UNIVERSITY OF NAPLES FEDERICO II

Ph.D. COURSE IN BIOLOGY

XXXVI CYCLE

Development of molecular and morphological markers for rapid sex screening in plants of agronomic and horticultural interest

Supervisor:	**Candidate:**
Prof	Ylenia De Luca
Donata Cafasso	DR995507

Co-Supervisor:
Prof
Salvatore Cozzolino

Years 2021-2024

1. Abstract

This study addresses the urgent need for early sex identification in dioecious plants, focusing on *Cycas revoluta*. For *C. revoluta*, a popular ornamental species with toxic seeds from female plants, seedling morphological traits were examined and correlated with molecular sex determination. The results revealed a sexually dimorphic trait, the distance between leaflets on the leaf rachis (IRD), showing clear distinctions between male and female individuals, particularly in the early developmental stage of the plant. This finding suggests the establishment of sexual dimorphism in leaf morphology very early in the life cycle of *C. revoluta*. Whenever available, morphological traits for sex differentiation in greenhouses and plant nurseries can represent a rapid and economic tool when coupled with automatized image recognition systems. Furthermore, we employed reduced representation genomic analysis using ddRAD sequencing and a dedicated bioinformatics pipeline to assess the feasibility of this approach as a rapid alternative to whole-genome sequencing in developing sex-linked molecular markers. This is particularly significant for non-model plant species characterized by large genomes, as exemplified by *C. revoluta*. This method identified male-specific contigs, indicative of an XY sex-determination system. BLASTn analysis and alignments confirmed the presence of contigs on both the X and Y chromosomes, with an SNP within a contig allowing for the development of a male-specific molecular marker. Validation of this marker confirmed its specificity for the Y chromosome in *C. revoluta*. The nearly equal sex ratio observed in F1 generation seedlings supported the XY sex-determination system in *C. revoluta*. The study highlights the effectiveness of ddRAD sequencing for the rapid development of sex-linked molecular markers in non-model plant species, with potential applications in agriculture, breeding, and conservation activities.

2. Introduzione

2.1 Premessa

Il concetto di "sistema sessuale" si riferisce alla distribuzione e alla funzione delle strutture coinvolte nella produzione dei gameti, con una vasta gamma di schemi riflessi nella grande varietà osservata (Barrett 2002; Yampolsky and Yampolsky 1922). Contrariamente al predominio della gonocoria negli animali, caratterizzata dalla netta separazione dei sessi in individui distinti, nelle piante questo fenomeno è sorprendentemente raro, con solo circa il 9-10% delle specie terrestri che mostra una separazione rigida dei sessi (dioecia), corrispondente a circa 29.000 specie su un totale di 300.000, con una distribuzione non uniforme.

Tra le angiosperme, approssimativamente il 6% delle specie, pari a circa 14.600 specie distribuite in 960 generi e 200 famiglie, presenta un sistema dioico (Renner and Ricklefs 1995). Tuttavia, l'ermafroditismo rappresenta la condizione predominante nelle piante terrestri, con la separazione dei sessi in individui distinti che è considerata una condizione derivata da uno stato ermafrodita ancestrale (Tandurdzic and Banks 2004). Di conseguenza, l'evoluzione dei cromosomi sessuali nel regno vegetale è un processo relativamente giovane, molto più recente rispetto a quanto avvenuto nei mammiferi o negli insetti (Drosophila), il che spiega l'attenzione crescente dedicata agli studi evolutivi e ai meccanismi di determinazione sessuale nelle piante (Ming and Moore 2007; Charlesworth 2015).

La comprensione dei sistemi riproduttivi, delle strategie di determinazione sessuale e dei geni collegati al sesso è cruciale non solo per la ricerca scientifica, ma anche per applicazioni pratiche come il miglioramento delle specie coltivate. Grazie allo sviluppo di tecniche avanzate in biologia molecolare e genomica, è stata resa possibile un'indagine più dettagliata della sessualità delle piante, che include il clonaggio dei geni coinvolti nella determinazione sessuale e il sequenziamento dei cromosomi sessuali, similmente a quanto fatto in passato per i mammiferi (Skalesky et al. 2003; Ross et al. 2005).

Geni responsabili della determinazione sessuale sono stati clonati in piante di interesse agronomico come il mais (DeLong et al. 1993; Bensen et al. 1995), e sono state costruite mappe dei cromosomi sessuali in diverse specie, tra cui *Carica papaya* (Zhang et al. 2008; Yu et al. 2008a), *Asparagus officinalis* (Telgmann-Rauber et al. 2007), *Fragaria virginiana* (Goldberg et al. 2010), e *Populus* *trichocarpa* (Tuskan et al. 2006). Tale disponibilità di nuovi dati genomici, genetici e molecolari ha permesso di far luce sul processo di evoluzione dei cromosomi sessuali anche su specie non coltivate.

2.2 Meccanismi di determinazione sessuale

Da un punto di vista strettamente biologico ed evolutivo è interessante osservare che nelle piante la maggior parte delle specie "attualmente dioiche" (caratterizzate dalla presenza di sessi separati su individui diversi) si siano evolute, da un ancestore ermafrodita o monoico, caratterizzato dal possedere sullo stesso individuo funzioni di entrambi i sessi. Da quasi un secolo, infatti, si è concordi nell'idea che l'ermafroditismo, come condizione sessuale, sia comparso molto prima della gonocoria o dioicismo con una serie di passaggi sequenziali, tra cui il monoicismo dove si osserva la presenza di fiori unisessuali maschili e femminili sullo stesso individuo (Freeman, 1997). Lo sviluppo di strutture riproduttive quali lo stame e il carpello coinvolge una serie di funzioni genetiche specializzate che sono richieste a diversi stadi di sviluppo della pianta, e l'occorrenza di mutazioni in uno qualsiasi dei geni regolatori tali meccanismi sessuali, può comportare l'aborto o la perdita delle funzioni degli organi riproduttivi maschili o femminili (Wellmer et al. 2004; Zhang et al. 2005). La comparsa di geni caratterizzanti strutture esclusivamente maschili e femminili ha determinato lo sviluppo delle strutture riproduttive unisessuali, segnando l'inizio dell'evoluzione dei cromosomi sessuali (Ming et al. 2011). Nel mondo vegetale la determinazione sessuale è un meccanismo complesso, che può essere conseguenza di diversi fattori, ed al contempo è un fenomeno molto più "flessibile" di quanto avvenga nel regno animale. Nelle piante si trovano infatti organismi con condizioni sessuali intermedie rispetto all'unisessualità e all'ermafroditismo: si pensi ad individui che presentano sullo stesso individuo fiori ermafroditi e fiori unisessuati (andromonoicimo e il ginodioicismo) o piante che recano fiori ermafroditi e fiori unisessuati su individui distinti (androdioicismo e ginodioicismo). In qualche specie può avvenire una determinazione fisiologica ovvero influenzata da ormoni, oppure una determinazione ambientale condizionata da fattori esterni come luce, temperatura, disponibilità idrica, disponibilità di nutrienti. In dettaglio, la determinazione del sesso può essere legata a fattori ambientali (Environmental Sex Determination), con proporzioni tra i sessi variabili e indeterminate, oppure può essere genotipica (Genetic Sex Determinazion), cioè, determinata da fattori genetici. Tale fenomeno è conosciuto nelle piante fin dall'inizio del secolo scorso (Freeman 1997).

Gli studi sui geni di determinazione sessuale recentemente clonati nella famiglia delle Cucurbitaceae forniscono un esempio di come tali geni possano essere influenzati da fattori ormonali. In molte specie di Cucurbitaceae, come il cetriolo (*Cucumis sativus*) e il melone (*Cucumis melo*), si osserva polimorfismo intraspecifico nei loro sistemi sessuali.

Nel melone, la determinazione sessuale è controllata dai geni andromonoici (a) e ginoici (g), con interazioni tra i loro diversi alleli che determinano una gamma di fenotipi sessuali (Kenigsbuch and Cohen 1990; Poole and Grimball 1939). La determinazione sessuale nel cetriolo è controllata da tre geni principali, tra cui il gene femminile (F), il gene androceo (a) e il gene monoico (M), che influenzano l'equilibrio tra i fiori maschili e femminili (Yamasaki et al. 2001).

