et al., 1992). However, the assembly and lack of evidence for homology between the pseudomolecules indicated that jojoba should be considered as a diploid ($n = 26$). Analysis of the syntenic regions in the genome identified 346 putative homologous genes and Ks analysis showed evidence of two whole genome duplication events: an ancient whole genome duplication and a more recent one (Figure S2). The jojoba genome has apparently gone through a diploidization process, as suggested by cytological evidence that the chromosomes behaved like those of a diploid (Tobe et al., 1992). Links between polyploidy and the evolution of dioecy have been suggested but remain unclear (Ashman et al., 2013).

The male and female genomes were compared by aligning the HiFi contigs produced by assembly of the PacBio reads with the Hi-C derived pseudomolecules (Figure 2). The results showed that chromosome 9 (chromosomes were numbered in order of decreasing size) differs between the male and female and corresponds to a sex chromosome. This comparative genomic analysis showed a significant difference in the length of chromosome 9 between male and female plants. This suggests an XY chromosome system, in which the Y chromosome (male-specific) is much longer than the X chromosome. The 2D dot plot of the assembled male contigs against the chromosome-level assembly showed that the total length of the Y chromosome was 37.6 Mb, while the X chromosome was only 26.9 Mb, that is, 10.7 Mb shorter than the Y chromosome (Figure 2). This was largely due to the presence of two large insertions in the Y chromosome, one 5.5 Mb (insertion Y₁) and the other 5.2 Mb (insertion Y₂) (Figure 2). These two insertions are genome regions found only in the male (male-specific). We also found insertions

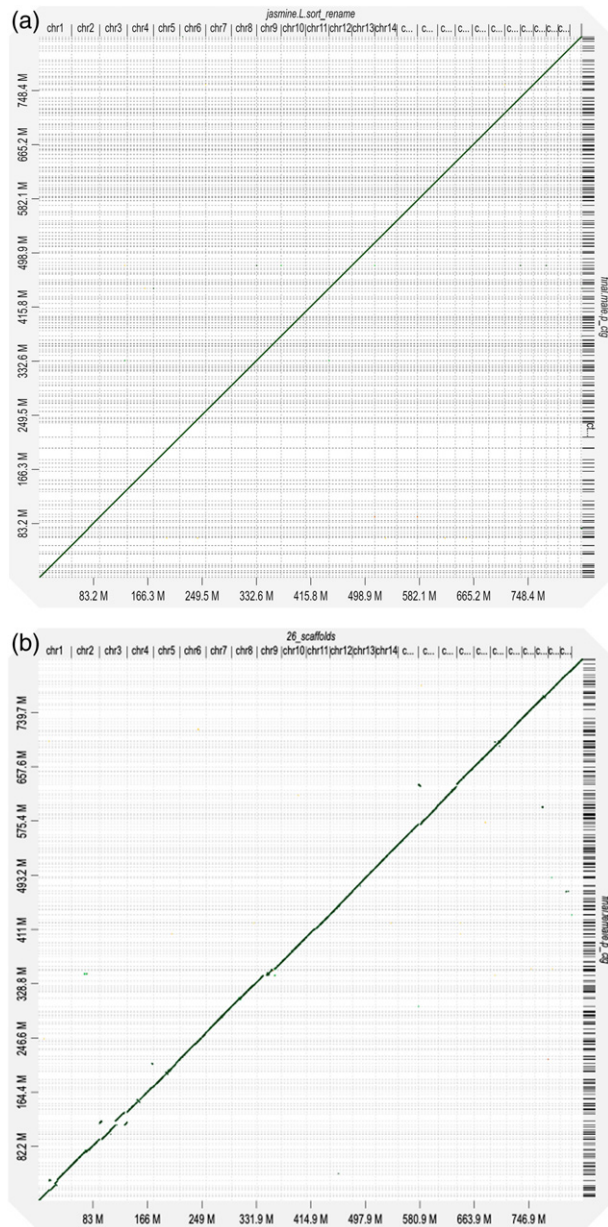


Figure 1. Comparison of (a) HiFi male assembly and (b) HiFi female assembly against male genome chromosome-level assembly (Omni-C scaffolding). The x-axis shows the 26 pseudomolecules representing the chromosomes in the male genome and the y-axis shows the contigs of the HiFi assembly of the (a) male and (b) female genome.

on the X chromosome within contig 265F (ctg.265f) that were not present in chromosome Y (Figure 2). The total length of these insertions was 887 kb, including a 713-kb insertions (insertion X₁) and a 174-kb insertion (insertion X₂). These specific insertions in the X and Y chromosomes together explain the genome size difference between the male and female, which was first detected in the long-read (IPA) assembly (Figure 1) and in the k-mer analysis of short reads (Figure S1). This is the first genomic sequencing of

an XY chromosome system in dioecious plants and suggests that these two large insertions (in the Y chromosome) may largely explain the wide divergence between the two sexes. This differs from the widely reported mammalian system, in which the Y chromosome is much shorter than the X chromosome. Mammalian sex chromosomes are usually thought to be derived from autosomes that have gone through many evolutionary processes, especially in speciation (Presgraves, 2008). However, the larger Y chromosome in *jojoba* contrasts with the classical view in animals that the Y chromosomes represent a highly degenerate version of the X chromosome, in which the Y chromosome has lost most of the ancestral genes (Bachtrog, 2013) and is supported by other reports of large X chromosomes in some plants (Charlesworth, 2016). We suggest that the Y and X insertions may have evolved by segmental genome duplication and divergence events during the evolution of dioecy in this species.

A total of 426,380 transcript isoforms were produced from 296.6 Mb of PacBio long-read cDNA data. These transcripts were mapped to both the male and female assemblies, 99.8% and 99.6% mapping successfully for male and female, respectively. The BUSCO and RNA alignment results indicated a high-quality genome assembly with excellent completeness and accuracy for both sexes.

Annotation of male and female genomes

Annotation of the whole genome shed more light on the characteristics of the sex chromosomes and the two insertions in the Y chromosome (Table 1). The whole genome annotation showed 1616 genes in the Y chromosome.

A large number of genes were found in the regions of the two insertions in the Y chromosome. Of the 429 genes in the two insertions in the Y chromosome, 248 and 181 genes belong to the Y_1 and Y_2 regions, respectively. Comparison of the genes in the male-specific inserts in the Y chromosome showed that only three genes (with no match) in the Y_1 region and none of the genes in the Y_2 region had 90% similarity, establishing that all of the genes in these regions were distinct and not duplicates. Only one gene (with no match) was common between the Y_1 and Y_2 regions. This suggests that the male-specific regions of the jojoba genome have more than 400 unique genes. Only one gene in these regions was also present on the other parts of the Y chromosome.

The identification of many sex-specific genes in jojoba will aid the development of tests to distinguish male and female plants for deployment in commercial crops (Waser, 1984).

