Elucidating the parallel origin of sex determination and homeotic sexual dimorphism in *Thalictrum*

By

Melody P. Sain

A dissertation submitted in partial fulfillment of

the requirements of the degree of

Doctor of Philosophy

(Botany)

at the

UNIVERSITY OF WISCONSIN-MADISON

2022

Date of final oral examination: 08/18/2022

The dissertation is approved by the following members of the Final Oral Committee:

David A. Baum, Professor, Botany
Kenneth J. Sytsma, Professor, Botany
Cécile Ané, Professor, Statistics
Anne Pringle, Professor, Botany
Verónica Di Stilio, Professor, Biology (University of Washington)

Dedication

I am forever grateful for my given and chosen family's unwavering and unconditional encouragement and love throughout my journey to PhDone!

Thereby I dedicate this dissertation to:

Bernetrice Sain – my supportive and loving mom
Dr. Cody Sain, MD – my big little brother
Keith Hunt, BSN, RN – my little brother from another
Dr. Kayla Key, PhD – my Ride or Die
"The Walkers" Erin, Jess, Hannah, and Owen – Best "neighbors" ever

Thank you ALL for being my ray of sunshine when my blue skies were grey…

Acknowledgements

Thank you to my committee members, Kenneth J. Sytsma, Cecile Ane, Anne Pringle and Veronica S. Di Stilio for all of your support and guidance throughout the completion of my dissertation.

I would like to acknowledge the funding sources that supported my dissertation research: The University of Wisconsin Botany Department Endowment funds and the Wisconsin Alumni Research Foundation (WARF) Fall Competition research and financial support. I also thank the Graduate Research Scholar Community for financial funding from the State of Wisconsin Advanced Opportunity Program and also for providing a supportive community of both graduate students and faculty during my time at UW-Madison. I would like to acknowledge the UW-Madison Biotechnology Center for genomic DNA extractions, and sequencing.

Another big thank you and acknowledgement to all members of the UW Botany department for making my time at UW enjoyable and getting me to the finish line. The support and friendship that my fellow graduate students, undergraduate mentees and students, office staff, research staff, and teaching staff have given me will carry me on throughout my next steps in life.

Last but most definitely not least to my advisor, David Baum, I must say you have been the most supportive mentor I could have possible had to guide me through the process of completing my doctoral degree. I thank you for all of your dedication to making me the best science researcher and teacher that I could be. You have taught me many invaluable lessons that I will carry with me throughout my career and life.

Thank you to everyone who made this dissertation possible.

Table of Contents

Abstract

The evolution of dioecy, the condition of having populations with distinct male (with only staminate flowers) and female (with only carpellate flowers) plants, has been widely studied in angiosperms. In particular, there is great interest in why there were so many independent transitions from the more common condition of cosexuality (flowers with both stamens and carpels) to dioecy and what selective pressures are involved. A key to understanding the evolution of dioecy is to understand the genetic mechanisms responsible for the differentiation of male versus female plants. My dissertation research builds on the existing knowledge of sexual system evolution, by investigating the genomic architecture of sex determination in the genus *Thalictrum* (the meadow-rues, Ranunculaceae). The goal was to build a foundation for determining whether similar or different genetic mechanisms were utilized in the two independent origins of dioecy within the genus. *Thalictrum* is a particularly interesting genus for the study of sex determination because flower developmental studies have suggested that it may use a homeotic mechanism to achieve unisexuality: numerous similar primordia in the center of each flower develop as either stamens or free carpels depending on the sex-determining genetic regions. Using Illumina short read sequencing from pooled same sex individuals of *T. dasycarpum* and *T. dioicum*, representatives of each dioecious clade, I developed a k-mer based pipeline to make inferences about the architecture of sex determination and evaluate the likelihood that similar or distinct genomic regions determine sex in each clade. I found evidence of XY sex determination systems in both *T. dasycarpum* and *T. dioicum*, but the Y-linked genomic region of *T. dasycarpum* appears to be very small and the two species seem to use different genetic mechanisms. Then, I used long read sequencing with low sequencing error rate to construct a high-quality draft genome assembly of a male *T. dioicum* individual to use in conjunction with my short read data to examine if the same or different regions of the genome are used in each clade. I identified a number of contigs that appear to be Y-linked in *T. dioicum.* None of the *T. dasycarpum* male-specific k-mers mapped to candidate Y-linked regions of the *T. dioicum* assembly further supporting the two species using independent mechanisms to regulate gender identity. This work advances our knowledge of how sex is determined in Ranunculaceae and serves to develop *Thalictrum* as a promising new model system for future research on the evolutionary origins of dioecy.

List of Figures

List of Tables

Glossary

Term	Definition
Androdioecy (androdioecious)	Flowering plant sexual system with populations having two kinds of plants: staminate and cosexual
Andromonoecy (andromonecious)	Flowering plant sexual system in which the population consists of one kind of plant that produces both staminate and cosexual flowers
Angiosperms	Flowering plants
Carpel (pistil)	Angiosperm female reproductive organ that produces the ovule/seed
Carpellate (pistillate, female)	Flower having functional female reproductive organs (carpels)
Cosexual (cosexuality, hermaphroditism, bisexuality)	A seed plant that produces both ovules (female) and pollen (male) bearing reproductive structures within a single individual flower or cone
Dioecy (dioecious)	Seed plant sexual system where two classes of sporophyte are present in the population, one (females) producing only carpels/ovules the other (males) producing on stamens/pollen [from the Greek for "two houses"]
Female	The sex of the organism that produces the ovules/eggs
Floral Identity Genes	Homeotic genes involved in floral structure development in flowers
Functional Dioecy	A dioecious plant population where some flowers have rudimentary/non-functional reproductive organs of the other sex
Gametophyte	Multicellular haploid stage in the life cycle of plants with alternating generations
Gynodioecy (gynodioecious)	Flowering plant sexual system with populations having two kinds of plants: carpellate and cosexual
Gynomonoecy (gynomonoecious)	Flowering plant sexual system in which the population consists of one kind of plant that produces both carpellate and cosexual flowers
Hermaphroditism	Former/past terminology used to describe cosexual plants. This term is not used in this thesis due to the negative connotation surrounding its past usage in describing human beings

Heteromorphic sex chromosomes	Sex chromosomes that show visible differences in size and structure
Homeosis	The developmental transformation of one organ into another homologous organ type by mutation or gene misexpression
Homeotic Gene	Genes responsible for determining organ identity
Homomorphic sex chromosomes	Sex chromosomes that lack visible differences in size and structure
Male	The sex of the organism that produces sperm/pollen
Monoecy (monoecious)	Seed plant sexual system where there is only one kind of sporophyte plant that bears both carpels/ovules and stamens/pollen [from the Greek for "one house"]
Sex	The differentiation of male and female organisms
Sex Chromosomes	A pair of differentiated chromosomes that contain genes that determine the sex of an organism
Sex- specific non-recombining region (SNR)	Regions of the sex-chromosomes that show suppressed recombination and include the sex-determining loci
Sporophyte	Multicellular diploid stage in the life cycle of plants with alternating generations
Stamen	Angiosperm male reproductive organ that produces the sperm/pollen
Staminate	Flower having functional male reproductive organs (stamens) and no female reproductive organs (carpels)
Unisexual (unisexuality)	An organism/flower having only one type of reproductive structure, producing either ovule/seed or sperm/pollen

Chapter I: Introduction to dioecy and the flowering plant genus *Thalictrum* L. (Ranunculaceae)

In the animal kingdom, organisms with separate sexes are the norm, but when thinking about plant sporophytes, unisexuality is seen in only a minority of species (Barrett & Hough, 2013). Angiosperms exhibit a diversity of sexual systems, of which the most common is cosexuality, historically referred to as hermaphroditism (Custódio et al., 2018). However, many taxa produce unisexual flowers, either carpellate or staminate, which are often informally called “female” and “male” flowers, respectively. Unisexual flowers can be distributed as in monoecy, where both flower types are present on the same individual plant, or as dioecy where female and males’ flowers occur on different plants. Additional sexual systems can be defined when populations contain both unisexual and cosexual flowers. For example, androdioecious populations consist of male and cosexual plants, whereas gynodioecious ones include female and cosexual plants. Similarly, andromonoecious plants have male and cosexual flowers, while gynomonoecy entails female and cosexual flowers on the same plant. The genus *Thalictrum* has hermaphroditic, andromonoecious, gynomonoecious and dioecious sexual systems (Boivin, 1944). The work described here focuses primarily on dioecy, a sexual system that comes closest to that seen in most animals.

Dioecy exists in ~43% of angiosperm families but only in ~6% of angiosperm species(Renner, 2014; Renner et al., 1995). Thus, the phenomenon of dioecy is rare, but has a phylogenetically widespread occurrence, which suggests multiple independent origins. The repeated evolution of dioecy raises important evolutionary questions such as: What factors might favor the evolution of dioecy? What alternative genetic and developmental changes underlie different origins of dioecy? (Ainsworth, 2000; Barrett, 1992; Käfer et al., 2017).

My research is focused on the evolution of dioecy and the genomic basis of sex determination in the plant genus *Thalictrum* (Ranunculaceae). In this chapter, I review some general patterns seen in the evolution of dioecy, introduce *Thalictrum*, and provide a brief summary of the remainder of this dissertation.

Ecological and morphological traits associated with dioecy

Many macroevolutionary correlations have been detected between the occurrence of dioecy and life-history and ecological attributes, such as latitudinal distribution (tropical vs temperate), growth form (woody vs. herbaceous), pollination mode (abiotic vs biotic), fruit type (fleshy vs dry), and flower size (Bawa, 1980a; Dufay et al., 2014; Renner et al., 1995; Vamosi & Vamosi, 2004). The majority of dioecious species have been noted to have higher frequencies of tropical distributions, woodiness (vines/trees), fleshy fruits, wind pollination, and small flowers (Givnish, 1980). Vamosi et al. (2003) examined many of these traits by assigning binary states to genera and looking for correlated evolution in a large-scale angiosperm phylogeny using a maximum-likelihood model of binary trait evolution (Pagel, 1994). They showed that many of the traits are correlated with dioecy, but also with one another, implying that many functional associations may influence the evolution of dioecy. As a result, the causal ecological drivers of the evolution of dioecy remain poorly understood.

Dioecious *Thalictrum* species only possess one of the well-known ecological traits that have been shown to have a correlation with dioecy, wind pollination(Soza et al., 2012). Although *Thalictrum* plants are not woody, they are perennials, and it has been shown that when herbaceous species are dioecious, they tend to be perennials (Conn et al., 1980). Renner (2014) conducted a very comprehensive analysis of traits associated with dioecy, monoecy, and gynodioecy among the angiosperms and was able to conclude that wind pollination occurs in at least 31.6% of all dioecious species versus only 5.5-6.4% of non-dioecious angiosperms. The proposed reasoning is that due to the high pollen production needed for wind-pollination, cosexual flowers are less functional for wind pollination due to their stigmas becoming congested with their own pollen (Lloyd & Webb, 1986). This model implies that a male-sterile mutant in a bisexual, wind-pollinated population might be favored because it would have less self-pollen-stigma interference that could block the stigmas from successful cross-pollination. Although it is still not certain whether wind pollination or dioecy evolve first, or if they tend to evolve simultaneously, it has been determined that within *Thalictrum* wind pollination precedes the evolution of dioecy in at least one of two independent origins of dioecy in the genus (Soza et al., 2012).

Along with the ecological trait of wind pollination *Thalictrum* has species that have undergone whole genome duplications, polyploidization, a genetic trait that has also been shown to correlate with the evolution of dioecy (Soza et al., 2013).

<u>Morphological distinctions in unisexual floral development</u>

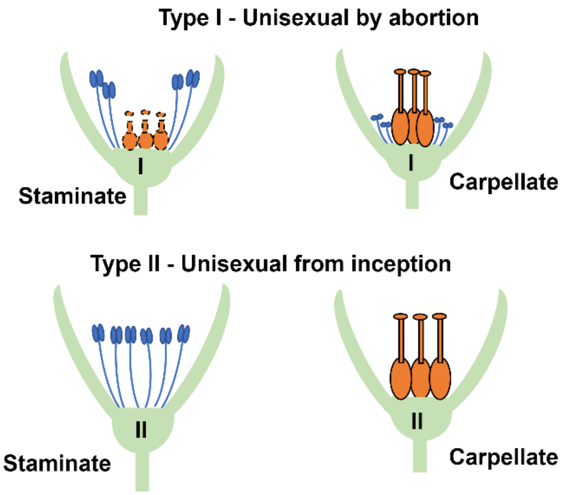

Figure 1 Comparison of Type I and Type II unisexual flowers (adapted from Mitchell & Diggle, 2005). Type I carpel and stamen abortion/non functionality represented by dotted outline and smaller size.

Darwin (1877) originally described two types of unisexuality in flowers: (Type I) “in the female flowers plain rudiments of male organs” or “in the male flowers rudiments of the female organs,” and (Type II) those that “exhibit no rudiments in the one sex of the organs proper to the other” (p.278). The significance of this distinction is still recognized today (Ainsworth, 2000; Dellaporta and Calderon-Urrea, 1993), and can be associated with the question of when in development sexual differentiation occurs (Diggle et al., 2011; Mitchell & Diggle, 2005). Type I flowers become unisexual by aborting either the carpellate or staminate reproductive structures late in development whereas true Type II flowers are unisexual at inception and do not possess any cosexual stage (Fig. 1; adapted from Mitchell and Diggle, 2005). The evolution of unisexuality from cosexuality can, thus, proceed via two developmental paths. There is, however, also the possibility of evolutionary shifts from Type I to Type II unisexuality when organ abortion moves progressively earlier in ontogeny (Mitchell and Diggle, 2005).

Type I dioecy is much more common and, as a result, mechanisms of sex-determination have been determined for several Type I dioecious species. In contrast, the only Type II system elucidated to date is spinach (Diggle et al. 2011; Golenberg and West, 2013). Many dioecious *Thalictrum* are of Type II, so it will be interesting to see if the genetic and developmental mechanisms involved resemble spinach.

Pathways to dioecy

There have been many theoretical and empirical studies conducted to understand the transitions between different sexual systems (Bawa, 1980; Charlesworth and Charlesworth, 1978, 1979; Darwin, 1877; Lloyd, 1975; Ross and Weir, 1976; Webb, 1999). Several different pathways have been suggested for the evolution of dioecy, of which I will focus only on pathways that could apply to *Thalictrum*, namely the evolution of dioecy directly from cosexuality, via gynodioecy, via androdioecy, or via monoecy.

The evolution of dioecy via gynodioecy has received the most attention historically, in part because gynodioecy and dioecy co-occur within the same genera in angiosperms more often than would be expected if each had evolved independently (Dufay et al., 2014). As described by Charlesworth and Charlesworth (1978), the transition to dioecy via gynodioecy can involve three phases. The first phase consists of the invasion of a male sterility mutation into a cosexual population. The mutants (females) must successfully reproduce at a frequency that is higher than the cosexual flowers in the population for the mutant allele to become established. This would be aided by the presence of inbreeding depression and selfing within the cosexual population. As a result of the increasing frequency of females, selection then favors alleles that cause cosexuals to increase investment in pollen production even if they lower investment in ovules. This is likely because cosexuals have more potential partners with which they can act as pollen parents than with which they can act as seed parents. This can result in a feedback loop, where male-sterile, plants function as females and provide most of the population's seed production. The next phase consists of the appearance of one or multiple female-sterilizing mutation(s), which generates male plants in the gynodioecious population. Such female-sterile mutations can be favored by selection if they result in an increase in male function sufficient to compensate for the loss of female function. The presence of male plants establishes a subdioecious (i.e., male + female + cosexual plants; also called trioecious) population. The final phase is the loss of cosexuals from the population, which is likely to entail a mutation that causes a suppression of recombination between the original male-sterility and later-arising female-sterility genes. In principle, the mirror image pathway is possible starting with female-sterility, i.e., androdioecy, although this generally gets less attention (Charlesworth and Charlesworth 1978)

Although the evolution of dioecy via monoecy has not received as much attention, Renner and Ricklefs (1995) found support for clustering of monoecy and dioecy within the angiosperm families. The genetic model for the monoecy pathway assumes that reduced selfing is the selective pressure acting on the population to favor plants becoming specialized on the production of primarily female or male flowers. Therefore, the evolution of dioecy via monoecy requires only one or multiple mutation(s) that alter the ratio between male and female flowers on each plant (Charlesworth and Charlesworth, 1979, Lloyd, 1980).

Developmental genetics of sexual dimorphism in dioecious plants

The evolution of dioecy and sex determination has been studied in select model plant species, especially *Silene latifolia* (white campion) and *Spinacia oleracea* (spinach). *Silene latifolia* has been a prominent model system for studying plant sex chromosome evolution because, like mammals it possesses a single visually distinctive Y chromosome that determines the production of stamens and suppression of carpels (Bergero et al., 2007; Westergaard, 1946). *Silene latifolia* exhibits Type I unisexual flowers, where differentiation entails the abortion of sexual organs (Hardenack et al., 1994). This contrasts with *Spinacia oleracea*, which have Type II unisexual flowers (Sherry et al., 1993).

Sex chromosome evolution is very well studied in model animal systems and now has begun to be examined in diverse plant systems, (Ming et al., 2011). Thus far, plant sex chromosomes have been documented and studied in ~40 plant taxa (Ming et al., 2011; Muyle, 2017). It appears that most dioecious species have homomorphic sex chromosomes, yet it is reasonable to assume that they have specific sex-determining loci organized into sex-linked regions with reduced recombination. These sex chromosomes can be XY, with male-heterogamety (e.g., Fujito et al., 2015; Harkess, 2017), or ZW, with female heterogamety (e.g., Käfer et al., 2022).

Two-gene models for the development of sex chromosomes in plants, of which the model of evolution proposed by Charlesworth and Charlesworth (1978; explained above) is the most famous, assumes there are two distinct sex determining genes, one for male sterility and one for female sterility, and these are located on a single chromosome and are tightly linked. Some support for this model has been found in *Silene latifolia,* although the exact sex determining genes are still unknown (Bergero et al., 2008). Under the "two-gene model," and assuming that most plants evolve dioecy via gynodioecy, most plants are predicted to have XY systems rather than ZW ones. This follows because the initial male-sterilizing mutation is likely to be recessive. The relative rarity of the ZW sex chromosome system could be because androdioecy provides a less common pathway to dioecy (Charlesworth & Charlesworth, 1978; Ming et al., 2011). It is unknown how prevalent this two-gene model is throughout angiosperms because of bias in taxon sampling for species with visible sex chromosomes (Charlesworth & Charlesworth, 1978; Ming et al., 2011; Renner, 2016).

Methods for identifying sex-determining genes in plants

Muyle et al. (2017) reviewed the various techniques that have been used to characterize dioecious plant sex determination. Sequencing the sex-specific nonrecombining regions (SNRs) of sex chromosomes can be difficult due to numerous repetitive sequences within SNRs that tend to accumulate in areas of low recombination (Wang et al. 2012). The studies reviewed in Muyle et al. (2017) were conducted using a combination of methods, including some that used next-generation sequencing with or without explicit genome assembly. There are now several low-cost genetic methods that can generate partial or full genome assemblies, which can provide a robust basis for characterizing sex-determining genes or chromosomes.

The first category of approaches requires sex chromosomes that are old and sufficiently diverged that reads from one sex chromosome do not map to the other. In such a case, comparison of draft male and female genome assemblies can be used to identify X-linked scaffolds based on the ratio of male over female reads being ~0.5, with males having about half as deep coverage of the X-link scaffolds than of autosomal contigs. Y-linked sequences can be detected by using an assembled genome and female short read sequences because female reads should only match autosomal, or X-scaffolds so long as shared male and female repeats have been removed prior to the analysis. In all cases, it is possible to increase the discovery efficiency by using a reference genome in combination with a male and female genome of the target species, plus perhaps transcriptomic data for male and female flowers.

The second category of approaches can be used with young chromosomes and works by detecting how markers segregate between the sexes. One method identifies sex-linked markers from polymorphism data between male and females, or uses markers obtained from genetic mapping to locate sex-linked scaffolds in an assembled genome. If the species does not have a closely related reference genome or has a large genome, individuals from an inbred population can be used to obtain RNA-seq data from males and females so as to identify genes showing sex-limited expression, which tend to be enriched on SNRs.

Thalictrum as a model system for studying the genetic basis of dioecy

Thalictrum (Ranunculaceae) is a genus of ~ 200 species that exhibits a variety of sexual systems, including cosexuality, andromonoecy, gynomonoecy, and dioecy (Boivin, 1944; Soza et al., 2012). This group is somewhat unusual among flowering plants with diverse sexual systems in that it does not have any representatives that exhibit gynodioecy, which is thought to be the most common intermediate condition on the path to dioecy (Charlesworth and Charlesworth, 1978; Spigler and Ashman, 2012; Dufay et al., 2014). Within the genus *Thalictrum* there are two distinct dioecious clades, suggesting two independent origins of dioecy (Soza et al., 2012, 2013; Wang et al., 2019). It is worth noting that in one of the dioecious clades, two individuals, *Thalictrum macrostylum* (Penny and Steven, 2009) and *Thalictrum pubescens* (Davis,1997,2001), seem to be androdioecious yet cosexual plants are cryptically female, meaning these species manifest functional dioecy.

Di Stilio et al. (2005) has shown that *Thalictrum dioicum* has a stable sex expression and that the sex ratios of families are 1:1, which helps support the inference of genetic sex determination. Inter-species crosses suggested that males are the heterogametic sex in some dioecious species of *Thalictrum* (Westergaard, 1958). *Thalictrum*, at least *T. dioicum,* which has been studied developmentally, exhibits Type II unisexuality, which occurs early in development, prior to sex organ initiation (Di Stilio et al., 2005).

An unusual feature of *Thalictrum dioicum* is that, since female flowers have numerous free carpels, whose early primordia resemble early stamens, sex determination may involve the homeotic conversion of the same set of sexual primordia into either stamens or carpels (Di Stilio et al., 2005) This suggests a possible one-locus genetic model for the origin of unisexuality, in the context of the ABCE model of floral organ identity determination (Bowman et al., 1991, 2012). This model identifies classes of MADS-box homeotic genes that function together to produce the characteristic distribution of floral organs in angiosperm flowers. Stamen identity arises from the co-expression of B, C, and E genes, whereas carpels identity is due to the co-expression of C and E genes and absence of B-gene expression. Using the names for the homologous *A. thaliana* genes: B function genes, which usually function as a heterodimer, are PISTILLATA (PI) and APETALA3 (AP3), and C function genes are AGAMOUS (AG) homologs (Bowman et al., 1991). It has been shown that many plant taxa have multiple copies of MADS-box genes, and in some cases, paralogs have diverged in expression and function. Within Ranunculales, the clade that includes *Thalictrum*, there are multiple paralogous copies of the B and C floral identity genes (Damerval & Becker, 2017; LaRue et al., 2013; Martínez-Gómez et al., 2021).

LaRue et al. (2013) used a combination of gene expression analyses and virus-induced gene silencing (VIGS) of B class MADS-box genes in *Thalictrum thalictroides* (cosexual) and *Thalictrum dioicum* (dioecious) to explore the genetic mechanisms underlying floral organ identity determination. They were able to show that silencing both paralogs of PISTILLATA (PI) caused conversion of male to female flowers in *T. dioicum* and cosexual to female flowers in *T. thalictroides*. Based on this result, LaRue et al. (2013) proposed a two-step, but single locus mutation process by which the transition to dioecy could have occurred in *Thalictrum*. This model imagines a cosexual ancestor gaining a dominant gain-of-function mutation in a B class gene locus (e.g., *Thd*PI-1/2) resulting in androdioecy followed by recessive loss-of-function mutation in the same locus replacing cosexual flowers with female flowers. This model is unusual in that it allows for a single locus to serve as the sex-determinant, avoiding the need to invoke tight genetic linkage of male- and female-promoting gene functions. Although there was evidence that the *Thd*PI-1/2 B class genes may be involved in sex determination, LaRue et al. (2013) did not specifically test AP3 genes and, therefore, they too may be plausible candidate genes for determining sex in *Thalictrum*.

Overview of this dissertation

This research seeks to elucidate the genomic architecture of *Thalictrum* sex determination in both dioecious clades to determine whether the transition from cosexuality to dioecy utilized similar mechanisms in each independent origin and whether it used the single-locus homeotic mechanism proposed by LaRue et al. (2013).

This first chapter has provided a brief introduction to dioecy and concludes by introducing the flowering plant genus *Thalictrum*, a member of the buttercup family Ranunculaceae, as a system to examine the evolution and genetic underpinnings of dioecy. Chapter two focuses on using a genomic k-mer pipeline to infer the sex determination system in *Thalictrum*'s two dioecious clades and examine if similar genomic motifs are used in each. Chapter three builds on chapter two by assembling a draft genome for a male *T. dioicum* and then mapping the sex specific k-mers from chapter II to identify candidate sex-chromosomes. Overall, this research helps to develop *Thalictrum* as a genomic model system for comparing the genetic architecture of sex-determination in two closely related clades that appear to have transitioned independently to dioecy. Ultimately, it should prove possible to determine whether the sex-determining loci in the two dioecious clades of *Thalictrum* are (1) homologous due to allopolyploidy, (2) independently co-opted from the same genomic region for sex-determination or (3) nonhomologous.