In entrambe le specie di *Cucumis*, l'espressione sessuale può essere modificata da fattori ormonali ed ambientali, con l'etilene che gioca un ruolo chiave (Byers et al. 1972; Yin and Quinn 1995). L'etilene agisce come agente femminizzante, con il gene ACS che regola la biosintesi dell'etilene e determina la formazione di organi riproduttivi maschili e femminili (Papadopoulou et al. 2005) (Fig.1). Il gene femminile nel cetriolo è presumibilmente coinvolto nella regolazione dell'ACS (Trebitsh et al. 1997).

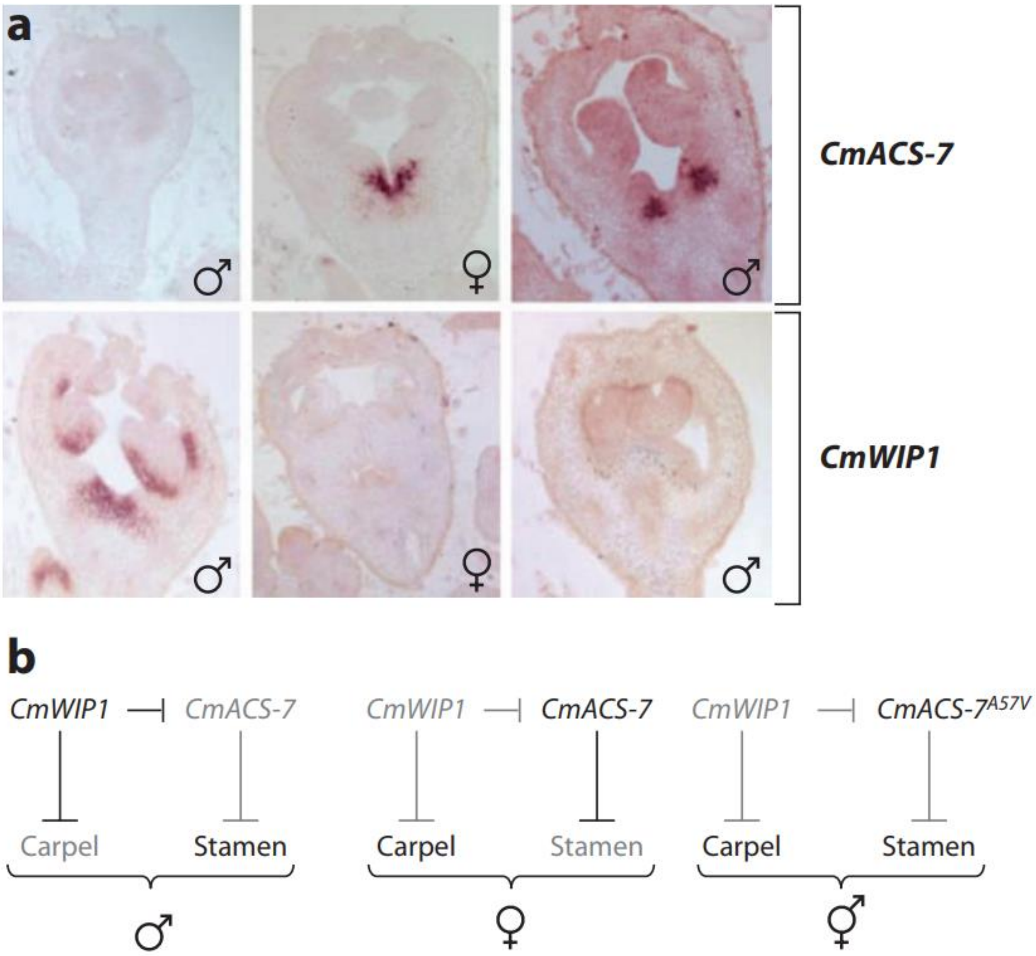

Fig.1 Modello di determinazione del sesso in *Cucumis melo*. (a) Analisi in situ dell'espressione di CmACS-7 e CmWIP1 nelle prime fasi dello sviluppo del fiore. L'espressione di CmACS-7 è rilevata nei primordi dei carpelli dei fiori femminili ed ermafroditi, mentre non viene rilevata espressione nel fiore maschile. Al contrario, CmWIP1 è espresso nei primordi dei carpelli del fiore maschile e non è espresso nei fiori femminili ed ermafroditi. (b) Modello per la determinazione del sesso nel melone: CmACS-7 e CmWIP1 sono repressori dello stame e del carpello, rispettivamente. I geni in grigio non sono espressi (Ming et al. 2011).

2.3 Cromosomi sessuali nelle piante, evoluzione e filogenesi

Nell'ambito della determinazione genetica, in molte tra le specie dioiche, si è riscontrato come il sesso venga controllato non solo da geni sex-linked, bensì da un intero cromosoma o da una coppia di omologhi. In queste specie i cromosomi sessuali possono presentarsi omomorfi o morfologicamente distinti. I cromosomi eteromorfi e si dividono in due tipi:

1) sistema XY in cui il sesso maschile è eterogametico è, e il sesso femminile è omogametico (XX), il più frequente tra i modelli di cromosomi sessuali eteromorfi (Ming et al. 2011).

2) sistema ZW, in cui il sesso femminile è eterogametico ed il sesso maschile omogametico (ZZ). Quest'ultima combinazione, comune nel regno degli Uccelli (Smith et al. 2009) si ritrova ad esempio in *Fragaria elatior* (Ahmadi et al. 1991). Inoltre, il sesso dell'individuo può ulteriormente essere determinato dal rapporto numerico tra il numero degli X e gli altri cromosomi non sessuali (es. *Humulus lupulus, Rumex acetosa*) (Grabowska-Joachimiak et al. 2006; Karlov et al. 2003; Blocka-Wandas et al. 2007; Cuñado et al. 2007).

Un evento cruciale nell'evoluzione dei cromosomi sessuali è la soppressione della ricombinazione tra i due geni di determinazione del sesso. Senza questa soppressione, le mutazioni maschio sterile o femmina sterile potrebbero causare il ritorno all'ermafroditismo o a individui neutri. La soppressione della ricombinazione avviene attraverso vari meccanismi, inclusi riarrangiamenti cromosomici e metilazione del DNA (Jaarola et al. 1998; Maloisel and Rossignol, 1998).

La soppressione della ricombinazione è un tratto distintivo dei cromosomi sessuali e consente loro di trasmettere i geni sessuali in modo coerente alle generazioni successive senza essere mescolati attraverso la ricombinazione genetica durante la meiosi. In piante come *Humulus lupulus* e il pioppo, è stata osservata una riduzione significativa della ricombinazione sul cromosoma Y o W rispetto al corrispondente cromosoma X o Z (Seefelder et al. 2000; Yin et al. 2008).

In definitiva, la soppressione della ricombinazione nei cromosomi sessuali è un fenomeno chiave nell'evoluzione dei sistemi sessuali delle piante, contribuendo alla stabilità dei geni sessuali attraverso le generazioni.

Approfonditi studi genomici hanno consentito di riconoscere sei stadi dell'evoluzione dei cromosomi sessuali. Nella Fig. 2 ciascuno dei sei stadi è rappresentato da una specie di piante terrestri come segue:

- Stadio 1: Una mutazione maschile e una femminile sterile con dominanza complementare si verificano in prossimità su un cromosoma. La sterilità maschile è dovuta a una mutazione di perdita di funzione in un sistema XY, mentre la sterilità femminile è dovuta a un allele dominante legato al cromosoma Y che causa l'aborto dei gameti femminili. La situazione si inverte in un sistema ZW, dove le femmine sono il sesso eterogametico. In questo caso, un allele femminile sterile legato al cromosoma Z causa lo sviluppo di maschi ZZ, mentre un allele maschile sterile dominante legato al cromosoma W causa lo sviluppo di femmine ZW. La ricombinazione tra i loci complementari maschili e femminili sterili non è soppressa, e individui ermafroditi e neutri esistono nella popolazione. *Fragaria virginiana* è un eccellente esempio di questo stadio (Sakai and Weller 1999).