Analysis of many of the genes in the Y₁ and Y₂ insertions in the whole genome annotation resulted in matches with known genes. Of a total of 429 genes, 147 (87 from Y₁ and 60 from Y₂) had no matches, 14 matched hypothetical proteins, and 26 matched genes for uncharacterised

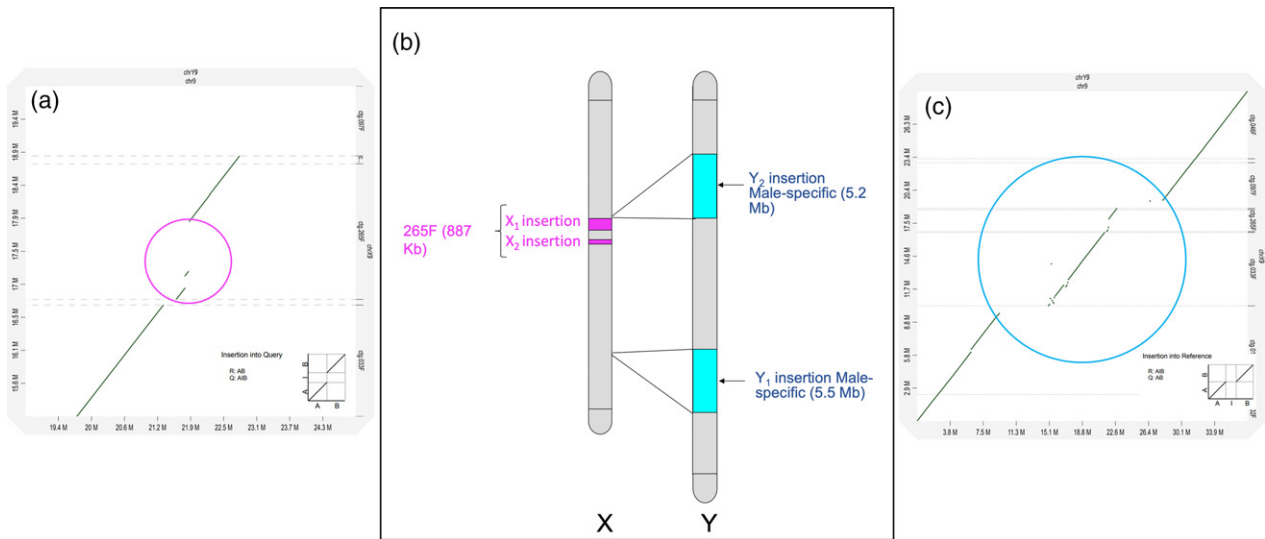


Figure 2. Illustration of sex chromosomes in jojoba. The numbers indicate the name and position of contigs. (a) 2D plot of alignment between the female-specific region within contig 265F and the Y chromosome assembly. The purple circle shows the two insertions in contig 265F. (b) The X chromosome includes 877 Kb in two insertions which are not found in the Y chromosome or other parts of the male genome (female-specific). The Y chromosome contains two male-specific insertions, which are not found in the X chromosome or elsewhere in the female genome. The total length of the male-specific insertions is approximately 10.7 Mb (5.5 Y₁ + 5.2 Mb Y₂). (c) 2D dot plot of alignment between the female genome assembly (contigs) and the Y chromosome assembly. The blue circle shows the two insertions into the Y chromosome (ChrY9).

Table 1 Annotation of chromosome-level assembly of male genomic sequences

Component	Gene#	Repeat elements (%)	LTRs (%)
Whole genome	35 089*	70.6	32.4
Chromosome Y	1616	62.9	28.2
Y ₁ insertion	248	59.7	40.3
Y ₂ insertion	181	48.0	30.6

LTR, long terminal repeat.

*31 848 mRNAs, 424 tRNAs, 34 625 proteins, and 35 089 genes.

proteins. This clearly indicates that many of the genes are poorly known and around one-third of these sequences were unique with homologues not found in other organisms despite the strong evidence from their expression in RNA sequencing (RNA-Seq) data (see details below), confirming they were not artifacts of annotation.

Many genes found in the Y₁ and Y₂ regions were associated with flowering, which might be expected to vary between male and female (Table 2). These included embryo sac development arrest (Pagnussat et al., 2005) and FTIP1 (Liu et al., 2012) in Y₁ and *HOTHEAD* (Kurdyukov et al., 2006), stamen-specific protein (Nacken et al., 1991), protein gamete expressed 2 (Mori et al., 2014), self-incompatibility S1 domain-containing protein (Williams et al., 2015), and F-box protein genes (Xu et al., 2009) in Y₂.

More focused annotations focused on just annotating the Y₁ and Y₂ regions rather than the whole genome

revealed additional genes related to flowering or floral development. A gene with homology to a gene regulating floral development, *APETALA2* (Okamuro et al., 1993), was found in Y₂. *Perpetual Flowering 2* (*PEP2*) (Lazaro et al., 2018) and a MADS-box transcription factor (Teo et al., 2019), both of which have been linked to control of flowering, were found in this region. MADS-box genes have been linked to control of the transition to flowering (Teo et al., 2019). Genes controlling flower development are to be expected on sex chromosomes of dioecious plants. A *Pollenless 3* gene (Sanders et al., 1999) was found to be located on the distal ends of both the Y and X chromosomes of jojoba. Male-specific expression of genes on the X chromosome has been found in humans (Lercher et al., 2003) and plants may also have genes with sex-specific expression on both or either of the X and Y chromosomes. All three male-specific inter-simple sequence repeat markers (Heikrujam et al., 2014a,b; Sharma et al., 2008) that have been reported for jojoba were also found to be specifically located in the Y₁ region. The sex-determining genes reported for poplar (Mueller et al., 2020) and date palm (*Phoenix dactylifera*; Torres et al., 2018) were found elsewhere in the jojoba genome -and appear to not have a role in sex determination in jojoba. Independent evolution of dioecious plants has not followed the same path.

The mapping of RNA-Seq reads to genes in these two inserts confirmed that they were expressed, with more than 94% of genes having 30% or more of their length covered by transcribed sequences. Long terminal repeat (LTR)

Table 2 Examples of notable male-specific genes associated with flowering or floral development and markers found in the Y₁ and Y₂ chromosome regions

Chromosome region	Gene or marker name	Reference
Y ₁	<i>Embryo sac development arrest</i>	Pagnussat <i>et al.</i> (2005)
	<i>FTIP1</i>	Liu <i>et al.</i> (2012)
	Jojoba male-specific marker ISSR-UBC-807	Sharma <i>et al.</i> (2008)
	Jojoba male-specific marker ISSR-VIS-11	Heikrujam <i>et al.</i> (2014a)
	Jojoba clone pkmssj male-specific	Heikrujam <i>et al.</i> (2014b)
Y ₂	<i>HOTHEAD</i>	Kurdyukov <i>et al.</i> (2006)
	<i>Stamen specific protein</i>	Nacken <i>et al.</i> (1991)
	<i>Protein gamete expressed 2</i>	Mori <i>et al.</i> (2014)
	<i>Self-incompatibility S1 domain containing protein</i>	Williams <i>et al.</i> (2015)
	<i>F-box protein genes</i>	Xu <i>et al.</i> (2009)
	<i>APETALA2</i>	Okamuro <i>et al.</i> (1993)
	<i>MADS-box transcription factor</i>	Teo <i>et al.</i> (2019)
	<i>Perpetual Flowering 2 (PEP2)</i>	Lazaro <i>et al.</i> (2018)

Abbreviations: ISSR, inter-simple sequence repeat.

elements were more abundant in the male-specific insertions on the Y chromosome (Table 1). The presence of repetitive elements is a characteristic of plant sex chromosomes (Hobza *et al.*, 2017).

To establish that these differences were consistently found in the genomes of male and female jojoba plants, analysis of the sex-specific regions of several male and female plants was achieved by mapping Illumina reads from three male and three female genotypes to the Y chromosome assembly (Figure 3). Illumina reads from female genotypes mapped only to the common regions and not to any significant extent to the Y₁ and Y₂ insertions. Illumina reads from the male genotypes mapped across the whole chromosome with higher coverage in the common regions due to contribution of reads from both the X and Y chromosomes and lower coverage in the inserts contributed by the Y chromosome. The distal parts of the chromosomes showed mapping of both X and Y chromosome reads with 21.9–23.1× coverage in all six genotypes. The male-specific insertion Y₁ (as marked on Figure 2) had 12.1–12.4× coverage for the three male genotypes and 0.4–0.6× coverage for the three female genotypes, demonstrating the presence of a single copy of these chromosome regions in the males only. The second male-specific region (Y₂ in Figure 2) had a slightly higher mapping from the male, suggesting some reads with homology in the female, and this was confirmed by the 6× coverage with female reads. The region between the two insertions (Y₁ and Y₂) was found to be more complex. This region had insertions specifically on the X chromosome (Figure 2) that did not show here with a male reference chromosome. Some smaller male-specific insertions were also present (as evident in Figure 2) resulting in an average coverage over the whole region of 11.5–11.9× for the males and 20.7–21.3× for the females. All three male and all three female genotypes had

a very similar mapping coverage, confirming the sex rather than genotype specificity of these structures. The overlay of repeat element data and coding regions showed a perfect packaging of repetitive elements and LTRs in intergenic regions and introns of the genes (Figure 3). This indicates that transposons and LTRs have been integrated into the landscape of the genes to create a high content of both genes and repetitive elements. The genome had a high repetitive sequence content being 70% in the male genome (Table S5). The proportions of the male-specific regions that were composed of repetitive elements were 59.7 and 48.0% for the Y₁ and Y₂ insertions, respectively. LTR Gypsy elements were more common in the male-specific inserts and were reported as sex-specific in poplar (Xue *et al.*, 2020).