References

Ainsworth, C. (2000). Boys and girls come out to play: The molecular biology of dioecious plants. *Annals of Botany*, *86*(2), 211–221. https://doi.org/10.1006/anbo.2000.1201

Barrett, S. C. H. (1992). Gender variation and the evolution of dioecy in Wurmbea dioica (Liliaceae). *Journal of Evolutionary Biology*, *5*, 423–444. https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1420-9101.1992.5030423.x

Barrett, S. C. H., & Hough, J. (2013). Sexual dimorphism in flowering plants. *Journal of Experimental Botany*, *64*(1), 67–82. https://doi.org/10.1093/jxb/ers308

Bawa, K. S. (1980). Evolution of Dioecy in Flowering Plants. *Source: Annual Review of Ecology and Systematics*, *11*, 15–39. https://www.jstor.org/stable/2096901?seq=1&cid=pdf-

Boivin, B. (1944). AMERICAN THALICTRA AND THEIR OLD WORLD ALLIES. *Contributions from the Gray Herbarium of Harvard University*, *152*(551), 337-377,391-445, 453–487. http://www.jstor.org/stable/417647

Bowman, J. L., Smyth, D. R., & Meyerowitzt, E. M. (1991). Genetic interactions among floral homeotic genes of Arabidopsis. *Development*, *112*, 1–20. http://dev.biologists.org/content/develop/112/1/1.full.pdf

Bowman, J. L., Smyth, D. R., & Meyerowitzt, E. M. (2012). The ABC model of flower development: then and now. *Development*, *139*, 4095–4098. https://doi.org/10.1242/dev.083972

Charlesworth, B., & Charlesworth, D. (1978). A Model for the Evolution of Dioecy and Gynodioecy. *The American Naturalist*, *112*(988), 975–997. http://www.jstor.org/stable/2460344

Conn, J. S., Wentworth, T. R., & Blum, U. (1980). Patterns of Dioecism in the Flora of the Carolinas. *The American Midland Naturalist*, *103*(2), 310–315.

Custódio, J., Cardoso, F., Lacerda Viana, M., Matias, R., Furtado, M. T., Paula De Souza Caetano, A., Consolaro, H., Lourenço, V., & de Brito, G. (2018). Towards a unified terminology for angiosperm reproductive systems. *Acta Botanica Brasilica*, *32*(2), 329–348. https://doi.org/10.1590/0102-33062018abb0124

Damerval, C., & Becker, A. (2017). Genetics of flower development in Ranunculales – a new, basal eudicot model order for studying flower evolution. *New Phytologist*, *216*(2), 361–366. https://doi.org/10.1111/nph.14401

di Stilio, V. S., Kramer, E. M., & Baum, D. A. (2005). Floral MADS box genes and homeotic gender dimorphism in Thalictrum dioicum (Ranunculaceae)-a new model for the study of dioecy. *The Plant Journal*, *41*(5), 755–766. https://doi.org/10.1111/j.1365-313X.2005.02336.x

Diggle, P. K., di Stilio, V. S., Gschwend, A. R., Golenberg, E. M., Moore, R. C., Russell, J. R. W., & Sinclair, J. P. (2011). Multiple developmental processes underlie sex differentiation in angiosperms. *Trends in Genetics*, *27*(9), 368–376. https://doi.org/10.1016/J.TIG.2011.05.003

Dufay, M., Champelovier, P., Kafer, J., Henry, J. P., Mousset, S., & Marais, G. A. B. (2014). An angiosperm-wide analysis of the gynodioecy-dioecy pathway. *Annals of Botany*, *114*(3), 539–548. https://doi.org/10.1093/aob/mcu134

Fujito, S., Takahata, S., Suzuki, R., Hoshino, Y., Ohmido, N., & Onodera, Y. (2015). *Evidence for a Common Origin of Homomorphic and Heteromorphic Sex Chromosomes in Distinct Spinacia Species*. *5*. https://doi.org/10.1534/g3.115.018671

Givnish, T. J. (1980). *Ecological Constraints on the Evolution of Breeding Systems in Seed Plants: Dioecy and Dispersal in Gymnosperms*. *34*(5), 959–972.

Harkess, A. (2017). The asparagus genome sheds light on the origin and evolution of a young Y chromosome. *Nature Communications*, *8*(1), 1279. https://doi.org/10.1038/s41467-017-01064-8

Käfer, J., Bewick, A., Andres-Robin, A., Lapetoule, G., Harkess, A., Caı¨uscaı¨us, J., Fogliani, B., Gâteblé, G., Ralph, P., dePamphilis, C. W., Picard, F., Scutt, C., Marais, G. A. B., & Leebens-Mack, J. (2022). A derived ZW chromosome system in Amborella trichopoda, representing the sister lineage to all other extant flowering plants. *New Phytologist*, *233*(4), 1636–1642. https://doi.org/10.1111/nph.17662

Käfer, J., Marais, G. A. B., & Pannell, J. R. (2017). On the rarity of dioecy in flowering plants. *Molecular Ecology*, *26*(5), 1225–1241. https://doi.org/10.1111/MEC.14020

LaRue, N. C., Sullivan, A. M., & di Stilio, V. S. (2013). Functional recapitulation of transitions in sexual systems by homeosis during the evolution of dioecy in Thalictrum. *Frontiers in Plant Science*, *4*(November), 1–10. https://doi.org/10.3389/fpls.2013.00487

Lloyd, D. G., & Webb, C. J. (1986). The avoidance of interference between the presentation of pollen and stigmas in angiosperms I. Dichogamy. *New Zealand Journal of Botany*, *24*(1), 135–162. https://doi.org/10.1080/0028825X.1986.10409725

Martínez-Gómez, J., Galimba, K. D., Coté, E. Y., Sullivan, A. M., & di Stilio, V. S. (2021). Spontaneous homeotic mutants and genetic control of floral organ identity in a ranunculid. *Evolution and Development*, *23*(3), 197–214. https://doi.org/10.1111/EDE.12357

Ming, R., Bendahmane, A., & Renner, S. S. (2011). Sex chromosomes in land plants. *Annual Review of Plant Biology*, *62*, 485–514. https://doi.org/10.1146/annurev-arplant-042110-103914

Mitchell, C. H., & Diggle, P. K. (2005). The evolution of unisexual flowers: morphological and functional convergence results from diverse developmental transitions. *American Journal of Botany*, *92*(7), 1068–1076. https://doi.org/10.3732/ajb.92.7.1068

Renner, S. S. (2014). The relative and absolute frequencies of angiosperm sexual systems: Dioecy, monoecy, gynodioecy, and an updated online database. *American Journal of Botany*, *101*(10), 1588–1596. https://doi.org/10.3732/AJB.1400196

Renner, S. S. (2016). Pathways for making unisexual flowers and unisexual plants:Moving beyond the “two mutations linked on one chromosome” model. *American Journal of Botany*, *103*(4), 587–589. https://doi.org/10.3732/AJB.1600029

Renner, S. S., Ricklefs, R. E., & Renner, S. S. (1995). Dioecy and Its Correlates in the Flowering Plants. *American Journal of Botany*, *82*(1), 596–606. http://www.jstor.org/stable/2445418

Sherry, R. A., Eckard, K. J., & Lord, E. M. (1993). Flower Development in Dioecious Spinacia oleracea (Chenopodiaceae). *American Journal of Botany*, *80*(3), 283–291. http://www.jstor.org/stable/2445351

Soza, V. L., Brunet, J., Liston, A., Salles Smith, P., & di Stilio, V. S. (2012). Phylogenetic insights into the correlates of dioecy in meadow-rues (Thalictrum, Ranunculaceae). *Molecular Phylogenetics and Evolution*, *63*(1), 180–192. https://doi.org/10.1016/j.ympev.2012.01.009

Soza, V. L., Haworth, K. L., & di Stilio, V. S. (2013). Timing and consequences of recurrent polyploidy in meadow-rues (Thalictrum, ranunculaceae). *Molecular Biology and Evolution*, *30*(8), 1940–1954. https://doi.org/10.1093/molbev/mst101

Vamosi, J. C., & Vamosi, S. M. (2004). THE ROLE OF DIVERSIFICATION IN CAUSING THE CORRELATES OF DIOECY. *Evolution*, *58*(4), 723–731. https://onlinelibrary-wiley-com.ezproxy.library.wisc.edu/doi/pdf/10.1111/j.0014-3820.2004.tb00405.x

Wang, T. N., Clifford, M. R., Martínez-Gómez, J., Johnson, J. C., Riffell, J. A., & di Stilio, V. S. (2019). Scent matters: differential contribution of scent to insect response in flowers with insect vs. wind pollination traits. *Annals of Botany*, *123*, 289–301. https://doi.org/10.1093/aob/mcy131

Westergaard, M. (1958). The Mechanism of Sex Determination in Dioecious Flowering Plants. *Advances in Genetics*, *9*(C), 217–281. https://doi.org/10.1016/S0065-2660(08)60163-7

Chapter II: Genomic evidence for XY sex determination and two independent origins of dioecy in *Thalictrum* (Ranunculaceae)

Abstract

Resolving the genetic mechanisms regulating sexual dimorphism in *Thalictrum* (Ranunculaceae) is of broad importance since the phylogeny suggests two independent origins of dioecy, which allows us to evaluate whether each origin utilized the same or distinct genetic mechanisms. Here, we explore the genetic underpinnings of unisexuality in *Thalictrum* using deep Illumina short read sequencing from one representative of each dioecious clade (*T. dasycarpum* from clade A and *T. dioicum* from clade C, Soza et al 2013). We use a k-mer-based method to make inferences about the architecture of sex determination in each species and to evaluate the likelihood that similar genomic regions determine sex in each dioecious clade. Our findings show that both *T. dasycarpum* and *T. dioicum* have XY sex determination systems while also suggesting that each has utilized different regions of the genome to achieve dioecy. These data also suggest that the Y-linked and X-linked regions may be of comparable size in *T. dioicum*, while the X-linked region might be markedly larger than the Y-linked region in *T. dasycarpum*.

Keywords

genomic; k-mer; short-reads; homeotic unisexuality; evolution; sexual system, eudicot, angiosperm

Introduction

Dioecy, the production of separate staminate/male and carpellate/female plants exists in ~43% of angiosperm families but is only represented in ~6% of angiosperm species (Renner, 2014; Renner et al., 1995). Although only a small fraction of species are dioecious, the fact that this sexual system occurs in nearly half of all families indicates that it has evolved hundreds to thousands of times independently in angiosperms (Renner, 2014). Many macroevolutionary correlations have been detected between the occurrence of dioecy and life-history and ecological attributes, such as tropical distribution, woody growth form, wind pollination, and fleshy fruit (Givnish, 1980; Goldberg et al., 2017; Renner et al., 1995). It has been suggested that independent origins of dioecy might be driven by one or a few adaptive mechanisms (Bawa, 1980). What is less clear is whether there are few or many genetic pathways that can be coopted to produce unisexual individuals, and what might determine which pathway a particular lineage follows.

The meadowrues, *Thalictrum* (Ranunculaceae), are a genus of ~ 200 species that display a variety of sexual systems, including cosexuality, andromonoecy, gynomonoecy, and dioecy (Boivin, 1944; Soza et al., 2012). *Thalictrum* has four features that make it an interesting system to use for understanding the evolution of dioecy in angiosperms. **(1)** It is a member of the Ranunculales clade, thus phylogenetically distant from current model systems for examining genetic sex determination mechanisms (Damerval & Becker, 2017) **(2)** Despite having several sexual systems, *Thalictrum* does not have any representatives that exhibit gynodioecy, which has been suggested to be a transitional stage on the most common pathway to dioecy in angiosperms (Charlesworth & Charlesworth, 1978; Dufay et al., 2014; Spigler & Ashman, 2012). **(3)** *Thalictrum* contains two non-sister dioecious clades, suggesting two independent origins of dioecy from otherwise quite closely related ancestors (Soza et al., 2013). **(4)** Lastly, at least one species, *T. dioicum*, possesses unisexual Type II flowers, which have no evidence of aborted organs of the non-expressed sex (Darwin, 1877; Mitchell & Diggle, 2005), and many free carpels or stamens, opening up the possibility that mutations at a single homeotic locus could account for the transition from cosexuality to dioecy (Di Stilio et al., 2005; LaRue et al., 2013).

To characterize sex determination in *Thalictrum*, we used deep short-read sequencing from *T. dasycarpum and T. dioicum,* representing the two non-sister dioecious clades. Then, we developed a simple k-mer-based method to make inferences about the architecture of sex determination in each species and evaluate the likelihood that they use similar or distinct genomic regions to determine sex. Our findings confirm previous inferences of an XY sex determination system in *T. dioicum using* new genomic (Di Stilio et al., 2005; Westergaard, 1958). We also provide new evidence suggesting that *T. dasycarpum* also has an XY sex determination system and infer that the two species use distinct regions of the genome in expressing dioecy. These results show that *Thalictrum* has great promise as a model system for studying sex-determination in angiosperms.

Materials and Methods

Sample collection and sequencing

Thalictrum dasycarpum and *T. dioicum* were selected to represent the two, non-sister dioecious clades in the genus because both species grow locally, and the latter has the most published resources while also being a low level polyploid with a relatively small genome (V. S. di Stilio et al., 2005; Soza et al., 2012, 2013; D. di Stilio et al., 2010). Short-read sequencing was conducted on the sex-sorted, pooled DNA from two distinct full sibling groups of *T. dioicum.* The full sibling groups were created by crossing a male and female plant sourced in Madison, Wisconsin. Genomic DNA for ten male and ten female individuals from each cross was provided by Verónica Di Stilio (University of Washington, Seattle). The extracted genomic DNA of same-sex full siblings from each cross was pooled together in equal amounts based on the DNA concentration of each sample and used for next-generation short read whole genome sequencing (WGS) (Figure 2). We reference these pooled samples hereafter as male or females of *T. dioicum* cross one and cross two.

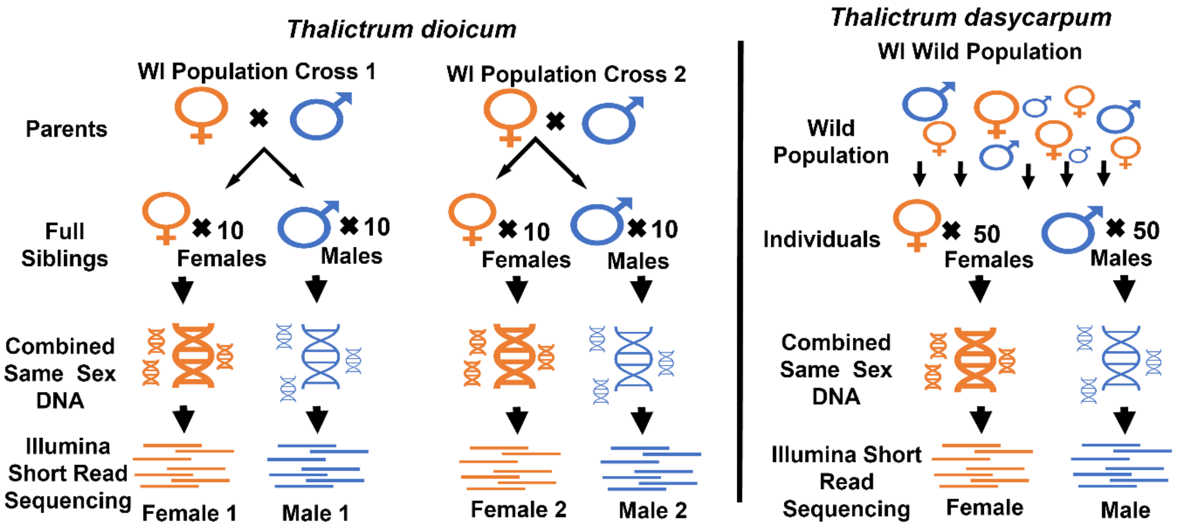

Figure 2 Species sampling. Representation of *Thalictrum* samples used in this study. *T. dioicum* male (blue) and female (orange) pooled samples were from two sets of full siblings resulting from crossing two separate male and female parent plants from a population in Wisconsin. Each sibling group had ten male and female individuals *T. dasycarpum* male and female pooled samples were from a wild population in Madison, WI, and consisted of 50 individuals per sex.

Thalictrum dasycarpum leaf samples from 50 male and 50 female plants were collected from a wild population located in Madison WI (43°02'42.6"N 89°27'16.2"W). Equal amounts (~32mm^2) of leaf tissue from each individual were combined based on sex to create a single male and female sex-specific pool of leaf tissue to use for High Molecular Weight (HMW) DNA extraction, which was performed by the University of Wisconsin-Madison Biotechnology Resource Center using their HMW Plant DNA extraction protocol.

Illumina short read DNA library preparation and whole genome sequencing for all samples were prepared by the University of Wisconsin-Madison Biotechnology Resource Center using their standard protocols. The four, pooled *T. dioicum* genomic DNA paired-end 125bp read libraries (one male and one female from each of two crosses) were sequenced on the Illumina HiSeq2500 High Throughput (V4) platform using the HiSeq SBS Kit V4 and a single flow cell. Pooled *T. dasycarpum (*one male and one female) genomic DNA paired-end 150bp read libraries were similarly prepared and were run on the Illumina NovaSeq 6000 sequencing system.

We used FastQC v 0.11.2 (Andrews, 2015) to assess the quality of all raw sequencing data, and preformed quality trimming and filtering using BBDuk from the BBTools V38. suite (BBMap-BushnellB.-sourceforge.net/projects/bbmap/), with the following settings: ktrim=r k=21 mink=11 hdist=1 hdist2=0 tpe tbo qtrim=r trimq=10 maq=10 minlength=75 maxns=0 forcetrimmod=5. Sequence coverage was estimated based on a genome size of 965-1012 Mb for *T. dioicum* and 3.74 Gb for *T. dasycarpum based* on flow cytometry (Di Stilio et al, 2005; Soza et al., 2013).

Reference Free K-mer Analysis to Distinguish XY vs ZW Sex Determination

We aimed to find k-mers present in one sex only. A flow chart of the k-mer sorting approach used to analyze both *T. dasycarpum and T. dioicum sequencing* datasets is shown in Figure 3.

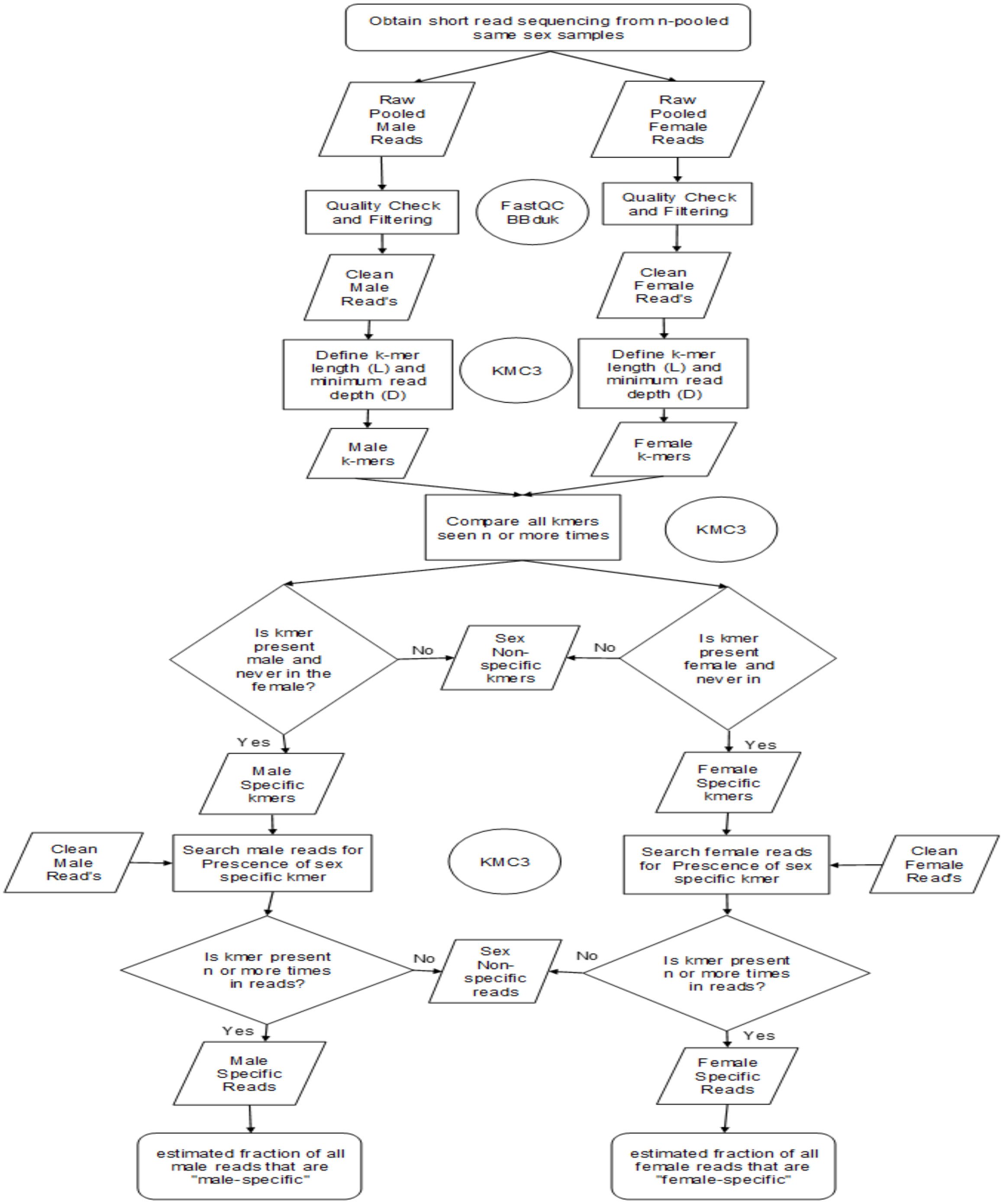

Figure 3 K-mer sorting workflow for sex-determination k-mer analysis. Flow chart of genomic k-mer sorting approach to distinguish between XY and ZW sex-determination systems. A distribution of user chosen k-mers was created for each male and female Illumina short-read sequencing data set. k-mers present N times in both male samples but zero times in females were considered male N-fold enriched K-mers, and vice-versa.

We used quality-filtered short-read Illumina WGS data from each male and female pooled sequencing dataset for each species to create distributions of 51-mer sequences (hereafter referred to as “k-mers”) using the program KMC3 (Kokot et al., 2017). The k-mer value of 51 was chosen based on the results from the program Kmergenie (Chikhi & Medvedev, 2014), which produces an automated best k-mer value to use for genome assemblies based on the sequencing read dataset. We chose 51-mers because it was the lowest best k-mer value that was shared for all genomic samples of *T. dioicum* and *T. dasycarpum* best k-mer ranges. Using KMC3’s kmc_toolbox (Kokot et al., 2017), we isolated sex-specific k-mers, namely those k-mers present in one sex’s k-mer distribution yet absent in the other sex. We explored different upper and lower bound of k-mer occurrence, including the range 15 to 2000 proposed by Neves et al., (2020). The lower bound is intended to avoid counting k-mers that result from sequencing errors, or k-mers that are present in only a small subset of the pooled individuals. The upper bound is there to avoid counting k-mers that likely occur in repetitive elements, since these could occur in only a small number of individuals in the pool yet be highly represented.

Starting with all sex specific k-mers occurring zero times in one sex and 15-2000 times in the other sex, we gradually increased the lower bound up from fifteen, in increments of five up to 200. The goal was to identify the lower threshold at which k-mer occurrences were either stable (not changing markedly with changing threshold) or close to zero. By incrementally increasing the lower-bound occurrence threshold, we reasoned that we are progressively lowering the number of false positives – k-mers only found in one sex that are actually present in both. The counts for specific k-mers contained in the genomic sequencing data for both sexes were plotted based on the determined bin thresholds to examine differences in the pattern of sex specific k-mer presence between male and female genomes.

Determining if non-sister dioecious clades share similar sex specific k-mers

We sought to determine if individual k-mers showed similar sex biases between the two *T. dioicum* crosses and between *T. dasycarpum* and each T. *dioicum* cross. For a given comparison, we restricted consideration to k-mers occurring at least fifteen times in one sample and no more than 2000 times in any sample. We deleted any k-mers that did not meet that threshold in both samples and then binned k-mers based on whether they were male-specific (never occurring in females), female-specific (never occurring in males), or non-sex-specific. A contingency table was then created to summarize the number of shared k-mers in each category.

To complement discrete scoring of k-mers based on sex-specificity, we developed a quantitative measure of sex-bias. We restricted the analysis to shared k-mers occurring at least ten times in at least one sample and no more than 200 times in either sample. The lower bound is approximately half of the total coverage, as inferred from genome size estimates (Table 1). This is warranted since sex-specific nonrecombining region of the putative haploid Y chromosome would have half the coverage of the rest of the genome. The upper bound is set to focus in on k-mers that occur at one or few loci. For each of these k-mers, a sex bias metric S was calculated as: $S = \log2(M+1/F+1)$, where M is the count in males and F is the count in females. S values of zero indicate identical counts in both sexes, whereas positive values indicate male bias and negative values indicate female bias.

Results

Thalictrum dioicum and *T. dasycarpum* contain many more male-specific than female-specific k-mers

We conducted short-read genome sequencing for sex-sorted *T. dioicum and T. dasycarpum,* which each represent different dioecious clades. For *T. dioicum we* sequenced a pool of ten same-sex full sibling individuals from two crosses (both derived from parents in the same Madison, Wisconsin population). In the case of *T. dasycarpum the* male and female pools each contained fifty individuals from a single wild population, also in Madison, WI (Figure *2*). A summary of sequencing results and read depth is in Table 1.

Table 1 Thalictrum short read sequencing results. Information for Illumina HiSeq2500 High Throughput (V4) 2x125bp read sequencing for sex pooled samples of ten males and ten females from each of two greenhouse crosses of *T. dioicum* and Illumina NovaSeq 2x150pb read sequencing for sex pooled samples of fifty female and fifty male T. dasycarpum from a wild population (43.0457463, -89.4526024). F=female and M=male; F1/M1 refer to cross 1; F2/M2 refer to cross 2.