- Stadio 2: La ricombinazione è soppressa tra i due loci determinanti il sesso e le regioni vicine, portando all'avvio del processo di degenerazione. Si forma una piccola regione specifica del maschio sul cromosoma Y nascente. La seconda caratteristica di questo stadio è che il genotipo YY è vitale ed esiste una popolazione interamente maschile. I cromosomi sessuali di *Asparagus officinalis* rappresentano questo stadio.

- Stadio 3: La soppressione della ricombinazione si estende alle regioni vicine permettendo a un gran numero di geni legati al cromosoma Y di degenerare e formare una regione specifica del maschio sul cromosoma Y nascente. La regione specifica del maschio si espande attraverso l'accumulo di retrotrasposoni e la traslocazione e duplicazione di frammenti genomici. I cromosomi XY sembrano essere omomorfi a livello citologico, ma sono eteromorfi a livello molecolare. La perdita di contenuto genico è sufficientemente estesa da causare la letalità del genotipo YY (Charlesworth and Charlesworth 2000). I cromosomi sessuali di *Carica papaya* illustrano questo stadio (Lewis 1942).

- Stadio 4: La regione specifica del maschio si estende alla maggior parte del cromosoma Y e continua la degenerazione. L'accumulo di elementi trasponibili e duplicazioni all'interno della regione specifica del maschio provoca un'espansione significativa del contenuto di DNA sul cromosoma Y. A questo stadio, i cromosomi X e Y sono eteromorfi, e il cromosoma Y può essere molto più grande del cromosoma X. I cromosomi sessuali di *Silene latifolia* possiedono queste proprietà (Delph et al. 2010).

- Stadio 5: Una grave degenerazione del cromosoma Y causa la perdita di funzione per la maggior parte dei geni, e la perdita delle sequenze non funzionanti del cromosoma Y porta a una diminuzione del cromosoma Y. Alcuni sistemi cromosomici sessuali potrebbero non attraversare questa fase di riduzione, ma invece continuano ad espandersi e degenerare fino a una completa perdita del cromosoma Y. In entrambi i casi, una piccola porzione del cromosoma Y continua ad accoppiarsi meioticamente con il cromosoma X consentendo una corretta disgiunzione. Non sono noti cromosomi sessuali di angiosperme a questo stadio, ma la gimnosperma *Cycas revoluta* ha cromosomi sessuali eteromorfi con un cromosoma Y ridotto.

- Stadio 6: La soppressione della ricombinazione si estende all'intero cromosoma Y. Si verifica un ulteriore ridimensionamento del cromosoma Y e la completa perdita della regione pseudoautosomica ricombinante. Il cromosoma Y è totalmente perso e si evolve un nuovo sistema di determinazione del sesso basato sul rapporto X-autosoma. Un nuovo cromosoma Y potrebbe formarsi ma non svolge alcun ruolo nella determinazione del sesso. I cromosomi sessuali di *Rumex acetosa* si trovano a questo stadio.

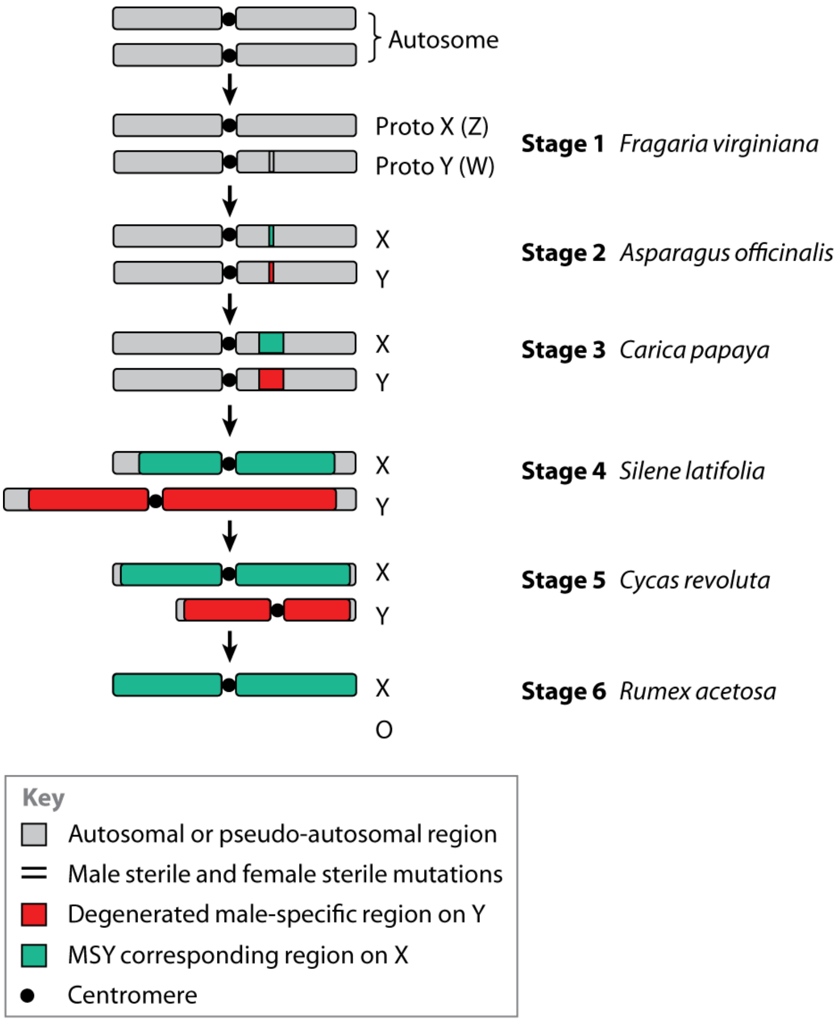

Fig. 2 Le sei fasi dell'evoluzione dei cromosomi sessuali. Stadio 1: Mutazione unisessuale di due geni per la determinazione del sesso con dominanza complementare. Stadio 2: Soppressione della ricombinazione tra i due geni di determinazione del sesso e il genotipo YY diviene praticabile. Stadio 3: La soppressione della ricombinazione si estende alle regioni adiacenti e si sviluppa una piccola regione specifica maschile del cromosoma Y (MSY). Il genotipo YY non è vitale. Stadio 4: L'MSY si espande in dimensioni e subisce degenerazione del contenuto genico attraverso l'accumulo di inserzioni di elementi trasponibili e riarrangiamenti intracromosomici. I cromosomi X e Y diventano eteromorfi. Stadio 5: Grave degenerazione del cromosoma Y. La delezione di sequenze di DNA non funzionali porta a una riduzione delle dimensioni del cromosoma Y. Stadio 6: La soppressione della ricombinazione si diffonde sull'intero cromosoma Y. Il cromosoma Y si perde e si evolve il sistema di determinazione del sesso basato sul rapporto X-autosoma (Ming et al. 2011).

Attualmente sono conosciute circa 40 specie in 21 generi appartenenti a 15 famiglie di piante a fiore che possiedono cromosomi sessuali (Fig.3) e di queste, solo circa 20 specie in 6 generi (*Cannabis, Humulus, Silene, Coccinia, Trichosanthes e Rumex*) presentano cromosomi sessuali eteromorfi. Questi numeri sono rimasti praticamente invariati fin dalla revisione di Westergaard negli anni '50 (Westergaard 1958), probabilmente perché la citologia botanica è una disciplina in declino. Di conseguenza, poche nuove specie vegetali vengono esaminate in termini di complementi cromosomici. Tuttavia, i metodi molecolari-citologici e molecolari-genetici stanno ora rivelando cromosomi sessuali (cioè, regioni con soppressione della ricombinazione con geni determinanti il sesso maschile e femminile) in specie in cui in precedenza erano stati solo sospettati sulla base dei rapporti sessuali dei progenitori o di prove citologiche deboli. Tuttavia, rimane incerto se tutte le specie dioiche siano su una traiettoria evolutiva verso l'evoluzione dei cromosomi sessuali.