Phylogenetic analysis has placed jojoba in the Caryophyllales, and this was confirmed by analysis of homology of jojoba gene sequences with those of other species. The most closely related sequences, based on a functional annotation analysis in OmicsBox v2.0 for the isoseq transcripts (Table S6) and coding sequences from the HiFi assembly (Table S7), were from beetroot (*Beta vulgaris* subsp. *vulgaris*), quinoa (*Chenopodium quinoa*), and spinach (*Spinacia oleracea*).

Evolution of male and female genomes

Dioecy has evolved many times in plants and is found in around 7% of genera and 6% of species of flowering plants (Renner and Ricklefs, 1995). Substantial morphological differences have been found between male and female jojoba plants, especially when grown under dry conditions (Inoti *et al.*, 2015; Kohorn, 1994). These differences can now be associated with the substantial number of additional genes located on the Y chromosome in male plants. Many different explanations of the causes of the

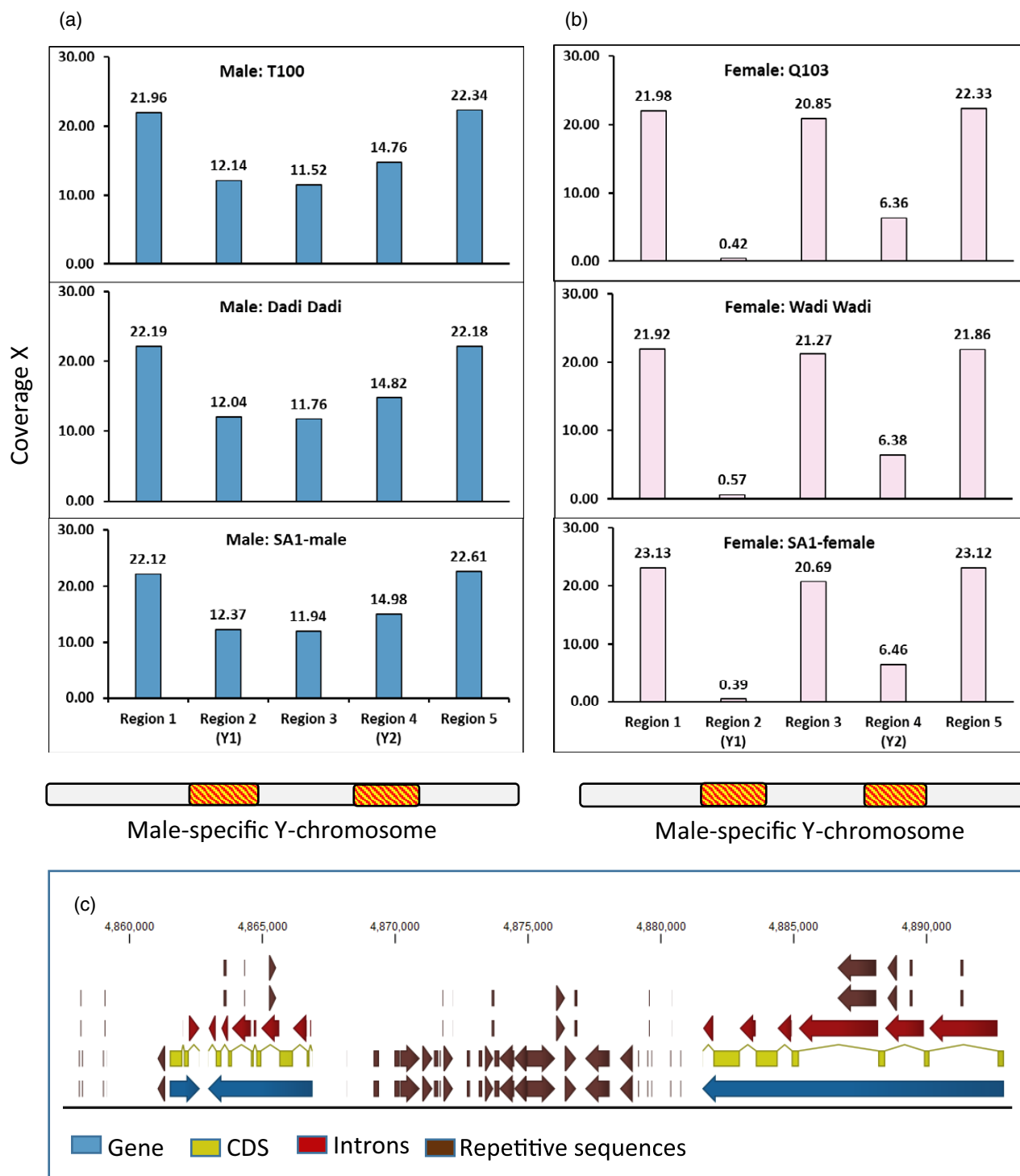


Figure 3. Characterisation of male-specific regions in the Y chromosomes.

Illumina paired-end read mapping coverage along the male-specific Y chromosome (Figure 2) as the reference. Region 1 (distal end); region 2 (male-specific Y₁); region 3 (region between Y₁ and Y₂); region 4 (male-specific Y₂); and region 5 (distal end). (a) SA1-Male (Saudi), Dadi Dadi, and T100: jojoba male genotypes. (b) SA1-Female (Saudi), Wadi Wadi, and Q103: Jojoba female genotypes. Region-wise average coverage values are normalised with respect to the total nucleotides of paired-end reads of the SA1-Female (Saudi) genotype mapped to the jojoba male-specific Y chromosome. (c) Repetitive element mapping track of a region of the jojoba male-specific Y₁ chromosome. The sex-specific regions are gene-rich. The space between the coding sequences (intergenic spaces and introns) is almost completely filled with repetitive sequences.

evolution of dioecious plants have been explored (Thomson and Brunet, 1990). Divergent evolution of males and females may be driven by selection for reproductive success in the desert environment (Thomson, 2006). Male seedlings are more tolerant of drought stress and their more rapid growth and development may allow them to reach flowering early. This may be due to stress response genes found in the male-specific insertions on the Y chromosome. For example, a gene encoding a protein dehydration response protein (Choudhary et al., 2009; Vessal et al., 2020) was located in Y₁. Female plants have been shown to devote significant proportions of their resources to seed production (Kohorn, 1994). Success may be enhanced by greater root growth, allowing the female plants to establish for the longer growth phase required to support seed production. The ratio of male to female plants in wild populations has only a slight bias towards more males (Waser, 1984), while in cultivation males may predominate when grown under stress (Inoti et al., 2015). Many studies on jojoba cultivation have noted this high male to female ratio (Agrawal et al., 2007). Hosseini et al. (2011) also cited Agrawal et al. (2007), indicating this abnormal ratio. So far, there is no evidence explaining how this ratio could be controlled genetically. The presence of an XY chromosome system suggests equal numbers of male and female seeds should be produced. However, Cole (1979) studied the effect of environmental factors on aberrant sex ratios in jojoba. They established a 3-month study and planted seeds to create 14 populations on north- and south-facing slopes in Arizona, US. They reported 51% of female jojoba on north-facing slopes but only 45% on south-facing (hotter and dryer) slopes. They suggested that there is some correlation between orientation of the slope and the male:female ratio, although many other environmental factors such as water stress and high temperature can increase female mortality and change the sex ratio in the population. The much higher abundance of males reported in cultivation may be due to higher survival rates of male seedlings that have a greater stress tolerance. In the wild, females that survive the seedling stage may survive longer with a better root system, balancing the greater early survival of the males as seedlings. The large number of male-specific genes in the male-specific parts of the Y chromosome may account for the very significant sexual dimorphism of morphological and growth traits in this species. This divergence of the sexes allowing adaptation to their sex-specific roles in reproduction may be the critical selective advantage that has led to the evolution of dioecious plants. The identification of large numbers of sex-specific genes in jojoba provides an opportunity to explore the functions that have diverged with sex separation.