Pooled Samples	Read Length (bp)	Number of Reads (Paired end)	Coverage	%>=Q30 bases	mean quality score	Yield (Mbp)
T. dioicum F1	126	211,150,830	~27x	88.36	34.02	26,605
T. dioicum M1	126	163,599,416	~22x	89.54	34.20	20,614
T. dioicum F2	126	229,251,176	~30x	89.18	34.22	28,886
T. dioicum M2	126	188,479,594	~24x	91.12	34.71	23,748
T. dasycarpum F	151	1,144,000,000	~58x	89.69	35.13	172,754
T. dasycarpum M	151	1,002,400,000	~50x	89.79	35.13	151,357

We determined the number of k-mers that were specific to the male or female pools. To be sex specific, a k-mer must occur a specific minimum number of times in one sex pool and never occur in the other sex pool. At low occurrence thresholds, a large number of sex-specific k-mers were detected, especially in cross 1. There were more female-specific k-mers at low occurrence thresholds for both species which may be an artifact reflecting the larger number of reads obtained for each female pool than for the corresponding male pools. The difference in sequencing results for females and males could be an artifact of the sequencing technology, because to obtain the guaranteed read numbers the male samples had to be resequenced whereas, by chance the female samples exceeded the guarantee on the first run. Nonetheless in each case, as we raised the occurrence threshold more male-specific than female-specific k-mers were detected, with female-specific k-mers tending to disappear (Figure 4).

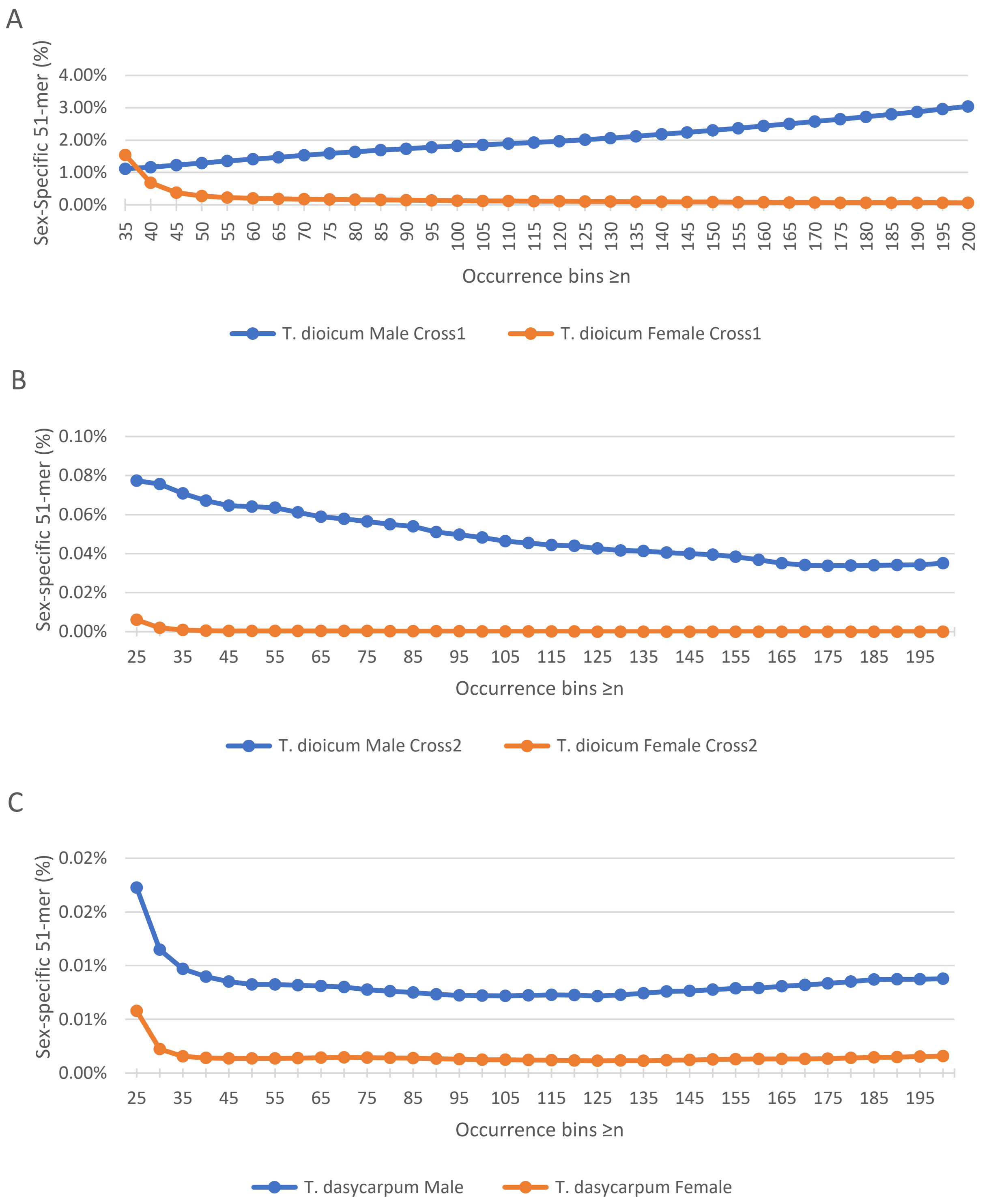

Figure 4 Confirmation of sex determination system based on the distribution of distinct sex specific 51-mers. Line graph of the presence of enriched T. dioicum (A, B) and T. dasycarpum (C) male (blue) and female (orange) 51-mers with a presence threshold of 15-2000 that appeared in the genome at lower bin thresholds beginning at n≥15 times and increasing in increments of five. T. dioicum cross1 (A), cross2 (B), and T. dasycarpum (C) sex-specific k-mer counts are shown as the percent of total possible k-mers within each threshold to allow for comparisons of the segregation of sex-biased representation of 51mers seen in the genome of each sample.

Sex specific k-mers differ between *T. dioicum and T. dasycarpum*

Each sample contained very different numbers of sex specific k-mers, when using an occurrence threshold of 15 ≤ n ≤2000 (Table 2). At this occurrence threshold, there are many more female-specific than male-specific k-mers, which likely reflects a large number of k-mers that are erroneously scored as sex-specific (Figure 4). Nonetheless, restricting analysis to k-mers shared between each pair of samples, it was notable that the two *T. dioicum crosses* have a large number of shared k-mers with the same sex specificity (Table 3A). Many shared-mers that are sex specific in one cross are not sex specific in the other, but shared k-mers that are male specific in one cross are 1.5-29 times more likely to be male specific than female specific in the other cross and female specific k-mers are 8.8-16.8 times more likely to be female-specific. Some shared female specific k-mers can be hypothesized to be X chromosome sequences that were not detected in males (which have half as many copies).

Table 2 Total 51-mer counts in *T. dasycarpum* and *T. dioicum* Cross 1 and 2 within the occurrence threshold 15≤ n ≤2000. The counts represent the number of 51-mers seen in each sex specificity with the total possible counts for each category in parentheses.

51-mer Category	*T. dasycarpum*	*T. dioicum* Cross 1	*T. dioicum* Cross 2
Male Specific	490,000 (1,210,000,000)	28,167,095 (117,294,099)	100,107 (135,396,793)
Female Specific	1,930,148 (1,400,000,000)	135,245,013 (224,372,017)	232,945 (257,621,434)
Nonspecific	2,607,579,852 (2,610,000,000)	178,254,008 (341,666,116)	392,685,175 (393,018,227)

Table 3 Shared 51-mers represented between all samples. Contingency table comparing shared sex specific and non-sex-specific 51-mer presence in the genome (A) *T. dioicum* cross one (yellow) vs cross two (red). (B) *T. dioicum* cross 1 and *T. dasycarpum* (grey) (C) *T. dioicum* cross 2 and *T. dasycarpum.* Male specific 51-mers shared between samples are highlighted in blue and the female specific 51-mers shared between samples are highlighted in orange.

A		***T. dioicum* Cross 1**			
		Male Specific	**Female Specific**	**Non-Sex Specific**	**Total Shared**
***T. dioicum* Cross 2**	**Male Specific**	5,299	3,495	23,237	32,031
	Female Specific	1,824	30,551	2,869	35,244
	Non-Sex Specific	4,374,251	21,885,361	54,022,383	80,281,995
	Total shared	4,381,374	21,919,407	54,048,489	80,349,270
B		***T. dioicum* Cross 1**			
		Male Specific	**Female Specific**	**Non-Sex Specific**	**Total shared**
T. dasycarpum	**Male Specific**	819	4,173	4,253	9,245
	Female Specific	1,156	5,549	4,480	11,185
	Non-Sex Specific	7,273,067	35,921,967	41,228,587	84,423,621
	Total Shared	7,275,042	35,931,689	41,237,320	84,444,051
C		***T. dioicum* Cross 2**			
		Male Specific	**Female Specific**	**Non-Sex Specific**	**Total Shared**
T. dasycarpum	**Male Specific**	0	3	4,728	4,731
	Female Specific	0	0	5,224	5,224
	Non-Sex Specific	3,277	18,934	46,388,380	46,410,591
	Total shared	3,277	18,937	46,398,332	46,420,546

In contrast to the many shared sex-specific k-mers between the two *T. dioicum crosses*, a comparison of *T. dasycarpum* and *T. dioicum crosses* suggests a lack of shared sex-specific k-mers (Table 3B, C). When comparing *T. dasycarpum* to *T. dioicum cross* 1, there are 32 million k-mers that are sex specific in either species, of which only 6,368 (5,549 + 819) show matching sex specificity, where 5,329 (4,173 + 1,156) show opposite-matching specificity, that is male specific k-mers in one species being female specific in the other (Table 3B). Interestingly, k-mers that are sex specific in *T. dasycarpum* but shared with *T. dioicum*, are about five times more likely to be female specific than male specific in *T. dioicum* cross 1, regardless of whether they are male or female specific in *T. dasycarpum.* Nonspecific *T. dasycarpum* k-mers shared with *T. dioicum* are also five times more likely to be female than male specific in *T. dioicum* cross 1. The lack of correlated sex specificity between the two species is even more obvious in cross 2: of the 32,166 k-mers that are sex specific in either species, none show matching sex specificity (Table 3C).

K-mer sex bias is not correlated between *T. dioicum and T. dasycarpum*

We examined the sex-bias metric (*S*) of k-mers for each sample. The two *T. dioicum* crosses each show an average k-mer bias greater than 0, consistent with an excess of putatively Y-chromosome k-mers Figure 5; (Table *4*). In contrast, *T. dasycarpum* k-mers have a negative mean and median, implying an excess of female-biased k-mers Figure 5; (Table *4*).This may appear in conflict with our detection of more male-specific than female specific k-mers in *T. dasycarpum (*Figure 4), but it is worth noting that k-mers restricted to X-chromosomes are not expected to be sex-specific but are expected to have a female-biased occurrence. This means that if there were many more X-linked than Y-linked k-mers, for example because the Y-specific region of the genome is much smaller than the X-specific region, the average sex-bias can be negative even under XY sex determination. Recall, that the expected value of *S* for an X-linked k-mer is -1.

To compare sex-bias between samples, we first restricted the comparison to shared k-mers that occurred at least ten times in one sample and no more than 200 times in any sample. The two *T. dioicum* crosses shared more k-mers with each other than either did with *T. dasycarpum* (Table *4*): *T. dioicum* cross one and two shared 346,163,900 k-mers which represents 72% and 67% of the total pool of k-mers in cross1 and cross2 respectively (Table *4*). Of the 2,359,706,727 k-mers in *T. dasycarpum,* less than 7% were shared with either *T. dioicum* cross (Table *4*).

Table 4. Sex bias statistics of shared 51-mers within *Thalictrum* species. Each sample represents the restricted data set with a lower threshold of occurring ≥ 10 in either sex per sample while not exceeding an upper threshold occurrence of 200 in both sexes. Descriptive statistics for each restricted sample of shared k-mers statistics for *S* are provided: median, mean (r=Pearson's correlation coefficient), standard error of the mean (SEM) standard deviation (SD). Information in each data row corresponds to the species listed first in the pair.

Sample	51-mer Occurrence	Median	Mean (r)	SEM	SD
Individual K-mer Sample Data					
T. dasycarpum	2,359,706,727	-0.26	-0.51	4.84E-05	2.35
T. dioicum Cross 1	482,197,730	0.42	0.41	3.58E-05	0.79
T. dioicum Cross 2	514,413,929	0.44	0.43	3.47E-05	0.79
Shared k-mer Sample Data					
T. dasycarpum x *T. dioicum* Cross 1	152,043,814	-0.20	-0.27 (-0.0012)	1.02E-04	1.25
T. dioicum Cross 1 x *T. dasycarpum*	152,043,814	0.39	0.40 (-0.0012)	5.27E-05	0.65
T. dasycarpum x T. dioicum Cross 2	157,000,339	-0.20	-0.27 (-0.0030)	1.01E-04	1.26
T. dioicum Cross 2 x *T. dasycarpum*	157,000,339	0.42	0.41 (-0.0030)	5.09E-05	0.64
T. dioicum Cross 1 x *T. dioicum* Cross 2	346,163,900	0.39	0.40 (0.0200)	3.55E-05	0.66
T. dioicum Cross 2 x *T. dioicum* Cross1	346,163,900	0.42	0.41 (0.0200)	3.45E-05	0.64

Visual examination of the density of *S* in the pairwise comparison suggests a positive correlation between the two *T. dioicum* crosses, primarily driven by shared male-biased k-mers (Figure 6). In contrast, there was no obvious positive correlation of sex-bias in k-mers shared between *T. dioicum* and *T. dasycarpum* – indeed there appears to be a negative correlation of sex bias between the two species, presumably driven primarily by the negative bias seen in *T. dasycarpum* k-mers (Figure 6).

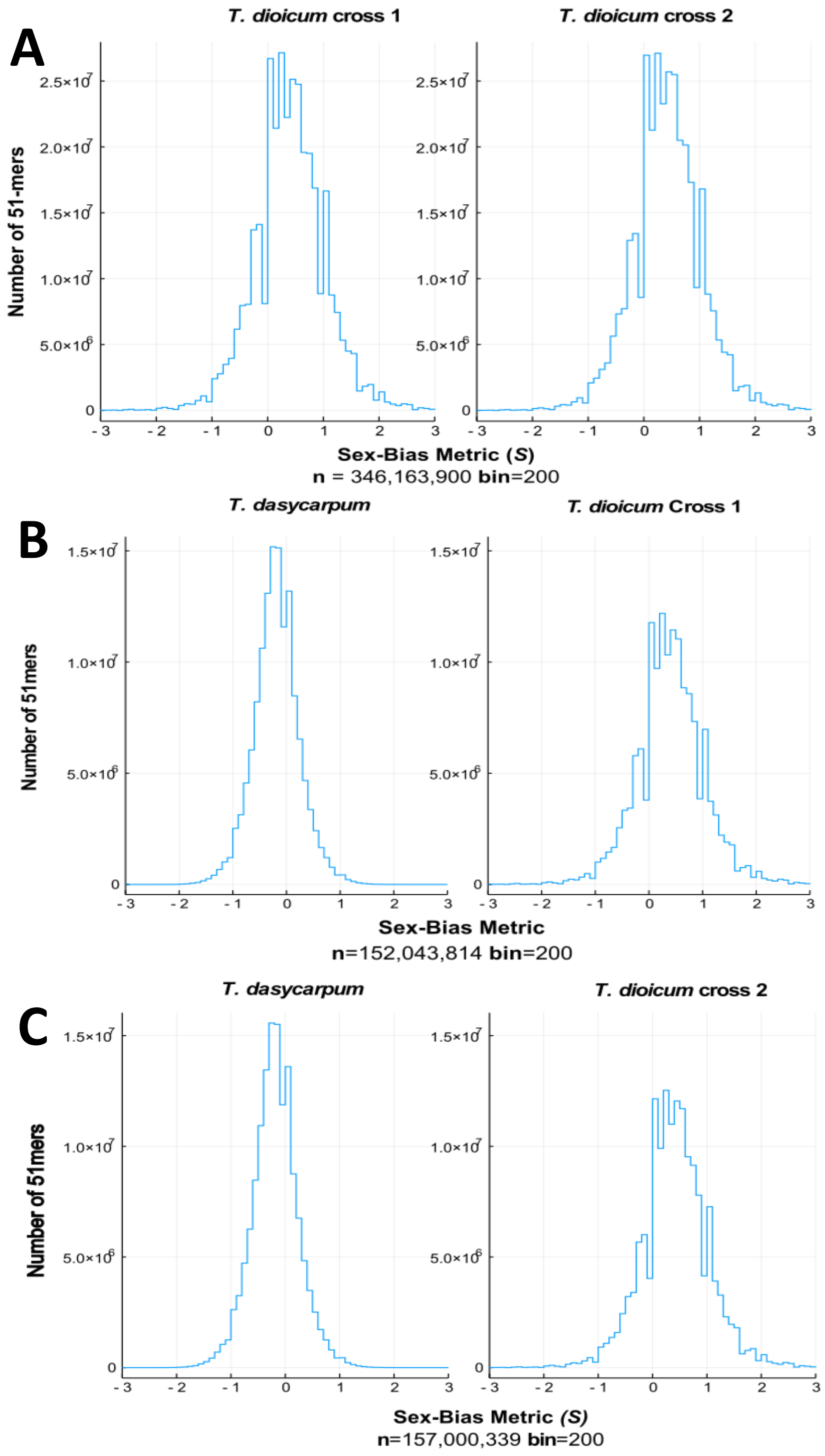

Figure 5 Distribution of sex-bias (S) in 51-mers shared between Thalictrum samples. Shared k-mers are based on an occurrence for each sample of 10 in either sex, and exclusion of k-mers that occur in both sexes with values above 200. Histograms show the sex bias of shared 51mers between T. dioicum cross1 and cross2 (A), *T. dioicum* cross1 and *T. dasycarpum* (B), and *T. dioicum* cross2 and *T. dasycarpum* (C) k-mer sex-bias metric (S) is the log 2-fold-change in the occurrence of male and female 51-mers present within each sample's genome (sex-bias(S) =log2((male+1)/(female+1)).

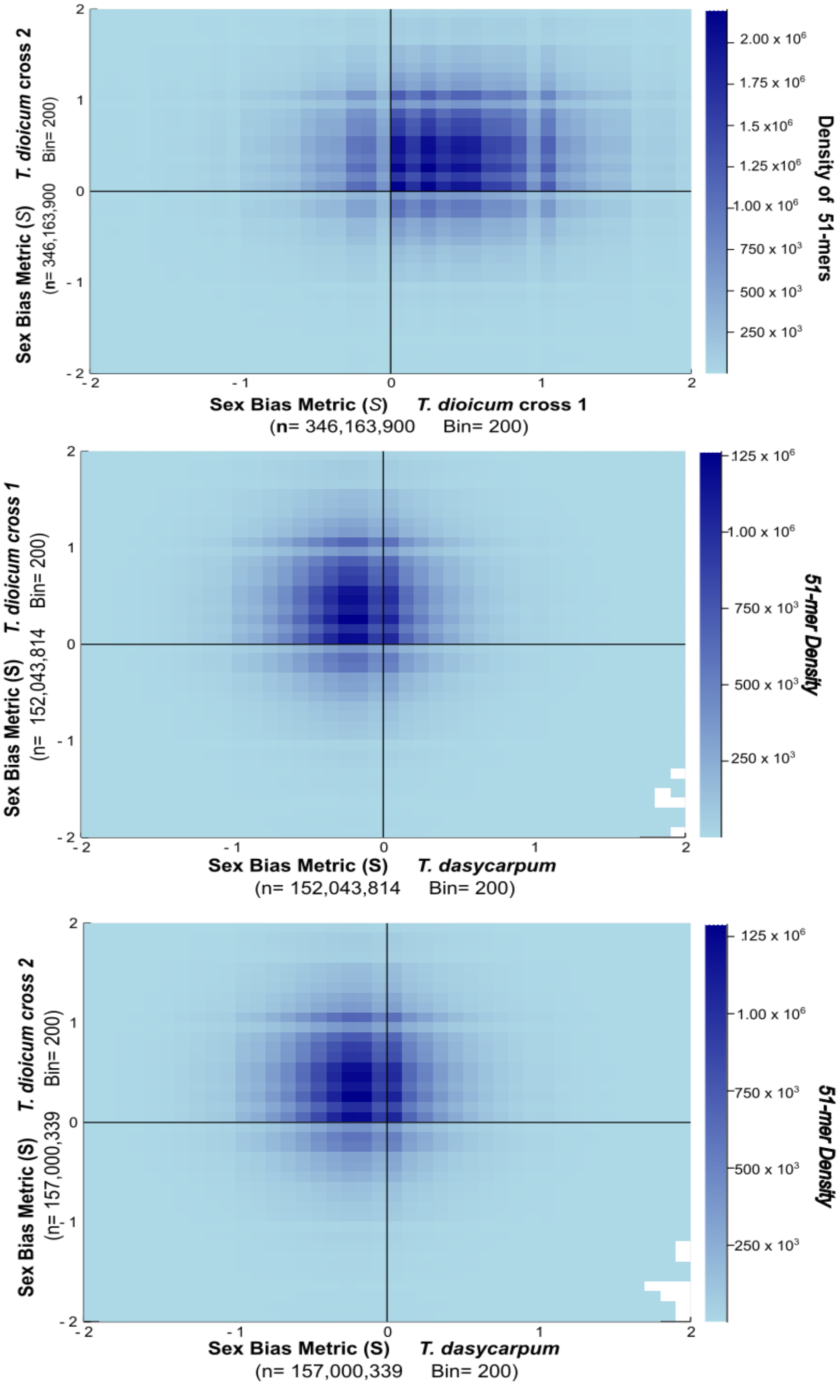

Figure 6 2D-histogram of sex-specific 51-mers shared within and between *Thalictrum* species. 2D-histogram showing the density of shared sex-biased 51-mers with an occurrence of 10 in either sex, and exclusion of k-mers that occur in both sexes with values above 200 for each sample. Each plot shows the density of shared k-mers shared between *T. dioicum* cross1 and cross2 (A); *T. dasycarpum* and *T. dioicum* cross1 (B); *T. dasycarpum* and *T. dioicum* cross2 (C)**.** The sex-bias metric (S) is the log2-fold change seen between male and female 51-mers seen in each sample (Plots are zoomed to 4x change, the *S*= $\log_2((M+1)/(F+1))$, M=male, F=female)

Discussion

We examined the genomic architecture of sex determination in the meadowrue genus *Thalictrum* by comparing the k-mers present in the sequencing reads of pools of same sex individuals from *T. dioicum* and *T. dasycarpum.* We hypothesized that, even without a genome assembly, we would be able to make robust inferences about the sex determination system in each species and the degree to which similar genomic regions play a role in sex determination. Using simple k-mer-based analyses it was possible to find k-mers with sex-specific and sex-biased occurrences and use that information to make inferences as to which sex is heterogametic and also to evaluate whether the two species have homologous male-specific (Y) genomic regions.

Genomic confirmation of an XY Sex Determination System in *Thalictrum*

Using the distribution of sex specific k-mers we expected that only the heterogametic sex (XY or ZW) would retain sex-specific k-mers once the threshold for occurrence was set high enough to exclude k-mers that are present in both sexes but were only detected in one. Applying this approach to both species of *Thalictrum*, we found an excess of male-specific k-mers at high occurrence thresholds. These data support male heterogamety and thus genetic XY sex determination in both species. While there are cases of ZW sex determination, the majority of previously studied dioecious species also have an XY system (Carey et al., 2021), and a few others are known to have an X-autosome ratio system similar to that found in *Drosophila melanogaster (e.g. Rumex;* Ainsworth, 2000; Ainsworth et al., 1999).

In a strictly XY system, only k-mers that are limited to Y-linked genomic regions are expected to be sex specific. However, k-mers restricted to both X- and Y-linked genome regions have a sex-biased occurrence. Y-linked genes should be absent in females meaning they will have very high positive S statistics. X-linked genes should have about half as many occurrences in males than in females, which corresponds to an S-statistic of about –1. This means that the average value of S across all k-mers contains information on the relative number of X- versus Y-linked k-mers. The fact that *T. dioicum* genomes have an average *S* that is positive, but significantly less than +1.0, is consistent with the X and Y being comparable in size. In contrast the negative value of *S* in *T. dasycarpum* suggests that the pool may include many more X-linked than Y-linked k-mers, which might imply that the Y-linked region is small relative to the X-linked regions.

It is worth considering the possibility that *T. dasycarpum* has an X:Autosomal sex system, where the effect of the Y-chromosome can be overcome by a sufficient dosage of X-chromosomes. An XO system, which lacks a Y-chromosome, seems to be ruled out by the presence of many male-specific k-mers. However, given polyploidy in *T. dasycarpum*, it is possible there is a Y-chromosome that is male determining but can be overcome when sufficient X chromosomes are present (Warmke, 1946; WesterGaard, 1938). This scenario cannot be ruled out but seems incompatible with a consistent 50:50 sex ratio. Thus, we consider it slightly more plausible that this species has an XY system with a relatively small Y-linked region.