Non esiste un modello chiaro nella distribuzione dei cromosomi sessuali tra le piante terrestri o tra le piante a fiore. Sono presenti sia nelle piante annuali che in quelle perenni, sia nelle specie tropicali che in quelle temperate. Ciò è in contrasto con la distribuzione della dioecia, che è sovrarappresentata tra alberi e specie rampicanti, rispetto alle specie erbacee o arbustive (Renner and Ricklefs 1995).

L'età dei cromosomi sessuali delle piante è stata dedotta utilizzando tecniche di datazione molecolare, che si basano sulla misurazione dei tassi di mutazione nei genomi delle specie strettamente imparentate. Queste tecniche confrontano le sequenze di DNA delle specie con e senza cromosomi sessuali per stimare quando tali cromosomi si sono evoluti e divergono geneticamente. Ad esempio, in *S. latifolia*, i cromosomi sessuali potrebbero avere un'età compresa tra 8 e 24 milioni di anni (Moore et al. 2003). In *C. papaya*, sono stati stimati 0,5-2,2 milioni di anni di divergenza per quattro coppie di geni X/Y distribuiti su più della metà del MSY, indicando che i cromosomi sessuali si sono probabilmente evoluti a livello di specie, molto tempo dopo la divergenza delle Caricaceae dai loro parenti più stretti, le Moringaceae (Yu et al. 2008a, 2008b). Per *B. dioica*, è stata inferita un'età inferiore a 10 milioni di anni (Volz and Renner 2008), mentre i cromosomi sessuali in *Rumex* potrebbero essere originati tra 15 e 16 milioni di anni fa.

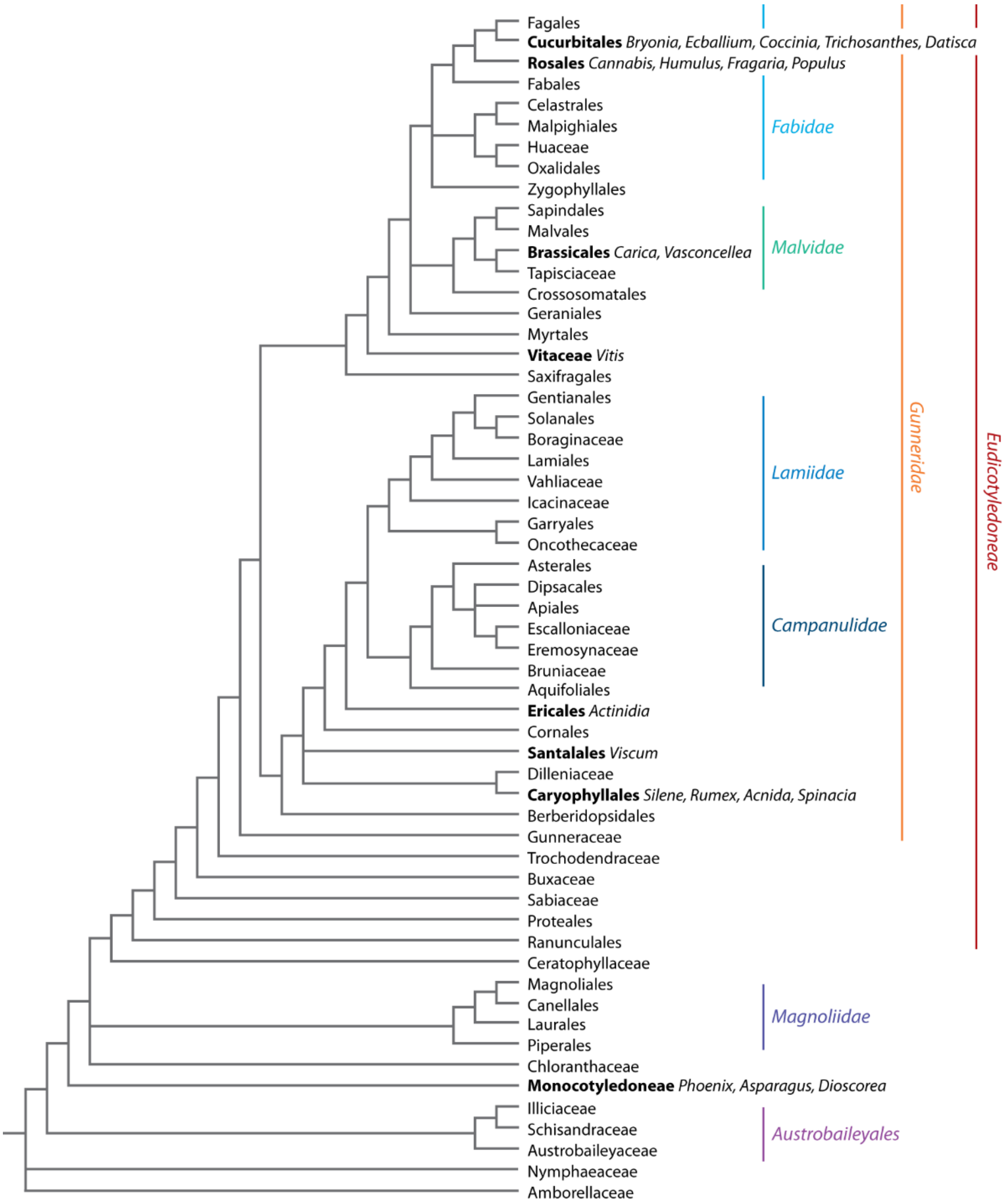

Fig. 3 Distribuzione dei cromosomi sessuali tra le piante a fiore (Bell et al. 2010).

2.4 Cromosomi sessuali nelle piante terrestri

Angiosperme

Le angiosperme sono le piante terrestri più di successo in termini di numero di specie. Tra le loro 250.000 specie esistenti, solo 37 (0,01%) appartenenti a 16 famiglie presentano evidenze citogenetiche e/o molecolari per la presenza di cromosomi sessuali, e molte altre specie con cromosomi sessuali rimangono ancora da scoprire. Tra le 37 specie, 17 appartenenti a 4 famiglie hanno cromosomi sessuali eteromorfi, mentre le altre 20 specie appartenenti a 12 famiglie hanno cromosomi sessuali omomorfi.

Actinidiaceae. *Actinidia chinensis*, il kiwi, ha un sistema di determinazione del sesso XX/XY e cromosomi omomorfi (Fraser et al. 2009). Linkage group 17 è stato identificato come un cromosoma sessuale iniziale dove il locus di determinazione del sesso si trova in una regione subtelomerica con ricombinazione soppressa (Fraser et al. 2009). Il genere Actinidia contiene 30 specie, ma oltre ad *A. chinensis* poche sono state analizzate in termini di cariotipo.

Asparagaceae. Gli incroci tra femmine e maschi di *Asparagus officinalis* danno origine o a progenie con un rapporto di segregazione sessuale maschio-femmina 1:1 o a progenie completamente maschili (Rich and Hanna 1943). La determinazione del sesso in questa specie è controllata da un sistema attivo di cromosomi Y. Le popolazioni interamente maschili sono composte da super maschi con un genotipo YY. Il cromosoma 5 in *A. officinalis* è stato identificato come il cromosoma sessuale tramite analisi trisomica, anche se i cromosomi X e Y sono omomorfi e non distinguibili dal punto di vista citologico (Loptien 1979). Il cromosoma Y contiene due geni strettamente legati, un attivatore maschile (M) e un soppressore femminile (F). I ricombinanti rari tra questi loci possono produrre progenie ermafrodite e sterili (Marks 1973). La determinazione del sesso nell'asparago è stata mappata su un singolo locus M (Reamon-Buttne et al. 1998), indicando la soppressione della ricombinazione tra i due geni di determinazione del sesso. Il mappaggio fisico ha ridotto il divario tra i marcatori flanking del locus M, ma il divario non è ancora stato colmato a causa della natura altamente ripetitiva della sequenza (Telgmann-Rauber et al. 2007). La vitalità del genotipo YY e l'occorrenza di rare ricombinazioni tra i due geni di determinazione del sesso indicano che i cromosomi sessuali dell'asparago sono tra gli stadi 1 e 2, e vicini allo stadio 2. Il genere *Asparagus* è composto da circa 300 specie, e l'evoluzione dei cromosomi sessuali da altre specie del genere non è ben studiata.