Climate change may adversely impact on the sex ratio (Hultine et al., 2016) in wild populations of jojoba but the

direction of any change is not clear. Dioecious plants may be at greater risk of extinction if climate change results in a rapid development of sex imbalance in their populations (Tognetti, 2012). Analysis of the sex-specific genes discovered in this study and determination of their functions may help determine the nature of the risk faced by being dioecious.

Dioecious plants have arisen many times independently in divergent lineages. This divergence may have followed similar (or parallel) paths in different species but may be different in differing environments posing divergent selection pressures. The extent of divergence discovered here is much greater than that reported in other plants or in animals. For example, the presence of 1616 genes on the jojoba Y chromosome with 429 genes found in Y chromosome-specific insertions can be compared with only 78 protein-coding genes on the human Y chromosome (Bachtrog, 2013). Differences in some species seem relatively small (Martine et al., 2016). The substantial divergence in size of the sex chromosome in jojoba contrasts with the view that these chromosomes are often not easily distinguished in the cytology of dioecious plants (Zhou et al., 2020). The extreme stress imposed by the desert environment may drive strong selection for diversification to achieve reproductive success for males and females. Many other dioecious desert plants have been the subject of genetic research (Case and Barrett, 2004; Wolfe and Shmida, 1997). The extent to which other desert plants have undergone parallel evolution and substantial sex divergence in this extreme environment will be revealed as more dioecious desert plants and dioecious plants in general are subjected to genome analysis (Charlesworth, 2013). The common presence of repeat elements, specifically LTRs, in sex chromosomes has been explained by their potential to block recombination and as a result drive chromosome divergence (Charlesworth, 2017). The presence of common or functionally similar sex-specific genes may also be found as we explore the genomes of dioecious plants and gene interactions in sex chromosomes (Harkess et al., 2020).

The analysis of the genomes of dioecious plants can be expected to reveal a great diversity of options for crop improvement, providing a rich source of novel genes for both the manipulation of reproduction and environmental adaptation in crops.

EXPERIMENTAL PROCEDURES

Plant materials

Jojoba seed were obtained from King Faisal University (KFU) Al-Hofuf, Saudi Arabia (25°16'15.1"N, 49°42'42.6"E) during July 2019. Two jojoba plants, male and female, were identified and assigned for leaf tissue collection. For each sex, 10 g of healthy plant leaf tissue was collected and manually ground under liquid N₂ using motor and pestle. The ground tissue was immediately suspended in a 50-ml Falcon tube containing 40 ml of 2% cetyl trimethylammonium bromide (CTAB) extraction buffer.

The leaf tissues from two different varieties per sex (Daddi-Daddi and T100 for male; Wadi-Wadi and Q103 for female) were collected randomly from the 'Chris-Egan' farm at Inglewood (151°4'.20'E, 28°25'13"S), Queensland, Australia. The collected leaf tissues were snap-frozen in liquid nitrogen and placed in dry ice followed by preservation at −80°C. The leaf tissues were ground to a fine powder using a Tissue Lyser-II (Qiagen, Valencia, CA, USA) at a frequency of 30 Hz for 30 sec prior to DNA extraction.

Genomic DNA isolation and sequencing

Genomic DNA from mature leaf tissue of jojoba male and female plants from Saudi Arabian and Australian varieties were extracted following a modified CTAB method that was published previously (Carroll et al., 1995; Furtado, 2014). The modification took place in several steps through the extraction protocol. In the step of adding nuclear lysis buffer and 5% Sarkosyl solution, 0.06 g of sodium sulphite was also added to the extraction mix. The mixture was incubated at 50°C for 45 min with repeated inversions (5–8 times) and was followed by adding 10 ml of chloroform. During the DNA washing steps, 5 ml of 70% ethanol was added. The extracted DNA was quantitatively and qualitatively evaluated using agarose gel electrophoresis (0.7%, 110 V for 45 min) and a NanoDrop 8000 Spectrophotometer (Thermo Scientific, Waltham, MA, USA), respectively. The DNA samples from both sexes (male and female) were analysed using two different sequencing platforms: PacBio (Menlo Park, CA, USA) and Illumina (San Diego, CA, USA). The sequencing libraries for short (NextSeq2000) and long reads required 0.025 and 20 µg of DNA, respectively. The sequencing libraries were generated based on the manufacturer's protocols. Illumina sequencing yielded the following amounts of sequence data: Saudi male 71.5 Gb, Saudi female 72.3 Gb, Daddi 34.3 Gb, T100 33.7 Gb, Wadi 38.8 Gb, Q103 34.5 Gb.

An Omni-C library was prepared by extracting chromatin after fixation in the nucleus using formaldehyde. The extracted chromatin ends were digested using DNase I and then repaired and re-ligated to a biotinylated bridge adapter followed by proximity ligation of adapter-containing ends. Furthermore, the crosslinks were reversed, and the DNA was purified. A sequencing library was prepared using NEBNext Ultra enzymes and Illumina-compatible adapters. The library was sequenced using an Illumina HiSeqX platform.

Genome size estimation

K-mer frequency distribution analysis was used to estimate jojoba male and female genome sizes and heterozygosity rates using KAT software v2.4.2 (Mapleson et al., 2017). K-mer frequency distribution was calculated and plotted using 71.5 and 72.3 Gb of Illumina short-read sequencing data for male and female to determine the total number of k-mers of length 27 by means of the count function in Jellyfish (Marçais and Kingsford, 2011). The length of k-mers (27-mers) was chosen based on jojoba genome characteristics. Moreover, the k-mer frequency peak of the reads (M) corresponded to the sequencing depth (N), read length (L), and k-mer length (K) and by the following formula: $M = N \times (L - K + 1) / L$ (Dong et al., 2010). The peak of the 27-mer frequency from the paired-end reads of jojoba was 63 and 65 for male and female, respectively (Figure S1). The genome sizes were estimated to be 724.9 and 727.4 Mb for male and female, respectively.

Total RNA isolation and sequencing

A mixture of total RNA was obtained from a pilot water stress experiment. The water stress experiment was designed to have 10

time-points and two replications, and one plant was designated to every replicate. The 10 time-points were as follows: day 0, day 4, day 7, day 9, day 11, day 14, day 16, day 18, day 22, and day 25. Twenty 3-month-old jojoba plants were used in this experiment. The experiment involved withholding water from initially well-watered plants for 25 days. Leaf tissue collection took place at every time-point, and 20 random leaves were collected into micro-perforated polyethylene (PET) bags from every replication and snap-frozen in liquid nitrogen. The collected leaves from every sample were ground to a fine powder using a Tissue Lyser-II (Qiagen) or a Mixer Mill-400 (Retsch, Haan, Germany) at a frequency of 30 Hz for 30 sec. The fine powder of each sample was transferred into 50-ml Falcon tubes and stored at −80°C.