Genomic analyses support two independent origins of XY-based dioecy in *Thalictrum*

The genetic architecture of angiosperm sex determination has only been established for a small fraction of all dioecious flowering plant species, and there has yet to emerge a consensus as to the degree to which similar genetic mechanisms are coopted in independent origins of dioecy (Bachtrog et al., 2014; Leite Montalvão et al., 2021; Moore et al., 2016). Mapping of sexual systems onto the phylogeny of *Thalictrum* has suggested there were at least two separate dioecious origins within the genus (Soza et al., 2012, 2013). However, ancestral state reconstruction is rarely definitive because of the possibility of undetected homoplasy, for example multiple reversals of dioecy to cosexuality. Additional insights can come from comparing the genetic basis of sex-determination in each clade to see if similar genomic regions are involved, which would support a shared origin of dioecy, or if distinct genomic regions are involved which would tend to confirm independent origins. Our analyses of sex-specific k-mers suggests that, indeed, the *T. dioicum* and *T. dasycarpum* clades might have evolved dioecy independently.

By comparing sex-specific k-mers between the two species, we failed to find evidence of the same k-mers showing the same sex-specificity pattern in the two species at a rate greater than expected by chance. Likewise, we did not detect a correlation between a k-mer's sex bias statistic in the two species. These facts imply either that the two species coopted distinct parts of the genome or different chromosomal mechanisms when their ancestors independently transitioned to dioecy or that the sex-determining regions of the genome have diverged too greatly for sex-biased k-mers to be conserved. To distinguish between these, we could look at shorter k-mers or alternatively create a genome assembly and look for homologous genes (Chapter III).

Our findings begin to explain how unisexuality may be determined in *Thalictrum*. However, to completely understand and describe the genetic architecture of sex determination in *Thalictrum*, it will be important to create a complete genome assembly and annotation from at least male representatives of each dioecious clade. These male genomes can be used in combination with the short read data discovered in this study to locate the sex determining region(s) in the genome. Once sex determining region(s) are located, characterization and evaluation of the expression of sex-linked genes can also shed light on whether the transition from cosexuality to dioecy in some *Thalictrum* species could have been triggered by mutations at a single homeotic locus (Di Stilio et al., 2005; LaRue et al., 2013).

Acknowledgements

The authors thank the University of Wisconsin Biotechnology Center DNA Sequencing Facility for high molecular weight DNA extractions and Illumina DNA sequencing facilities and services including short read library preparation and sequencing. We thank the Wisconsin Alumni Research Foundation- Fall

Competition Grant and the University of Wisconsin Botany Department Endowment funds for helping fund this project.

References

Ainsworth, C. (2000). Boys and girls come out to play: The molecular biology of dioecious plants. *Annals of Botany*, *86*(2), 211–221. https://doi.org/10.1006/anbo.2000.1201

Ainsworth, C. C., Lu, J., Winfield, M., & Parker, J. S. (1999). Sex determination by X: autosome dosage: Rumex acetosa (sorrel). In *Sex determination in plants* (pp. 124–136).

Bachtrog, D., Mank, J. E., Peichel, C. L., Kirkpatrick, M., Otto, S. P., Ashman, T.-L., Hahn, M. W., Kitano, J., Mayrose, I., Ming, R., Perrin, N., Ross, L., Valenzuela, N., & Vamosi, J. C. (2014). Sex Determination: Why So Many Ways of Doing It? *PLOS Biology*, *12*(7). https://doi.org/10.1371/journal.pbio.1001899

Bawa, K. S. (1980). Evolution of Dioecy in Flowering Plants. *Annual Review of Ecology and Systematics*, *11*, 15–39. https://doi.org/10.1146/annurev.es.11.110180.000311

Boivin, B. (1944). AMERICAN THALICTRA AND THEIR OLD WORLD ALLIES. *Contributions from the Gray Herbarium of Harvard University*, *152*(551), 337-377,391-445, 453–487. http://www.jstor.org/stable/417647

Carey, S., Yu, Q., & Harkess, A. (2021). The Diversity of Plant Sex Chromosomes Highlighted through Advances in Genome Sequencing. *Genes 2021, Vol. 12, Page 381*, *12*(3), 381. https://doi.org/10.3390/GENES12030381

Charlesworth, B., & Charlesworth, D. (1978). A Model for the Evolution of Dioecy and Gynodioecy. *The American Naturalist*, *112*(988), 975–997. http://www.jstor.org/stable/2460344

Chikhi, R., & Medvedev, P. (2014). Informed and automated k-mer size selection for genome assembly. *Bioinformatics*, *30*(1), 31–37. https://doi.org/10.1093/BIOINFORMATICS/BTT310

Damerval, C., & Becker, A. (2017). Genetics of flower development in Ranunculales – a new, basal eudicot model order for studying flower evolution. *New Phytologist*, *216*(2), 361–366. https://doi.org/10.1111/nph.14401

Darwin, C. (1877). *The Different Forms of Flowers on Plants of the Same Species* (J. Murray, Ed.). WILLIAM CLOWES AND SONS.

di Stilio, V. S., Kramer, E. M., & Baum, D. A. (2005). Floral MADS box genes and homeotic gender dimorphism in Thalictrum dioicum (Ranunculaceae)-a new model for the study of dioecy. *The Plant Journal*, *41*(5), 755–766. https://doi.org/10.1111/j.1365-313X.2005.02336.x

Dufay, M., Champelovier, P., Kafer, J., Henry, J. P., Mousset, S., & Marais, G. A. B. (2014). An angiosperm-wide analysis of the gynodioecy-dioecy pathway. *Annals of Botany*, *114*(3), 539–548. https://doi.org/10.1093/aob/mcu134

Givnish, T. J. (1980). *Ecological Constraints on the Evolution of Breeding Systems in Seed Plants: Dioecy and Dispersal in Gymnosperms*. *34*(5), 959–972.

Goldberg, E. E., Otto, S. P., Vamosi, J. C., Mayrose, I., Sabath, N., Ming, R., & Ashman, T.-L. (2017). Macroevolutionary synthesis of flowering plant sexual systems. *Evolution*, *71*(4), 898–912. https://doi.org/10.1111/evo.13181

Kokot, M., Dlugosz, M., & Deorowicz, S. (2017). KMC 3: counting and manipulating k-mer statistics. *Bioinformatics (Oxford, England)*, *33*(17), 2759–2761. https://doi.org/10.1093/bioinformatics/btx304

LaRue, N. C., Sullivan, A. M., & di Stilio, V. S. (2013). Functional recapitulation of transitions in sexual systems by homeosis during the evolution of dioecy in Thalictrum. *Frontiers in Plant Science*, *4*(November), 1–10. https://doi.org/10.3389/fpls.2013.00487

Leite Montalvão, A. P., Kersten, B., Fladung, M., & Müller, N. A. (2021). The Diversity and Dynamics of Sex Determination in Dioecious Plants. *Frontiers in Plant Science*, *0*, 2280. https://doi.org/10.3389/FPLS.2020.580488

Mitchell, C. H., & Diggle, P. K. (2005). The evolution of unisexual flowers: morphological and functional convergence results from diverse developmental transitions. *American Journal of Botany*, *92*(7), 1068–1076. https://doi.org/10.3732/ajb.92.7.1068

Moore, R. C., Harkess, A. E., & Weingartner, L. A. (2016). How to be a seXY plant model: A holistic view of sex-chromosome research. *American Journal of Botany*, *103*(8), 1379–1382. https://doi.org/10.3732/ajb.1600054

Neves, C. J., Matzrafi, M., Thiele, M., Lorant, A., Mesgaran, M. B., & Stetter, M. G. (2020). Male Linked Genomic Region Determines Sex in Dioecious Amaranthus palmeri. *Journal of Heredity*, *111*(7), 606–612. https://doi.org/10.1093/jhered/esaa047

Renner, S. S. (2014). The relative and absolute frequencies of angiosperm sexual systems: Dioecy, monoecy, gynodioecy, and an updated online database. *American Journal of Botany*, *101*(10), 1588–1596. https://doi.org/10.3732/AJB.1400196

Renner, S. S., Ricklefs, R. E., & Renner, S. S. (1995). Dioecy and Its Correlates in the Flowering Plants. *American Journal of Botany*, *82*(1), 596–606. http://www.jstor.org/stable/2445418

Soza, V. L., Brunet, J., Liston, A., Salles Smith, P., & di Stilio, V. S. (2012). Phylogenetic insights into the correlates of dioecy in meadow-rues (Thalictrum, Ranunculaceae). *Molecular Phylogenetics and Evolution*, *63*(1), 180–192. https://doi.org/10.1016/j.ympev.2012.01.009

Soza, V. L., Haworth, K. L., & di Stilio, V. S. (2013). Timing and consequences of recurrent polyploidy in meadow-rues (Thalictrum, ranunculaceae). *Molecular Biology and Evolution*, *30*(8), 1940–1954. https://doi.org/10.1093/molbev/mst101

Spigler, R. B., & Ashman, T. L. (2012). Gynodioecy to dioecy: Are we there yet? *Annals of Botany*, *109*(3), 531–543. https://doi.org/10.1093/aob/mcr170

Stilio, D. di, Kumar, V. S., Oddone, R. A., Tolkin, A. M., & Salles, T. R. (2010). Virus-Induced Gene Silencing as a Tool for Comparative Functional Studies in Thalictrum. *PLoS ONE*, *5*(8), 12064. https://doi.org/10.1371/journal.pone.0012064

Warmke, H. E. (1946). Sex Determination and Sex Balance in Melandrium. *Source: American Journal of Botany*, *33*(8), 648–660.

WesterGaard, M. (1938). Induced Tetraploidy in Melandrium album. *Nature*, *142*, 917.

Westergaard, M. (1958). The Mechanism of Sex Determination in Dioecious Flowering Plants. *Advances in Genetics*, *9*(C), 217–281. https://doi.org/10.1016/S0065-2660(08)60163-7

Chapter III: Draft genome assembly and investigation of sex-specific genomic regions in the two independent origins of dioecy in the meadow rues, *Thalictrum* (Ranunculaceae)

Abstract

The advancement of sequencing technology from obtaining only a single gene at a time to achieving high quality long reads, sometimes the length of complete chromosomes, has allowed for successful assembly of high quality *de novo* draft genomes even for non-model organisms. Therefore, using PacBio High Fidelity long read sequencing technology we have assembled a draft genome for a single *T. dioicum* male individual. We used our *de novo* draft genome to gain insight into sex determination in the non-model genus *Thalictrum* (Ranunculaceae)*. Thalictrum* has several interesting characteristics that make it useful for studying the evolution of sex determination in angiosperms and more specifically the evolution of unisexuality (refer to Chapter I and II for more detail). By using inferred sex specific genomic regions (Chapter II) and our high quality *de novo* male *T. dioicum* draft genome, we provide preliminary insight for putative Y-chromosomal contigs in determining unisexuality in some *Thalictrum* species. We found no B- or C-class MADS-box genes that confidently mapped to the candidate *T. dioicum* Y-linked contigs, but we did find one gene (AP3-1) whose best match was located on a contig that is enriched for *T. dasycarpum* male-specific k-mer containing reads.

Keywords

Genome assembly; PacBio HiFi reads; homeotic unisexuality; evolution; sexual system; eudicot; angiosperm; Thalictrum dioicum

Introduction

With the increasing advances in sequencing and assembly it is now possible to create assemblies of organisms for non-model organisms. One such group of plants that is interesting to study is the genus

Thalictrum (refer to Chapter II for more detail). *Thalictrum* is a genus of flowering plants that exhibit full and functional dioecy, with parsimony suggesting two independent origins, one in a clade that includes *T. dioicum* and the other in a clade that includes *T. dasycarpum*.

Here we assemble a draft genome for a single male *T. dioicum based* on Pacific Biosciences (PacBio) high fidelity (HiFi) reads and use it to detect genomic regions that are significantly enriched for sex specific sequences identified using the Illumina short read sequencing of bulked males and females (from Chapter II). We show that using only HiFi read sequencing it is possible to create a good draft assembly used to identify several large contigs that are highly enriched for male-specific k-mer-containing reads, suggesting that they are likely to be parts of the Y-linked genomic regions. Mapping sex-specific k-mers from *T. dasycarpum* does not support the two species having homologous Y-linked regions.

Materials and Methods

DNA isolation and sequencing

Thalictrum dioicum DNA was extracted from leaf tissue from young leaves of a single male individual from a wild population in Madison, Wisconsin. The University of Wisconsin-Madison Biotechnology Resource Center extracted high molecular weight DNA and performed PacBio high fidelity (HiFi) long read library preparation and PacBio SMRT (single molecule real time) sequencing. The Sequel® II Binding Kit 2.0 was used for barcode binding and sequencing was performed using the Sequel® II Sequencing Plate 2.0. We obtained 2 SMRT Cell 8M HiFi sequencing for *T. dioicum.* Genome assembly coverage was estimated based on a genome size of 965Mb (Soza et al., 2013).

Sequence assembly and quality evaluation

HiFi reads were assembled using a PacBio HiFi specific assembler, HiFiasm v.0.16.1 (Cheng et al., 2021) includes a data processing step to purge duplicate assembly contigs to improve assembly accuracy. Purging duplicate assembly contigs reduces the rate of redundant and possibly false duplicated assembly contigs. Duplicate assembly contigs were purged within HiFiasm using a default similarity threshold (0.75) for duplicate haplotigs.

BUSCO (Benchmarking Universal Single-Copy Orthologs) v5.3.2 (Manni, et al., 2021) was used to assess genome content and completeness by recording the presence of conserved genes in the assembly based on the embryophyta-specific dataset embryophyta_odb10(Creation date: 2022-09-10, number of species: , number of BUSCOs: 1614) and more closely related eudicot-specific dataset eudicot_odb10 (Creation date: 2022-09-10, number of species: 31, number of BUSCOs: 2326). BUSCO was run in genome mode using hmmsearch v 3.1 and metaeuk v 5.34c21f2 gene predictor (Manni, 2021).

The contigs identified in the genome assembly were tentatively ordered by mapping them to the *Aquilegia oxysepala* var. *kansenansis* chromosome level assembly (Xie et al., 2020) using BBMap aligner for DNA/RNAseq version 38.84 within Geneious Prime 2022.1.1. The ordered *Thalictrum dioicum* male genome assembly was used as a reference for mapping the sex-specific k-mer containing reads.

Sex-specific read extraction and annotation of sex specific genomic regions

Thalictrum dioicum crosses one and two sex-specific k-mers (from Chapter II) were used to extract short read sequences that contained at least one sex-specific k-mer for each sex using the KMC3 toolbox filter command (Kokot et al., 2017). To be included, sex-specific k-mers had to appear at least 40 times in the genome but no more than 200 times. Based on our k-mer sex-bias graphs (from Chapter II Figure 4), we selected the lower threshold to minimize the number of false positive sex specific reads without excluding too many true sex specific reads. The upper threshold was used to remove k-mers that are likely to map to paralogs or repetitive elements.

The male-specific reads for male cross 1, denoted msk-1, were analyzed separated from msk-2, the male specific k-mer containing reads from cross 2. Likewise, female specific k-mer containing reads from the two crosses were analyzed in parallel as fsk-1 and fsk-2.

Sex-specific reads were imported into Geneious Prime 2022.1.1. Duplicate reads were removed from each dataset using Dedupe duplicate read remover with a k-mer setting of 31. The ordered *Thalictrum dioicum* male genome assembly was used as a reference for mapping the sex-specific k-mer containing reads using BBMap aligner for DNA/RNAseq version 38.84 by Brian Bushnell within Geneious Prime 2022.1.1. We calculated the sex specific read enrichment for each contig by dividing the number of observed sex-specific reads mapping to the assembly contig by the expected number of reads. The expected number of reads per contig was calculated as the length (L) of the contig divided by the sum of all contigs times the total number of sex-specific k-mers. We also calculated the ratio of the number of fsk-reads and msk-reads for each assembly contigs. Contigs were considered to be SNRs candidates if the number of msk-reads mapping to them was at least 50 times greater than expected and the contig was at least ten times more male enriched than female enriched.

We used glimmer gene prediction within the QUAST-LG package to locate open reading frames and label hypothetical genes for each contig in the *T. dioicum* assembly. We downloaded all published records of *Thalictrum* B- and C-class MADS-box genes from GenBank to determine their location within the *T. dioicum* assembly. We used the BBMap alinger for DNA/RNAseq version 38.84 by Brian Bushnell within Geneious Prime 2022.1.1 to map all *Thalictrum* B- and C-class genes along with coding to the *T. dioicum* assembly to determine whether or not any were located on male enriched *T. dioicum* contigs. Genes were mapped using the highest sensitivity settings and set to map onto all best locations to find mappings with the highest percent similarity.

Results

Thalictrum dioicum (male) draft genome assembly and quality assessment

A total of 9,092,644 ZMWs (zero-mode waveguides) were input for the two PacBio SMRT cells with a total of 2,443,674 ZMWs passing the filtering with an average size of 15.2 kb. The two cells of HiFi reads had a joint average read quality (RQ) of 30.1 and a total length of 18.5Gb. Detailed results for ZMWs not passing the read filtering stage are in Appendix Table *12*. HiFi reads will be deposited at GenBank Sequencing Read Archive (SRA).

The Hifiasm primary genome assembly was 1.17Gb contained within 612 contigs representing 121.2-115.6% of the flow –cytometry-estimated genome size for *T. dioicum* of (965-1012Mbp; (Di Stilio et al., 2005; Soza et al., 2013). The assembly had an N50 value of 30.8Mb with half the assembly (L50) contained within 15 assembly contigs (**Table *5***).

Table 5 Basic genome assembly statistics for *T. dioicum* male individual genome assembly using PacBio HiFi Reads.

Assembly statistics	*T. dioicum v0.16.1(Quast-LG)*
Number of contigs	612
Total assembly length (bp)	1,173,583,115
Largest contig (bp)	76,726,413
GC (%)	36.72%
N50 value (bp)	30,814,064
NG50 (bp)	33,073,770
N75 (bp)	12,763,841
NG75 (bp)	29,033,338
L50	15
LG50	12
L75	28
LG75	19

The N50 genome assembly quality metric only provides information about the assembly contig size but does not give an estimate of genome completeness in relation to gene content. Therefore, to obtain a metric to assess genome assembly completeness we used BUSCO to provide a quantitative assessment of the percentage of conserved single copy orthologs in the assembly. We performed a BUSCO analysis using both the Embryophyta and Eudicot BUSCO groups. There were 1600 (99.1%) complete BUSCOs of the 1614 Embryophyta BUSCO groups identified in the assembly with 7 (0.4%) fragmented BUSCOs and 7 (0.5%) missing BUSCOs (**Table *6*** and **Figure *7***). Of the complete Embryophyta BUSCO groups 184 (11.4%) were single copy and 1416 (87.7%) were duplicated copies.

Table 6 Male *Thalictrum dioicum* genome assembly BUSCO scores for completeness using the Embryophyta conserved gene database.

BUSCO Embryophyta Results	
Complete BUSCOs (C)	1600 (99.1%)
Complete and single-copy BUSCOs (S)	184 (11.4%)
Complete and duplicated BUSCOs (D)	1416 (87.7%)
Fragmented BUSCOs (F)	7 (0.4%)
Missing BUSCOs (M)	7 (0.5%)
Total BUSCO groups searched	1614

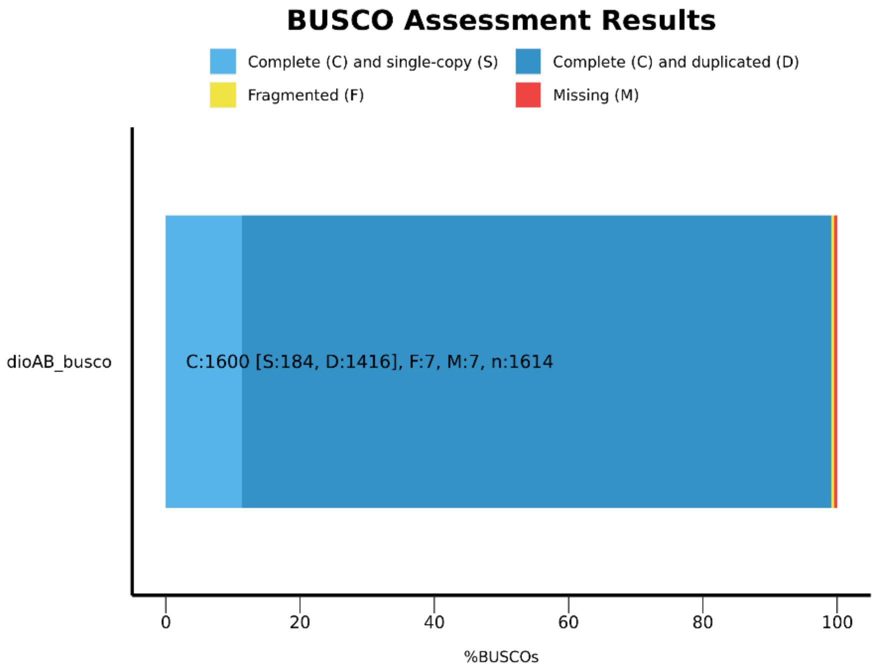

Figure 7 BUSCO conserved gene assessment of *Thalictrum dioicum* (male) genome assembly using the Embryophyta conserved gene database.

When using the Eudicot conserved genes there were 2256 (97.0%) complete BUSCOs of the 2326 Eudicot BUSCO groups identified in the assembly with 20 (0.9%) fragmented BUSCOs and 50 (2.1%) missing BUSCOs (**Table *7*** and **Figure *8***). Of the complete Eudicot BUSCO groups 193 (8.3%) were single copy and 2063 (88.7%) were duplicated copies.

Table 7 Male *T. dioicum* genome assembly BUSCO scores for completeness using the eudicot conserved gene database eudicot_odb10 (Creation date: 2022-09-10, number of species: 31, number of BUSCOs: 2326).

BUSCO Eudicots Results	
Complete BUSCOs (C)	2256 (97.0%)
Complete and single-copy BUSCOs (S)	193 (8.3%)
Complete and duplicated BUSCOs (D)	2063 (88.7%)
Fragmented BUSCOs (F)	20 (0.9%)
Missing BUSCOs (M)	50 (2.1%)
Total BUSCO groups searched	2326

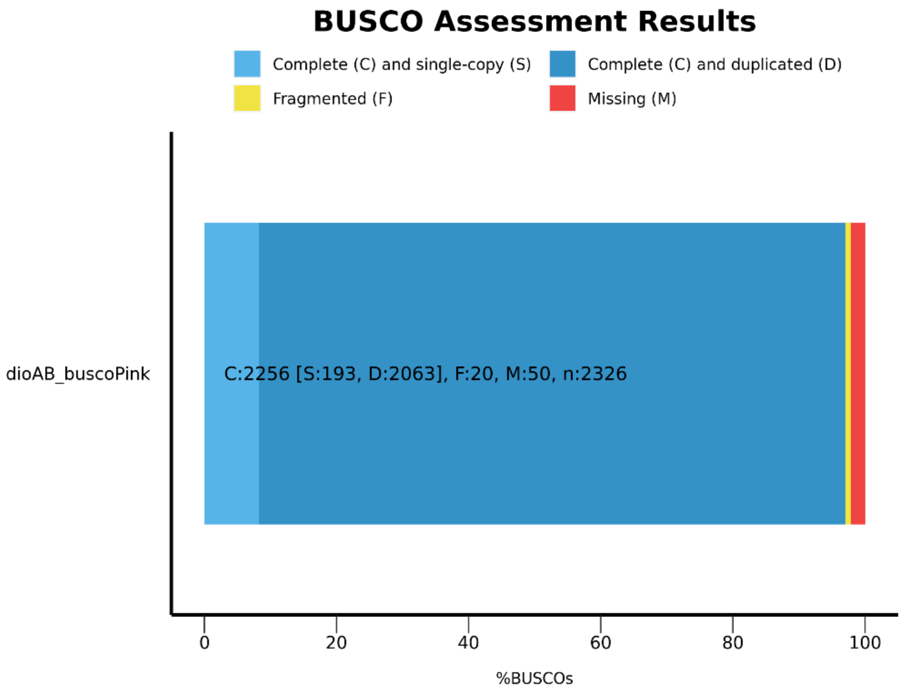

Figure 8 BUSCO conserved gene assessment of *Thalictrum dioicum* (male) genome assembly using the eudicot conserved gene database eudicot_odb10 (Creation date: 2022-09-10, number of species: 31, number of BUSCOs: 2326*)*

Male and Female short read enrichment in the male *Thalictrum dioicum* genome assembly

The enrichment of male specific k-mer containing reads (msk-reads) and female specific k-mer containing reads (fsk-reads) mapped to the male *Thalictrum dioicum* genome assembly in each cross found many contigs with a higher enrichment of msk-reads than fsk-reads. The pattern of enrichment was similar in the two crosses (**Table *8***).

For each contig we calculated the degree of enrichment of mapped reads relative to a null expectation based on the contig length. While several contigs have fsk-read enrichment greater than 10-fold, in all cases these contigs are even more enriched for msk-reads. This is expected, since the sequenced individual is male, so it should carry no female-specific genome sequences. In contrast, many contigs show very high enrichment for msk-reads, and many of these do not also show enrichment for fsk-reads. **Table *8*** lists all contigs that have either a 10-fold or greater msk-read enrichment, or a greater than 5-fold higher msk-enrichment than fsk-enrichment. The large number of contigs showing male enrichment is consistent with *T. dioicum* having XY sex-determination. Thus, these highly msk-enriched contigs could be located on the SNRs. Especially strong candidates (marked with an asterisk) are those that are at least 10-fold msk-enriched and show a greater than 5-fold higher msk- than fsk-enrichment in both crosses individually, as well as the combined pool. Additionally, since the contig numbers are based on mapping to the *Aquilegia* chromosome assemblies, adjacent contigs pairs 108-110, 186-187, and 329-330 are potentially interesting since they likely map to a contiguous chromosomal region. Similarly, the fact that many of the strongest candidates appear in relative proximity between contigs 71 to 79 and 234 and 291 might point to the SNRs sharing long areas of synteny with *Aquilegia*.

