

Sex is determined by XY chromosomes across the radiation of dioecious *Nepenthes* pitcher plants

Mathias Scharmann,^{1,2,3}  T. Ulmar Grafe,⁴  Faizah Metali,⁴  and Alex Widmer¹ 

¹Institute of Integrative Biology, ETH Zurich, Zürich 8092, Switzerland

²Department of Ecology and Evolution, University of Lausanne, Lausanne 1015, Switzerland

³E-mail: mathias.scharmann@env.ethz.ch

⁴Faculty of Science, Universiti Brunei Darussalam, Gadong BE 1410, Brunei Darussalam

Received October 24, 2018

Accepted September 3, 2019

Species with separate sexes (dioecy) are a minority among flowering plants, but dioecy has evolved multiple times independently in their history. The sex-determination system and sex-linked genomic regions are currently identified in a limited number of dioecious plants only. Here, we study the sex-determination system in a genus of dioecious plants that lack heteromorphic sex chromosomes and are not amenable to controlled breeding: *Nepenthes* pitcher plants. We genotyped wild populations of flowering males and females of three *Nepenthes* taxa using ddRAD-seq and sequenced a male inflorescence transcriptome. We developed a statistical tool (privacy rarefaction) to distinguish true sex specificity from stochastic noise in read coverage of sequencing data from wild populations and identified male-specific loci and XY-patterned single nucleotide polymorphisms (SNPs) in all three *Nepenthes* taxa, suggesting the presence of homomorphic XY sex chromosomes. The male-specific region of the Y chromosome showed little conservation among the three taxa, except for the essential pollen development gene *DYT1* that was confirmed as male specific by PCR in additional *Nepenthes* taxa. Hence, dioecy and part of the male-specific region of the *Nepenthes* Y-chromosomes likely have a single evolutionary origin.

KEY WORDS: Carnivorous plant, dioecy, molecular sexing, plant sex chromosome, privacy rarefaction, sex-determination, sex-specific loci.

Impact Summary

One of the most striking polymorphisms observed in organismal populations is the existence of male and female individuals. In contrast to animals, where this condition is common, plants are usually functional hermaphrodites. Some plants, however, are dioecious, that is, individuals are either of male or female sex. Dioecy has evolved hundreds of times independently in plants, which offers the potential for comparative studies of sex chromosome evolution and for investigating the genetic basis of transitions between hermaphroditism

and dioecy (Charlesworth 2015). Yet empirical data to test hypotheses about why some species are dioecious and others hermaphroditic, and how such transitions are achieved, are lacking. Despite their potential, the sex-determination mechanisms of most dioecious plants are not known, and few new species have been investigated since the seminal review by Westergaard (1958). Also, markers for molecular sexing have important applications in agriculture, horticulture, and conservation. Historically, the identification of sex-determination systems was limited to species that can be bred in controlled

conditions or have heteromorphic sex chromosomes. This is now changing with genome-scale sequencing technology. Here, we investigated the sex-determination system of carnivorous pitcher plants in the genus *Nepenthes*. We surveyed wild populations of three species by genotyping a large number of loci throughout their genomes. Because such data are noisy, we propose a solution to the common problem of distinguishing true signal from noise in presence–absence data by generating null distributions through permutations of the observed data. We discovered loci that occur only in males and reveal an XY sex chromosome system. One gene on the *Nepenthes* Y chromosome is particularly interesting, because it is presumably essential for pollen development and present only in males, and thus can be used to diagnose the sex of nonflowering plants

Although the majority of flowering plant species are functional hermaphrodites, plant sexual systems and sex-determination mechanisms are highly diverse (Charlesworth 2002; Bachtrog et al. 2014). Only 5–6% of species have female and male flowers on separate individuals (dioecy), but the evolutionary transition to dioecy may have occurred as many as 800 times independently in angiosperms (Renner 2014). In contrast to outcrossing–selfing transitions due to loss of self-incompatibility, for some of which the underlying genetic changes have recently been uncovered (e.g., Shimizu and Tsuchimatsu 2015), relatively little is known about the genes involved in transitions from hermaphroditism to dioecy and in sex determination in plants (Charlesworth 2016), although sex-determining genes have been identified in three dioecious plant species: persimmon (*Diospyros lotus*, Akagi et al. 2014), *Asparagus officinalis* (Harkess et al. 2017; Murase et al. 2017), and kiwifruit (*Actinidia*, Akagi et al. 2018). The main hypotheses for the evolution of separate sexes in plants involve a combination of trade-offs between the sex functions, plus disadvantage of inbreeding (Charlesworth and Charlesworth 1978).

Many dioecious plants have genetic sex determination, which may involve sex chromosomes. Sex chromosomes differ from autosomes by having suppressed meiotic recombination around the sex-determining genes. These form a fully sex-linked chromosomal region whose transmission is limited to one sex. When this sex is male, the system is referred to as male heterogamety (male genotype XY, female XX), and female heterogamety when the fully sex-linked region is transmitted via females (male ZZ, female ZW). The sex-specific, fully nonrecombining regions (male-specific region of the Y, termed MSY, or female-specific

region of the W) show a number of special properties. First, an MSY may contain sequences that are absent from the X, and thus male specific (Y-hemizygous, transmitted only from fathers to sons). As X and Y chromosomes are thought to evolve from a pair of autosomes, the gain of male-specific sequences can be explained by several mechanisms, such as the rise of a new male-determining mutation, or the translocation of a male-determining cassette (Tennesen et al. 2018), or the localized expansion of repetitive sequences due to the lack of recombination. Second, over evolutionary time, the MSY may undergo genetic degeneration and lose functional genes that were initially shared with the X chromosome (Bachtrog 2013). Sex chromosomes have evolved independently many times in plants, and will therefore probably have diverse ages and levels of degeneration. Heteromorphic sex chromosomes have diverged sufficiently in size or structure to be distinguished optically with a microscope, whereas homomorphic sex chromosomes may have more subtle differences that can be detected only by molecular genetic methods. Despite their great potential for comparative studies, few plant sex chromosomes have been studied in detail (Ming et al. 2011; Harkess and Leebens-Mack 2017; Muyle et al. 2017). Knowledge of sex-determination systems and the identification of fully sex-linked genetic markers are important for molecular sexing of juveniles or nonflowering adults in agriculture, breeding, and conservation.

Cytogenetics and linkage analysis in families are established methods to study sex determination and discover sex linkage of genes (Charlesworth and Mank 2010). However, these strategies fail in many dioecious organisms because their karyotypes are homomorphic (Filatov 2015), or because controlled breeding is difficult, since many dioecious plants are woody and reproduce only after many years (Renner and Ricklefs 1995). Several next-generation sequencing techniques have now greatly increased knowledge about sex-linked genes (reviewed by Muyle et al. 2017). However, they require either prior knowledge of heterogamety, controlled breeding, or whole-genome sequencing, which remains expensive and time consuming. An alternative class of strategies uses population polymorphism to infer sex linkage of loci (reviewed by Muyle et al. 2017), even without pedigrees. These strategies can potentially allow sex-linked regions to be discovered by cheaper reduced-representation sequencing (RRS) methods such as RAD-seq (Baird et al. 2008; Elshire et al. 2011; Peterson et al. 2012), although the gained information will remain incomplete, because typically only a few percent of a genome is covered. Nevertheless, the discovery of sex-linked markers by RRS has been successful in organisms such as Crustaceans (Carmichael et al. 2013), *Anolis* lizards (Gamble and Zarkower 2014), geckos (Gamble et al. 2015a), and frogs (Brelsford et al. 2017; Jeffries et al. 2018).