Arecaceae. I cromosomi sessuali del dattero (*Phoenix dactylifera*) non sono distinguibili mediante microscopia ottica. Tuttavia, le intensità fluorescenti dei punti di eterocromatina sono maggiori sul cromosoma Y (Siljak-Yakovlev et al. 1996). L'identificazione di marcatori del DNA specifici per il sesso fornisce ulteriore supporto all'esistenza di cromosomi sessuali nel dattero (Younis et al. 2008). Phoenix comprende altre 12 specie oltre a *P. dactylifera*, la maggior parte delle quali coltivate come piante ornamentali, rendendo questo un sistema promettente per lo studio dell'evoluzione dei cromosomi sessuali.

Cannabaceae. Le Cannabaceae comprendono solo sei specie, *Humulus lupulus, H. japonicus, H. yunnanensis, Cannabis sativa, C. indica e C. ruderalis*. La determinazione del sesso è simile a quella di *Rumex* e coinvolge un sistema di bilanciamento X/autosome (X/A) o cromosomi sessuali multipli (XX/XY1Y2). Le analisi genetiche si sono concentrate principalmente sul luppolo comune, *H. lupulus*, e sulla canapa, *C. sativa*, per le quali sono stati descritti marcatori specifici per il maschio (Danilova and Karlov 2006; Jakse et al. 2008; Polley et al 1997; Seefelder et al 2000). Non sono ancora stati condotti studi molecolari su *H. japonicus*, ma i cromosomi sessuali possono essere facilmente identificati dopo la colorazione C-banding/DAPI, suggerendo un'arricchimento in sequenze ripetitive (Grabowska-Joachimiak et al. 2011). Infatti, diverse centinaia di retrotrasposoni sembrano essersi accumulati nella regione terminale del braccio lungo del cromosoma Y della canapa (Sakamoto et al. 1995). In *H. lupulus* i cromosomi sessuali sono eteromorfi e possono essere distinti dalle dimensioni leggermente più piccole del cromosoma Y e dalla colorazione differenziale con la colorazione DAPI (Karlov et al 2003). L'analisi dei marcatori intorno al locus determinante il sesso indica che la soppressione della ricombinazione è limitata a una piccolissima regione (Seefelder et al. 2000).

Caryophyllaceae. *Silene* rimane il sistema modello vegetale più studiato per l'evoluzione dei cromosomi sessuali, anche se le dimensioni ampie del genere (750 specie) e la filogenesi poco compresa ostacolano l'interpretazione delle direzioni evolutive. La dioecia e i cromosomi sessuali si sono evoluti diverse volte nel genere: alcune specie hanno cromosomi sessuali eteromorfi e altre omomorfi. Una clade dioica è costituita dalle specie *S. diclinis, S. dioica, S. heuffelii, S. latifolia e S. marizii*, tutte precedentemente considerate in passato con cromosomi sessuali simili (Nicolas et al. 2005). Tuttavia, dati recenti rivelano che *S. diclinis* ha cromosomi sessuali che hanno avuto origine da una traslocazione reciproca tra il cromosoma Y originale (omologo al cromosoma Y di *S. latifolia*) e un autosoma (Howel et al. 2009; Janousek and Mrackova 2010).

Cucurbitaceae. Nelle Cucurbitaceae, si sono evoluti cromosomi sessuali specializzati in *Bryonia* ed *Ecballium*, dove sono omomorfi, e in *Coccinia* e *Trichosanthes*, dove sono eteromorfi. *Bryonia dioica* è stato il primo organismo per il quale è stata fornita evidenza sperimentale di un sistema di determinazione del sesso XY attraverso incroci tra il dioico *B. dioica* e il monoico *B. alba* (Correns 1903). Lo studio filogenetico delle dieci specie di *Bryonia* indica che i cromosomi sessuali di *B. dioica* sono giovani, cioè di pochi milioni di anni. Finora non ci sono prove di poliploidizzazione o altre ristrutturazioni cromosomiche importanti, e il numero di cromosomi di *B. dioica* [$2n = 2x = 20$ (n = numero di cromosomi in un gamete aploide e x = numero di cromosomi monoploidi)] è lo stesso di quello della maggior parte delle altre specie monoiche e dioiche di *Bryonia* (Volz and Renner 2008). Le evidenze molecolari di due marcatori SCAR legati al maschio e il walking cromosomico hanno rivelato una disposizione complessa di elementi trasponibili simili a Copia sul cromosoma Y (Oyama et al. 2010). *Ecballium elaterium* ha popolazioni dioiche e monoiche presenti in tutta la regione mediterranea. Le due forme sono interfertili, e gli incroci tra *Ecballium* monoici e dioici hanno dimostrato l'eterogametia maschile (Westergaard 1958; Galan 1946).

Polygonaceae. I sistemi riproduttivi nel genere *Rumex*, che conta 200 specie, includono ermafroditismo, poligamia, ginodioecia, monoecia e dioecia. Nelle specie di *Rumex* dioiche, sono stati descritti due sistemi cromosomici sessuali e meccanismi di determinazione del sesso: XX/XY con un cromosoma Y attivo (ad esempio, *R. acetosella*) e XX/XY1Y2 con determinazione del sesso basata sul rapporto X/A (ad esempio, *R. acetosa*). I cromosomi sessuali sono i più grandi del cariotipo, con il cromosoma X leggermente più grande di ciascun cromosoma Y.I diploidi o i poliploidi con un rapporto X/A di 6 a 1 sono femmine, mentre quelli con un rapporto X/A di 0,5 sono maschi.

Rosaceae. I cromosomi ZW della fragola selvatica, *Fragaria virginiana*, si trovano allo stadio più precoce dell'evoluzione dei cromosomi sessuali (Spigler et al. 2010). I due geni di determinazione del sesso, che causano rispettivamente sterilità maschile recessiva e fertilità femminile dominante, sono stati mappati sul linkage group 6 e stanno ancora ricombinando con una distanza genetica di 5,6 cM.

Salicaceae. Il sequenziamento del genoma di *Populus trichocarpa* ha rivelato la presenza di cromosomi sessuali (Tuskan et al. 2006; Yin et al. 2008). I cromosomi ZW mostrano soppressione della ricombinazione intorno alla regione determinante del sesso, che si trova su un'estremità (regione telomerica) del cromosoma 19. La FSW comprende 706 kb e mostra una completa soppressione della ricombinazione. Nella regione immediatamente vicina, un segmento di 6,8 cM del cromosoma W è composto da 257 kb (37,8 kb/cM) di sequenza, mentre un segmento di 15,8 cM della corrispondente regione sul cromosoma Z consiste di 231 kb (14,6 kb/cM), una riduzione del tasso di ricombinazione del 155% sul cromosoma W.

Gimnosperme

Tra le 1010 specie di Gimnosperme conosciute, sono stati identificati i cromosomi sessuali solo per sei specie (0.6%) provenienti da 3 famiglie differenti, le quali presentano tutte cromosomi sessuali eteromorfi (Ming et al. 2011). Tale caratterizzazione di eteromorfismo si basa principalmente sulla rilevazione di deboli satelliti situati al termine dei bracci dei cromosomi sessuali. Nel caso specifico, il cromosoma X evidenzia un satellite a livello di entrambe le estremità dei bracci cromosomici, mentre il cromosoma Y presenta un satellite ad una sola estremità. Ad esempio, in *Cycas pectinata*, il cromosoma X mostra un debole satellite ad una estremità, del tutto assente invece sul cromosoma Y (Abraham and Mathew 1962). Un'osservazione simile è stata riportata per *Ginkgo biloba*, in un paio di brevi regioni subterminali (Lee 1954; Pollock 1957), ma questo dato è stato contestato da studi più recenti (Chen et al. 1987; Lan et al. 2008).