Table 8 ***T. dioicum*** **cross 1 and 2 sex-specific k-mer containing reads with a high male or female enrichment when mapped to the male** ***T. dioicum*** **genome assembly contigs.** Msk- and fsk-reads mapping to multiple locations were mapped randomly to the male *T. dioicum* genome assembly. The table is ordered by the assembly contig number. Contigs shaded gray are ten times enriched and five times higher than females. Contigs that appeared consistently msk-read enriched regardless of cross are marked in bold text. Sex-specific enriched contigs that appeared consecutively are highlighted in yellow.

Assembly		**Cross1**				**Cross2**				**Combined**		
contig	**Contig Length (bp)**	**msk-1 reads**	**fsk-1 reads**	**msk-1 enrichment**	**c1 m/f ratio**	**msk-2 reads**	**fsk-2 reads**	**msk-2 enrichment**	**c2 m/f ratio**	**combined enrichment**	**combine m/f ratio**	**predicted genes**
7	15800358	576	11653	0.52	0.48	508	3	1.15	5.05	0.71	0.66	3311
13	31995408	236	19038	0.10	0.12	4646	2	5.20	69.31	1.63	1.86	6471
20	5277996	100	7120	0.27	0.14	2889	0	19.60	n/a	6.08	3.07	569
35	36127415	236	20353	0.09	0.11	1962	8	1.95	7.32	0.64	0.77	7373
43	8150080	534	4108	0.93	1.26	377	2	1.66	5.62	1.06	1.39	2220
64	452587	4859	11168	152.61	4.20	43	23	3.40	0.06	110.23	3.08	3
71	1817328	1076	6376	8.42	1.63	604	2	11.90	9.01	9.57	1.87	299
73	**192046**	**6765**	**2356**	**500.73**	**27.75**	**500**	**0**	**93.24**	**n/a**	**379.28**	**22.45**	**0**
75	3949091	12151	21088	43.74	5.57	2793	2	25.33	41.67	38.30	4.99	595
79	89027	646	36	103.15	173.42	0	0	0.00	0.00	75.87	137.77	8
91	194166	559	2838	40.92	1.90	220	0	40.58	n/a	42.25	2.03	2
95	138491	110	51	11.29	20.84	6	0	1.55	0.00	7.30	13.05	171
98	1148870	6593	21013	81.57	3.03	3090	6	96.33	15.37	84.48	3.25	20
102	341623	1057	2267	43.98	4.51	296	0	31.03	n/a	41.06	4.27	7
108	2476811	2208	3643	12.67	5.86	109	0	1.58	0.00	9.60	4.44	19
109	**258234**	**2007**	**2215**	**110.48**	**8.76**	**286**	**0**	**39.67**	**n/a**	**89.31**	**7.22**	**1**
110	1432496	5561	26143	55.18	2.06	8951	0	223.79	n/a	101.02	3.95	40
124	79091	1182	1275	212.44	8.96	3	0	1.36	0.00	155.46	6.77	0
138	**146626**	**2465**	**949**	**238.97**	**25.10**	**301**	**0**	**73.52**	**n/a**	**188.25**	**19.79**	**7**
144	821352	9846	15054	170.40	6.32	19	43	0.83	0.01	121.96	4.67	7
186	87458	53	24	8.61	21.34	3	0	1.23	0.00	5.01	11.09	113
187	320975	142	130	6.29	10.56	10	0	1.12	0.00	4.55	7.27	3
205	472102	3072	5311	92.50	5.59	274	0	20.79	n/a	72.73	4.49	2
219	167626	6	320	0.51	0.18	155	0	33.12	n/a	11.13	4.74	10

234	113795	679	422	84.82	15.55	0	2	0.00	0.00	61.53	11.71	7
248	49057	534	1	154.73	5160.75	0	0	0.00	0.00	114.55	1776.52	2
255	**22933**	**282**	**92**	**174.79**	**29.62**	**77**	**0**	**120.25**	**n/a**	**156.82**	**27.73**	**0**
264	**36961**	**377**	**257**	**144.99**	**14.18**	**11**	**0**	**10.66**	**n/a**	**104.30**	**10.24**	**1**
271	24207	238	0	139.76	n/a	0	0	0.00	0.00	98.89	n/a	0
291	**33885**	**870**	**515**	**364.97**	**16.33**	**259**	**0**	**273.75**	**n/a**	**347.60**	**16.26**	**1**
293	58627	271	1083	65.71	2.42	14	0	8.55	0.00	46.39	1.78	0
299	43715	25	23	8.13	10.50	0	0	0.00	0.00	6.68	9.24	49
329	90077	50	43	7.89	11.24	6	0	2.39	0.00	5.24	8.29	101
330	103685	180	59	24.68	29.48	2	0	0.69	0.00	15.61	21.77	121
360	**57232**	**262**	**344**	**65.07**	**7.36**	**33**	**0**	**20.65**	**n/a**	**55.77**	**6.68**	**0**
363	32781	261	582	113.18	4.33	0	0	0.00	0.00	87.08	3.41	0
431	51457	20	18	5.52	10.74	2	0	1.39	0.00	3.06	4.52	59
450	44614	17	10	5.42	16.43	2	0	1.61	0.00	4.03	12.63	0
522	38463	24	14	8.87	16.57	0	0	0.00	0.00	5.26	12.79	43
559	26553	86	14	46.04	59.37	0	0	0.00	0.00	34.71	44.82	0
568	38171	0	8	0.00	0.00	19	0	17.83	n/a	7.66	46.19	5
573	21112	1	6	0.67	1.61	2	0	3.39	0.00	2.13	9.47	26
610	24295	7	4	4.10	16.91	2	0	2.95	0.00	1.39	3.05	31

Mapping *Thalictrum dasycarpum* sex-specific k-mers onto the *T. dioicum* genome assembly

A total of *T. dasycarpum* 977 fsk-reads and 32,743 msk-reads could be mapped on the *T. dioicum* genome assembly. The low number of fsk-reads mapping makes the msk to fsk enrichment uninformative. However, two contigs, (578 and 612) showed a greater than 20-fold enrichment of msk-reads relative to expectations, and a further one (605) showed a 12-fold enrichment. None of these contigs had any *T. dioicum* msk-reads mapped to them (Appendix **Table *15***). Although there were not shared contigs between *T. dioicum* and *T. dasycarpum* conting 578 did contain B-Class floral homeotic gene AP3-1 and AP3-2b (Table 9**Error! Reference source not found.**).

Mapping *Thalictrum* B- and C-class MADS-box genes onto the *T. dioicum* genome assembly

We mapped all published *Thalictrum* B- and C-Class genes and coding sequences obtained form Genbank to the genome assembly to see if any mapped to the candidate y-linked contigs. As summarized in Table 9, The best hits in the assembly that contained *T. dioicum* coding sequences had surprisingly low similarity scores, of 52.8-80.0%. These differences most likely represent subspecific variation in this widespread species and combined mapping of all *Thalictrum* species. Many additional B-gene candidate homologs that were from *T. thalictroides* were found, some of which had much higher sequence identities than those that also contained T. dioicum sequences (Table 9). None of the best hits map to any of the *T. dioicum* msk-enriched contigs. However, the closest match to AP3-1 maps to a *T. dioicum* contig that is msk-enriched in *T. dasycarpum*. This finding leaves open the possibility that *AP3-1* may be playing a sex-determining role in *T. dasycarpum*, but this inference should be tested using a *T. dasycarpum* genome assembly.

***Table 9 Mapping of Thalictrum B and C class genes to the T. dioicum male genome assembly contigs**. Highlighted in grey are contigs that were also found to be male enriched in T. dioicum and those highlighted in yellow were male enriched in T. dasycarpum. Species names refer to T. thalictroides (Tht), T. dioicum (Thdi), and T. delavayi (Thde)*

Contig Name	Gene	Coding Sequence	% Pairwise Identity	% Identical Sites	Species
1	AG-2	AGAMOUS-like 2 protien	51.20%	13.20%	Tht
3	AG-2	AGAMOUS-like 2 protien	74.00%	33.90%	Tht
4	AP3-2a	APETALA3-2A	90.30%	60.50%	Tht
4	AG-1	similar to AGAMOUS1(MISC)	77.60%	77.50%	Tht
5	AG-1	AGAMOUS1 and AGAMOUS-like protien 1	84.80%	59.00%	Tht
7	AG-1	similar to AGAMOUS1(MISC)	81.30%	81.20%	Tht
8	AP3-2b	APETALA3-2B	78.80%	78.80%	Tht
9	AP3-2b	APETALA3-2B	78.80%	78.80%	Tht
14	AP3-1	APETALA3-1	59.90%	48.00%	Tht
14	AP3-2b	APETALA3-2B	68.00%	67.90%	Tht
16	AP3-2b	APETALA3-2B	83.30%	83.30%	Tht
16	AG1	AGAMOUS1 and AGAMOUS-like protien 1	80.00%	61.20%	Tht-Thdi
21	AP3-1	APETALA3-1	63.00%	39.50%	Tht
21	AP3-2b	APETALA3-2B	63.70%	43.60%	Tht-Thdi
25	PI-1	PISTILLATA-like	60.80%	38.60%	Tht-Thdi
25	PI-2	PISTILLATA-like 2	81.80%	81.70%	Tht
28	AG1	AGAMOUS1 and AGAMOUS-like protien 1	62.00%	64.20%	Tht-Thde
36	AP3-2b	APETALA3-2B	75.00%	75.00%	Tht
38	AP3-2a	APETALA3-2A	95.20%	75.10%	Tht
41	AP3-2a	APETALA3-2A	72.00%	63.60%	Tht
43	AG-1	similar to AGAMOUS1(MISC)	69.00%	69.00%	Tht
44	AG-1	AGAMOUS1 and AGAMOUS-like protien 1	67.60%	61.50%	Tht
46	AG-2	AGAMOUS-like 2 protien	73.40%	33.20%	Tht
61	PI	PISTILLATA-like	75.30%	46.10%	Tht
61	PI-2	PISTILLATA-like 2	52.80%	49.40%	Tht-Thdi
62	PI	PISTILLATA-like	96.00%	88.90%	Tht
63	AP3-2b	APETALA3-2B	78.80%	78.80%	Tht
67	PI	PISTILLATA-like	87.90%	87.80%	Tht
70	AG-1	AGAMOUS1 and AGAMOUS-like protien 1	67.80%	62.90%	Tht
77	PI	PISTILLATA-like	92.80%	92.80%	Tht
83	PI	PISTILLATA-like	92.80%	92.80%	Tht
86	AP3-2a	APETALA3-2A	85.10%	67.80%	Tht
94	PI	PISTILLATA-like	92.80%	92.80%	Tht
100	PI	PISTILLATA-like	92.80%	92.80%	Tht
103	PI	PISTILLATA-like	97.80%	97.80%	Tht
133	AG-2	AGAMOUS-like 2 protien	73.90%	39.80%	Tht
578	AP3-1	APETALA3-1 & APETALA3-1 like	59.90%	48.00%	Tht-Thdi
578	AP3-2b	APETALA3-2B	34.80%	34.80%	Tht-Thdi

Discussion

We provided the first draft genome of a dioecious meadowrue and showed that this genome is of high quality and can be used to aid further investigations of the genomic basis of sex determination. The only other *Thalictrum* species with a draft genome is the hermaphroditic species, *T. thalictroides* (Arias et al. 2021). Moreover, that genome assembly was based on Illumina short-reads resulting in many more and smaller contigs. Consequently, our work will provide a useful reference point for future comparative genomic and phylogenetic research in *Thalictrum* and close relatives.

A key motivation for this work was to aid in the interpretation of Illumina short-read data for sex-specific pools of *T. dioicum* (Chapter II). Analysis in the absence of a genome assembly suggested that there are more male-specific k-mers than female-specific k-mers, implying that males are the heterogametic sex and there is XY sex-determination. However, without further knowledge of the genomic context of these k-mers, interpretation is difficult. Here, we identified reads containing sex-specific k-mers and mapped them onto the draft genome of a single male to identify contigs that are enriched for male-specific k-mers reads (msk-reads), using female-specific k-mers reads (fsk-reads) as an internal control. As expected, we found no evidence of contigs enriched for fsk-reads, but we found many contigs that showed a greater than 10-fold enrichment of msk-reads or a 10-fold higher enrichment of msk- than fsk-reads, or both. Future work, ideally including the generation of a complementary, female genome assembly, should be able to test the hypothesis that these are situated on the Y-linked genomic region. Mapping the *T. dasycarpum* msk-reads on the *T. dioicum* draft genome detected three contigs showing msk enrichment, all in close proximity in the genome assembly. None of these regions are msk-enriched in *T. dioicum,* supporting our prior inference (Chapter II) that different regions of the genome were coopted for sex-determination in the two clades.

Acknowledgements

The authors thank the University of Wisconsin Biotechnology Center DNA Sequencing Facility for providing high molecular weight DNA extractions and Illumina and PacBio DNA sequencing facilities and services. We thank the Wisconsin Alumni Research Foundation- Fall Competition Grant and the University of Wisconsin Botany Department Endowment funds for helping fund this project.

References

Cheng, H., Concepcion, G. T., Feng, X., Zhang, H., & Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nature Methods*, *18*(2), 170–175. https://doi.org/10.1038/S41592-020-01056-5

di Stilio, V. S., Kramer, E. M., & Baum, D. A. (2005). Floral MADS box genes and homeotic gender dimorphism in Thalictrum dioicum (Ranunculaceae)-a new model for the study of dioecy. *The Plant Journal*, *41*(5), 755–766. https://doi.org/10.1111/j.1365-313X.2005.02336.x

Kokot, M., Dlugosz, M., & Deorowicz, S. (2017). KMC 3: counting and manipulating k-mer statistics. *Bioinformatics (Oxford, England)*, *33*(17), 2759–2761. https://doi.org/10.1093/bioinformatics/btx304

Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., & Zdobnov, E. M. (2021). BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Molecular Biology and Evolution*, *38*(10), 4647–4654. https://doi.org/10.1093/molbev/msab199

Manni, M., Berkeley, M. R., Seppey, M., & Zdobnov, E. M. (2021). BUSCO: Assessing Genomic Data Quality and Beyond. *Current Protocols*, *1*, e323. https://doi.org/10.1002/CPZ1.323

Soza, V. L., Haworth, K. L., & di Stilio, V. S. (2013). Timing and consequences of recurrent polyploidy in meadow-rues (Thalictrum, ranunculaceae). *Molecular Biology and Evolution*, *30*(8), 1940–1954. https://doi.org/10.1093/molbev/mst101

Xie, J., Zhao, H., Li, K., Zhang, R., Jiang, Y., Wang, M., Guo, X., Yu, B., Kong, H., Jiao, Y., & Xu, G. (2020). A chromosome-scale reference genome of Aquilegia oxysepala var. kansuensis. *Horticulture Research*, *7*, 113. https://doi.org/10.1038/s41438-020-0328-y

Appendix

Table 10 Total 51-mer Counts with an occurrence of 15-2000 for each *T. dioicum* cross and *T. dasycarpum* male and female samples.

51mer bins = counts ≥n	*T. dioicum* Male Cross 1	*T. dioicum* Male Cross 2	*T. dioicum* Female Cross 1	*T. dioicum* Female Cross 2	*T. dasycarpum* Male	*T. dasycarpum* Female
15	117,294,099	135,396,793	224,372,017	257,621,434	1,205,334,398	1,397,502,159
20	52,601,242	60,399,998	98,938,805	120,136,902	872,289,938	1,019,480,028
25	33,022,509	37,524,452	53,695,603	62,791,202	671,226,032	788,271,040
30	23,304,966	26,541,523	37,079,746	42,359,493	535,648,869	634,676,932
35	17,903,179	20,253,925	27,443,309	31,435,190	434,790,278	523,341,113
40	14,609,369	16,400,529	21,406,169	24,362,087	354,967,502	436,612,155
45	12,364,041	13,787,286	17,608,252	19,840,683	290,044,849	365,775,610
50	10,722,454	11,905,286	15,009,033	16,814,112	237,502,145	306,596,334
55	9,483,831	10,503,217	13,075,830	14,613,516	196,019,917	257,047,591
60	8,514,538	9,413,875	11,585,276	12,924,915	164,236,926	216,184,639
65	7,736,197	8,543,556	10,406,289	11,594,968	140,411,826	183,307,554
70	7,094,542	7,832,831	9,453,002	10,519,050	122,582,443	157,537,448
75	6,556,365	7,239,475	8,663,606	9,634,059	108,854,445	137,662,411
80	6,101,122	6,736,959	8,002,321	8,895,230	97,841,823	122,300,526
85	5,707,324	6,305,355	7,441,453	8,266,710	88,618,380	110,146,332
90	5,365,297	5,930,220	6,961,384	7,728,733	80,640,877	100,194,613
95	5,063,747	5,601,518	6,541,909	7,266,594	73,629,103	91,728,550
100	4,793,702	5,309,442	6,171,565	6,860,793	67,416,177	84,324,479
105	4,551,411	5,048,518	5,843,646	6,501,190	61,918,639	77,727,612
110	4,333,262	4,816,606	5,550,959	6,179,921	57,060,758	71,800,888
115	4,137,694	4,606,161	5,287,106	5,890,012	52,788,011	66,459,057
120	3,957,051	4,415,664	5,047,778	5,627,811	49,031,689	61,652,306
125	3,791,824	4,240,437	4,830,118	5,388,125	45,709,930	57,359,966
130	3,641,268	4,079,463	4,633,302	5,170,090	42,754,577	53,524,557
135	3,501,384	3,930,246	4,453,236	4,969,350	40,109,794	50,101,780
140	3,371,498	3,792,090	4,285,154	4,784,583	37,716,015	47,042,399
145	3,250,405	3,664,932	4,128,429	4,614,055	35,535,695	44,291,510
150	3,137,179	3,545,846	3,982,621	4,455,117	33,537,413	41,802,263
155	3,029,723	3,434,662	3,845,503	4,305,965	31,691,256	39,534,266
160	2,928,319	3,330,099	3,716,908	4,165,775	29,986,688	37,453,332
165	2,832,079	3,231,504	3,597,113	4,034,909	28,416,127	35,533,985
170	2,740,822	3,138,209	3,483,924	3,912,277	26,965,856	33,759,441
175	2,654,241	3,049,528	3,377,694	3,797,788	25,618,623	32,105,572
180	2,570,692	2,964,823	3,277,241	3,689,916	24,371,780	30,565,658
185	2,490,718	2,882,419	3,181,611	3,589,093	23,211,502	29,126,760
190	2,413,904	2,803,624	3,090,197	3,492,755	22,134,335	27,785,411
195	2,339,950	2,726,770	3,003,061	3,400,861	21,131,967	26,530,980
200	2,268,077	2,652,818	2,919,111	3,311,711	20,192,637	25,358,113

Table 11 Total Sex-specific 51-mer counts with an occurrence of 15-2000 for each *T. dioicum* cross and *T. dasycarpum* male and female samples. T. dioicum crosses were analyzed individually for sex-specific k-mers.

Sex Specific 51-mer Counts with an occurrence of 15-2000

51mer Occurrence bins (≥n)	*T. dioicum Male Cross 1*	*T. dioicum Male Cross 2*	*T. dioicum Female Cross 1*	*T. dioicum Female Cross 2*	*T. dasycarpum Male*	*T. dasycarpum Female*
15	28,167,095	100,107	135,245,013	232,945	489,630	1,930,148
20	3,255,301	44,823	35,376,214	25,021	343,009	264,298
25	537,598	29,036	6,656,900	3,794	115,837	45,540
30	262,432	20,088	1,471,011	837	61,466	14,151
35	199,688	14,361	423,092	275	42,112	8,065
40	170,172	11,014	145,463	114	31,817	6,107
45	151,481	8,902	65,798	76	24,685	4,970
50	138,264	7,631	40,710	64	19,605	4,132
55	128,194	6,672	29,489	58	16,152	3,482
60	120,267	5,758	23,260	52	13,443	3,009
65	113,754	5,033	19,379	45	11,388	2,635
70	108,373	4,533	16,480	43	9,819	2,282
75	103,833	4,090	14,317	28	8,449	1,963
80	99,918	3,709	12,688	20	7,464	1,721
85	96,345	3,402	11,174	20	6,634	1,531
90	93,037	3,024	9,903	18	5,917	1,337
95	90,024	2,786	8,873	14	5,317	1,183
100	87,126	2,561	8,000	8	4,853	1,045
105	84,356	2,339	7,201	6	4,451	951
110	81,863	2,188	6,562	2	4,124	863
115	79,611	2,043	6,006	2	3,843	786
120	77,750	1,941	5,579	2	3,554	719
125	76,278	1,806	5,146	1	3,276	648
130	75,157	1,697	4,729	1	3,111	618
135	74,211	1,626	4,356	1	2,977	576
140	73,454	1,538	4,045	-	2,865	553
145	72,832	1,466	3,739	-	2,718	536
150	72,271	1,399	3,436	-	2,607	524
155	71,770	1,319	3,147	-	2,495	507
160	71,268	1,226	2,895	-	2,371	486
165	70,853	1,135	2,679	-	2,295	463
170	70,487	1,071	2,394	-	2,212	444
175	70,172	1,030	2,201	-	2,135	431
180	69,864	1,003	2,044	-	2,075	427
185	69,592	982	1,951	-	2,021	421
190	69,327	957	1,902	-	1,932	412
195	69,132	935	1,825	-	1,846	405
200	68,981	931	1,753	-	1,773	402

Table 12 PacBio high fidelity (HiFi) closed consensus sequencing (CCS) results for a male *T. dioicum* individual. Results for zero-mode waveguides (ZMWs) input and output for sequencing of two PacBio 8M SMRT cells, 1_A01 and 2_B01.

***T. dioicum 1_A01* HiFi CCS report**	
ZMWs input	4654443
ZMWs pass filters	1228702 (26.40%)
ZMWs fail filters	3425741 (73.60%)
ZMWs shortcut filters	0 (0.00%)
ZMWs with tandem repeats	5757 (0.12%)
Exclusive counts for ZMWs failing filters:	
Below SNR threshold	142098 (4.15%)
Median length filter	0 (0.00%)
Lacking full passes	2889812 (84.36%)
Heteroduplex insertions	61310 (1.79%)
Coverage drops	4223 (0.12%)
Insufficient draft cov	23354 (0.68%)
Draft too different	2622 (0.08%)
Draft generation error	15287 (0.45%)
Draft above --max-length	556 (0.02%)
Draft below --min-length	0 (0.00%)
Reads failed polishing	2 (0.00%)
Empty coverage windows	5131 (0.15%)
CCS did not converge	25 (0.00%)
CCS below minimum RQ	281877 (8.23%)
Unknown error	0 (0.00%)
***T. dioicum 2_B01* HiFi CCS report**	
ZMWs input	4437126
ZMWs pass filters	1214972 (27.38%)
ZMWs fail filters	3222154 (72.62%)
ZMWs shortcut filters	0 (0.00%)
ZMWs with tandem repeats	6185 (0.14%)
Exclusive counts for ZMWs failing filters:	
Below SNR threshold	115606 (3.59%)
Median length filter	0 (0.00%)
Lacking full passes	2709040 (84.08%)
Heteroduplex insertions	61272 (1.90%)
Coverage drops	4615 (0.14%)
Insufficient draft cov	26039 (0.81%)
Draft too different	2890 (0.09%)
Draft generation error	15862 (0.49%)
Draft above --max-length	519 (0.02%)
Draft below --min-length	0 (0.00%)
Reads failed polishing	2 (0.00%)
Empty coverage windows	5482 (0.17%)
CCS did not converge	38 (0.00%)
CCS below minimum RQ	281308 (8.73%)
Unknown error	0 (0.00%)

Table 13 Hifiasm genome assembler summary for de-novo genome assembly of T. dioicum single male individual.

HiFiasm summary output:	***T. dioicum* (male)**
[M::ha_analyze_count] left	none
[M::ha_analyze_count] right	none
[M::ha_pt_gen] peak_hom	38
[M::ha_pt_gen] peak_het	19
[M::ha_ct_shrink:: 59020.479*9.65] [M::ha_ct_shrink59534.159*7.75]	==> counted 31,612,201 distinct minimizer k-mers
[M::ha_pt_gen::]	counting in normal mode
[M::yak_count]	collected 998,478,423 minimizers
[M::ha_pt_gen:: 59416.776*9.64] [M::ha_pt_gen60458.471*7.73]	==> indexed 997,048,369 positions,
[M::ha_pt_gen:: 59416.776*9.64] [M::ha_pt_gen60458.471*7.73]	counted 31,612,201 distinct minimizer k-mers
[M::ha_assemble::63422.475*9.66@51.036GB] [M::ha_assemble64481.494*7.74@49.071GB]	==> found overlaps for the final round
# overlaps	119,272,458
# strong overlaps	54,361,934
# weak overlaps	64,910,524
# exact overlaps	114,681,360
# inexact overlaps	4,591,098
# overlaps without large indels	11,9024,160
# reverse overlaps	38,481,412
[M::purge_dups] homozygous read coverage threshold	37
[M::purge_dups] purge duplication coverage threshold	47
[M::purge_dups] homozygous read coverage threshold	37
[M::purge_dups] purge duplication coverage threshold	47
[M::adjust_utg_by_primary] primary contig coverage range	[31, infinity]
Inconsistency threshold for low-quality regions in BED files	70%
Version	0.16.1-r375
CMD	./hifiasm-0.16.1/hifiasm -o dioBA_8.asm --primary -t 10 --hg-size 965m --n-hap 4 -s 0.75 dioBA.fq.gz
Real time: (sec)	64333.202 sec
CPU (sec)	615600.107
Peak RSS (GB)	51.036

Table 14 All *T. dioicum* cross 1 and 2 sex-specific k-mer containing reads mapped to the male *T. dioicum* genome assembly contigs. Msk- and fsk-reads mapping to multiple locations were mapped randomly to the male *T. dioicum* genome assembly. The table is ordered by the assembly contig name.