A major problem faced by approaches that use population polymorphism to infer sex-linkage, and sex specificity in

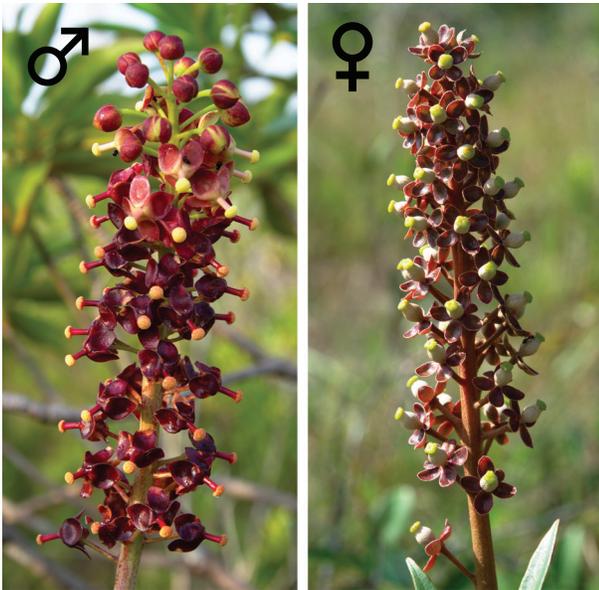


Figure 1. Sexual dimorphism in *Nepenthes* inflorescences. Left: male inflorescence of *N. rafflesiana* s.l. Right: Female inflorescence of *N. mirabilis* var. *globosa*. Photos: M. Scharmann

particular, is error in the measurement of locus presence and absence (Text S4). Presence–absence error has long been recognized as a problem in fragment length genotyping methods, but it is exacerbated in RRS data, in which missing loci occur in a highly stochastic manner (Mastretta-Yanes et al. 2015; Bresadola et al. 2019), and can make sex-specific sequences appear in both sexes (Bewick et al. 2013; Gamble and Zarkower 2014; Heikrujam et al. 2015; Breilsford et al. 2017), and probably represent false positive results. One suggested solution to reduce the number of false positives is to compare increasing numbers of males and females (Gamble and Zarkower 2014; Gamble et al. 2015b). Unfortunately, in RAD data the number of shared loci decreases with increasing sample numbers (Mastretta-Yanes et al. 2015). Beyond a number that is unpredictable and specific to each dataset, true sex-specific loci may be missed because they are too rarely sequenced.

We developed a statistical procedure to deal with this problem, and applied it to characterize the sex-determination system of *Nepenthes* pitcher plants. *Nepenthes* (Nepenthaceae, Caryophyllales) includes at least 160 species of perennial vines and shrubs occurring mostly in Southeast Asia (Cheek and Jebb 2001; Clarke et al. 2018). They are carnivorous plants that supplement their nutrition by killing and digesting animals in their modified pitcher leaves (Juniper et al. 1989; Moran and Clarke 2010; Pavlovič and Saganová 2015). All *Nepenthes* are dioecious, whereas close relatives (families Ancistrocladaceae, Dioncophyllaceae, Droseraceae, and Drosophyllaceae; Cuénoud et al. 2002; Renner and Specht 2011; Walker et al. 2017) are hermaphroditic. The individual male and female flowers (Fig. 1) are readily recognized be-

cause reproductive organs of the other sex abort early in development (Subramanyam and Narayana 1971). We hypothesized that sex in *Nepenthes* has a genetic basis, or is determined during early life stages, because there are no reports of sexual plasticity (Clarke 2001), or functional hermaphroditism. *Nepenthes* karyotypes ($2n = 80$, Heubl and Wistuba 1997) do not suggest heteromorphic sex chromosomes.

Here, we investigated the previously undescribed sex-determination system of multiple *Nepenthes* species. Because controlled breeding of these slow growing plants faces many challenges, we sampled wild populations. We used *Silene latifolia* to test our method, as this species has well-studied heteromorphic sex chromosomes. Specifically, we asked the following questions: (1) Are there sex-linked loci in *Nepenthes*? (2) Are the same sex-linked loci shared among different *Nepenthes* species? (3) Which expressed genes are sex linked? We discovered fully male-specific and XY-patterned loci and developed a molecular sexing assay for *Nepenthes*. The identified markers include two candidate sex-determination genes, and these suggest that part of the Y chromosome is ancestral in this genus.

Methods

SAMPLING, ddRAD-seq, AND GENOTYPING

Natural populations of *Nepenthes* were sampled in Brunei Darussalam (Borneo), Singapore, and the Seychelles. Fresh leaf material was stored in a nucleic acid preserving buffer (Camacho-Sanchez et al. 2013). The sexes of *Nepenthes* plants were recorded from fresh or dry inflorescences. Scans for sex-linked loci were conducted separately on three *Nepenthes* taxa: *N. pervillei* Blume, *N. gracilis* Korth., and *N. rafflesiana* sensu lato (Table 1). We extracted DNA from leaves using silica column kits (Nucleospin Plant II; Macherey Nagel, Düren, Germany) and prepared sequencing libraries following the ddRAD-seq protocol (Peterson et al. 2012) using the enzymes *EcoRI* and *TaqI*. Library pools (84 or 96-plex) were sequenced on an Illumina HiSeq 2500. Bioinformatic data filtering, *de novo* assembly of reference contigs (“RAD-tags,” very short contigs with a mean length of c. 96 bases), read mapping, genotype calling, and quality filtering followed a modified dDocent pipeline (Puritz et al. 2014) and code is deposited at <https://github.com/mscharmann>.

The exploration of sex-specific markers in *N. rafflesiana* s.l. followed an iterative strategy with two rounds of sexing, genotyping, bioinformatic analysis, and PCR validation (Text S1). To increase the phylogenetic range of our study and to validate molecular sexing, we also included individuals of known sex for additional species (Table S2-1 in Text S2). To validate our method for detection of sex-specific loci, we also genotyped populations of a species with a well-known, heteromorphic XY sex-determination

Table 1. Sample sizes and origin for the taxa sequenced in this study.

Taxon	Sampling location	Number of males	Number of females	Sequencing method
<i>Nepenthes pervillei</i> Blume	Seychelles, Mahé	28	22	ddRAD-seq (Peterson et al. 2012)
<i>Nepenthes gracilis</i> Korth.	Brunei Darussalam, Borneo	10	10	ddRAD-seq (Peterson et al. 2012)
<i>Nepenthes rafflesiana</i> sensu lato, here defined as:		39	22	ddRAD-seq (Peterson et al. 2012)
<i>Nepenthes rafflesiana</i> “typical form” (Clarke 1992, 1997)	Brunei Darussalam, Borneo	13	7	
<i>Nepenthes rafflesiana</i> “giant form” (Clarke 1992, 1997)	Brunei Darussalam, Borneo	5	3	
<i>Nepenthes rafflesiana</i> Jack	Singapore	10	4	
<i>Nepenthes hemsleyana</i> Macfarl.	Brunei Darussalam, Borneo	11	8	
<i>Nepenthes khasiana</i> Hook.f.	cultivated/artificially prop.	1	–	RNA-seq
<i>Silene latifolia</i> Poiret	Switzerland	27	32	GBS (Elshire et al. 2011)

system, *S. latifolia* Poiret, using a single-digest GBS protocol (Elshire et al. 2011). Details of the *Silene* samples and genotyping are provided in Text S3.

DETECTION OF SEX-LINKAGE

We distinguished between sex-linked loci showing sex-specificity, meaning without homology between the two sex chromosomes (e.g., Y-linked loci whose fully X-linked copy is absent or undetectable by our methods), and loci that are present on both sex chromosomes but whose allele frequencies diverged between the sex chromosomes, called ZW- or XY-patterned variants (Gammerdinger and Kocher 2018). We define sex-specific loci as contigs to which sequencing reads can be aligned from only one of the sexes (presence–absence polymorphism). However, this need not reflect true presence–absence, because observed absence may be due to methodological factors (deliberate or random) specific to each dataset. Imperfect detection may be caused by many underlying factors, including the number of male and female individuals investigated, their genetic relatedness, the species’ genome size and structure, the library preparation method, sequencing depth, and bioinformatic processing. Consider, for example, a particular genetic marker scored “present” in nine out of ten males and “absent” in five out of five females. True complete male specificity cannot be distinguished from technical artifacts due to random variation, for example, in sequencing depth or sampling bias. To improve the inference of sex specificity and facilitate comparisons of datasets from different species, sampling schemes, and sequencing runs, we used a resampling strategy that evaluates the biological signal among dataset-specific artifacts and uncertainties. The procedure is implemented in python and named “privacy rarefaction” (<https://github.com/mschermann>). Further details, including a performance analysis with simulations, are presented in Texts S4 and S5.