A two-step protocol, including a cetyltrimethylammonium bromide (CTAB) method followed by a Qiagen RNeasy Plant mini kit (#74134, Qiagen), was used to ensure the complete removal of contaminating genomic DNA (Yang et al., 2008). The qualitative and quantitative evaluation of the total RNA extracts were accomplished using a NanoDrop 8000 Spectrophotometer (ThermoFisher Scientific, Wilmington, DE, USA) and a 2100 Agilent Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA), respectively. Samples from day 0 to day 14 were mixed to use in jojoba genome annotation. RNA integrity number (RIN) for the mixed RNA samples ranged from 2.20 (day 14) to 5.30 (day 0) due to possible contaminants. The mixed RNA sample was sequenced using an Illumina platform.

Genome assembly

The final *de novo* assembly of PacBio reads was conducted using IPA v:1.3.1 with default parameters. Assembly statistics were calculated using the Quality Assessment Tool (QUAST) v5.1.0 based on a contig size of ≥500 bp (Gurevich et al., 2013). The completeness of the reference genome assembly was assessed using BUSCO v4.1.2 with the viridiplantae_odb10 database (Gurevich et al., 2018).

Assembly scaffolding using Hi-Rise

The jojoba male *de novo* assembly and Omni-C library reads were inputted into HiRise to scaffold the genome assemblies using proximity ligation data (Putnam et al., 2016). Sequences from the Dovetail Omni-C library were mapped to the draft assembly using a modified SNAP read aligner (<http://snap.cs.berkeley.edu>). The separations of Dovetail Omni-C read pairs aligned within the draft assembly were examined by HiRise to create a likelihood model for genomic distance between read pairs. The generated model was employed to determine and separate putative mis-joins, to score prospective joins, and to make joins above a threshold.

Transcriptome sequencing

Equal volumes of five RNA samples representing three time-points (day 0, day 8, and day 14) of a water stress experiment were mixed and sequenced using PacBio long-read cDNA sequences. The cDNA was amplified and filtered following the Iso-Seq protocol into standard and long transcript libraries. The two libraries were quantified using two SMRT cells of the Pacific Bioscience (PacBio) Sequel II platform, generating 296.6 Mb of data (426 380 isoforms).

Assessment of genome assembly

Jojoba genome assembly completeness was assessed using BUSCO and a jojoba transcriptomic database. The jojoba genome assembly was searched against 425 conserved single-copy

orthologs in the viridiplantae_odb10 BUSCO v4.1.2 dataset with default parameters. Furthermore, transcript isoforms that were produced by PacBio long-read sequencing of cDNA were used to confirm the high quality of the genome assembly. The isoforms were aligned to the final assembly using CLC Genomics Workbench v20.0.4 (CLC-GWB, CLC Bio- Qiagen, Aarhus, Denmark) with default parameters. High-molecular-weight genomic DNA from three male jojoba genotypes (Male-SA1, Dadi Dadi, and T-100) and from three female genotypes (Female-SA1, Wadi Wadi, and Q103) were subjected to 150-bp paired-end read sequencing on the Illumina NovaSeq 6000 platform. The sequence reads were trimmed at 0.01 quality limit (equivalent to a Fred Score of 30 or above) and sequence read data between 22 and 33 million paired-end reads from the male and female genotypes were used for mapping to the Saudi jojoba male assembly as the reference. Mapping parameters included: match score 1, mismatch cost 2, insertion/deletion cost 3, length and similarity fraction of 0.8 each, and ignoring non-specific matches (CLC Genomics Workbench, Qiagen). Total nucleotides of the paired-end reads mapped to the Saudi jojoba female assembly (as reference) were used to normalise the nucleotide coverage of specific regions (regions 1 to 4) for each of the jojoba genotypes.

Genome annotation

Whole genome annotation was conducted by Dovetail Genomics. Coding sequences from *S. chinensis* (Sturtevant et al., 2020), *Theobroma cacao*, *Arabidopsis thaliana*, and *Oryza sativa* were used to train the initial *ab initio* model for *S. chinensis* using AUGUSTUS software (version 2.5.5). Six rounds of prediction optimisation were done with the software package provided by AUGUSTUS. The same coding sequences were also used to train a separate *ab initio* model for *S. chinensis* using SNAP (version 2006-07-28). RNA-Seq reads were mapped onto the genome using the STAR aligner software (version 2.7) and intron hints were generated with the bam2hints tools within AUGUSTUS. MAKER, SNAP, and AUGUSTUS (with intron–exon boundary hints provided from RNA-Seq data) were then used to predict genes in the repeat-masked reference genome. To help guide the prediction process, Swiss-Prot peptide sequences from the UniProt database were downloaded and used in conjunction with the protein sequences from *S. chinensis*, *T. cacao*, *A. thaliana*, and *O. sativa* to generate peptide evidence in the MAKER pipeline. Only genes that were predicted by both SNAP and AUGUSTUS software were retained in the final gene sets. To help assess the quality of the gene prediction, AED scores were generated for each of the predicted genes as part of the MAKER pipeline. Genes were further characterised for their putative function by performing a homology search of the peptide sequences against the UniProt database. tRNAs were predicted using the software tRNAscan-SE (version 2.05).

An alternate male and female genome annotation and annotation of male-specific parts of the genome was conducted using the HiFi assemblies for both sexes. The web-based Genome Sequence Annotation Server (GeneSAS) was employed to annotate the draft genome assemblies (Humann et al., 2019). GeneSAS is a platform that combines multiple annotation tools for whole genome structural and functional annotations.

Repeat annotation

Repeat families found in the genome assemblies of *S. chinensis* were identified *de novo* and classified using the software package RepeatModeler (version 2.0.1). RepeatModeler depends on the programmes RECON (version 1.08) and RepeatScout (version

1.0.6) for the *de novo* identification of repeats within the genome. The custom repeat library obtained from RepeatModeler were used to discover, identify, and mask the repeats in the assembly file using RepeatMasker (Version 4.1.0). The repetitive elements of the male and female jojoba genomes including tandem repeats and transposable elements were identified using RepeatModeler, which uses two *de novo* repeat finding programmes (Recon and RepeatScout) (Bao et al., 2015).

ACKNOWLEDGEMENTS

This work was supported by King Faisal University, Saudi Arabia (Grant KFU-14400005199) and the Queensland Alliance for Agriculture and Food Innovation (QAAFI), University of Queensland (UQ), Australia. We thank the Research Computing Centre (RCC), University of Queensland for support and providing high-performance computing resources. Dovetail Genomics conducted the Omni-C assembly and whole genome annotation. PacBio Sequencing was conducted by the University of Queensland sequencing facility. We thank Fred Ghirardello for provision of jojoba samples.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Study conception and design: all authors. Data collection: OA and BA (BA contributed most to the annotation and OA to the assembly). All authors contributed to analysis and interpretation of results and drafting the manuscript, reviewed the results, and approved the final version of the manuscript.

DATA AVAILABILITY STATEMENT

Jojoba reference genome sequence data (Illumina, PacBio, and Hi-C interaction reads), the final genome assemblies, structural and functional annotations for both male and female, and the long-read transcriptome are available at the NCBI website under Bioproject ID PRJNA694450. The data is also available the Genome Warehouse website under bioProject number PRJCA006974

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Table S1. Jojoba male and female genome assemblies using Improved Phase Assembly (IPA) v:1.3.1.

Table S2. Quality Assessment Tool (QUAST) statistics of the jojoba (*Simmondsia chinensis*) male genome assembly using Dovetail HiRise v2.0.

Table S3. The lengths of 26 jojoba pseudochromosomes based on chromosome-level Dovetail HiRise assembly.