Assembly Contig		Expected Number of Sequences		Observed Number of Sequences		Sex-Specific Read Enrichment		Enrichment Difference	Enrichment Ratio	
Assembly Contig	Contig Length (bp)	fsk-expected	msk-expected	fsk-reads	msk-reads	msk-enrichment	fsk-enrichment	fsk - msk	msk/fsk	fsk/msk
ptg000001l	61,382,776	38,809.00	5,461.45	40,216	5,201	0.95	1.04	0.08	0.92	1.09
ptg000002l	927,027	586.11	82.48	260	2	0.02	0.44	0.42	0.05	18.29
ptg000003l	45,649,582	28,861.76	4,061.61	24,843	371	0.09	0.86	0.77	0.11	9.42
ptg000004l	34,209,567	21,628.85	3,043.75	19,828	1,708	0.56	0.92	0.36	0.61	1.63
ptg000005l	5,417,901	3,425.45	482.05	2,412	29	0.06	0.70	0.64	0.09	11.70
ptg000006l	10,696,020	6,762.51	951.66	4,358	50	0.05	0.64	0.59	0.08	12.27
ptg000007l	15,800,358	9,989.71	1,405.82	10,811	999	0.71	1.08	0.37	0.66	1.52
ptg000008l	24,015,344	15,183.60	2,136.73	16,888	1,106	0.52	1.11	0.59	0.47	2.15
ptg000009l	31,118,052	19,674.26	2,768.69	14,958	92	0.03	0.76	0.73	0.04	22.88
ptg000010l	33,073,770	20,910.75	2,942.69	15,265	186	0.06	0.73	0.67	0.09	11.55
ptg000011l	30,380,337	19,207.84	2,703.05	32,005	607	0.22	1.67	1.44	0.13	7.42
ptg000012l	29,237,344	18,485.19	2,601.35	10,842	84	0.03	0.59	0.55	0.06	18.16
ptg000013l	31,995,408	20,228.96	2,846.75	17,724	4,647	1.63	0.88	(0.76)	1.86	0.54
ptg000014l	12,763,841	8,069.88	1,135.65	5,609	158	0.14	0.70	0.56	0.20	5.00
ptg000015l	8,363,160	5,287.57	744.10	2,762	7	0.01	0.52	0.51	0.02	55.53
ptg000016l	12,080,268	7,637.70	1,074.83	6,916	293	0.27	0.91	0.63	0.30	3.32
ptg000017l	76,715	48.50	6.83	1	0	0.00	0.02	0.02	0.00	0.00
ptg000018l	44,476,517	28,120.09	3,957.24	24,024	449	0.11	0.85	0.74	0.13	7.53
ptg000019l	76,726,413	48,509.95	6,826.63	43,160	441	0.06	0.89	0.83	0.07	13.77
ptg000020l	5,277,996	3,336.99	469.60	6,605	2,853	6.08	1.98	(4.10)	3.07	0.33
ptg000021l	36,227,159	22,904.47	3,223.26	12,428	105	0.03	0.54	0.51	0.06	16.66
ptg000022l	36,900,733	23,330.33	3,283.19	18,858	138	0.04	0.81	0.77	0.05	19.23
ptg000023l	4,408,031	2,786.96	392.20	1,802	4	0.01	0.65	0.64	0.02	63.40
ptg000024l	17,852,609	11,287.24	1,588.41	8,109	83	0.05	0.72	0.67	0.07	13.75

ptg000025l	9,746,350	6,162.09	867.17	6,343	77	0.09	1.03	0.94	0.09	11.59
ptg000026l	30,814,064	19,482.06	2,741.64	17,163	304	0.11	0.88	0.77	0.13	7.95
ptg000027l	6,415,627	4,056.25	570.82	1,997	68	0.12	0.49	0.37	0.24	4.13
ptg000028l	18,597,651	11,758.29	1,654.70	8,267	228	0.14	0.70	0.57	0.20	5.10
ptg000029l	14,286,320	9,032.47	1,271.11	6,821	60	0.05	0.76	0.71	0.06	16.00
ptg000030l	9,549,700	6,037.76	849.67	21,379	426	0.50	3.54	3.04	0.14	7.06
ptg000031l	10,022,985	6,336.99	891.78	3,762	73	0.08	0.59	0.51	0.14	7.25
ptg000032l	10,054,820	6,357.12	894.61	2,918	27	0.03	0.46	0.43	0.07	15.21
ptg000033l	36,043,935	22,788.63	3,206.96	20,342	203	0.06	0.89	0.83	0.07	14.10
ptg000034l	1,015,632	642.13	90.36	845	307	3.40	1.32	(2.08)	2.58	0.39
ptg000035l	36,127,415	22,841.41	3,214.39	18,981	2,059	0.64	0.83	0.19	0.77	1.30
ptg000036l	34,095,065	21,556.46	3,033.56	17,655	390	0.13	0.82	0.69	0.16	6.37
ptg000037l	10,310,964	6,519.06	917.40	3,997	120	0.13	0.61	0.48	0.21	4.69
ptg000038l	7,548,890	4,772.75	671.65	3,623	19	0.03	0.76	0.73	0.04	26.83
ptg000039l	59,673	37.73	5.31	6	0	0.00	0.16	0.16	0.00	0.00
ptg000040l	8,911,710	5,634.39	792.91	3,663	222	0.28	0.65	0.37	0.43	2.32
ptg000041l	30,393,699	19,216.29	2,704.24	14,119	169	0.06	0.73	0.67	0.09	11.76
ptg000042l	4,512,657	2,853.11	401.51	1,600	23	0.06	0.56	0.50	0.10	9.79
ptg000043l	8,150,080	5,152.85	725.14	3,936	768	1.06	0.76	(0.30)	1.39	0.72
ptg000044l	4,551,151	2,877.45	404.93	1,969	11	0.03	0.68	0.66	0.04	25.19
ptg000045l	73,557	46.51	6.54	21	0	0.00	0.45	0.45	0.00	0.00
ptg000046l	3,375,395	2,134.08	300.32	1,849	14	0.05	0.87	0.82	0.05	18.59
ptg000047l	16,521,110	10,445.40	1,469.94	6,977	89	0.06	0.67	0.61	0.09	11.03
ptg000048l	7,754,013	4,902.44	689.90	3,973	54	0.08	0.81	0.73	0.10	10.35
ptg000049l	3,908,735	2,471.28	347.77	3,134	9	0.03	1.27	1.24	0.02	49.00
ptg000050l	47,471	30.01	4.22	2	0	0.00	0.07	0.07	0.00	0.00
ptg000051l	8,626,767	5,454.24	767.56	4,250	13	0.02	0.78	0.76	0.02	46.01
ptg000052l	3,780,431	2,390.16	336.36	1,262	176	0.52	0.53	0.00	0.99	1.01
ptg000053l	1,324,218	837.23	117.82	317	1	0.01	0.38	0.37	0.02	44.61
ptg000054l	10,183,987	6,438.78	906.11	3,450	15	0.02	0.54	0.52	0.03	32.37
ptg000055l	32,538	20.57	2.90	8	0	0.00	0.39	0.39	0.00	0.00
ptg000056l	6,199,021	3,919.30	551.55	3,200	141	0.26	0.82	0.56	0.31	3.19

ptg000057l	2,258,311	1,427.81	200.93	832	0	0.00	0.58	0.58	0.00	0.00
ptg000058l	15,612,692	9,871.06	1,389.12	9,595	366	0.26	0.97	0.71	0.27	3.69
ptg000059l	4,107,956	2,597.24	365.50	2,200	752	2.06	0.85	(1.21)	2.43	0.41
ptg000060l	2,673,336	1,690.21	237.86	1,188	58	0.24	0.70	0.46	0.35	2.88
ptg000061l	11,754,656	7,431.83	1,045.85	7,482	82	0.08	1.01	0.93	0.08	12.84
ptg000062l	19,808,607	12,523.91	1,762.44	8,886	99	0.06	0.71	0.65	0.08	12.63
ptg000063l	29,033,338	18,356.21	2,583.20	17,416	355	0.14	0.95	0.81	0.14	6.90
ptg000064l	452,587	286.15	40.27	10,234	4,439	110.24	35.76	(74.47)	3.08	0.32
ptg000065l	3,225,988	2,039.62	287.03	855	25	0.09	0.42	0.33	0.21	4.81
ptg000066l	530,107	335.16	47.17	274	0	0.00	0.82	0.82	0.00	0.00
ptg000067l	93,435	59.07	8.31	41	0	0.00	0.69	0.69	0.00	0.00
ptg000068l	81,502	51.53	7.25	1	0	0.00	0.02	0.02	0.00	0.00
ptg000069l	164,882	104.25	14.67	21	0	0.00	0.20	0.20	0.00	0.00
ptg000070l	9,167,258	5,795.96	815.64	4,527	33	0.04	0.78	0.74	0.05	19.31
ptg000071l	1,817,328	1,149.00	161.69	5,887	1,548	9.57	5.12	(4.45)	1.87	0.54
ptg000072l	171,885	108.67	15.29	89	31	2.03	0.82	(1.21)	2.48	0.40
ptg000073l	192,046	121.42	17.09	2,051	6,481	379.29	16.89	(362.40)	22.45	0.04
ptg000074l	5,925,831	3,746.58	527.24	3,029	27	0.05	0.81	0.76	0.06	15.79
ptg000075l	3,949,091	2,496.80	351.37	19,177	13,459	38.30	7.68	(30.62)	4.99	0.20
ptg000076l	30,687	19.40	2.73	1	0	0.00	0.05	0.05	0.00	0.00
ptg000077l	51,654	32.66	4.60	18	0	0.00	0.55	0.55	0.00	0.00
ptg000078l	1,575,749	996.26	140.20	2,294	312	2.23	2.30	0.08	0.97	1.03
ptg000079l	89,027	56.29	7.92	31	601	75.87	0.55	(75.32)	137.76	0.01
ptg000080l	821,389	519.32	73.08	304	3	0.04	0.59	0.54	0.07	14.26
ptg000081l	3,577,722	2,262.00	318.32	1,285	15	0.05	0.57	0.52	0.08	12.06
ptg000082l	2,252,127	1,423.90	200.38	663	73	0.36	0.47	0.10	0.78	1.28
ptg000083l	37,560	23.75	3.34	10	0	0.00	0.42	0.42	0.00	0.00
ptg000084l	2,739,728	1,732.18	243.76	1,099	1	0.00	0.63	0.63	0.01	154.66
ptg000085l	3,906,051	2,469.58	347.54	2,623	282	0.81	1.06	0.25	0.76	1.31
ptg000086l	4,153,077	2,625.77	369.51	1,479	2	0.01	0.56	0.56	0.01	104.07
ptg000087l	1,236,065	781.50	109.98	452	7	0.06	0.58	0.51	0.11	9.09
ptg000088l	42,049	26.59	3.74	4	0	0.00	0.15	0.15	0.00	0.00

ptg000089l	57,227	36.18	5.09	5	0	0.00	0.14	0.14	0.00	0.00
ptg000090l	560,204	354.19	49.84	150	0	0.00	0.42	0.42	0.00	0.00
ptg000091l	194,166	122.76	17.28	2,556	730	42.26	20.82	(21.43)	2.03	0.49
ptg000092l	47,404	29.97	4.22	28	7	1.66	0.93	(0.73)	1.78	0.56
ptg000093l	118,640	75.01	10.56	32	21	1.99	0.43	(1.56)	4.66	0.21
ptg000094l	50,297	31.80	4.48	20	0	0.00	0.63	0.63	0.00	0.00
ptg000095l	138,491	87.56	12.32	49	90	7.30	0.56	(6.74)	13.05	0.08
ptg000096l	30,513	19.29	2.71	5	0	0.00	0.26	0.26	0.00	0.00
ptg000097l	1,453,588	919.02	129.33	440	8	0.06	0.48	0.42	0.13	7.74
ptg000098l	1,148,870	726.37	102.22	18,887	8,636	84.49	26.00	(58.48)	3.25	0.31
ptg000099l	76,509	48.37	6.81	6	0	0.00	0.12	0.12	0.00	0.00
ptg000100l	30,352	19.19	2.70	15	0	0.00	0.78	0.78	0.00	0.00
ptg000101l	3,292,698	2,081.79	292.96	2,181	2	0.01	1.05	1.04	0.01	153.46
ptg000102l	341,623	215.99	30.40	2,078	1,248	41.06	9.62	(31.44)	4.27	0.23
ptg000103l	52,831	33.40	4.70	11	0	0.00	0.33	0.33	0.00	0.00
ptg000104l	48,301	30.54	4.30	12	0	0.00	0.39	0.39	0.00	0.00
ptg000105l	36,091	22.82	3.21	10	0	0.00	0.44	0.44	0.00	0.00
ptg000106l	35,981	22.75	3.20	5	0	0.00	0.22	0.22	0.00	0.00
ptg000107l	30,245	19.12	2.69	8	0	0.00	0.42	0.42	0.00	0.00
ptg000108l	2,476,811	1,565.95	220.37	3,387	2,115	9.60	2.16	(7.43)	4.44	0.23
ptg000109l	258,234	163.27	22.98	2,019	2,052	89.31	12.37	(76.94)	7.22	0.14
ptg000110l	1,432,496	905.69	127.45	23,135	12,876	101.02	25.54	(75.48)	3.95	0.25
ptg000111l	37,518	23.72	3.34	21	0	0.00	0.89	0.89	0.00	0.00
ptg000113l	40,564	25.65	3.61	4	0	0.00	0.16	0.16	0.00	0.00
ptg000114l	35,738	22.60	3.18	2	0	0.00	0.09	0.09	0.00	0.00
ptg000115l	727,267	459.81	64.71	181	11	0.17	0.39	0.22	0.43	2.32
ptg000116l	37,959	24.00	3.38	1	0	0.00	0.04	0.04	0.00	0.00
ptg000117l	230,628	145.81	20.52	112	123	5.99	0.77	(5.23)	7.80	0.13
ptg000118l	34,544	21.84	3.07	2	0	0.00	0.09	0.09	0.00	0.00
ptg000119l	37,845	23.93	3.37	31	5	1.48	1.30	(0.19)	1.15	0.87
ptg000120l	36,368	22.99	3.24	11	0	0.00	0.48	0.48	0.00	0.00
ptg000121l	43,577	27.55	3.88	5	0	0.00	0.18	0.18	0.00	0.00

ptg000122l	39,183	24.77	3.49	1	0	0.00	0.04	0.04	0.00	0.00
ptg000123l	127,969	80.91	11.39	54	20	1.76	0.67	(1.09)	2.63	0.38
ptg000124l	79,091	50.00	7.04	1,148	1,094	155.46	22.96	(132.51)	6.77	0.15
ptg000125c	134,426	84.99	11.96	68	32	2.68	0.80	(1.88)	3.34	0.30
ptg000126l	39,092	24.72	3.48	17	0	0.00	0.69	0.69	0.00	0.00
ptg000127l	38,118	24.10	3.39	1	0	0.00	0.04	0.04	0.00	0.00
ptg000128l	16,137	10.20	1.44	13	1	0.70	1.27	0.58	0.55	1.83
ptg000129l	1,103,165	697.47	98.15	716	328	3.34	1.03	(2.32)	3.26	0.31
ptg000130l	429,041	271.26	38.17	13	14	0.37	0.05	(0.32)	7.65	0.13
ptg000131l	1,186,096	749.90	105.53	1,790	393	3.72	2.39	(1.34)	1.56	0.64
ptg000132l	34,885	22.06	3.10	5	0	0.00	0.23	0.23	0.00	0.00
ptg000133l	38,414,622	24,287.48	3,417.89	19,966	113	0.03	0.82	0.79	0.04	24.86
ptg000134l	240,433	152.01	21.39	93	40	1.87	0.61	(1.26)	3.06	0.33
ptg000135l	67,315	42.56	5.99	7	0	0.00	0.16	0.16	0.00	0.00
ptg000137l	26,477	16.74	2.36	14	0	0.00	0.84	0.84	0.00	0.00
ptg000138l	146,626	92.70	13.05	882	2,456	188.26	9.51	(178.74)	19.79	0.05
ptg000139l	60,710	38.38	5.40	23	0	0.00	0.60	0.60	0.00	0.00
ptg000140l	468,248	296.05	41.66	136	0	0.00	0.46	0.46	0.00	0.00
ptg000141l	48,011	30.35	4.27	8	0	0.00	0.26	0.26	0.00	0.00
ptg000142l	41,220	26.06	3.67	2	0	0.00	0.08	0.08	0.00	0.00
ptg000143l	28,292	17.89	2.52	2	0	0.00	0.11	0.11	0.00	0.00
ptg000144l	821,352	519.30	73.08	13,561	8,913	121.96	26.11	(95.85)	4.67	0.21
ptg000145l	40,172	25.40	3.57	24	0	0.00	0.94	0.94	0.00	0.00
ptg000146c	215,847	136.47	19.20	10	0	0.00	0.07	0.07	0.00	0.00
ptg000147l	35,691	22.57	3.18	1	0	0.00	0.04	0.04	0.00	0.00
ptg000148l	92,461	58.46	8.23	10	0	0.00	0.17	0.17	0.00	0.00
ptg000149l	83,013	52.48	7.39	34	21	2.84	0.65	(2.20)	4.39	0.23
ptg000150l	63,166	39.94	5.62	43	7	1.25	1.08	(0.17)	1.16	0.86
ptg000151l	25,480	16.11	2.27	1	0	0.00	0.06	0.06	0.00	0.00
ptg000152l	40,112	25.36	3.57	19	4	1.12	0.75	(0.37)	1.50	0.67
ptg000153l	41,312	26.12	3.68	24	4	1.09	0.92	(0.17)	1.18	0.84
ptg000154l	55,288	34.96	4.92	17	0	0.00	0.49	0.49	0.00	0.00

ptg000155l	26,643	16.84	2.37	3	0	0.00	0.18	0.18	0.00	0.00
ptg000156l	78,902	49.89	7.02	12	0	0.00	0.24	0.24	0.00	0.00
ptg000157l	94,777	59.92	8.43	31	0	0.00	0.52	0.52	0.00	0.00
ptg000158l	42,293	26.74	3.76	3	0	0.00	0.11	0.11	0.00	0.00
ptg000159c	99,123	62.67	8.82	3	0	0.00	0.05	0.05	0.00	0.00
ptg000160l	237,487	150.15	21.13	289	57	2.70	1.92	(0.77)	1.40	0.71
ptg000161l	97,380	61.57	8.66	356	9	1.04	5.78	4.74	0.18	5.57
ptg000162l	36,428	23.03	3.24	9	0	0.00	0.39	0.39	0.00	0.00
ptg000164l	55,647	35.18	4.95	18	0	0.00	0.51	0.51	0.00	0.00
ptg000165l	52,461	33.17	4.67	6	0	0.00	0.18	0.18	0.00	0.00
ptg000166l	65,517	41.42	5.83	15	7	1.20	0.36	(0.84)	3.32	0.30
ptg000167l	46,773	29.57	4.16	22	0	0.00	0.74	0.74	0.00	0.00
ptg000168l	72,469	45.82	6.45	6	0	0.00	0.13	0.13	0.00	0.00
ptg000169l	30,789	19.47	2.74	16	0	0.00	0.82	0.82	0.00	0.00
ptg000170l	33,551	21.21	2.99	8	0	0.00	0.38	0.38	0.00	0.00
ptg000171l	188,253	119.02	16.75	85	0	0.00	0.71	0.71	0.00	0.00
ptg000172l	386,802	244.55	34.42	141	0	0.00	0.58	0.58	0.00	0.00
ptg000173l	68,446	43.27	6.09	41	21	3.45	0.95	(2.50)	3.64	0.27
ptg000174l	25,068	15.85	2.23	7	0	0.00	0.44	0.44	0.00	0.00
ptg000175l	92,712	58.62	8.25	37	10	1.21	0.63	(0.58)	1.92	0.52
ptg000176l	39,254	24.82	3.49	22	0	0.00	0.89	0.89	0.00	0.00
ptg000177l	36,261	22.93	3.23	5	0	0.00	0.22	0.22	0.00	0.00
ptg000179l	48,323	30.55	4.30	18	0	0.00	0.59	0.59	0.00	0.00
ptg000180l	36,976	23.38	3.29	2	0	0.00	0.09	0.09	0.00	0.00
ptg000181l	35,114	22.20	3.12	20	5	1.60	0.90	(0.70)	1.78	0.56
ptg000182l	37,339	23.61	3.32	3	0	0.00	0.13	0.13	0.00	0.00
ptg000183l	34,109	21.57	3.03	7	0	0.00	0.32	0.32	0.00	0.00
ptg000184l	35,363	22.36	3.15	5	0	0.00	0.22	0.22	0.00	0.00
ptg000185l	124,274	78.57	11.06	36	0	0.00	0.46	0.46	0.00	0.00
ptg000186l	87,458	55.29	7.78	25	39	5.01	0.45	(4.56)	11.09	0.09
ptg000187l	320,975	202.94	28.56	127	130	4.55	0.63	(3.93)	7.27	0.14
ptg000188l	32,919	20.81	2.93	19	0	0.00	0.91	0.91	0.00	0.00

ptg000189l	56,619	35.80	5.04	19	0	0.00	0.53	0.53	0.00	0.00
ptg000190l	70,077	44.31	6.24	16	0	0.00	0.36	0.36	0.00	0.00
ptg000191l	47,809	30.23	4.25	2	0	0.00	0.07	0.07	0.00	0.00
ptg000192l	41,232	26.07	3.67	33	9	2.45	1.27	(1.19)	1.94	0.52
ptg000193l	272,731	172.43	24.27	126	11	0.45	0.73	0.28	0.62	1.61
ptg000194l	58,339	36.88	5.19	26	0	0.00	0.70	0.70	0.00	0.00
ptg000195l	89,081	56.32	7.93	33	16	2.02	0.59	(1.43)	3.45	0.29
ptg000196l	44,000	27.82	3.91	15	0	0.00	0.54	0.54	0.00	0.00
ptg000197l	36,159	22.86	3.22	6	0	0.00	0.26	0.26	0.00	0.00
ptg000198l	45,499	28.77	4.05	11	0	0.00	0.38	0.38	0.00	0.00
ptg000199l	43,232	27.33	3.85	10	0	0.00	0.37	0.37	0.00	0.00
ptg000200l	32,619	20.62	2.90	1	0	0.00	0.05	0.05	0.00	0.00
ptg000201l	32,569	20.59	2.90	16	0	0.00	0.78	0.78	0.00	0.00
ptg000202l	65,238	41.25	5.80	3	0	0.00	0.07	0.07	0.00	0.00
ptg000203l	37,951	23.99	3.38	4	0	0.00	0.17	0.17	0.00	0.00
ptg000204l	39,404	24.91	3.51	12	0	0.00	0.48	0.48	0.00	0.00
ptg000205l	472,102	298.48	42.00	4,838	3,055	72.73	16.21	(56.52)	4.49	0.22
ptg000206l	71,668	45.31	6.38	16	0	0.00	0.35	0.35	0.00	0.00
ptg000207l	40,222	25.43	3.58	16	0	0.00	0.63	0.63	0.00	0.00
ptg000208l	55,255	34.93	4.92	22	0	0.00	0.63	0.63	0.00	0.00
ptg000209l	41,868	26.47	3.73	4	0	0.00	0.15	0.15	0.00	0.00
ptg000210l	43,933	27.78	3.91	24	4	1.02	0.86	(0.16)	1.18	0.84
ptg000211l	72,557	45.87	6.46	6	0	0.00	0.13	0.13	0.00	0.00
ptg000212l	32,914	20.81	2.93	15	0	0.00	0.72	0.72	0.00	0.00
ptg000213l	33,465	21.16	2.98	11	1	0.34	0.52	0.18	0.65	1.55
ptg000214l	25,858	16.35	2.30	8	0	0.00	0.49	0.49	0.00	0.00
ptg000215l	38,752	24.50	3.45	25	0	0.00	1.02	1.02	0.00	0.00
ptg000216l	35,198	22.25	3.13	1	0	0.00	0.04	0.04	0.00	0.00
ptg000217l	39,589	25.03	3.52	16	0	0.00	0.64	0.64	0.00	0.00
ptg000218l	60,328	38.14	5.37	8	0	0.00	0.21	0.21	0.00	0.00
ptg000219c	167,626	105.98	14.91	249	166	11.13	2.35	(8.78)	4.74	0.21
ptg000220l	32,435	20.51	2.89	10	0	0.00	0.49	0.49	0.00	0.00