We tested for completely or partially XY- or ZW-patterned SNPs, that is, ones with different allele and heterozygote frequencies in the two sexes, using genotypes and PLINK version 1.07 (Purcell et al. 2007). Associations between biallelic SNPs and sex were analyzed by chi-squared tests, and candidate SNPs were accepted as sex associated at a false discovery rate of 5% (Benjamini and Hochberg 1995), and then classified as XY-patterned if males were predominantly heterozygous, or as ZW-patterned if females were predominantly heterozygous. To perform this test on a reasonable number of SNPs, we allowed data from up to 25% of individuals to be missing per SNP.

POPULATION GENETICS OF CANDIDATE SEX-LINKED LOCI

We tested whether linkage disequilibrium (LD), that is, the non-random association of alleles at separate loci, differed in male samples between sex-linked regions and the genomic background (represented by 15,000 randomly selected pairs of nonsex-linked SNPs). Stringent quality filters were applied: singleton SNPs were removed, SNPs in male-specific contigs were excluded if any heterozygous genotypes were called in the males (because Y-hemizygosity implies that heterozygosity is not true, so these were probably paralogous sequences), and for non-sex-linked contigs any excessively heterozygous SNPs (Hardy–Weinberg test with significance level 5%) were excluded. Excessively heterozygous SNPs were retained for XY-patterned contigs. LD (r^2) was calculated exclusively for SNPs from different contigs using PLINK. The same contrasts were made for nucleotide diversity π , which was averaged per contig for all SNP-containing contigs (including singleton SNPs) in VCFtools version 0.1.15 (Danecek et al. 2011) and for contigs without SNPs taken directly from bam alignments. The same filters were applied to both data (minimum read depth 3, maximum read depth 75, and minimum population presence

0.75). The significance of differences of the means was evaluated by randomization (10,000 rounds of re-sampling without replacement from the two groups, randomizing group membership).

COMPARISON OF CANDIDATE LOCI TO A MALE INFLORESCENCE TRANSCRIPTOME

We sequenced and assembled the transcriptome of a single developing male inflorescence of *Nepenthes khasiana* Hook.f. (Text S6) to identify and annotate sex-linked candidate loci. Fresh inflorescences of the species used for ddRAD-seq were not available in cultivation. The transcriptome was searched (a) by BLAST for similarity to candidate contigs (thresholds ≥ 90 aligned bases and $\geq 75\%$ identity) and (b) by repeating privacy rarefaction with ddRAD-seq reads directly mapped to the transcriptome rather than the RAD-tag reference (bwa mem; Li 2013; retaining multiple mappings).

Candidate transcripts from both approaches were annotated by BLAST search against the NCBI Genbank nucleotide collection (November 7, 2016 version) and the nonredundant protein collection (March 26, 2016 version). Transposable elements (TEs) were detected using RepeatMasker 4.0.6 (Smit et al. 2013) version 20150807 (eukaryota). Proteins with at least 50 amino acids were predicted by TransDecoder (Trinity package) and annotated against nr, UniProt Swiss-Prot (August 17, 2016 version), and *Arabidopsis thaliana* proteins in UniProtKB (April 3, 2016 version). PFAM domains were detected using hmmer 3.1b1 (Eddy et al. 2016), accepting hits at e -value $\leq 10^{-5}$.

PCR VALIDATION

Candidate sex-specific contigs were chosen for PCR validation based on a ranking of the privacy rarefaction results (using the highest stringency level reached and the bootstrap support), taxonomic overlap, and the quality of annotation of matching transcripts. PCR primers were designed in Geneious R6 (Biomatters Ltd., Auckland, New Zealand), and tested according to the protocol described in Text S2.

Results

SEX-LINKED LOCI

We first searched for sex-specific contigs in the illustrative example of *S. latifolia* GBS data using privacy rarefaction. When small numbers of individuals of each sex were analyzed, the procedure yielded similar numbers of male- and female-specific candidates (Table S8; Fig. 2, top, dark gray zone), which decreased monotonically with greater numbers of individuals, as expected. When the numbers of males and females analyzed were increased, a clear signal of male heterogamety emerged. With four or more individuals of each sex analyzed, the proportion of male-specific candidate loci increased and finally became significantly greater

than the number of female-specific candidates (Fig. 2, top, light gray zone). At 11 and more individuals of each sex analyzed, the number of female-specific candidates dropped to zero, whereas the number of male-specific candidates remained high (Fig. 2, top, white zone). Hence, these curves correctly diagnosed an XY-system for *S. latifolia*, and rejected a ZW-system. Due to this characteristic drop-out of false-positives, we also refer to the numbers of analyzed individuals of each sex as “privacy rarefaction stringency.” Some of the herein identified male-specific *S. latifolia* contigs were previously reported to be sex-linked (Text S3, Table S11).

Qualitatively consistent signatures of male-specific contigs were detected independently in *N. pervillei*, *N. gracilis*, and *N. rafflesiana s.l.* (Fig. 2, Table S8). The proportion of male-specific loci among all loci was, however, about 10-fold lower in *Nepenthes* than *S. latifolia*. Estimates based on subsamples of 10 males and 10 females were only 0.02% for *N. pervillei* (11.8/51,002), 0.11% for *N. rafflesiana s.l.* (43.7/40,508.7), and 0.06% for *N. gracilis* (13/22,789), versus 1.52% for *S. latifolia* (586.4/38,455).

In *N. pervillei* and *N. rafflesiana s.l.*, as well as in *S. latifolia*, we also detected XY-patterned SNPs, but not in *N. gracilis*, while none of the species yielded any ZW-patterned SNPs. The XY-patterned SNPs for *S. latifolia* recovered 289 contigs already known to be sex linked in that species, which, however, represent only a small fraction of the known and theoretically expected *S. latifolia* sex-linked sequences (c. 1/7 of the genome; Text S3, Table S11). This low power was expected for a sequencing strategy that targets only a small subset of the genome. Almost all XY-patterned SNPs had an allele frequency close to 0.5 and near-complete heterozygosity in males, but were homozygous in females (Table S9). The proportions of XY-patterned SNPs were also much lower in *Nepenthes* than in *S. latifolia* (2376/149,311 = 1.6%, similar to the proportion of male-specific loci): the estimates were, respectively, 0.25% and 0.017% for *N. pervillei* (97/38,783) and *N. rafflesiana s.l.* (37/222,188).

POPULATION GENETICS OF SEX-LINKED LOCI

Fully Y-linked loci experience no recombination, which should lead to increased population LD between different male-specific contigs. All three testable male-specific contigs of *N. pervillei* showed perfect LD ($r^2 = 1$, which was c. 0.7 units higher than the genomic background; $N = 3$ SNP pairs, $P = 0.06$). Likewise, the male-specific contigs of *N. rafflesiana s.l.* (containing 82 SNP pairs) had an r^2 that was on average c. 0.4 above the genomic mean ($P < 10^{-5}$); the median value was complete LD ($r^2 = 1$). Among contigs with XY-patterned SNPs, r^2 was 0.15 higher than the background mean in *N. pervillei* ($P = 10^{-5}$, $N = 147$), but was no different from the mean in *N. rafflesiana s.l.* ($P = 0.56$, $N = 11$). Our observation of some low LD values between sex-specific contigs is likely due to the discreteness of