Cycadaceae. Le piante femminili e maschili di *Cycas revoluta* hanno cromosomi eteromorfi distinti XY (Segawa et al. 1971). Questa specie ha 10 paia di autosomi e un paio di cromosomi sessuali. Il cromosoma X è il quarto più grande cromosoma e il cromosoma Y il nono in ordine crescente di lunghezza.

Ginkgoaceae. In *Ginkgo biloba*, la ricerca dei cromosomi sessuali viene condotta da più di 50 anni. Le piante di *Ginkgo* maschio e femmina hanno 22 autosomi e una coppia di cromosomi sessuali ZW. Newcomer (1954) è stato il primo a riconoscere che i cromosomi sessuali sono i più grandi nel corredo cromosomico della specie e che presentano un dimorfismo sulla base della posizione del centromero: uno è metacentrico e l'altro sub metacentrico. Tuttavia, Newcomer aveva erroneamente ipotizzato un sistema del tipo XY, con il maschio eterogametico. Solo nel 1987 è stato riconosciuto in *G. biloba* un sistema ZW, caratterizzato da una femmina eterogametica (Chen et al. 1987; Lan et al. 2008).

Podocarpaceae. Le piante diploidi femminili e maschili di *Podocarpus macrophyllus* hanno un diverso numero di cromosomi con 2n = 38 nelle femmine e 2n = 37 nei maschi (Hizume 1988). Tutti i 38 cromosomi nelle femmine sono telocentrici, mentre 36 cromosomi nei maschi sono telocentrici e il 37° è un grande cromosoma submetacentrico.

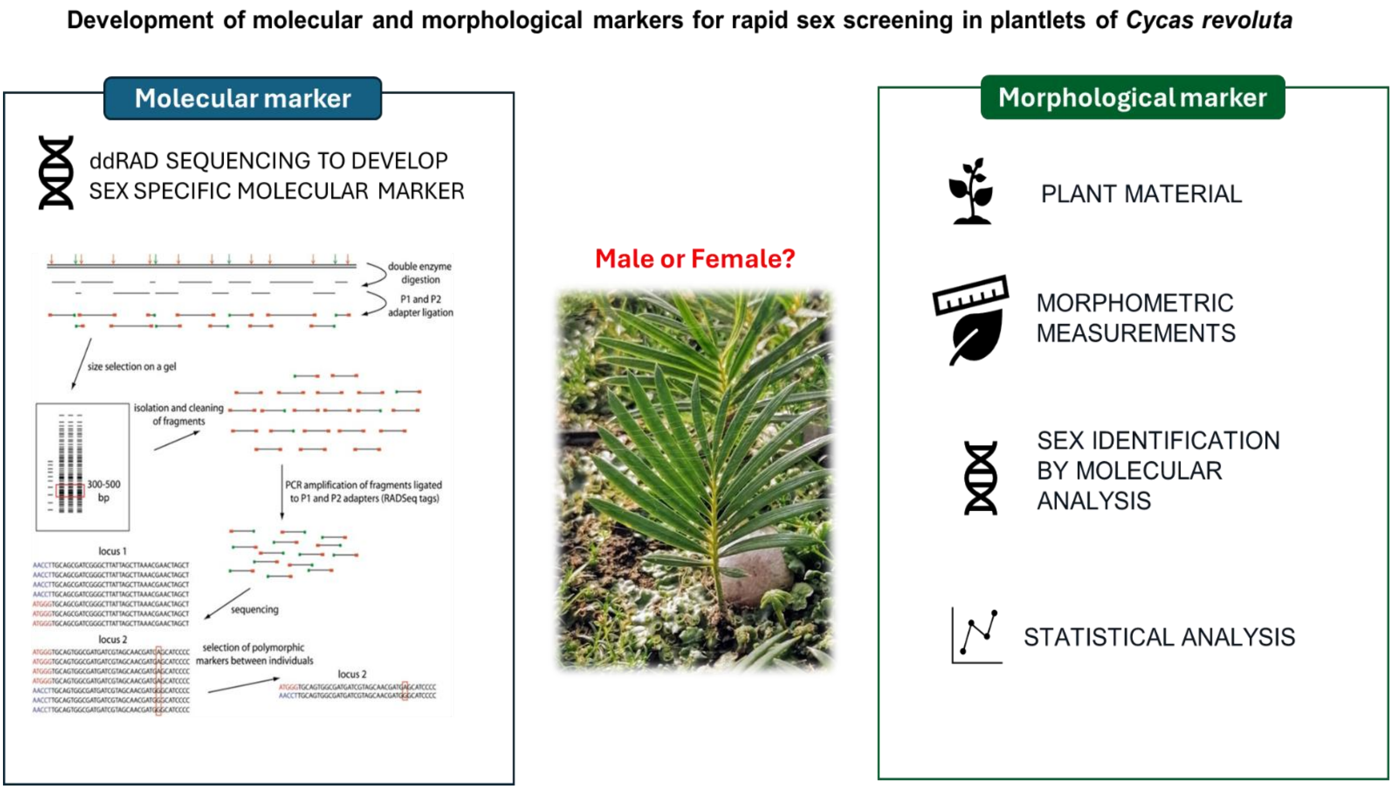

Fig. 4 General experimental design of thesis

3. Early sex identification by leaflet distance in plantlets of *Cycas revoluta*

Ylenia De Luca · Salvatore Cozzolino · Antonia Cristaudo · Alex Widmer · Donata Cafasso

Y. De Luca · S. Cozzolino (✉) · D. Cafasso
Department of Biology, University of Naples Federico II, Complesso Universitario MSA, 80126 Naples, Italy
Y. De Luca
e-mail: ylenia.deluca@unina.it
S. Cozzolino
e-mail: cozzolin@unina.it
D. Cafasso
e-mail: donata.cafasso@unina.it

A. Cristaudo
Department of Biological, Geological and Environmental Sciences, University of Catania, 58 Empedocle St., 95128 Catania, Italy
e-mail: acristau@unict.it

A. Widmer
Department of Environmental Systems Science, ETH Zurich, 16, Universitätstrasse St., 8092 Zurich, Switzerland
e-mail: alex.widmer@usys.ethz.ch

Acknowledgments: We are very grateful to Piante Faro (Giarre, Italy) for providing us with the greenhouse structures and plants. We are also grateful to Delia Terranova for her help in collecting the plant material.

Funding: YDL was funded by PON Dottorato Industriale «Ricerca e innovazione 2014-2020» – Asse "Capitale Umano" (DOT1318787). SC and DC were funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4 - NextGenerationEU; Award Number: Project code CN_00000033, "National Biodiversity Future Center - NBFC."

Abstract

In dioecious plants, the process of sex determination usually occurs during the reproductive stage. However, it can be challenging to determine the sex of long-lived plants with long generation times. This is true for *Cycas revoluta*, which produces toxic seeds in female plants, leading to a preference for male plants in public green spaces. In this study we aim to identify a morphological trait that can be used to distinguish between the sexes in *C. revoluta* at a very early stage of plant growth. We sampled *C. revoluta* seedlings/plantlets at three different early growth stages and identified the sex of individuals by PCR amplification of a male-specific molecular marker. On the same plants, we measured morphological traits, including the distance between leaflets on the leaf rachis, perimeter, area, number of indents, and leaf complexity, and evaluated their correlation with the sex of the individuals. Among all measured traits, the medium distance between leaflets on the leaf rachis was found to be a sexually dimorphic morphological trait in *Cycas* plantlets in all three growth stages. In particular, the distance between leaflets on the leaf rachis in the upper part of the leaf in the first stage revealed no overlap between the sexes. Whenever available, morphological traits for sex differentiation in greenhouses and plant nurseries can represent a rapid and economic tool when coupled with automatized image recognition systems.