Table S4. Benchmarking Universal Single-Copy Orthologs (BUSCO) analysis of jojoba male results for both assemblers: HiRise and Improved Phase Assembly (IPA).

Table S5. The percentage of repeat family sequences in male and female jojoba genomes v1.0.11.

Table S6. The top four jojoba closely related species based on a functional annotation analysis in OmicsBox v2.0 for the jojoba

isoseq reference. A close relationship was shown between jojoba and the following species: quinoa (*Chenopodium quinoa*), beetroot (*Beta vulgaris* subsp. *vulgaris*), spinach (*Spinacia oleracea*), and grape vine (*Vitis vinifera*).

Table S7. The top three jojoba closely related species based on a functional annotation analysis in OmicsBox v2.0 for coding sequences (CDSs) from the jojoba male HiFi assembly. A close relationship was shown between jojoba and the following species: beetroot (*Beta vulgaris* subsp. *vulgaris*), quinoa (*Chenopodium quinoa*), and spinach (*Spinacia oleracea*).

Figure S1. The k-mer distribution and coverage of sequencing reads at K = 27 for (a) male and (b) female. Peaks with single and double asterisks were evaluated as k-mer species derived from heterozygous (k-mer frequency = 13) and homozygous (k-mer frequency = 63) sequences for male and heterozygous (k-mer frequency = 12) and homozygous (k-mer frequency = 65) sequences for female.

Figure S2. Syntenic comparison of the jojoba genome. A dot plot showing the syntenic gene pairs (totally 346) in the jojoba genome. Each dot shows one syntenic gene pair (left: colour-coded based on Ks rates in Figure 2). The grey lines (on the x and y-axes) represent chromosomes.

Figure S3. Ks analysis (synonymous distribution) of the jojoba genome. The median peak (orange) with a Ks value of 0.2 shows an early whole genome duplication (WGD) event (syntenic orthologs). The red peak with a Ks value of 0.5 shows the younger (second) WGD (syntenic out-paralogs). The green and blue columns are noises.

REFERENCES

- Agrawal, V., Sharma, K., Gupta, S., Kumar, R. & Prasad, M. (2007) Identification of sex in *Simmondsia chinensis* (Jojoba) using RAPD markers. *Plant Biotechnology Reports*, **1**, 207–210. <https://doi.org/10.1007/s11816-007-0031-6>
- Almeida, P., Proux-Wera, E., Churcher, A., Soler, L., Dainat, J., Pucholt, P. et al. (2020) Genome assembly of the basket willow, *Salix viminalis*, reveals earliest stages of sex chromosome expansion. *BMC Biology*, **18** (1), 1–78. <https://doi.org/10.1186/s12915-020-00808-1>
- Alotaibi, S.S., Elseehy, M.M., Aljuaid, B.S. & El-Shehawi, A.M. (2020) Transcriptome analysis of jojoba (*Simmondsia chinensis*) during seed development and liquid wax ester biosynthesis. *Plants*, **9**(5), 588. <https://doi.org/10.3390/plants9050588>
- Arya, D. & Khan, S. (2016) A review of *Simmondsia chinensis* (Jojoba)* The Desert Gold*: a multipurpose oil seed crop for industrial uses. *Journal of Pharmaceutical Science and Research*, **8**(6), 381.
- Ashman, T.-L., Kwok, A. & Husband, B.C. (2013) Revisiting the Dioecy-Polyploidy association: alternate pathways and research opportunities. *Cytogenetic and Genome Research*, **140**(2–4), 241–255. <https://doi.org/10.1159/000353306>
- Atanassov, I., Delichere, C., Filatov, D.A., Charlesworth, D., Negrutiu, I. & Monéger, F. (2001) Analysis and evolution of two functional Y-linked loci in a plant sex chromosome system. *Molecular Biology and Evolution*, **18** (12), 2162–2168. <https://doi.org/10.1093/oxfordjournals.molbev.a003762>
- Bachtrog, D. (2013) Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nature Reviews Genetics*, **14**(2), 113–124. <https://doi.org/10.1038/nrg3366>
- Bao, W., Kojima, K.K. & Kohany, O. (2015) Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA*, **6**(1), 11. <https://doi.org/10.1186/s13100-015-0041-9>
- Baránková, S., Pascual-Díaz, J.P., Sultana, N., Alonso-Lifante, M.P., Balant, M., Barros, K. et al. (2020) Sex-chrom, a database on plant sex chromosomes. *New Phytologist*, **227**(6), 1594–1604. <https://doi.org/10.1111/nph.16635>
- Carroll, B.J., Klimyuk, V.I., Thomas, C.M., Bishop, G.J., Harrison, K., Scofield, S.R. et al. (1995) Germinal transpositions of the maize element Disso from T-DNA loci in tomato. *Genetics*, **139**(1), 407.
- Case, A.L. & Barrett, S.C.H. (2004) Environmental stress and the evolution of dioecy: *Wurmbea dioica* (Colchicaceae) in Western Australia. *Evolutionary Ecology*, **18**(2), 145–164. <https://doi.org/10.1023/B:EVEC.0000021152.34483.77>
- Charlesworth, D. (2013) Plant sex chromosome evolution. *Journal of Experimental Botany*, **64**(2), 405–420. <https://doi.org/10.1093/jxb/ers322>
- Charlesworth, D. (2016) Plant sex chromosomes. *Annual Review of Plant Biology*, **67**(67), 397–420. <https://doi.org/10.1146/annurev-arplant-043015-111911>
- Charlesworth, D. (2017) Evolution of recombination rates between sex chromosomes. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **372**(1736), 20160456. <https://doi.org/10.1098/rstb.2016.0456>
- Choudhary, M.K., Basu, D., Datta, A., Chakraborty, N. & Chakraborty, S. (2009) Dehydration-responsive nuclear proteome of rice (*Oryza sativa* L.) illustrates protein network, novel regulators of cellular adaptation, and evolutionary perspective. *Molecular & Cellular Proteomics*, **8**(7), 1579–1598. <https://doi.org/10.1074/mcp.M800601-MCP200>
- Cole, S.L. (1979) Aberrant sex-ratios in jojoba associated with environmental factors. Available from: <https://repository.arizona.edu/handle/10150/528199>
- Dong, D., Tian, G., Ruan, J., Li, H., Guo, Y. & Bolund, L. et al. (2010) The sequence and de novo assembly of the giant panda genome. *Nature*, **463** (7284), 1106. <https://doi.org/10.1038/nature08846>
- Furtado, A. (2014). DNA extraction from vegetative tissue for next-generation sequencing. In: Henry, R.J. & Furtado, A. (Eds.) *Cereal genomics*. Totowa, NJ: Springer, pp. 1–5.
- Guo, L.T., Wang, S.L., Wu, Q.J., Zhou, X.G., Xie, W. & Zhang, Y.J. (2015) Flow cytometry and K-mer analysis estimates of the genome sizes of *Bemisia tabaci* B and Q (Hemiptera: Aleyrodidae). *Frontiers in Physiology*, **6**, 144.
- Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., Afgan, E., Baker, D. et al. (2018). *The Galaxy platform for accessible, reproducible and collaborative biomedical analyses*. *Nucleic Acids Research*, **46**(W1, 2 July 2018), W537–W544. <https://doi.org/10.1093/nar/gky379>
- Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. (2013) QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, **29**(8), 1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>
- Harkess, A., Huang, K., van der Hulst, R., Tissen, B., Caplan, J.L., Koppula, A. et al. (2020) Sex determination by two Y-linked genes in garden asparagus. *The Plant Cell*, **32**(6), 1790–1796. <https://doi.org/10.1105/tpc.19.00859>
- Heikrujam, M., Kumar, J. & Agrawal, V. (2015a) Genetic diversity analysis among male and female Jojoba genotypes employing gene targeted molecular markers, start codon targeted (SCoT) polymorphism and CAAT box-derived polymorphism (CBDP) markers. *Meta Gene*, **5**, 90–97. <https://doi.org/10.1016/j.mgene.2015.06.001>
- Heikrujam, M., Sharma, K., Kumar, J. & Agrawal, V. (2014a) Generation and validation of unique male sex-specific sequence tagged sites (STS) marker from diverse genotypes of dioecious Jojoba-*Simmondsia chinensis* (Link) Schneider. *Euphytica*, **199**(3), 363–372. <https://doi.org/10.1007/s10681-014-1136-y>
- Heikrujam, M., Sharma, K., Kumar, J. & Agrawal, V. (2014b) Validation of male sex-specific UBC-80712001SSR marker and its conversion into sequence tagged sites marker in Jojoba: a high precision oil yielding dioecious shrub. *Plant Breeding*, **133**(5), 666–671. <https://doi.org/10.1111/pbr.12199>
- Heikrujam, M., Sharma, K., Prasad, M. & Agrawal, V. (2015b) Review on different mechanisms of sex determination and sex-linked molecular markers in dioecious crops: a current update. *Euphytica*, **201**(2), 161–194. <https://doi.org/10.1007/s10681-014-1293-z>
- Hobza, R., Cegan, R., Jesionek, W., Kejnovsky, E., Vyskot, B. & Kubat, Z. (2017) Impact of repetitive elements on the Y chromosome formation in plants. *Genes*, **8**(11), 302. <https://doi.org/10.3390/genes8110302>
- Hosseini, F.S., Shahsav, H., Arvin, M.J., Baghizadeh, A. & Mohammadi-Nejad, G. (2011) Sex determination of jojoba (*Simmondsia chinensis* cv. Arizona) by random amplified polymorphic DNA (RAPD) molecular markers. *African Journal of Biotechnology*, **10**(4), 470–474.
- Hultine, K.R., Grady, K.C., Wood, T.E., Shuster, S.M., Stella, J.C. & Whitham, T.G. (2016) Climate change perils for dioecious plant species. *Nature Plants*, **2**(8), 16109. <https://doi.org/10.1038/nplants.2016.109>
- Humann, J.L., Lee, T., Ficklin, S. & Main, D. (2019). Gene prediction. In: Kollmar, M.(Ed.) *Gene prediction methods and protocols*. New York, NY: Springer New York, pp. 29–51.