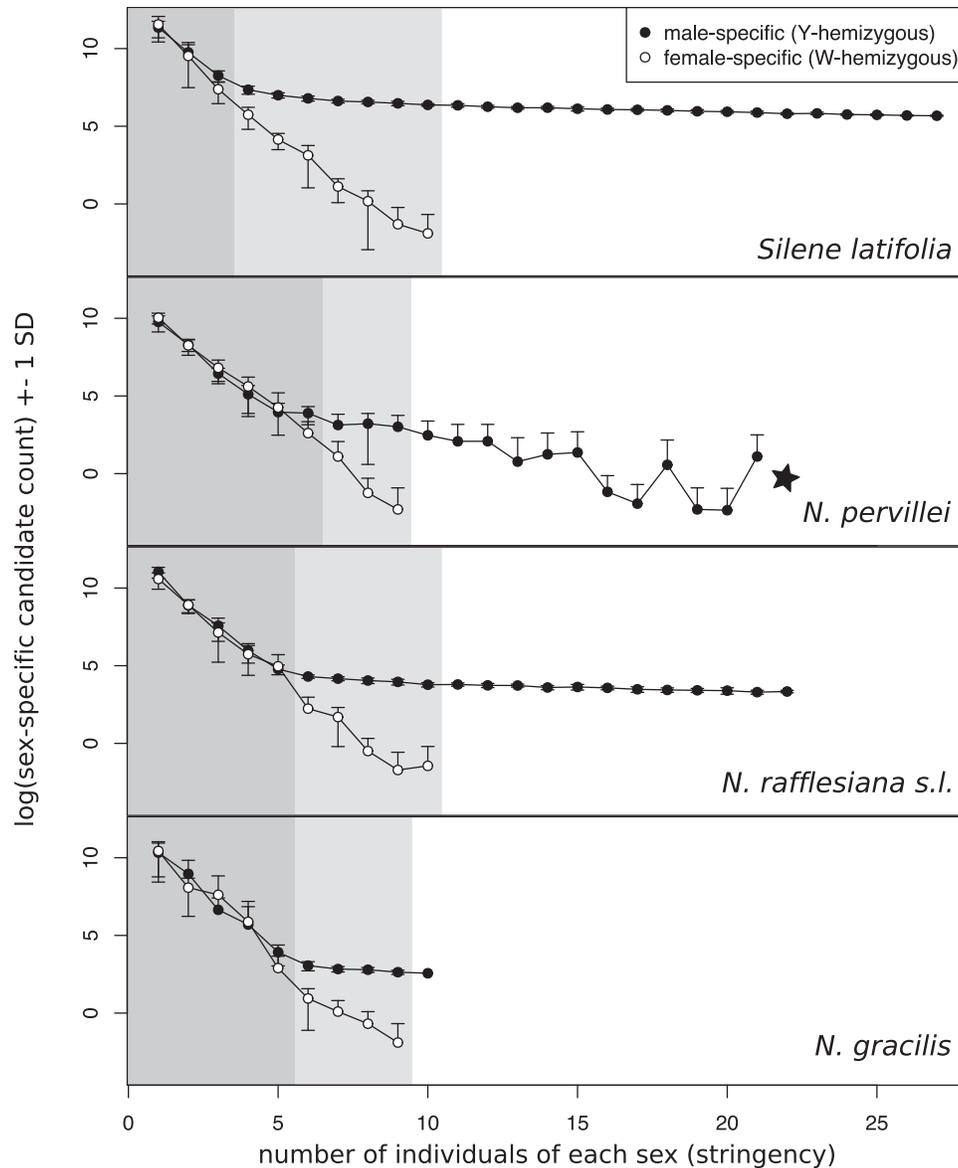


Figure 2. Evidence for male-specific loci and XY sex-determination systems in *Silene latifolia* and three *Nepenthes* spp. (privacy rarefaction curves). Shown are counts of sex-specific contigs (y-axis) as a function of the number of individuals of each sex sampled to score sex specificity (x-axis, stringency). Sex-specific contigs are defined as those to which sequencing reads from only one sex can be aligned. Dots represent averages, and whiskers one standard deviation of 200 bootstrapped combinations of males and females. Note natural log-scale of y-axis and hence undefined zero and negative values in the SD ranges. The background shading of the plots indicates three relevant zones that are directly informative on the sex-determination system: a dark gray zone (low stringency) indicates no difference between the sexes, the light gray zone (intermediate stringency) highlights where significant differences between sexes are found, and white background (highest stringency) shows the biologically plausible zone where sex-specific markers are obtained in only one sex. Male-specific candidates were found in all species up to the maximum possible stringency (the minimum number of male individuals and female individuals), except in *N. pervillei* (asterisk).

allele frequency estimates from small sample sizes, and the allele frequency dependence of LD metrics, whose maximum possible value is frequently much less than 1.0 (VanLiere and Rosenberg 2008).

The mean nucleotide diversity π in male-specific (putatively Y-linked) contigs tended to be lower than the genomic background

in all three taxa (Fig. 3), consistent with theoretical expectations for Y-specific loci, whose effective population size is only 1/4 of that of autosomal genes, and can be reduced much further by genetic hitchhiking and high variance among males in the number of sired offspring (Wilson Sayres et al. 2014). This difference was significant for *N. rafflesiana s.l.* ($P < 10^{-3}$), but not for *N. pervillei*

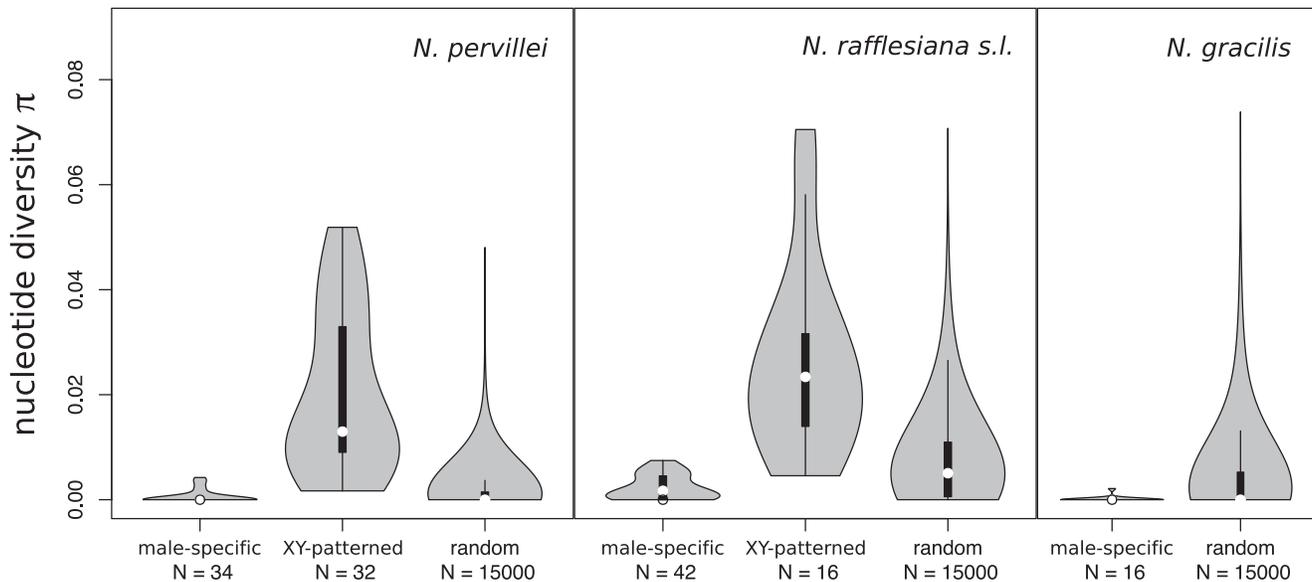


Figure 3. Mean per-site nucleotide diversity π of contigs in male *Nepenthes* of three taxa for male-specific, XY-patterned, and random nonsex-linked contigs. All contigs mapping 3–75 reads in $\geq 75\%$ of males per population were included. The same sets of individuals are considered in each category. No XY- or ZW-patterned contigs were found in *N. gracilis*. Median = white dot, box = 25–75% quartiles, whiskers = 1.5* interquartile range, violin = estimated kernel density.

or *N. gracilis* ($P = 0.13$ and $P = 0.055$, respectively). In contrast, mean π of males in contigs with XY-patterned SNPs was higher than the genomic mean for both *N. pervillei* ($P = 10^{-5}$, Fig. 3) and *N. rafflesiana s.l.* ($P = 0.002$; Fig. 3), a consequence of high male heterozygosity in the XY-patterned SNPs.