Keywords Cycads · plant sex determination · plant sexual dimorphism · secondary sexual characters· sex-linked molecular marker

3.1 Introduction

The difference in the prevalence of hermaphroditism and monoecy, where a single organism expresses both sexes, versus dioecy, where an individual is either male or female, is one of the most important factors distinguishing plants and animals in terms of sexual evolution (Mank 2022). Whereas most animals and the presumed ancestor are dioecious, this condition is derived and relatively rare in most land plants (Käfer et al. 2017; Renner and Müller 2021; Renner and Ricklefs 1995). While dioecy is rare in plants, it is nevertheless an evolutionary outcome of (basal) hermaphroditism in some lineages (Renner 2014). The origin of separate sexes from a hermaphrodite ancestor is commonly associated with the evolution of sexual dimorphism, and this has occurred to varying degrees in many dioecious plants (Correns 1929; Dawson and Geber 1999; Lloyd and Webb 1977). Sexual dimorphism refers to differences between the sexes in primary and secondary sex characters. The former refers directly to male (androecium) and female (gynoecium) sex organs, and the latter refers to differences between the sexes in structures other than the sex organs themselves, including any aspect of morphology or physiology (Barett and Hough 2012). There are plenty of examples of sexual dimorphism in reproductive traits in dioecious species (Delph 1999; Eckhart 1999; Lloyd and Webb 1977). Sex-specific differences include flowering phenology and periodicity (Thomas and LaFrankie 1993), flower size (Delph et al. 1996), number of flowers per plant (Delph et al. 2005), flower longevity (Primack 1985), flower scent (Ashman 2009), flower defence against herbivory (Cornelissen and Stiling 2005) and various inflorescence traits including total flower number (Barrett 2012), daily floral display size (Yakimowski et al. 2011), and inflorescence architecture (e.g. Rourke 1989). In contrast, differences in secondary (vegetative) sex characters are less common and evident. This is partly because plant sexual dimorphism is usually less obvious than in most animals.

At the seed or seedling stage, there are few reports of differences in sexual secondary characters. In *Rumex nivalis*, male seeds are heavier and germinate earlier than female seeds, but overall germination levels do not differ between the sexes (Stehlik and Barrett 2005). Male seeds are also heavier than female seeds in *Spinacia oleracea* (Freeman et al. 1994). On the other hand, several reports of size, morphology (leaf form, stem features, etc.), growth rate, and physiological differences between the sexes during the vegetative phase of growth have been reported (Dawson and Geber 1999; Lloyd and Webb 1977). However, for most dioecious plant species, sex differences only become evident in reproductive traits. Consequently, and in contrast to many animal groups, the sex of an individual in plants cannot usually be determined before the beginning of its reproductive stage (Garcia and Antor 1995), a major complication for early sex identification in long-lived plants with a long generation time. These become particularly relevant for those cultivated plants that require sex information well before the development of the reproductive organs. For instance, females of *Ginkgo biloba*, a dioecious tree that typically does not reach sexual maturity for 20 to 30 years (van Beek et al. 1998), produce seeds that emit a noxious, foul odour on falling to the ground (Wada and Haga 1997). These features caused city governments to remove and ban females from being planted (Echenard et al. 2008) and plant growers to have a strong interest in determining the sex of plants at an early stage of development. A similar case is that of Cycads, the most primitive extant seed-producing plants. These 'living fossils' are known to have existed at least since the Jurassic (Zhou and Zheng 2003) and, with their fascinating forms, Cycads are popular in landscape gardening. Among others, *Cycas revoluta* is used worldwide as an ornamental plant. *C. revoluta* has a distinctive trunk-like structure, resembling a palm tree and its foliage contains a cluster of long pinnate leaves that give it an exotic, tropical appearance. Although *C. revoluta* is an attractive ornamental plant, its seeds contain several toxic glycosides, as cycasin and neocycasin, which belong to the family of cyanogenic glycosides that pose a serious risk of poisoning through accidental ingestion by kids and pets (Nishida et al. 1956). A study highlighted the clinical cases of *Cycas* seed poisoning in Taiwan and its cyanogenic potential (Chang et al. 2004). Thus, it is desirable to introduce only *C. revoluta* male plants in public green areas and to have tools for their early sex identification.

Traditionally, sex identification of *C. revoluta* is done by visual inspection of the reproductive structures of mature adult plants, i.e., after the formation of male and female cones. Cycads plants only reach sexual maturity after ca. 15-20 years (Segalla et al. 2021) and this can potentially result in significant costs for the cultivation of female plants, which should subsequently be discarded. Therefore, identifying the sex of *C. revoluta* plants at an early stage of plant development could be valuable for breeding and planning of plant production.

Once sex-linked genes/markers are known, plants, even at the seedling stage, can be genotyped for their sexes (Heikrujam et al. 2015). However, the availability of morphological tools would be preferable over molecular tools for their efficient application in nurseries and greenhouses (Inoti et al. 2015). Attempts have been made to identify dimorphic secondary sexual characters in Cycads, particularly in leaf morphology, but significant differences between sexes have never been reported at the juvenile stage (Newell 1985; Niklas and Marler 2008; Ornuduff 1996).

More recently, Kaviani et al. (2014) compared a set of morphological traits between male and female two-year-old off-shoots of adult *C. revoluta* plants whose sex was previously determined (i.e. in plants that have already reached sexual maturity). Among others, they found the distance between leaflets 29 and 30 on the leaf rachis to be significantly different between sexes in male and female off-shoots. In the present work, we examined whether this trait and other related leaf traits such as area, perimeter, number of indents, and leaf complexity were different at the juvenile stage (from seedlings to plantlets) in *Cycas revoluta* plants whose sex was previously determined with a male sex-specific PCR marker (Liu et al. 2022). We aimed to find a vegetative morphological trait that could be used in greenhouses and plant nurseries as a rapid tool for easy identification of sex in *C. revoluta* plants at the earliest stage of development.

3.2 Materials and methods

3.2.1 Plant material

We sampled *C. revoluta* plantlets in March 2022 at the greenhouse Piante Faro (Giarre, Sicily, Italy). Seedlings of *C. revoluta* were randomly sampled from the nursery greenhouse. Specifically, thirty individuals for three different growth stages were sampled from a pool of approximately 1000 individuals. We defined individuals belonging to stage 0 as those with leaf lengths varying between 20 and 50 mm, individuals belonging to stage 1 as those with leaf lengths varying between 50 and 100 mm, and individuals belonging to stage 2 as those with leaf lengths varying between 100 and 170 mm. These three stages roughly corresponded to seedlings of approx. 12 months (stage 0), 24 months (stage 1), and 36 months (stage 2) after seed germination, respectively (Fig. S1). We collected the oldest leaf from each individual (the single leaf at stage 0).

In April 2024, we sampled an additional batch of 40 plants (all belonging to stage 0) at the greenhouse La Floricola (Ispica, Sicily, Italy).

3.2.2 Morphometric measurements

To measure the morphological traits, collected leaves were pressed and scanned with a 600-dpi resolution on graph paper. ImageJ software was used to measure the distance between leaflets on the leaf rachis for each leaflet pair (i.e. Inter Rachis Distance, IRD). To further characterize leaf morphology, leaf scans were also analyzed using Lamina software (Bylesjö et al. 2008), which generates estimates of leaf area, perimeter, and the number of indentations (indentations and leaf lobes). To estimate leaf complexity, we calculated perimeter^2 by area, where lower numbers indicate less complexity (i.e. more intact) leaves. It was impossible to measure these traits for stage 2 individuals because their large leaves were leathery, preventing complete relaxation on the flat surface of the scanner.