- Inoti, S.K. (2018) Effect of browsing on the growth of jojoba seedlings in the dry-lands of Kenya. *Journal of Agricultural Science and Technology B*, **8**, 320.
- Inoti, S.K., Chamshama, S.A.O., Thagana, W.M., Lulandala, L.L.L. & Dodson, R. (2015) Sex determination of young nursery Jojoba (*Simmondsia chinensis* L.) plants using morphological traits in semi arid areas of Voi, Kenya. *Computer*, **5**(16), 2224–3208.
- Kohorn, L.U. (1994) Shoot morphology and reproduction in jojoba: advantages of sexual dimorphism. *Ecology*, **75**(8), 2384–2394. <https://doi.org/10.2307/1940892>
- Kumar, J., Heikrujam, M. & Agrawal, V. (2016) Characterization of male and female jojoba plants employing antioxidants, lipid peroxidation and physiological studies. *Journal of the American Oil Chemists' Society*, **93** (7), 911–920. <https://doi.org/10.1007/s11746-016-2831-8>
- Kumar, J., Heikrujam, M., Sharma, K. & Agrawal, V. (2019) SRAP and SSR marker-assisted genetic diversity, population structure analysis and sex identification in Jojoba (*Simmondsia chinensis*). *Industrial Crops and Products*, **133**, 118–132. <https://doi.org/10.1016/j.indcrop.2019.03.006>
- Kurdyukov, S., Faust, A., Trenkamp, S., Bär, S., Franke, R., Efremova, N. *et al.* (2006) Genetic and biochemical evidence for involvement of HOT-HEAD in the biosynthesis of long-chain α -, ω -dicarboxylic fatty acids and formation of extracellular matrix. *Planta*, **224**(2), 315–329. <https://doi.org/10.1007/s00425-005-0215-7>
- Lazaro, A., Obeng-Hinne, E. & Albani, M.C. (2018) Extended vernalization regulates inflorescence fate in Arabidopsis by stably silencing PERPETUAL FLOWERING1. *Plant Physiology*, **176**(4), 2819–2833. <https://doi.org/10.1104/pp.17.01754>
- Leite Montalvão, A.P., Kersten, B., Fladung, M. & Müller, N.A. (2020) The diversity and dynamics of sex determination in dioecious plants. *Frontiers in Plant Science*, **11**, 580488. <https://doi.org/10.3389/fpls.2020.580488>
- Lercher, M.J., Urrutia, A.O. & Hurst, L.D. (2003) Evidence that the human X chromosome is enriched for male-specific but not female-specific genes. *Molecular Biology and Evolution*, **20**(7), 1113–1116. <https://doi.org/10.1093/molbev/msg131>
- Liu, L.U., Liu, C., Hou, X., Xi, W., Shen, L., Tao, Z. *et al.* (2012) FTIP1 is an essential regulator required for florigen transport. *PLoS Biology*, **10**(4), e1001313. <https://doi.org/10.1371/journal.pbio.1001313>
- Mapleson, D., Garcia Accinelli, G., Kettleborough, G., Wright, J. & Clavijo, B.J. (2017) KAT: a K-mer analysis toolkit to quality control NGS datasets and genome assemblies. *Bioinformatics*, **33**(4), 574–576.
- Marçais, G. & Kingsford, C. (2011) A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*, **27**(6), 764–770. <https://doi.org/10.1093/bioinformatics/btr011>
- Martine, C., Cantley, J., Frawley, E., Butler, A. & Jordon-Thaden, I. (2016) New functionally dioecious bush tomato from northwestern Australia, *Solanum ossicrueum*, may utilize "trample burr" dispersal. *PhytoKeys*, **63**(63), 19–29. <https://doi.org/10.3897/phytokeys.63.7743>
- Massonnet, M., Cochetel, N., Minio, A., Vondras, A.M., Lin, J., Muyle, A. *et al.* (2020) The genetic basis of sex determination in grapes. *Nature Communications*, **11**(1), 2902. <https://doi.org/10.1038/s41467-020-16700-z>
- Mori, T., Igawa, T., Tamiya, G., Miyagishima, S.-Y. & Berger, F. (2014) Gamete attachment requires GEX2 for successful fertilization in Arabidopsis. *Current Biology*, **24**(2), 170–175. <https://doi.org/10.1016/j.cub.2013.11.030>
- Müller, N.A., Kersten, B., Leite Montalvão, A.P., Mähler, N., Bernhardsson, C., Bräutigam, K. *et al.* (2020) A single gene underlies the dynamic evolution of poplar sex determination. *Nature Plants*, **6**(6), 630–637. <https://doi.org/10.1038/s41477-020-0672-9>
- Murigneux, V., Rai, S.K., Furtado, A., Bruxner, T.J.C., Tian, W., Harliwong, I. *et al.* (2020) Comparison of long-read methods for sequencing and assembly of a plant genome. *GigaScience*, **9**(12), <https://doi.org/10.1093/gigascience/giaa146>
- Nacken, W.K.F., Huijser, P., Beltran, J.-P., Saedler, H. & Sommer, H. (1991) Molecular characterization of two stamen-specific genes, tap1 and fil1, that are expressed in the wild type, but not in the deficient mutant of *Antirrhinum majus*. *Molecular and General Genetics*, **229**(1), 129–136.
- Negrutiu, I.O., Vyskot, B., Barbacar, N., Georgiev, S. & Moneger, F. (2001) Dioecious plants. A key to the early events of sex chromosome evolution. *Plant Physiology*, **127**(4), 1418–1424. <https://doi.org/10.1104/pp.127.4.1418>
- Okumuro, J.K., Caster, B., Villarreal, R., VanMontagu, R. & Jofuku, K.D. (1993) The AP2 domain of APETALA2 defines a large new family of DNA binding proteins of Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, **94**(13), 7076–7081. <https://doi.org/10.1073/pnas.94.13.7076>
- Pagnussat, G.C., Yu, H.J., Ngo, Q.A., Rajani, S., Mayalagu, S. & Johnson, C.S. *et al.* (2005) Genetic and molecular identification of genes required for female gametophyte development and function in Arabidopsis. *Development*, **132**(3), 603–614. <https://doi.org/10.1242/dev.01595>
- Panfilio, K.A., Jentzsch, I.M.V., Benoit, J.B., Erezylmaz, D., Suzuki, Y. *et al.* (2019) Molecular evolutionary trends and feeding ecology diversification in the Hemiptera, anchored by the milkweed bug genome. *Genome Biology*, **20**, 64. <https://doi.org/10.1186/s13059-019-1660-0>
- Presgraves, D.C. (2008) Sex chromosomes and speciation in Drosophila. *Trends in Genetics*, **24**(7), 336–343. <https://doi.org/10.1016/j.tig.2008.04.007>
- Putnam, N.H., O'Connell, B.L., Stites, J.C., Rice, B.J., Blanchette, M., Calaf, R. *et al.* (2016) Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. *Genome Research*, **26**(3), 342–350.
- Renner, S.S. & Ricklefs, R.E. (1995) Dioecy and its correlates in the flowering plants. *American Journal of Botany*, **82**(5), 596. <https://doi.org/10.2307/2445418>
- Sanders, P.M., Bui, A.Q., Weterings, K., McIntire, K.N., Hsu, Y.-C., Lee, P.Y. *et al.* (1999) Anther developmental defects in Arabidopsis thaliana male-sterile mutants. *Sexual Plant Reproduction*, **11**(6), 297–322.
- Sandouqa, A. & Al-Hamamre, Z. (2019) Energy analysis of biodiesel production from jojoba seed oil. *Renewable Energy*, **130**, 831–842. <https://doi.org/10.1016/j.renene.2018.07.015>
- Sharma, K., Agrawal, V., Gupta, S., Kumar, R. & Prasad, M. (2008) ISSR marker-assisted selection of male and female plants in a promising dioecious crop: jojoba (*Simmondsia chinensis*). *Plant Biotechnology Reports*, **2**(4), 239–243. <https://doi.org/10.1007/s11816-008-0070-7>
- Sharma, P., Aldossary, O., Alsubaie, B., Al-Mssalleem, I., Nath, O. & Mitter, N. *et al.* (2021). Improvements in the sequencing and assembly of plant genomes. *bioRxiv*.
- Smith, B. & Smith, M.T. (1947) Sex chromosomes in Rumex-Hastatulus Baldwin with Xy1y2 pairing. *Genetics*, **32**(1), 104–105.
- Sturtevant, D., Lu, S., Zhou, Z.W., Shen, Y., Wang, S. & Song, J.M. *et al.* (2020) The genome of jojoba (*Simmondsia chinensis*): a taxonomically isolated species that directs wax ester accumulation in its seeds. *Science Advances*, **6**(11), eaay3240. <https://doi.org/10.1126/sciadv.aay3240>
- Teo, Z.W.N., Zhou, W. & Shen, L. (2019) Dissecting the function of MADS-box transcription factors in orchid reproductive development. *Frontiers in Plant Science*, **10**, 1474. <https://doi.org/10.3389/fpls.2019.01474>
- Thomas, S.C. & LaFrankie, J.V. (1993) Sex, size and interyear variation in flowering among dioecious trees of the Malayan rain forest. *Ecology*, **74** (5), 1529–1537.
- Thomson, J.D. (2006) Tactics for male reproductive success in plants: contrasting insights of sex allocation theory and pollen presentation theory. *Integrative and Comparative Biology*, **46**(4), 390–397. <https://doi.org/10.1093/icb/icj046>
- Thomson, J.D. & Brunet, J. (1990) Hypotheses for the evolution of dioecy in seed plants. *Trends in Ecology & Evolution*, **5**(1), 11–16. [https://doi.org/10.1016/0169-5347\(90\)90006-Y](https://doi.org/10.1016/0169-5347(90)90006-Y)
- Tobe, H., Yasuda, S. & Oginuma, K. (1992) Seed coat anatomy, karyomorphology, and relationships of *Simmondsia* (*Simmondsiaceae*). *The Botanical Magazine Tokyo*, **105**(4), 529–538. <https://doi.org/10.1007/BF02489427>
- Tognetti, R. (2012) Adaptation to climate change of dioecious plants: does gender balance matter? *Tree Physiology*, **32**(11), 1321–1324. <https://doi.org/10.1093/treephys/tps105>
- Torres, M.F., Mathew, L.S., Ahmed, I., Al-Azwani, I.K., Krueger, R., Rivera-Núñez, D. *et al.* (2018) Genus-wide sequencing support a two-locus model for sex-determination in Phoenix. *Nature Communications*, **9**(1), 5219–5211. <https://doi.org/10.1038/s41467-018-07742-5>
- Vessal, S., Arefian, M. & Siddique, K.H.M. (2020) Proteomic responses to progressive dehydration stress in leaves of chickpea seedlings. *BMC Genomics*, **21**(1), 523. <https://doi.org/10.1038/s41467-018-07742-5>