SHARED CANDIDATE LOCI BETWEEN SPECIES, FUNCTIONAL ANNOTATIONS, AND PCR VALIDATION

Six candidate sex-specific contigs were found at privacy rarefaction stringency level ≥ 5 that were shared between *N. gracilis* and *N. rafflesiana s.l.* No sex-specific candidates were shared between *N. pervillei* and the other species (Table S10). There was no overlap in XY-patterned SNPs between the *Nepenthes* species, and no direct overlap between male-specific contigs and XY-patterned ones. However, one male-specific contig of *N. gracilis* and one XY-patterned contig of *N. rafflesiana s.l.* both matched (full length alignment, e -value $\leq 1 \times 10^{-19}$) to the same inflorescence transcript containing a DUF4283 (domain of unknown function, <http://pfam.xfam.org/family/PF14111>, November 9, 2016).

One male-specific contig of *N. pervillei* aligned to the transcript of a bHLH transcription factor, and the best matches in all accessed databases were consistently to predicted orthologs of the *Arabidopsis* gene *DYSFUNCTIONAL TAPETUM1 (DYTI)*. A further XY-patterned contig of *N. pervillei* matched a transcript annotated as *A. thaliana SEPALLATA-1 (SEPI)*, which aligned to the predicted 3'-UTR of the putative *N. pervillei* *SEPI*-ortholog, and contained two SNPs that were both homozygous in 95% of females and heterozygous in 96% of males. No

estimate of *SEPI* X–Y divergence was possible because the male inflorescence transcriptome reads were not heterozygous. In *N. gracilis*, a male-specific contig matched a long transcript similar to a mitochondrial NADH-ubiquinone oxidoreductase from *Beta vulgaris* (Swiss-Prot). This finding was unexpected and may represent either an unspecific match of the short (96 bp) contig to the inflorescence transcript, or else a cyto-nuclear transfer to the sex chromosomes. The occurrence of organellar genes on plant sex chromosomes has been documented in other species (Steflova et al. 2013).

All other candidate loci either included traces of TEs, or had no known sequence motifs (Table S10). In particular, of the 38 sex-linked inflorescence transcripts identified here (by 41 matching sex-linked contigs), 34 (89%) could not be annotated, or contained TEs. TEs were commoner than in nonsex-linked transcripts ($\chi^2_{df=1} = 5.2$, $P = 0.02$). Nine out of 13 sex-linked TE-transcripts annotated as gypsy-like retrotransposons, a significant overrepresentation relative to nonsex-linked TE transcripts ($\chi^2_{df=1} = 8.85$, $P = 0.003$).

Complementary to the sex-specificity scan on the ddRAD-de novo reference, we repeated privacy rarefaction by directly mapping the ddRAD reads to the male inflorescence transcriptome to find further annotated sex-linked genes. This identified seven transcripts (Table S10) that map male-specific regions of the genomes. We recorded only high-confidence male-specific candidate transcripts present in at least four males and absent in at least four females, and bootstrap support greater 0.5 in at least one species. No female-specific transcripts (false positives)

reached these support levels. A single transcript was male-specific in *N. rafflesiana s.l.* but could not be annotated. Four close transcripts “isoforms” (contigs that share sequence but differ slightly in structure, as assembled by Trinity; Haas et al. 2013) were male specific in both *N. gracilis* and *N. rafflesiana s.l.*, but they lacked similarity to any known motif except for one isoform similar to a Jockey-1_Drh retrotransposon. However, two transcripts were male specific in both *N. pervillei* and *N. rafflesiana s.l.*, and one of these also matched a *N. pervillei* male-specific contig (see above). These two transcripts appear to be close isoforms (putative intron presence-absence), and both annotated as *DYT1* (see above).

We tested by PCR whether the putative *DYT1*-ortholog is male specific in a broad range of *Nepenthes* species. A single PCR product of approximately 290 bp length was observed exclusively and consistently in phenotypically sexed male *Nepenthes* but never in females (Text S2). Multiple males and females were screened in eight taxa, and 1–2 individuals from 14 further taxa. Presence-absence of the PCR product was fully consistent with the phenotypic sex of all 56 individuals. Sanger sequencing of the PCR product confirmed the identity of the target region. Hence, this locus is male specific across a phylogenetically broad range of *Nepenthes* species and can be used for molecular sexing.

PERFORMANCE OF THE RESAMPLING STRATEGY

We tested the performance of privacy rarefaction on simulated datasets resembling typical RAD-seq experiments under a range of missing data levels, sampling schemes, and sizes of the sex-specific region (Text S5). The procedure correctly identified the heterogametic sex in >90% of simulations when the proportion of sex-specific contigs in the genome was at least one permil, and virtually all contigs classified as sex specific were true positives. A naive scoring method, in contrast, failed to detect the heterogametic sex in most scenarios because sex-specific contigs appeared in both sexes, and it typically reported many false positives. However, most of the true sex-specific contigs were not detected in simulated RAD-seq data because of the missing data inherent to this sequencing method (low sensitivity). Nevertheless, the relative size of the sex-specific regions was usually estimated to the true order of magnitude (Table S5-1 in Text S5).

Discussion

THE NEPENTHES SEX-DETERMINATION SYSTEM

Our findings reveal that sex determination has a genetic basis in *Nepenthes* and involves a nonrecombining region in males. *Nepenthes* karyotypes suggest that the sex chromosomes are homomorphic (Heubl and Wistuba 1997), consistent with the lower proportions of sex-linked contigs in *Nepenthes* compared to *S. latifolia* with its large and heteromorphic Y-chromosome. The proportions of male-specific contigs allow us to hypothesize

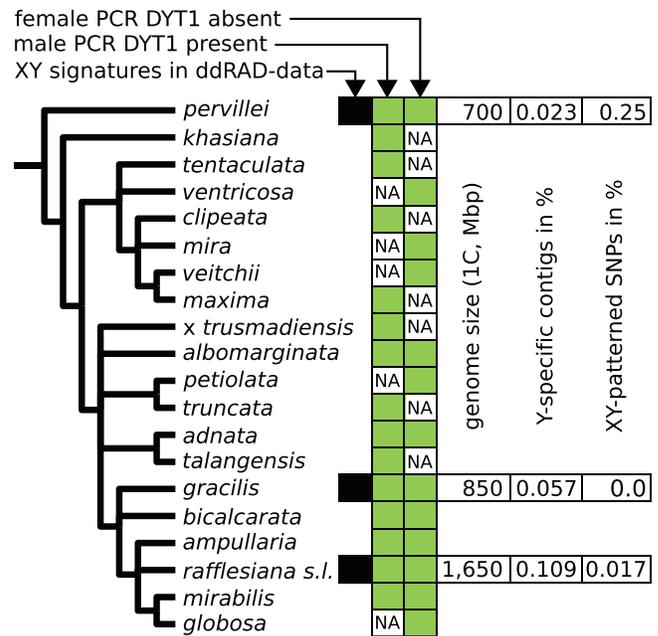


Figure 4. Summary of results on the sex-determination system for *Nepenthes*, annotated on a plastid phylogeny (after Meimberg and Heubl 2006). The crown of the genus is c. 17.7 (CI 11.0–24.3) million year old (Text S7). It constrains the minimum age at which dioecy evolved and *DYT1* became a male-specific gene. NA = not available/not tested. Genome sizes were quantified by flow cytometry. The proportion of Y-specific contigs is given at 10 individuals of each sex (stringency). *Nepenthes rafflesiana s.l.* contains several entities, for which the PCRs were conducted separately.

that the size of the MSY relative to the whole genome is about 10-fold smaller in *Nepenthes* than in *S. latifolia*, and that within the genus, it is smallest in *N. pervillei*. However, we note that the characterization of sex chromosomes via reduced-representation sequencing methods necessarily remains incomplete (Text S3), and very strict analyses, such as the resampling procedure we propose here, are required to avoid false inferences (Text S5). Furthermore, our sampling included several subpopulations in *S. latifolia* and *N. rafflesiana s.l.*, which may have impeded the detection of deme-specific sex-linked loci.