ptg000221l	49,685	31.41	4.42	3	0	0.00	0.10	0.10	0.00	0.00
ptg000222l	45,327	28.66	4.03	23	0	0.00	0.80	0.80	0.00	0.00
ptg000223l	46,242	29.24	4.11	16	0	0.00	0.55	0.55	0.00	0.00
ptg000224l	48,290	30.53	4.30	32	13	3.03	1.05	(1.98)	2.89	0.35
ptg000225l	40,689	25.73	3.62	21	0	0.00	0.82	0.82	0.00	0.00
ptg000226l	37,213	23.53	3.31	10	0	0.00	0.43	0.43	0.00	0.00
ptg000227l	42,649	26.96	3.79	17	0	0.00	0.63	0.63	0.00	0.00
ptg000228c	34,655	21.91	3.08	76	0	0.00	3.47	3.47	0.00	0.00
ptg000229l	27,339	17.28	2.43	11	3	1.23	0.64	(0.60)	1.94	0.52
ptg000230l	56,613	35.79	5.04	5	0	0.00	0.14	0.14	0.00	0.00
ptg000231l	75,816	47.93	6.75	28	9	1.33	0.58	(0.75)	2.28	0.44
ptg000233l	47,448	30.00	4.22	1	0	0.00	0.03	0.03	0.00	0.00
ptg000234l	113,795	71.95	10.12	378	623	61.53	5.25	(56.28)	11.71	0.09
ptg000235l	33,729	21.33	3.00	12	2	0.67	0.56	(0.10)	1.18	0.84
ptg000236l	59,180	37.42	5.27	30	0	0.00	0.80	0.80	0.00	0.00
ptg000237l	83,262	52.64	7.41	11	0	0.00	0.21	0.21	0.00	0.00
ptg000238l	31,323	19.80	2.79	11	0	0.00	0.56	0.56	0.00	0.00
ptg000239l	38,157	24.12	3.39	4	0	0.00	0.17	0.17	0.00	0.00
ptg000240l	301,786	190.80	26.85	6	0	0.00	0.03	0.03	0.00	0.00
ptg000241l	41,553	26.27	3.70	25	0	0.00	0.95	0.95	0.00	0.00
ptg000242l	31,512	19.92	2.80	13	0	0.00	0.65	0.65	0.00	0.00
ptg000243l	29,228	18.48	2.60	7	0	0.00	0.38	0.38	0.00	0.00
ptg000244l	47,193	29.84	4.20	13	0	0.00	0.44	0.44	0.00	0.00
ptg000245l	32,087	20.29	2.85	14	0	0.00	0.69	0.69	0.00	0.00
ptg000246l	121,432	76.77	10.80	39	19	1.76	0.51	(1.25)	3.46	0.29
ptg000247l	46,776	29.57	4.16	24	0	0.00	0.81	0.81	0.00	0.00
ptg000248l	49,057	31.02	4.36	2	500	114.55	0.06	(114.49)	1,776.50	0.00
ptg000249l	29,514	18.66	2.63	13	8	3.05	0.70	(2.35)	4.37	0.23
ptg000250l	228,542	144.49	20.33	96	49	2.41	0.66	(1.75)	3.63	0.28
ptg000251l	77,839	49.21	6.93	5	0	0.00	0.10	0.10	0.00	0.00
ptg000252l	102,337	64.70	9.11	60	18	1.98	0.93	(1.05)	2.13	0.47
ptg000253l	29,939	18.93	2.66	12	4	1.50	0.63	(0.87)	2.37	0.42

ptg000254l	37,302	23.58	3.32	4	0	0.00	0.17	0.17	0.00	0.00
ptg000255l	22,933	14.50	2.04	82	320	156.83	5.66	(151.17)	27.73	0.04
ptg000256l	80,263	50.75	7.14	11	0	0.00	0.22	0.22	0.00	0.00
ptg000257l	30,721	19.42	2.73	12	0	0.00	0.62	0.62	0.00	0.00
ptg000258l	25,996	16.44	2.31	1	0	0.00	0.06	0.06	0.00	0.00
ptg000259l	67,526	42.69	6.01	4	0	0.00	0.09	0.09	0.00	0.00
ptg000261l	34,695	21.94	3.09	5	0	0.00	0.23	0.23	0.00	0.00
ptg000262l	194,521	122.99	17.31	64	35	2.02	0.52	(1.50)	3.89	0.26
ptg000263l	47,196	29.84	4.20	19	0	0.00	0.64	0.64	0.00	0.00
ptg000264l	36,961	23.37	3.29	238	343	104.30	10.18	(94.12)	10.24	0.10
ptg000265l	34,848	22.03	3.10	18	0	0.00	0.82	0.82	0.00	0.00
ptg000266l	50,290	31.80	4.47	20	0	0.00	0.63	0.63	0.00	0.00
ptg000267l	38,883	24.58	3.46	14	7	2.02	0.57	(1.45)	3.55	0.28
ptg000268l	32,937	20.82	2.93	8	0	0.00	0.38	0.38	0.00	0.00
ptg000269l	38,571	24.39	3.43	8	0	0.00	0.33	0.33	0.00	0.00
ptg000270l	139,565	88.24	12.42	42	28	2.25	0.48	(1.78)	4.74	0.21
ptg000271l	24,207	15.30	2.15	0	213	98.90	0.00	(98.90)	0.00	0
ptg000272l	78,039	49.34	6.94	37	9	1.30	0.75	(0.55)	1.73	0.58
ptg000273l	35,348	22.35	3.15	15	0	0.00	0.67	0.67	0.00	0.00
ptg000274l	34,526	21.83	3.07	2	0	0.00	0.09	0.09	0.00	0.00
ptg000275l	31,721	20.06	2.82	9	0	0.00	0.45	0.45	0.00	0.00
ptg000276l	39,203	24.79	3.49	5	0	0.00	0.20	0.20	0.00	0.00
ptg000277l	38,076	24.07	3.39	21	4	1.18	0.87	(0.31)	1.35	0.74
ptg000278l	52,947	33.48	4.71	13	0	0.00	0.39	0.39	0.00	0.00
ptg000279l	42,983	27.18	3.82	15	6	1.57	0.55	(1.02)	2.84	0.35
ptg000280l	41,574	26.28	3.70	3	0	0.00	0.11	0.11	0.00	0.00
ptg000281l	28,424	17.97	2.53	1	0	0.00	0.06	0.06	0.00	0.00
ptg000282l	34,187	21.61	3.04	20	0	0.00	0.93	0.93	0.00	0.00
ptg000283l	52,101	32.94	4.64	31	4	0.86	0.94	0.08	0.92	1.09
ptg000284l	46,836	29.61	4.17	26	0	0.00	0.88	0.88	0.00	0.00
ptg000285l	26,236	16.59	2.33	23	0	0.00	1.39	1.39	0.00	0.00
ptg000286l	61,050	38.60	5.43	14	0	0.00	0.36	0.36	0.00	0.00

ptg000287l	40,984	25.91	3.65	2	0	0.00	0.08	0.08	0.00	0.00
ptg000288l	42,557	26.91	3.79	17	0	0.00	0.63	0.63	0.00	0.00
ptg000289l	34,075	21.54	3.03	7	0	0.00	0.32	0.32	0.00	0.00
ptg000290l	26,956	17.04	2.40	6	0	0.00	0.35	0.35	0.00	0.00
ptg000291l	33,885	21.42	3.01	458	1,048	347.61	21.38	(326.23)	16.26	0.06
ptg000292l	32,949	20.83	2.93	9	0	0.00	0.43	0.43	0.00	0.00
ptg000293l	58,627	37.07	5.22	967	242	46.39	26.09	(20.31)	1.78	0.56
ptg000294l	34,309	21.69	3.05	1	0	0.00	0.05	0.05	0.00	0.00
ptg000295l	35,113	22.20	3.12	16	0	0.00	0.72	0.72	0.00	0.00
ptg000296l	45,711	28.90	4.07	16	0	0.00	0.55	0.55	0.00	0.00
ptg000297l	147,938	93.53	13.16	40	13	0.99	0.43	(0.56)	2.31	0.43
ptg000298l	87,463	55.30	7.78	26	0	0.00	0.47	0.47	0.00	0.00
ptg000299l	43,715	27.64	3.89	20	26	6.68	0.72	(5.96)	9.24	0.11
ptg000300l	26,400	16.69	2.35	7	5	2.13	0.42	(1.71)	5.08	0.20
ptg000301l	31,793	20.10	2.83	8	0	0.00	0.40	0.40	0.00	0.00
ptg000302l	39,633	25.06	3.53	21	3	0.85	0.84	(0.01)	1.02	0.99
ptg000303l	40,595	25.67	3.61	13	0	0.00	0.51	0.51	0.00	0.00
ptg000304l	40,831	25.82	3.63	27	0	0.00	1.05	1.05	0.00	0.00
ptg000305l	32,092	20.29	2.86	9	0	0.00	0.44	0.44	0.00	0.00
ptg000306l	37,062	23.43	3.30	1	0	0.00	0.04	0.04	0.00	0.00
ptg000307l	26,319	16.64	2.34	6	0	0.00	0.36	0.36	0.00	0.00
ptg000308l	35,094	22.19	3.12	8	7	2.24	0.36	(1.88)	6.22	0.16
ptg000309l	139,705	88.33	12.43	60	30	2.41	0.68	(1.73)	3.55	0.28
ptg000310l	103,709	65.57	9.23	17	0	0.00	0.26	0.26	0.00	0.00
ptg000311l	33,942	21.46	3.02	2	0	0.00	0.09	0.09	0.00	0.00
ptg000312l	36,441	23.04	3.24	18	0	0.00	0.78	0.78	0.00	0.00
ptg000313l	32,179	20.35	2.86	5	0	0.00	0.25	0.25	0.00	0.00
ptg000314l	51,209	32.38	4.56	20	0	0.00	0.62	0.62	0.00	0.00
ptg000315l	64,887	41.02	5.77	35	0	0.00	0.85	0.85	0.00	0.00
ptg000316l	45,062	28.49	4.01	5	0	0.00	0.18	0.18	0.00	0.00
ptg000317l	63,374	40.07	5.64	8	0	0.00	0.20	0.20	0.00	0.00
ptg000318l	86,272	54.55	7.68	1	0	0.00	0.02	0.02	0.00	0

ptg000319l	26,991	17.06	2.40	3	0	0.00	0.18	0.18	0.00	0.00
ptg000320l	59,418	37.57	5.29	11	0	0.00	0.29	0.29	0.00	0.00
ptg000321l	50,753	32.09	4.52	20	0	0.00	0.62	0.62	0.00	0.00
ptg000322l	49,605	31.36	4.41	20	0	0.00	0.64	0.64	0.00	0.00
ptg000323l	33,536	21.20	2.98	4	0	0.00	0.19	0.19	0.00	0.00
ptg000324l	27,958	17.68	2.49	10	0	0.00	0.57	0.57	0.00	0.00
ptg000325l	44,581	28.19	3.97	19	0	0.00	0.67	0.67	0.00	0.00
ptg000326l	37,300	23.58	3.32	19	0	0.00	0.81	0.81	0.00	0.00
ptg000327l	31,929	20.19	2.84	10	0	0.00	0.50	0.50	0.00	0.00
ptg000328l	57,416	36.30	5.11	5	0	0.00	0.14	0.14	0.00	0.00
ptg000329l	90,077	56.95	8.01	36	42	5.24	0.63	(4.61)	8.29	0.12
ptg000330l	103,685	65.55	9.23	47	144	15.61	0.72	(14.89)	21.77	0.05
ptg000331l	20,901	13.21	1.86	4	5	2.69	0.30	(2.39)	8.88	0.11
ptg000332l	31,910	20.17	2.84	13	0	0.00	0.64	0.64	0.00	0.00
ptg000333l	56,112	35.48	4.99	1	1	0.20	0.03	(0.17)	7.11	0.14
ptg000334l	39,207	24.79	3.49	7	0	0.00	0.28	0.28	0.00	0.00
ptg000335l	47,613	30.10	4.24	12	0	0.00	0.40	0.40	0.00	0.00
ptg000337l	28,348	17.92	2.52	12	0	0.00	0.67	0.67	0.00	0.00
ptg000338l	38,744	24.50	3.45	21	0	0.00	0.86	0.86	0.00	0.00
ptg000339l	40,619	25.68	3.61	3	0	0.00	0.12	0.12	0.00	0.00
ptg000340l	16,186	10.23	1.44	5	0	0.00	0.49	0.49	0.00	0.00
ptg000341l	80,868	51.13	7.20	19	0	0.00	0.37	0.37	0.00	0.00
ptg000342l	57,208	36.17	5.09	10	0	0.00	0.28	0.28	0.00	0.00
ptg000343l	28,650	18.11	2.55	1	0	0.00	0.06	0.06	0.00	0.00
ptg000344l	49,579	31.35	4.41	5	0	0.00	0.16	0.16	0.00	0.00
ptg000345l	29,907	18.91	2.66	3	0	0.00	0.16	0.16	0.00	0.00
ptg000346l	53,167	33.61	4.73	6	0	0.00	0.18	0.18	0.00	0.00
ptg000347l	37,319	23.59	3.32	11	0	0.00	0.47	0.47	0.00	0.00
ptg000348l	82,290	52.03	7.32	12	0	0.00	0.23	0.23	0.00	0.00
ptg000349l	30,644	19.37	2.73	11	0	0.00	0.57	0.57	0.00	0.00
ptg000350l	25,954	16.41	2.31	2	0	0.00	0.12	0.12	0.00	0.00
ptg000351l	21,807	13.79	1.94	12	0	0.00	0.87	0.87	0.00	0.00

ptg000352l	66,473	42.03	5.91	30	0	0.00	0.71	0.71	0.00	0.00
ptg000353l	94,741	59.90	8.43	31	17	2.02	0.52	(1.50)	3.90	0.26
ptg000354l	46,472	29.38	4.13	15	0	0.00	0.51	0.51	0.00	0.00
ptg000355l	40,332	25.50	3.59	32	0	0.00	1.25	1.25	0.00	0.00
ptg000356l	115,918	73.29	10.31	30	0	0.00	0.41	0.41	0.00	0.00
ptg000357l	34,974	22.11	3.11	6	0	0.00	0.27	0.27	0.00	0.00
ptg000358l	43,393	27.44	3.86	12	0	0.00	0.44	0.44	0.00	0.00
ptg000359l	37,319	23.59	3.32	20	0	0.00	0.85	0.85	0.00	0.00
ptg000360l	57,232	36.18	5.09	302	284	55.77	8.35	(47.43)	6.68	0.15
ptg000362l	64,324	40.67	5.72	26	0	0.00	0.64	0.64	0.00	0.00
ptg000363l	32,781	20.73	2.92	529	254	87.09	25.52	(61.56)	3.41	0.29
ptg000364l	31,270	19.77	2.78	3	0	0.00	0.15	0.15	0.00	0.00
ptg000365l	39,280	24.83	3.49	11	0	0.00	0.44	0.44	0.00	0.00
ptg000366l	56,948	36.01	5.07	3	0	0.00	0.08	0.08	0.00	0.00
ptg000367l	44,577	28.18	3.97	4	0	0.00	0.14	0.14	0.00	0.00
ptg000368l	42,213	26.69	3.76	11	0	0.00	0.41	0.41	0.00	0.00
ptg000369l	32,965	20.84	2.93	13	0	0.00	0.62	0.62	0.00	0.00
ptg000370l	44,007	27.82	3.92	16	0	0.00	0.58	0.58	0.00	0.00
ptg000371l	34,518	21.82	3.07	2	0	0.00	0.09	0.09	0.00	0.00
ptg000372l	57,504	36.36	5.12	22	0	0.00	0.61	0.61	0.00	0.00
ptg000373l	41,001	25.92	3.65	5	0	0.00	0.19	0.19	0.00	0.00
ptg000374l	26,354	16.66	2.34	13	6	2.56	0.78	(1.78)	3.28	0.30
ptg000375l	31,341	19.82	2.79	18	0	0.00	0.91	0.91	0.00	0.00
ptg000376l	44,260	27.98	3.94	16	0	0.00	0.57	0.57	0.00	0.00
ptg000377l	20,931	13.23	1.86	16	2	1.07	1.21	0.14	0.89	1.13
ptg000378l	37,637	23.80	3.35	17	0	0.00	0.71	0.71	0.00	0.00
ptg000379l	32,251	20.39	2.87	5	0	0.00	0.25	0.25	0.00	0.00
ptg000380l	37,813	23.91	3.36	6	0	0.00	0.25	0.25	0.00	0.00
ptg000381l	46,818	29.60	4.17	16	5	1.20	0.54	(0.66)	2.22	0.45
ptg000382l	29,658	18.75	2.64	3	0	0.00	0.16	0.16	0.00	0.00
ptg000383l	43,898	27.75	3.91	14	0	0.00	0.50	0.50	0.00	0.00
ptg000384l	30,815	19.48	2.74	2	0	0.00	0.10	0.10	0.00	0.00

ptg000385l	31,571	19.96	2.81	11	4	1.42	0.55	(0.87)	2.58	0.39
ptg000386l	43,664	27.61	3.88	14	0	0.00	0.51	0.51	0.00	0.00
ptg000388l	89,265	56.44	7.94	29	0	0.00	0.51	0.51	0.00	0.00
ptg000389l	35,137	22.22	3.13	5	0	0.00	0.23	0.23	0.00	0.00
ptg000390l	28,597	18.08	2.54	8	0	0.00	0.44	0.44	0.00	0.00
ptg000391l	44,478	28.12	3.96	1	0	0.00	0.04	0.04	0.00	0.00
ptg000392l	30,032	18.99	2.67	10	0	0.00	0.53	0.53	0.00	0.00
ptg000393l	41,261	26.09	3.67	1	0	0.00	0.04	0.04	0.00	0.00
ptg000394l	44,294	28.00	3.94	18	0	0.00	0.64	0.64	0.00	0.00
ptg000395l	29,037	18.36	2.58	2	0	0.00	0.11	0.11	0.00	0.00
ptg000396l	144,237	91.19	12.83	60	16	1.25	0.66	(0.59)	1.89	0.53
ptg000397l	49,495	31.29	4.40	15	0	0.00	0.48	0.48	0.00	0.00
ptg000398l	20,658	13.06	1.84	11	0	0.00	0.84	0.84	0.00	0.00
ptg000399l	49,499	31.30	4.40	25	0	0.00	0.80	0.80	0.00	0.00
ptg000400l	29,399	18.59	2.62	1	0	0.00	0.05	0.05	0.00	0.00
ptg000401l	34,674	21.92	3.09	6	0	0.00	0.27	0.27	0.00	0.00
ptg000402l	43,920	27.77	3.91	25	0	0.00	0.90	0.90	0.00	0.00
ptg000403l	90,262	57.07	8.03	37	12	1.49	0.65	(0.85)	2.30	0.43
ptg000404l	44,690	28.26	3.98	4	0	0.00	0.14	0.14	0.00	0.00
ptg000405l	26,291	16.62	2.34	1	0	0.00	0.06	0.06	0.00	0.00
ptg000406l	43,392	27.43	3.86	16	4	1.04	0.58	(0.45)	1.78	0.56
ptg000407l	39,410	24.92	3.51	13	0	0.00	0.52	0.52	0.00	0.00
ptg000408l	36,403	23.02	3.24	20	0	0.00	0.87	0.87	0.00	0.00
ptg000409l	40,921	25.87	3.64	17	0	0.00	0.66	0.66	0.00	0.00
ptg000410l	23,331	14.75	2.08	9	0	0.00	0.61	0.61	0.00	0.00
ptg000411l	62,086	39.25	5.52	31	0	0.00	0.79	0.79	0.00	0.00
ptg000412l	37,214	23.53	3.31	26	0	0.00	1.11	1.11	0.00	0.00
ptg000413l	72,700	45.96	6.47	38	12	1.86	0.83	(1.03)	2.24	0.45
ptg000414l	56,995	36.03	5.07	19	0	0.00	0.53	0.53	0.00	0.00
ptg000415l	33,785	21.36	3.01	2	0	0.00	0.09	0.09	0.00	0.00
ptg000417l	38,249	24.18	3.40	9	0	0.00	0.37	0.37	0.00	0.00
ptg000418l	82,912	52.42	7.38	18	0	0.00	0.34	0.34	0.00	0.00

ptg000419l	20,770	13.13	1.85	5	0	0.00	0.38	0.38	0.00	0.00
ptg000420l	33,269	21.03	2.96	4	0	0.00	0.19	0.19	0.00	0.00
ptg000421l	30,000	18.97	2.67	16	0	0.00	0.84	0.84	0.00	0.00
ptg000422l	45,587	28.82	4.06	4	0	0.00	0.14	0.14	0.00	0.00
ptg000423l	24,647	15.58	2.19	17	0	0.00	1.09	1.09	0.00	0.00
ptg000424l	64,701	40.91	5.76	3	0	0.00	0.07	0.07	0.00	0.00
ptg000425l	47,792	30.22	4.25	24	6	1.41	0.79	(0.62)	1.78	0.56
ptg000427l	29,807	18.85	2.65	1	0	0.00	0.05	0.05	0.00	0.00
ptg000428l	46,120	29.16	4.10	22	0	0.00	0.75	0.75	0.00	0.00
ptg000430l	74,252	46.95	6.61	8	0	0.00	0.17	0.17	0.00	0.00
ptg000431l	51,457	32.53	4.58	22	14	3.06	0.68	(2.38)	4.52	0.22
ptg000432l	65,954	41.70	5.87	20	0	0.00	0.48	0.48	0.00	0.00
ptg000433l	51,059	32.28	4.54	16	0	0.00	0.50	0.50	0.00	0.00
ptg000434l	56,610	35.79	5.04	28	0	0.00	0.78	0.78	0.00	0.00
ptg000435l	38,237	24.18	3.40	9	0	0.00	0.37	0.37	0.00	0.00
ptg000436l	30,713	19.42	2.73	11	0	0.00	0.57	0.57	0.00	0.00
ptg000437l	75,862	47.96	6.75	24	0	0.00	0.50	0.50	0.00	0.00
ptg000438l	33,249	21.02	2.96	8	0	0.00	0.38	0.38	0.00	0.00
ptg000439l	22,766	14.39	2.03	6	6	2.96	0.42	(2.55)	7.11	0.14
ptg000440l	41,117	26.00	3.66	15	0	0.00	0.58	0.58	0.00	0.00
ptg000441l	50,144	31.70	4.46	1	0	0.00	0.03	0.03	0.00	0.00
ptg000442l	35,955	22.73	3.20	3	0	0.00	0.13	0.13	0.00	0.00
ptg000443l	58,297	36.86	5.19	19	0	0.00	0.52	0.52	0.00	0.00
ptg000445l	41,195	26.05	3.67	19	9	2.46	0.73	(1.73)	3.37	0.30
ptg000446l	66,655	42.14	5.93	12	0	0.00	0.28	0.28	0.00	0.00
ptg000447l	44,812	28.33	3.99	11	0	0.00	0.39	0.39	0.00	0.00
ptg000448l	45,044	28.48	4.01	25	9	2.25	0.88	(1.37)	2.56	0.39
ptg000449l	40,127	25.37	3.57	13	0	0.00	0.51	0.51	0.00	0.00
ptg000450l	44,614	28.21	3.97	9	16	4.03	0.32	(3.71)	12.63	0.08
ptg000451l	30,456	19.26	2.71	9	0	0.00	0.47	0.47	0.00	0.00
ptg000452l	33,644	21.27	2.99	13	6	2.00	0.61	(1.39)	3.28	0.30
ptg000454l	27,717	17.52	2.47	2	0	0.00	0.11	0.11	0.00	0.00

ptg000455l	39,979	25.28	3.56	4	0	0.00	0.16	0.16	0.00	0.00
ptg000456l	36,971	23.37	3.29	9	0	0.00	0.39	0.39	0.00	0.00
ptg000457l	47,288	29.90	4.21	3	0	0.00	0.10	0.10	0.00	0.00
ptg000458l	57,571	36.40	5.12	13	0	0.00	0.36	0.36	0.00	0.00
ptg000459l	16,821	10.64	1.50	4	4	2.67	0.38	(2.30)	7.11	0.14
ptg000460l	34,055	21.53	3.03	3	0	0.00	0.14	0.14	0.00	0.00
ptg000461l	50,196	31.74	4.47	5	0	0.00	0.16	0.16	0.00	0.00
ptg000462l	37,990	24.02	3.38	3	0	0.00	0.12	0.12	0.00	0.00
ptg000463l	33,584	21.23	2.99	1	0	0.00	0.05	0.05	0.00	0.00
ptg000464l	31,161	19.70	2.77	13	0	0.00	0.66	0.66	0.00	0.00
ptg000465l	46,829	29.61	4.17	11	0	0.00	0.37	0.37	0.00	0.00
ptg000466l	20,617	13.04	1.83	8	0	0.00	0.61	0.61	0.00	0.00
ptg000467l	31,323	19.80	2.79	10	0	0.00	0.50	0.50	0.00	0.00
ptg000468l	38,492	24.34	3.42	6	0	0.00	0.25	0.25	0.00	0.00
ptg000469l	27,335	17.28	2.43	16	0	0.00	0.93	0.93	0.00	0.00
ptg000470l	42,112	26.63	3.75	15	0	0.00	0.56	0.56	0.00	0.00
ptg000471l	28,868	18.25	2.57	11	5	1.95	0.60	(1.34)	3.23	0.31
ptg000472l	38,267	24.19	3.40	14	0	0.00	0.58	0.58	0.00	0.00
ptg000473l	45,755	28.93	4.07	7	0	0.00	0.24	0.24	0.00	0.00
ptg000474l	30,677	19.40	2.73	10	0	0.00	0.52	0.52	0.00	0.00
ptg000475l	45,741	28.92	4.07	15	6	1.47	0.52	(0.96)	2.84	0.35
ptg000476l	45,348	28.67	4.03	16	0	0.00	0.56	0.56	0.00	0.00
ptg000477l	32,114	20.30	2.86	15	0	0.00	0.74	0.74	0.00	0.00
ptg000478l	21,082	13.33	1.88	5	0	0.00	0.38	0.38	0.00	0.00
ptg000479l	37,477	23.69	3.33	23	3	0.90	0.97	0.07	0.93	1.08
ptg000481l	27,523	17.40	2.45	17	0	0.00	0.98	0.98	0.00	0.00
ptg000482l	33,710	21.31	3.00	17	0	0.00	0.80	0.80	0.00	0.00
ptg000483l	47,877	30.27	4.26	1	0	0.00	0.03	0.03	0.00	0.00
ptg000484l	33,507	21.18	2.98	15	0	0.00	0.71	0.71	0.00	0.00
ptg000485l	47,586	30.09	4.23	10	0	0.00	0.33	0.33	0.00	0.00
ptg000486l	40,317	25.49	3.59	16	0	0.00	0.63	0.63	0.00	0.00
ptg000487l	42,988	27.18	3.82	7	0	0.00	0.26	0.26	0.00	0.00