- Wang, X., Gao, F., Bing, J., Sun, W., Feng, X., Ma, X. *et al.* (2019) Overexpression of the Jojoba aquaporin gene, ScPIP1, enhances drought and salt tolerance in transgenic Arabidopsis. *International Journal of Molecular Sciences*, **20**(1), 153. <https://doi.org/10.3390/ijms20010153>
- Waser, N.M. (1984) Sex ratio variation in populations of a dioecious desert perennial, *Simmondsia chinensis* [jojoba, Sonoran desert]. *Oikos*, **42**(3), 343–348. <https://doi.org/10.2307/3544403>
- Williams, J.S., Wu, L., Li, S., Sun, P. & Kao, T.-H. (2015) Insight into S-RNase-based self-incompatibility in *Petunia*: recent findings and future directions. *Frontiers in Plant Science*, **6**, 41. <https://doi.org/10.3389/fpls.2015.00041>
- Wolfe, L.M. & Shmida, A. (1997) The ecology of sex expression in a gynodioecious Israeli Desert Shrub (*Ochradenus Baccatus*). *Ecology*, **78**(1), 101–110.
- Xu, G., Ma, H., Nei, M. & Kong, H. (2009) Evolution of F-box genes in plants: different modes of sequence divergence and their relationships with functional diversification. *Proceedings of the National Academy of Sciences of the United States of America*, **106**(3), 835–840. <https://doi.org/10.1073/pnas.0812043106>
- Xue, L., Wu, H., Chen, Y., Li, X., Hou, J., Lu, J. *et al.* (2020) Evidences for a role of two Y-specific genes in sex determination in *Populus deltoides*. *Nature Communications*, **11**(1), 5893. <https://doi.org/10.1038/s41467-020-19559-2>
- Yang, G., Zhou, R., Tang, T. & Shi, S. (2008) Simple and efficient isolation of high-quality total RNA from *Hibiscus tiliaceus*, a Mangrove associate and its relatives. *Preparative Biochemistry and Biotechnology*, **38**(3), 257–264. <https://doi.org/10.1080/10826060802164991>
- Zhou, R., Macaya-Sanz, D., Carlson, C.H., Schmutz, J., Jenkins, J.W., Kudrna, D. *et al.* (2020) A willow sex chromosome reveals convergent evolution of complex palindromic repeats. *Genome Biology*, **21**(1), 1–38. <https://doi.org/10.1186/s13059-020-1952-4>