The MSY of *Nepenthes* appears to contain a “core region” that is conserved throughout the genus. The *DYT1* gene was male specific in both *N. pervillei* and *N. rafflesiana s.l.*, and part of it was consistently PCR amplified in known males but never in females of 22 *Nepenthes* species (Text S2), representing all major clades (Fig. 4; Mullins 2000; Meimberg et al. 2001; Meimberg and Heubl 2006; Scharmann et al. unpubl. data). The shared MSY locus therefore suggests a single origin of dioecy in *Nepenthes* that most likely predates the most recent common ancestor of extant *Nepenthes* at 17.7 (CI 11.0–24.3) million years ago but followed the split between Nepenthaceae and

hermaphroditic Droseraceae at least 44.2 million years ago (average 71.1 with CI 44.2–98.0; Text S7). However, the age of the shared MSY core does not necessarily reflect the age of the sex chromosomes: their identity could have changed over time and may also differ between *Nepenthes* species because the ancestral MSY could have been translocated to other chromosomes in a process called “sex-chromosome turnover” (Blaser et al. 2014; Jeffries et al. 2018; Tennessen et al. 2018). Nevertheless, the *Nepenthes* MSY core is probably older than the heteromorphic *S. latifolia* sex chromosomes (~11 million years, Krasovec et al. 2018). These alternatives can be explored in future comparative whole genome sequencing or mapping studies.

During the radiation of *Nepenthes*, the MSY has diverged between species, as is expected over such long divergence times, particularly for noncoding sequences. Only six out of 135 male-specific contigs were shared between *N. rafflesiana s.l.* and *N. gracilis*, and none were shared with the more distant *N. pervillei* (Fig. 4). Male-specific loci shared between *N. pervillei* and *N. rafflesiana s.l.* were only recovered with the help of longer, transcriptome contigs to align ddRAD reads. Absence of shared male-specific contigs should not, therefore, be interpreted as evidence for independent origins of sex chromosomes, but rather reflects sequence divergence between species. Further evidence for a common origin followed by interspecific divergence is found in a DUF4283 transcript, which is male specific in *N. gracilis* but XY-patterned in *N. rafflesiana s.l.*, suggesting X and MSY alleles (i.e., gametologs) have lost sequence similarity in the former but not in the latter species.

NONCODING DNA AND SPECIAL SIGNIFICANCE OF *DYTI* AND *SEPI*

Nonrecombining regions of sex chromosomes accumulate repetitive, noncoding sequences and TEs in species with both heteromorphic or largely homomorphic sex chromosomes (Čermák et al. 2008; Wang et al. 2012). In *Nepenthes*, most sex-linked genomic regions detected by our approach were noncoding sequences and TEs and only a few genes with putative developmental functions were identified. Of these, a *Nepenthes* homolog of *DYTI* appears to be located in the MSY of all *Nepenthes* species. *DYTI* is essential for tapetum development and thus pollen fertility in *A. thaliana* (Zhang et al. 2006), rice (Jung et al. 2005; Wilson and Zhang 2009; Cai et al. 2015), and tomato (Jeong et al. 2014). Given this gene’s functional conservation in these distantly related Angiosperms, we speculate that its function is the same in *Nepenthes*, and future work could validate this hypothesis, for example, via transient transformation of *Nepenthes* (Miguel et al. 2019). Our analysis suggests that *DYTI* is absent from *Nepenthes* females and must thus be absent from the X chromosome. Such a deletion of *DYTI* from the X chromosome would constitute a recessive male-sterility mutation, as required early in the evolution

of dioecy for the transition from a hermaphroditic to a gynodioecious mating system (Charlesworth and Charlesworth 1978). It is notable that in *Arabidopsis*, *DYTI* directly regulates the expression of TDF1 (Gu et al. 2014), a gene that in dioecious asparagus is essential for male fertility and, like *DYTI* in *Nepenthes*, is located in the MSY (Harkess et al. 2017; Murase et al. 2017). Apparently, this pollen development pathway was involved twice independently in the evolution of angiosperm XY chromosomes, and possibly in the transition to dioecy.

The second *Nepenthes* gene of interest is a homolog of the homeotic MADS box gene *SEPI*, an early-acting regulator of floral organ identity in *A. thaliana* (Pelaz et al. 2000), which was XY-patterned in *N. pervillei*. Two *SEPI*-linked SNPs were heterozygous in 27 of 28 males, whereas 21 of 22 females were homozygous, consistent with the existence of strongly X- and Y-linked copies. If *SEP* homologs in *Nepenthes* are involved in the determination of floral organ identity (as in *A. thaliana*, Theißen et al. 2016), the sex-linked *Nepenthes SEPI* homolog could be involved in unisexual flower development. In particular, sequence differences between the *Nepenthes SEPI* X- and Y-linked copies might modify their functions such that they suppress the development of either carpels or stamens. In *S. latifolia*, however, *SEPI* homologs are not directly involved in sex determination and are not located on the sex chromosomes (Matsunaga et al. 2004).

The possible roles of *DYTI* and *SEPI* in the origin of dioecy in *Nepenthes* require further attention. Even if these are not primary sex-determining genes in extant *Nepenthes*, they might have been under sexually antagonistic selection during the evolution of dioecy because loss of function of *DYTI* or *SEPI* alleles might abort nonfunctional organs at early developmental stages, thus saving resources. The fully unisexual morphology of extant *Nepenthes* flowers (Subramanyam and Narayana 1971) implies further developmental genetic differences between males and females.

Conclusion

This study reports the discovery of an XY sex-determination system in dioecious pitcher plants (*Nepenthes* spp). The sex chromosomes are homomorphic with a small Y-specific region, which has a relatively old core that is shared between distinct species. The nonrecombining region is enriched for noncoding sequences and TEs, but also contains several expressed genes with putative developmental functions.

ACKNOWLEDGMENTS

We thank N. Zemp for data of *Silene latifolia*. We are indebted to H. Luqman, C. Küffer, J. Mougat, K. Beaver, C. Morel, E. Nancy, A. Street, and T. Kropf for accessing *N. pervillei*. A.H.H. Tinggal and I. Daud with family are thanked for hospitality in Brunei, and L.W. Ngai for assistance

with field work in Singapore. The Seychelles Bureau of Standards, the Seychelles Ministry of Environment and Energy, the Ministry of Industry and Primary Resources of Brunei Darussalam, the Agri-Food and Veterinary Authority, and the National Parks Board of Singapore kindly granted permits for research, sampling, and export of plant material. C. Michel is thanked for preparing sequencing libraries. S. Hartmeyer and U. Zimmermann kindly donated samples of sexed *Nepenthes* from their private collections. Sequencing and computation were carried out with the Genetic Diversity Center Zürich, the Quantitative Genomics Facility Basel, and the Functional Genomics Center Zürich. We thank the numerous colleagues and anonymous reviewers for valuable comments on earlier versions of the manuscript. This work was supported by the ETH Zürich and the Rübél foundation. Keygene N.V. owns patents and patent applications protecting its Sequence Based Genotyping technologies. The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

MS performed the research including data collection and analysis, TUG and FM contributed logistic support and materials, and MS and AW designed and interpreted the research and wrote the manuscript.

DATA ARCHIVING

Sequencing reads of ddRAD-seq and RNA-seq are deposited at the European Nucleotide Archive under projects PRJEB20488 and PRJEB22838. Bioinformatic and statistical code used for this study is available at <https://github.com/mscharmann>. VCF files, data used for statistics, and assembled *Nepenthes* contigs in fasta format including the sex-linked loci are deposited at Dryad (<https://doi.org/10.5061/dryad.8gs9388>).