ptg000488l	32,508	20.55	2.89	1	0	0.00	0.05	0.05	0.00	0.00
ptg000489l	36,145	22.85	3.22	12	4	1.24	0.53	(0.72)	2.37	0.42
ptg000490l	38,908	24.60	3.46	2	0	0.00	0.08	0.08	0.00	0.00
ptg000491l	36,185	22.88	3.22	16	0	0.00	0.70	0.70	0.00	0.00
ptg000492l	42,792	27.06	3.81	33	0	0.00	1.22	1.22	0.00	0.00
ptg000493l	29,449	18.62	2.62	3	0	0.00	0.16	0.16	0.00	0.00
ptg000494l	36,272	22.93	3.23	22	0	0.00	0.96	0.96	0.00	0.00
ptg000496l	31,865	20.15	2.84	9	5	1.76	0.45	(1.32)	3.95	0.25
ptg000497l	30,864	19.51	2.75	12	0	0.00	0.61	0.61	0.00	0.00
ptg000499l	28,395	17.95	2.53	18	5	1.98	1.00	(0.98)	1.97	0.51
ptg000500l	34,163	21.60	3.04	8	0	0.00	0.37	0.37	0.00	0.00
ptg000501l	33,824	21.39	3.01	8	0	0.00	0.37	0.37	0.00	0.00
ptg000502l	28,189	17.82	2.51	6	0	0.00	0.34	0.34	0.00	0.00
ptg000503l	33,599	21.24	2.99	20	0	0.00	0.94	0.94	0.00	0.00
ptg000504l	81,548	51.56	7.26	20	8	1.10	0.39	(0.71)	2.84	0.35
ptg000505l	28,861	18.25	2.57	16	0	0.00	0.88	0.88	0.00	0.00
ptg000506l	35,688	22.56	3.18	16	0	0.00	0.71	0.71	0.00	0.00
ptg000507l	54,349	34.36	4.84	14	0	0.00	0.41	0.41	0.00	0.00
ptg000508l	40,716	25.74	3.62	14	6	1.66	0.54	(1.11)	3.05	0.33
ptg000509l	29,070	18.38	2.59	9	0	0.00	0.49	0.49	0.00	0.00
ptg000510l	40,624	25.68	3.61	3	0	0.00	0.12	0.12	0.00	0.00
ptg000511l	37,625	23.79	3.35	13	0	0.00	0.55	0.55	0.00	0.00
ptg000512l	36,279	22.94	3.23	5	0	0.00	0.22	0.22	0.00	0.00
ptg000513l	44,919	28.40	4.00	6	0	0.00	0.21	0.21	0.00	0.00
ptg000514l	60,672	38.36	5.40	2	0	0.00	0.05	0.05	0.00	0.00
ptg000515l	31,167	19.71	2.77	7	0	0.00	0.36	0.36	0.00	0.00
ptg000516l	43,883	27.74	3.90	15	0	0.00	0.54	0.54	0.00	0.00
ptg000517l	29,084	18.39	2.59	11	0	0.00	0.60	0.60	0.00	0.00
ptg000518l	70,689	44.69	6.29	29	0	0.00	0.65	0.65	0.00	0.00
ptg000519l	28,562	18.06	2.54	1	0	0.00	0.06	0.06	0.00	0.00
ptg000520l	32,207	20.36	2.87	8	5	1.74	0.39	(1.35)	4.44	0.23
ptg000521l	36,149	22.86	3.22	2	0	0.00	0.09	0.09	0.00	0.00

ptg000522l	38,463	24.32	3.42	10	18	5.26	0.41	(4.85)	12.79	0.08
ptg000523l	28,829	18.23	2.57	16	0	0.00	0.88	0.88	0.00	0.00
ptg000524l	57,599	36.42	5.12	3	0	0.00	0.08	0.08	0.00	0.00
ptg000525l	38,237	24.18	3.40	19	0	0.00	0.79	0.79	0.00	0.00
ptg000526l	77,102	48.75	6.86	28	0	0.00	0.57	0.57	0.00	0.00
ptg000527l	42,863	27.10	3.81	26	5	1.31	0.96	(0.35)	1.37	0.73
ptg000528l	63,676	40.26	5.67	21	25	4.41	0.52	(3.89)	8.46	0.12
ptg000529l	32,922	20.81	2.93	4	0	0.00	0.19	0.19	0.00	0.00
ptg000531l	33,771	21.35	3.00	10	2	0.67	0.47	(0.20)	1.42	0.70
ptg000532l	41,667	26.34	3.71	1	0	0.00	0.04	0.04	0.00	0.00
ptg000533l	29,970	18.95	2.67	10	0	0.00	0.53	0.53	0.00	0.00
ptg000534l	26,629	16.84	2.37	1	0	0.00	0.06	0.06	0.00	0.00
ptg000535l	52,894	33.44	4.71	27	0	0.00	0.81	0.81	0.00	0.00
ptg000536l	41,069	25.97	3.65	26	11	3.01	1.00	(2.01)	3.01	0.33
ptg000537l	34,378	21.74	3.06	3	0	0.00	0.14	0.14	0.00	0.00
ptg000538l	29,503	18.65	2.62	1	0	0.00	0.05	0.05	0.00	0.00
ptg000539l	24,950	15.77	2.22	7	0	0.00	0.44	0.44	0.00	0.00
ptg000540l	34,871	22.05	3.10	9	0	0.00	0.41	0.41	0.00	0.00
ptg000541l	42,690	26.99	3.80	8	0	0.00	0.30	0.30	0.00	0.00
ptg000542l	48,749	30.82	4.34	20	0	0.00	0.65	0.65	0.00	0.00
ptg000543l	26,360	16.67	2.35	5	0	0.00	0.30	0.30	0.00	0.00
ptg000544l	32,216	20.37	2.87	15	0	0.00	0.74	0.74	0.00	0.00
ptg000545l	44,158	27.92	3.93	18	0	0.00	0.64	0.64	0.00	0.00
ptg000546l	23,861	15.09	2.12	6	4	1.88	0.40	(1.49)	4.74	0.21
ptg000547l	47,258	29.88	4.20	2	0	0.00	0.07	0.07	0.00	0.00
ptg000548l	55,205	34.90	4.91	14	0	0.00	0.40	0.40	0.00	0.00
ptg000550l	34,331	21.71	3.05	12	0	0.00	0.55	0.55	0.00	0.00
ptg000551l	50,622	32.01	4.50	2	0	0.00	0.06	0.06	0.00	0.00
ptg000552l	30,470	19.26	2.71	13	0	0.00	0.67	0.67	0.00	0.00
ptg000553l	41,914	26.50	3.73	1	0	0.00	0.04	0.04	0.00	0.00
ptg000554l	36,057	22.80	3.21	36	0	0.00	1.58	1.58	0.00	0.00
ptg000555l	39,207	24.79	3.49	16	0	0.00	0.65	0.65	0.00	0.00

ptg000556l	32,093	20.29	2.86	6	6	2.10	0.30	(1.81)	7.11	0.14
ptg000557l	48,829	30.87	4.34	17	0	0.00	0.55	0.55	0.00	0.00
ptg000559l	26,553	16.79	2.36	13	82	34.71	0.77	(33.93)	44.82	0.02
ptg000560l	43,905	27.76	3.91	4	0	0.00	0.14	0.14	0.00	0.00
ptg000561l	38,630	24.42	3.44	15	0	0.00	0.61	0.61	0.00	0.00
ptg000562l	33,765	21.35	3.00	13	0	0.00	0.61	0.61	0.00	0.00
ptg000563l	61,008	38.57	5.43	14	0	0.00	0.36	0.36	0.00	0.00
ptg000564l	44,331	28.03	3.94	20	0	0.00	0.71	0.71	0.00	0.00
ptg000565l	34,221	21.64	3.04	12	0	0.00	0.55	0.55	0.00	0.00
ptg000566l	60,699	38.38	5.40	3	0	0.00	0.08	0.08	0.00	0.00
ptg000567l	60,587	38.31	5.39	12	0	0.00	0.31	0.31	0.00	0.00
ptg000568l	38,171	24.13	3.40	4	26	7.66	0.17	(7.49)	46.19	0.02
ptg000569l	51,827	32.77	4.61	19	0	0.00	0.58	0.58	0.00	0.00
ptg000570l	34,315	21.70	3.05	18	3	0.98	0.83	(0.15)	1.18	0.84
ptg000572l	27,316	17.27	2.43	1	0	0.00	0.06	0.06	0.00	0.00
ptg000573l	21,112	13.35	1.88	3	4	2.13	0.22	(1.90)	9.47	0.11
ptg000574l	24,715	15.63	2.20	15	0	0.00	0.96	0.96	0.00	0.00
ptg000575l	44,979	28.44	4.00	6	0	0.00	0.21	0.21	0.00	0.00
ptg000576l	33,467	21.16	2.98	10	0	0.00	0.47	0.47	0.00	0.00
ptg000577l	30,124	19.05	2.68	10	0	0.00	0.53	0.53	0.00	0.00
ptg000578l	46,255	29.24	4.12	19	0	0.00	0.65	0.65	0.00	0.00
ptg000579l	35,015	22.14	3.12	21	0	0.00	0.95	0.95	0.00	0.00
ptg000580l	30,183	19.08	2.69	12	0	0.00	0.63	0.63	0.00	0.00
ptg000581l	38,895	24.59	3.46	7	0	0.00	0.28	0.28	0.00	0.00
ptg000582l	31,755	20.08	2.83	1	0	0.00	0.05	0.05	0.00	0.00
ptg000583l	25,829	16.33	2.30	16	0	0.00	0.98	0.98	0.00	0.00
ptg000584l	47,761	30.20	4.25	10	0	0.00	0.33	0.33	0.00	0.00
ptg000585l	39,777	25.15	3.54	5	0	0.00	0.20	0.20	0.00	0.00
ptg000586l	51,386	32.49	4.57	20	0	0.00	0.62	0.62	0.00	0.00
ptg000587l	45,816	28.97	4.08	23	0	0.00	0.79	0.79	0.00	0.00
ptg000588l	27,840	17.60	2.48	13	0	0.00	0.74	0.74	0.00	0.00
ptg000589l	46,197	29.21	4.11	6	0	0.00	0.21	0.21	0.00	0.00

ptg000590l	34,527	21.83	3.07	10	0	0.00	0.46	0.46	0.00	0.00
ptg000591l	40,515	25.62	3.60	4	0	0.00	0.16	0.16	0.00	0.00
ptg000592l	43,813	27.70	3.90	2	0	0.00	0.07	0.07	0.00	0.00
ptg000594l	38,383	24.27	3.42	7	0	0.00	0.29	0.29	0.00	0.00
ptg000595l	24,644	15.58	2.19	13	0	0.00	0.83	0.83	0.00	0.00
ptg000596l	40,357	25.52	3.59	2	0	0.00	0.08	0.08	0.00	0.00
ptg000597l	38,181	24.14	3.40	3	0	0.00	0.12	0.12	0.00	0.00
ptg000598l	44,801	28.33	3.99	26	0	0.00	0.92	0.92	0.00	0.00
ptg000599l	44,936	28.41	4.00	22	0	0.00	0.77	0.77	0.00	0.00
ptg000600l	31,689	20.04	2.82	4	0	0.00	0.20	0.20	0.00	0.00
ptg000601l	36,421	23.03	3.24	19	0	0.00	0.83	0.83	0.00	0.00
ptg000602l	20,852	13.18	1.86	1	0	0.00	0.08	0.08	0.00	0.00
ptg000603l	24,411	15.43	2.17	19	0	0.00	1.23	1.23	0.00	0.00
ptg000604l	10,905	6.89	0.97	1	0	0.00	0.15	0.15	0.00	0.00
ptg000606l	19,199	12.14	1.71	2	0	0.00	0.16	0.16	0.00	0.00
ptg000607l	19,189	12.13	1.71	16	0	0.00	1.32	1.32	0.00	0.00
ptg000608l	29,632	18.73	2.64	5	6	2.28	0.27	(2.01)	8.53	0.12
ptg000609l	35,500	22.44	3.16	24	0	0.00	1.07	1.07	0.00	0.00
ptg000610l	24,295	15.36	2.16	7	3	1.39	0.46	(0.93)	3.05	0.33
ptg000611l	18,857	11.92	1.68	2	0	0.00	0.17	0.17	0.00	0.00
Total	1,172,729,033	741,403	104,335	741,453	104,342	2,881	643	(2,238)	2,569	
Max	76,726,413	48,510	6,827	43,160	13,459	379	36	5	1,776	

Table 15 All *T. dasycarpum* sex-specific k-mer containing reads mapped to the male T. dioicum genome assembly contigs. Msk- and fsk-reads mapping to multiple locations were mapped randomly to the male T. dioicum genome assembly. The table is ordered by the assembly contig name.

T. dioicum male assembly Contig		Expected Number of Sequences		Observed Number of Sequences		Contig Sex-Specific Read Enrichment		Enrichment Difference	Enrichment Ratio
Name	Ref Seq Length	Female	Male	Female	Male	Female	Male	Female-Male	Male/Female
ptg000001l	61382776	880.65	2987.21	948	2968	1.08	0.99	0.08	0.92
ptg000002l	927027	13.30	45.11	7	113	0.53	2.50	(1.98)	4.76
ptg000003l	45649582	654.93	2221.55	825	7718	1.26	3.47	(2.21)	2.76
ptg000004l	34209567	490.80	1664.82	369	958	0.75	0.58	0.18	0.77
ptg000005l	5417901	77.73	263.66	104	564	1.34	2.14	(0.80)	1.60
ptg000006l	10696020	153.46	520.52	112	301	0.73	0.58	0.15	0.79
ptg000007l	15800358	226.69	768.93	225	549	0.99	0.71	0.28	0.72
ptg000008l	24015344	344.55	1168.71	459	882	1.33	0.75	0.58	0.57
ptg000009l	31118052	446.45	1514.37	482	1252	1.08	0.83	0.25	0.77
ptg000010l	33073770	474.51	1609.54	445	1703	0.94	1.06	(0.12)	1.13
ptg000011l	30380337	435.86	1478.47	325	958	0.75	0.65	0.10	0.87
ptg000012l	29237344	419.47	1422.84	327	862	0.78	0.61	0.17	0.78
ptg000013l	31995408	459.04	1557.07	371	1582	0.81	1.02	(0.21)	1.26
ptg000014l	12763841	183.12	621.16	174	793	0.95	1.28	(0.33)	1.34
ptg000015l	8363160	119.99	407.00	45	373	0.38	0.92	(0.54)	2.44
ptg000016l	12080268	173.31	587.89	185	1284	1.07	2.18	(1.12)	2.05
ptg000018l	44476517	638.10	2164.46	1092	1808	1.71	0.84	0.88	0.49
ptg000019l	76726413	1100.79	3733.91	819	1979	0.74	0.53	0.21	0.71
ptg000020l	5277996	75.72	256.86	93	188	1.23	0.73	0.50	0.60
ptg000021l	36227159	519.75	1763.00	318	949	0.61	0.54	0.07	0.88
ptg000022l	36900733	529.41	1795.78	537	1077	1.01	0.60	0.41	0.59
ptg000023l	4408031	63.24	214.52	61	174	0.96	0.81	0.15	0.84
ptg000024l	17852609	256.13	868.80	214	480	0.84	0.55	0.28	0.66
ptg000025l	9746350	139.83	474.31	207	413	1.48	0.87	0.61	0.59
ptg000026l	30814064	442.09	1499.57	494	803	1.12	0.54	0.58	0.48
ptg000027l	6415627	92.04	312.22	86	112	0.93	0.36	0.58	0.38
ptg000028l	18597651	266.82	905.06	210	475	0.79	0.52	0.26	0.67

ptg000029l	14286320	204.96	695.25	215	401	1.05	0.58	0.47	0.55
ptg000030l	9549700	137.01	464.74	79	280	0.58	0.60	(0.03)	1.04
ptg000031l	10022985	143.80	487.77	178	419	1.24	0.86	0.38	0.69
ptg000032l	10054820	144.26	489.32	69	160	0.48	0.33	0.15	0.68
ptg000033l	36043935	517.12	1754.09	479	1060	0.93	0.60	0.32	0.65
ptg000034l	1015632	14.57	49.43	2	0	0.14	0.00	0.14	n/a
ptg000035l	36127415	518.32	1758.15	351	1138	0.68	0.65	0.03	0.96
ptg000036l	34095065	489.16	1659.25	462	1471	0.94	0.89	0.06	0.94
ptg000037l	10310964	147.93	501.79	124	880	0.84	1.75	(0.92)	2.09
ptg000038l	7548890	108.30	367.37	90	210	0.83	0.57	0.26	0.69
ptg000040l	8911710	127.86	433.69	106	280	0.83	0.65	0.18	0.78
ptg000041l	30393699	436.06	1479.12	358	1386	0.82	0.94	(0.12)	1.14
ptg000042l	4512657	64.74	219.61	106	196	1.64	0.89	0.74	0.55
ptg000043l	8150080	116.93	396.63	75	831	0.64	2.10	(1.45)	3.27
ptg000044l	4551151	65.30	221.48	77	256	1.18	1.16	0.02	0.98
ptg000046l	3375395	48.43	164.26	56	78	1.16	0.47	0.68	0.41
ptg000047l	16521110	237.03	804.00	326	444	1.38	0.55	0.82	0.40
ptg000048l	7754013	111.25	377.35	58	138	0.52	0.37	0.16	0.70
ptg000049l	3908735	56.08	190.22	39	77	0.70	0.40	0.29	0.58
ptg000051l	8626767	123.77	419.82	93	208	0.75	0.50	0.26	0.66
ptg000052l	3780431	54.24	183.98	125	177	2.30	0.96	1.34	0.42
ptg000053l	1324218	19.00	64.44	13	13	0.68	0.20	0.48	0.29
ptg000054l	10183987	146.11	495.61	162	385	1.11	0.78	0.33	0.70
ptg000056l	6199021	88.94	301.68	74	357	0.83	1.18	(0.35)	1.42
ptg000057l	2258311	32.40	109.90	35	23	1.08	0.21	0.87	0.19
ptg000058l	15612692	223.99	759.80	164	619	0.73	0.81	(0.08)	1.11
ptg000059l	4107956	58.94	199.91	16	122	0.27	0.61	(0.34)	2.25
ptg000060l	2673336	38.35	130.10	0	302	0.00	2.32	(2.32)	n/a
ptg000061l	11754656	168.64	572.04	155	484	0.92	0.85	0.07	0.92
ptg000062l	19808607	284.19	963.99	217	429	0.76	0.45	0.32	0.58
ptg000063l	29033338	416.54	1412.92	403	1050	0.97	0.74	0.22	0.77
ptg000065l	3225988	46.28	156.99	19	81	0.41	0.52	(0.11)	1.26

ptg000066l	530107	7.61	25.80	18	7	2.37	0.27	2.10	0.11
ptg000069l	164882	2.37	8.02	4	5	1.69	0.62	1.07	0.37
ptg000070l	9167258	131.52	446.13	202	371	1.54	0.83	0.70	0.54
ptg000071l	1817328	26.07	88.44	13	26	0.50	0.29	0.20	0.59
ptg000073l	192046	2.76	9.35	7	19	2.54	2.03	0.51	0.80
ptg000074l	5925831	85.02	288.38	99	131	1.16	0.45	0.71	0.39
ptg000075l	3949091	56.66	192.18	30	155	0.53	0.81	(0.28)	1.52
ptg000078l	1575749	22.61	76.68	7	14	0.31	0.18	0.13	0.59
ptg000080l	821389	11.78	39.97	106	96	8.99	2.40	6.59	0.27
ptg000081l	3577722	51.33	174.11	77	140	1.50	0.80	0.70	0.54
ptg000082l	2252127	32.31	109.60	30	85	0.93	0.78	0.15	0.84
ptg000084l	2739728	39.31	133.33	10	62	0.25	0.47	(0.21)	1.83
ptg000085l	3906051	56.04	190.09	242	417	4.32	2.19	2.12	0.51
ptg000086l	4153077	59.58	202.11	25	132	0.42	0.65	(0.23)	1.56
ptg000087l	1236065	17.73	60.15	1	22	0.06	0.37	(0.31)	6.49
ptg000090l	560204	8.04	27.26	7	12	0.87	0.44	0.43	0.51
ptg000097l	1453588	20.85	70.74	76	79	3.64	1.12	2.53	0.31
ptg000098l	1148870	16.48	55.91	30	51	1.82	0.91	0.91	0.50
ptg000100l	30352	0.44	1.48	0	1	0.00	0.68	(0.68)	n/a
ptg000101l	3292698	47.24	160.24	27	639	0.57	3.99	(3.42)	6.98
ptg000108l	2476811	35.53	120.53	85	150	2.39	1.24	1.15	0.52
ptg000110l	1432496	20.55	69.71	124	80	6.03	1.15	4.89	0.19
ptg000115l	727267	10.43	35.39	17	14	1.63	0.40	1.23	0.24
ptg000117l	230628	3.31	11.22	0	3	0.00	0.27	(0.27)	n/a
ptg000123l	127969	1.84	6.23	1	0	0.54	0.00	0.54	n/a
ptg000125c	134426	1.93	6.54	0	12	0.00	1.83	(1.83)	n/a
ptg000129l	1103165	15.83	53.69	49	153	3.10	2.85	0.25	0.92
ptg000130l	429041	6.16	20.88	25	77	4.06	3.69	0.37	0.91
ptg000131l	1186096	17.02	57.72	0	1	0.00	0.02	(0.02)	n/a
ptg000133l	38414622	551.13	1869.46	882	7402	1.60	3.96	(2.36)	2.47
ptg000134l	240433	3.45	11.70	2	1	0.58	0.09	0.49	0.15
ptg000138l	146626	2.10	7.14	1	1	0.48	0.14	0.34	0.29

ptg000140l	468248	6.72	22.79	3	12	0.45	0.53	(0.08)	1.18
ptg000146c	215847	3.10	10.50	8	48	2.58	4.57	(1.99)	1.77
ptg000149l	83013	1.19	4.04	0	1	0.00	0.25	(0.25)	n/a
ptg000159c	99123	1.42	4.82	0	12	0.00	2.49	(2.49)	n/a
ptg000166l	65517	0.94	3.19	1	0	1.06	0.00	1.06	n/a
ptg000171l	188253	2.70	9.16	6	11	2.22	1.20	1.02	0.54
ptg000172l	386802	5.55	18.82	0	45	0.00	2.39	(2.39)	n/a
ptg000193l	272731	3.91	13.27	0	1	0.00	0.08	(0.08)	n/a
ptg000195l	89081	1.28	4.34	0	1	0.00	0.23	(0.23)	n/a
ptg000219c	167626	2.40	8.16	0	2	0.00	0.25	(0.25)	n/a
ptg000235l	33729	0.48	1.64	1	0	2.07	0.00	2.07	n/a
ptg000240l	301786	4.33	14.69	0	25	0.00	1.70	(1.70)	n/a
ptg000250l	228542	3.28	11.12	1	0	0.30	0.00	0.30	n/a
ptg000255l	22933	0.33	1.12	0	1	0.00	0.90	(0.90)	n/a
ptg000262l	194521	2.79	9.47	1	1	0.36	0.11	0.25	0.29
ptg000291l	33885	0.49	1.65	0	10	0.00	6.06	(6.06)	n/a
ptg000308l	35094	0.50	1.71	1	0	1.99	0.00	1.99	n/a
ptg000310l	103709	1.49	5.05	0	3	0.00	0.59	(0.59)	n/a
ptg000329l	90077	1.29	4.38	0	1	0.00	0.23	(0.23)	n/a
ptg000330l	103685	1.49	5.05	0	4	0.00	0.79	(0.79)	n/a
ptg000333l	56112	0.81	2.73	1	0	1.24	0.00	1.24	n/a
ptg000381l	46818	0.67	2.28	0	2	0.00	0.88	(0.88)	n/a
ptg000399l	49499	0.71	2.41	0	2	0.00	0.83	(0.83)	n/a
ptg000403l	90262	1.29	4.39	0	6	0.00	1.37	(1.37)	n/a
ptg000416l	91580	1.31	4.46	9	13	6.85	2.92	3.93	0.43
ptg000426l	112968	1.62	5.50	0	12	0.00	2.18	(2.18)	n/a
ptg000429l	50886	0.73	2.48	0	5	0.00	2.02	(2.02)	n/a
ptg000439l	22766	0.33	1.11	1	0	3.06	0.00	3.06	n/a
ptg000508l	40716	0.58	1.98	0	1	0.00	0.50	(0.50)	n/a
ptg000522l	38463	0.55	1.87	0	5	0.00	2.67	(2.67)	n/a
ptg000528l	63676	0.91	3.10	0	3	0.00	0.97	(0.97)	n/a
ptg000531l	33771	0.48	1.64	0	1	0.00	0.61	(0.61)	n/a

ptg000546l	23861	0.34	1.16	1	0	2.92	0.00	2.92	n/a
ptg000559l	26553	0.38	1.29	0	2	0.00	1.55	(1.55)	n/a
ptg000578l	46255	0.66	2.25	0	55	0.00	24.43	(24.43)	n/a
ptg000605l	15065	0.22	0.73	0	9	0.00	12.28	(12.28)	n/a
ptg000611l	18857	0.27	0.92	0	2	0.00	2.18	(2.18)	n/a
ptg000612l	10210	0.15	0.50	0	11	0.00	22.14	(22.14)	n/a