LITERATURE CITED

- Akagi, T., I. M. Henry, R. Tao, and L. Comai. 2014. A Y-chromosome-encoded small RNA acts as a sex determinant in persimmons. *Science* 346:646–650.
- Akagi, T., I. M. Henry, H. Ohtani, T. Morimoto, K. Beppu, I. Kataoka, et al. 2018. A Y-encoded suppressor of feminization arose via lineage-specific duplication of a cytokinin response regulator in kiwifruit. *Plant Cell* 30:780–795.
- Bachtrog, D. 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nat. Rev. Genet.* 14:113–124.
- Bachtrog, D., J. E. Mank, C. L. Peichel, M. Kirkpatrick, S. P. Otto, T.-L. Ashman, et al. 2014. Sex determination: why so many ways of doing it? *PLoS Biol.* 12:e1001899.
- Baird, N. A., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver, Z. A. Lewis, et al. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3:e3376.
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B Stat. Methodol.* 57:289–300.
- Bewick, A. J., F. J. J. Chain, L. B. Zimmerman, A. Sesay, M. J. Gilchrist, N. D. L. Owens, et al. 2013. A large pseudoautosomal region on the sex chromosomes of the frog *Silurana tropicalis*. *Genome Biol. Evol.* 5:1087–1098.
- Blaser, O., S. Neuenschwander, N. Perrin, A. E. A. Gardner, and E. T. Day. 2014. Sex-chromosome turnovers: the hot-potato model. *Am. Nat.* 183:140–146.
- Brelsford, A., G. Lavanchy, R. Sermier, A. Rausch, and N. Perrin. 2017. Identifying homomorphic sex chromosomes from wild-caught adults with limited genomic resources. *Mol. Ecol. Resour.* 17:752–759.
- Bresadola, L., V. Link, C. A. Buerkle, C. Lexer, and D. Wegmann. 2019. Estimating and accounting for genotyping errors in RAD-seq experiments. *bioRxiv* 587428.
- Cai, C.-F., J. Zhu, Y. Lou, Z.-L. Guo, S.-X. Xiong, K. Wang, Y. Zhong-Nan. 2015. The functional analysis of OsTDF1 reveals a conserved genetic pathway for tapetal development between rice and *Arabidopsis*. *Sci. Bull.* 60:1073–1082.
- Camacho-Sanchez, M., P. Burraco, I. Gomez-Mestre, and J. A. Leonard. 2013. Preservation of RNA and DNA from mammal samples under field conditions. *Mol. Ecol. Resour.* 13:663–673.
- Carmichael, S. N., M. Bekaert, J. B. Taggart, H. R. L. Christie, D. I. Bassett, J. E. Bron, et al. 2013. Identification of a sex-linked SNP marker in the Salmon Louse (*Lepeophtheirus salmonis*) using RAD sequencing. *PLoS One* 8:e77832.
- Čermák, T., Z. Kubát, R. Hobza, A. Koblizkova, A. Widmer, J. Macas, et al. 2008. Survey of repetitive sequences in *Silene latifolia* with respect to their distribution on sex chromosomes. *Chromosome Res.* 16:961–976.
- Charlesworth, B., and D. Charlesworth. 1978. A model for the evolution of dioecy and gynodioecy. *Am. Nat.* 112:975–997.
- Charlesworth, D. 2002. Plant sex determination and sex chromosomes. *Heredity* 88:94–101.
- . 2015. Plant contributions to our understanding of sex chromosome evolution. *New Phytol.*, 208:52–65.
- . 2016. Plant sex chromosomes. *Annu. Rev. Plant Biol.* 67:397–420.
- Charlesworth, D., and J. E. Mank. 2010. The birds and the bees and the flowers and the trees: lessons from genetic mapping of sex determination in plants and animals. *Genetics* 186:9–31.
- Cheek, M. R., and M. H. P. Jebb. 2001. *Nepenthaceae*. *Flora Malesiana* 15:1–157.
- Clarke, C. M. 1992. The ecology of metazoan communities in *Nepenthes* pitcher plants in Borneo, with special reference to the community of *Nepenthes bicalcarata* Hook. f. PhD Thesis. University of New England, Armidale, Australia.
- . 1997. *Nepenthes* of Borneo. Natural History Publications (Borneo), Kota Kinabalu, Malaysia.
- . 2001. *Nepenthes* of Sumatra and Peninsular Malaysia. Natural History Publications (Borneo), Kota Kinabalu, Malaysia.
- Clarke, C. M., J. Schlauer, J. A. Moran, and A. S. Robinson. 2018. Systematics and evolution of *Nepenthes*. Pp. 58–69 in A. M. Ellison and L. Adamec, eds. *Carnivorous plants: physiology, ecology, and evolution*. Oxford Univ. Press, Oxford, U.K.
- Cuénoud, P., V. Savolainen, L. W. Chatrou, M. Powell, R. J. Grayer, and M. W. Chase. 2002. Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid rbcL, atpB, and matK DNA sequences. *Am. J. Bot.* 89:132–144.
- Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, et al. 2011. The variant call format and VCFtools. *Bioinformatics* 27:2156–2158.
- Eddy, S. R., T. J. Wheeler, and the HMMER development team. 2016. HMMER. <http://hmmerr.org>.
- Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto, E. S. Buckler, et al. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6:e19379.
- Filatov, D. A. 2015. Homomorphic plant sex chromosomes are coming of age. *Mol. Ecol.* 24:3217–3219.
- Gamble, T., and D. Zarkower. 2014. Identification of sex-specific molecular markers using restriction site-associated DNA sequencing. *Mol. Ecol. Resour.* 14:902–913.
- Gamble, T., J. Coryell, T. Ezaz, J. Lynch, D. P. Scantlebury, and D. Zarkower. 2015a. Restriction site-associated DNA sequencing (RAD-seq) reveals

- an extraordinary number of transitions among gecko sex-determining systems. *Mol. Biol. Evol.* 32:1296–1309.
- Gamble, T., J. Coryell, T. Ezaz, J. Lynch, D. P. Scantlebury, and D. Zarkower. 2015b. Restriction site-associated DNA sequencing (RAD-seq) reveals an extraordinary number of transitions among gecko sex-determining systems. *Mol. Biol. Evol.* 32:1296–1309.
- Gammerding, W. J., and T. D. Kocher. 2018. Unusual diversity of sex chromosomes in African cichlid fishes. *Genes* 9:480.
- Gu, J.-N., J. Zhu, Y. Yu, X.-D. Teng, Y. Lou, X.-F. Xu, et al. 2014. DYT1 directly regulates the expression of TDF1 for tapetum development and pollen wall formation in *Arabidopsis*. *Plant J.* 80:1005–1013.
- Haas, B. J., A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, et al. 2013. De novo transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. *Nat. Protoc.* 8:1494–1512.
- Harkess, A., and J. Leebens-Mack. 2017. A century of sex determination in flowering plants. *J. Hered.* 108:69–77.
- Harkess, A., J. Zhou, C. Xu, J. E. Bowers, R. Hulst, S. Ayyampalayam, et al. 2017. The asparagus genome sheds light on the origin and evolution of a young Y chromosome. *Nat. Commun.* 8:1279.
- Heikrujam, M., K. Sharma, M. Prasad, and V. Agrawal. 2015. Review on different mechanisms of sex determination and sex-linked molecular markers in dioecious crops: a current update. *Euphytica* 201:161–194.
- Heubl, G., and A. Wistuba. 1997. A cytological study of the genus *Nepenthes* L. (Nepenthaceae). *Sendtnera* 4:169–172.
- Jeffries, D. L., G. Lavanchy, R. Sermier, M. J. Sredl, I. Miura, A. Borzée, et al. 2018. A rapid rate of sex-chromosome turnover and non-random transitions in true frogs. *Nat. Commun.* 9:4088.
- Jeong, H.-J., J.-H. Kang, M. Zhao, J.-K. Kwon, H.-S. Choi, J. H. Bae, et al. 2014. Tomato male sterile 1035 is essential for pollen development and meiosis in anthers. *J. Exp. Bot.* 65:6693–6709.
- Jung, K.-H., M.-J. Han, Y.-S. Lee, Y.-W. Kim, I. Hwang, M.-J. Kim, et al. 2005. Rice undeveloped Tapetum1 is a major regulator of early tapetum development. *Plant Cell* 17:2705–2722.
- Juniper, B. E., R. J. Robins, and D. M. Joel. 1989. *The carnivorous plants*. Academic Press, London.
- Krasovec, M., M. Chester, K. Ridout, and D. A. Filatov. 2018. The mutation rate and the age of the sex chromosomes in *Silene latifolia*. *Curr. Biol.* 28:1832–1838.e4.
- Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997 [q-bio].
- Mastretta-Yanes, A., N. Arrigo, N. Alvarez, T. H. Jorgensen, D. Piñero, and B. C. Emerson. 2015. Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Mol. Ecol. Resour.* 15:28–41.
- Matsunaga, S., W. Uchida, E. Kejnovsky, E. Isono, F. Moneger, B. Vyskot, et al. 2004. Characterization of two SEPALLATA MADS-box genes from the dioecious plant *Silene latifolia*. *Sex. Plant Reprod.* 17:189–193.
- Meimberg, H., and G. Heubl. 2006. Introduction of a nuclear marker for phylogenetic analysis of Nepenthaceae. *Plant Biol.* 8:831–840.
- Meimberg, H., A. Wistuba, P. Dittrich, and G. Heubl. 2001. Molecular phylogeny of Nepenthaceae based on cladistic analysis of plastid trnK intron sequence data. *Plant Biol.* 3:164–175.
- Miguel, S., E. Nisse, F. Biteau, S. Rottloff, B. Mignard, E. Gontier, et al. 2019. Assessing carnivorous plants for the production of recombinant proteins. *Front. Plant Sci.* 10:793.
- Ming, R., A. Bendahmane, and S. S. Renner. 2011. Sex chromosomes in land plants. *Annu. Rev. Plant Biol.* 62:485–514.
- Moran, J. A., and C. M. Clarke. 2010. The carnivorous syndrome in *Nepenthes* pitcher plants. *Plant Signal. Behav.* 5:644–648.
- Mullins, J. T. 2000. *Molecular systematics of Nepenthaceae*. PhD Thesis. University of Reading, Reading, U.K.
- Murase, K., S. Shigenobu, S. Fujii, K. Ueda, T. Murata, A. Sakamoto, et al. 2017. MYB transcription factor gene involved in sex determination in *Asparagus officinalis*. *Genes Cells* 22:115–123.
- Muyle, A., R. Shearn, and G. A. Marais. 2017. The evolution of sex chromosomes and dosage compensation in plants. *Genome Biol. Evol.* 9:627–645.
- Pavlovič, A., and M. Saganová. 2015. A novel insight into the cost–benefit model for the evolution of botanical carnivory. *Ann. Bot.* 115:1075.
- Pelaz, S., G. S. Ditta, E. Baumann, E. Wisman, and M. F. Yanofsky. 2000. B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature* 405:200–203.
- Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* 7:e37135.
- Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81:559–575.
- Puritz, J. B., C. M. Hollenbeck, and J. R. Gold. 2014. dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ* 2:e431.
- Renner, S. 2014. The relative and absolute frequencies of angiosperm sexual systems: dioecy, monoecy, gynodioecy, and an updated online database. *Am. J. Bot.* 101:1588–1596.
- Renner, S., and R. Ricklefs. 1995. Dioecy and its correlates in the flowering plants. *Am. J. Bot.* 82:596–606.
- Renner, T., and C. D. Specht. 2011. A sticky situation: assessing adaptations for plant carnivory in the Caryophyllales by means of stochastic character mapping. *Int. J. Plant. Sci.* 172:889–901.
- Scharmann, M., and T. U. Grafe. 2013. Reinstatement of *Nepenthes hemsleyana* (Nepenthaceae), an endemic pitcher plant from Borneo, with a discussion of associated *Nepenthes* taxa. *Blumea* 58:8–12.
- Shimizu, K. K., and T. Tsuchimatsu. 2015. Evolution of selfing: recurrent patterns in molecular adaptation. *Annu. Rev. Ecol. Evol. Syst.* 46:593–622.
- Smit, A., R. Hubley, and P. Green. 2013. RepeatMasker Open-4.0. 2013–2016.
- Steflova, P., R. Hobza, B. Vyskot, and E. Kejnovsky. 2013. Strong accumulation of chloroplast DNA in the Y chromosomes of *Rumex acetosa* and *Silene latifolia*. *Cytogenet. Genome Res.* 142:59–65.
- Subramanyam, K., and L. L. Narayana. 1971. A contribution to the floral anatomy of *Nepenthes khasiana* Hook f. *Proc. Indian Acad. Sci. B* 73:124–131.
- Tennessen, J. A., N. Wei, S. C. K. Straub, R. Govindarajulu, A. Liston, and T.-L. Ashman. 2018. Repeated translocation of a gene cassette drives sex-chromosome turnover in strawberries. *PLoS Biol.* 16:e2006062.
- Theißen, G., R. Melzer, and F. Rümpler. 2016. MADS-domain transcription factors and the floral quartet model of flower development: linking plant development and evolution. *Development* 143:3259–3271.
- VanLiere, J. M., and N. A. Rosenberg. 2008. Mathematical properties of the r2 measure of linkage disequilibrium. *Theor. Popul. Biol.* 74:130–137.
- Walker, J. F., Y. Yang, M. J. Moore, J. Mikenas, A. Timoneda, S. F. Brockington, et al. 2017. Widespread paleopolyploidy, gene tree conflict, and recalcitrant relationships among the carnivorous Caryophyllales. *Am. J. Bot.* 104:858–867.
- Wang, J., J.-K. Na, Q. Yu, A. R. Gschwend, J. Han, F. Zeng, et al. 2012. Sequencing papaya X and Yh chromosomes reveals molecular basis of incipient sex chromosome evolution. *Proc. Natl. Acad. Sci. USA* 109:13710–13715.

- Westergaard, M. 1958. The mechanism of sex determination in dioecious flowering plants. Pp. 217–281 in M. Demerec, ed. *Advances in Genetics*. Academic Press, London.
- Wilson Sayres, M. A., K. E. Lohmueller, and R. Nielsen. 2014. Natural selection reduced diversity on human Y chromosomes. *PLoS Genet.* 10:e1004064.
- Wilson, Z. A., and D.-B. Zhang. 2009. From *Arabidopsis* to rice: pathways in pollen development. *J. Exp. Bot.* 60:1479–1492.
- Zhang, W., Y. Sun, L. Timofejeva, C. Chen, U. Grossniklaus, and H. Ma. 2006. Regulation of *Arabidopsis* tapetum development and function by *DYSFUNCTIONAL TAPETUM1 (DYT1)* encoding a putative bHLH transcription factor. *Development* 133:3085–3095.

Associate Editor: S. Wright

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Text S1. Preliminary molecular sexing assay for *Nepenthes rafflesiana s.l.*

Text S2. A molecular sexing assay for the genus *Nepenthes*.

Text S3. Analyses of sex-linkage in *Silene latifolia*.

Text S4. Privacy rarefaction.

Text S5. Performance analysis of privacy rarefaction on simulated RAD data.

Text S6. Male inflorescence transcriptome of *N. khasiana*.

Text S7. Phylogenetic dating of *Nepenthes*.

Table S8. Result tables of privacy rarefaction for three *Nepenthes* and *Silene*.

Table S9. Result tables of association tests for SNPs with sex and heterozygosities.

Table S10. Summary table of sex-linked loci including overlap between species and annotations for RAD-tag and transcriptome references.

Table S11. Data table for *S. latifolia* sex-linkage comparison of our results to previous studies.