3.2.3 Sex identification by molecular analysis

DNA from all individuals was extracted using a cetyltrimethylammonium bromide extraction protocol (Doyle and Doyle 1990) and quality and quantity were evaluated with the Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies, US). To determine the sex of individuals, a male-specific MADS-Y/CYCAS_034085 region was amplified by PCR using a sex-specific primer pair as described in Liu et al. (2022). With this primer pair, we expected to amplify a 720 bp fragment only in male individuals. We also amplified the autosome CYCAS_010388 region as a control both in males and females (Liu et al. 2022). As a preliminary control for PCR primers and amplification, we first examined 30 DNA samples from adult *C. revoluta* plants whose sex was already known by visual observation of their reproductive traits. All PCR reactions (25 µl final volume), with 10 ng of DNA as a template, 50 nM of each primer, 1× Taq polymerase buffer, 1.5 mM MgCl2, 200 µM dNTP, and 0.5 unit Taq polymerase (Life Technologies, US), were conducted in an Applied Biosystems 2720 thermal cycler (Thermo Fisher, US) as follows: 3 minutes at 94° followed by 35 cycles of 30 s at 94°, 45 s at 54°, 50 s at 72° and 5 min at 72° for final extension. PCR products were then examined by electrophoresis in 1.5% agarose gel with a DNA molecular size marker.

3.2.4 Statistical analyses

The Shapiro-Wilk test was applied to all morphological traits to determine whether the values were normally distributed. If the p-value was less than p = 0.05, there was sufficient evidence that the values were from a non-normally distributed population. In that case, we normalized the data by transforming them to a logarithmic scale (log). For normally distributed values, the Linear Regression Model (LM) and Linear Mixed Model (LMM) were applied using the lme4 package (Bates et al. 2015) in R (R Core Team 2023), where the fixed variable is the sex of the two groups (males and females), the dependent variable corresponds to the measure of the morphological traits, and the random factor for the LMM analysis corresponds to the individuals. The Akaike Information Criterion (AIC) was used to choose the model with higher relative quality (package stats in R; Bertrand 1988). To test whether the morphological differences between the two sexes were significant, we used emmeans version 1.4 (Lenth et. al 2019) to conduct pairwise t-tests for multiple comparisons. For non-normally distributed values we used the generalized linear mixed model (GLMER), using the GAMMA function as a family, the morphological traits being quantitative, continuous, and discrete values. In this case, to assess whether there was a significant difference between the two sexes, we conducted the Mann-Whitney-U test in R. All statistical analyses were conducted with R version 4.2.3.

3.3 Results

A positive amplification of both male-specific MADS-Y/CYCAS_034085 region and autosome CYCAS_010388 was an indication of the male sex of the plantlets while the absence of the male-specific PRC product (and a positive amplification of the control target) was taken as an indication of the female sex (Fig. S2). Regarding the plantlets sampled at the greenhouse Piante Faro, for stage 0 we were able to assess the sex of 12 males and 16 females, (two samples failed PCR amplification), for stage 1 we assessed the sex of 20 females and 10 males, and for stage 2 we assessed the sex of 16 males and 14 females. For plantlets sampled at the greenhouse La Floricola, we assessed the sex of 18 females and 22 males. Once the sex was determined, in the first analysis, we compared the IRD between each pair of leaflets on the leaf rachis for all plants and stages (Table S1). We found different values for the two sexes but always with overlapping values. We also assessed whether medium leaf IRD (as the medium of all IRD measurements of the leaf) correlated with the sex of the plant. The results show that there is a statistically significant difference between the two sexes but with an overlapping distribution of values (Fig. 1). Finally, we analysed separate portions of the leaf by dividing it into three equal sections, namely the proximal, the median, and the distal part of the leaf petiole.

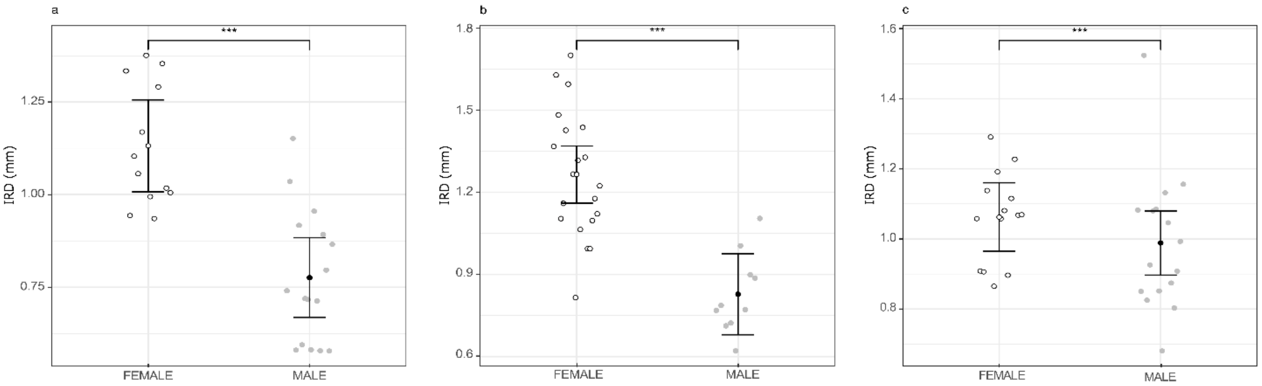

Fig. 1 Dimension of leaf IRD (as the medium of all measurements for each leaf with error bar, i.e. 95% confidence interval of the mean) in female and male plants at a) stages 0 b) stage 1 and c) stage 2. White dots correspond to female individuals and grey dots represent male individuals. *** (p-value <0.01)

Our analyses revealed a significant correlation between the sex and IRD for all three leaf sections at stage 0. However, only for the section corresponding to the distal part of the leaf petiole, male and female trait measurements did never overlap and were discontinuous (t = 6.8996, df = 73.907, p-value = 1.519e-09) (Fig. 2a). Instead, for the sections corresponding to the median and proximal part of the leaf petiole, the correlation was significant (t = 4.8991, df = 72.974, p-value = 5.624e-06; t = 4.8684, df = 56.434, p-value = 9.454e-06, respectively) but trait distribution between sexes overlapped due the presence of some outliers (Fig. 2b and 2c).

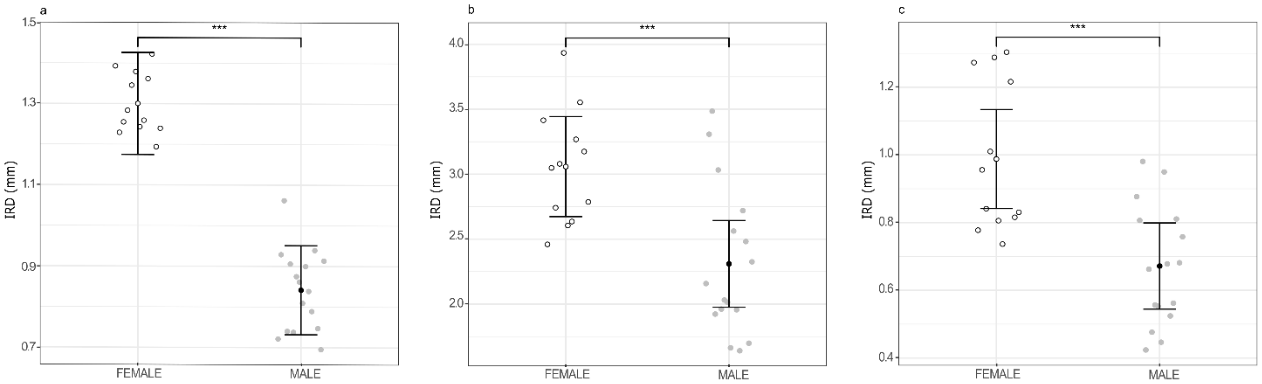

Fig. 2 Dimension of leaf IRD (as the medium of all measurements for each section with error bar, i.e. 95% confidence interval of the mean) in female and male plants at stage 0. White dots correspond to female individuals and grey dots represent male individuals. a) distal part of the leaf petiole, b) median part of the leaf petiole, and c) proximal part of the leaf petiole. *** (p-value <0.01)

At plant stages 1 and 2, statistical analyses for the same three sections found significant correlations between IRD and the sex of the individuals but always with an overlapping distribution of trait measurements (Table 1 and Fig. S3). We did not find significant differences between sexes for the other measured morphological traits, such as area, perimeter, number of indents, and leaf complexity (Table S2, Fig. S4, Fig. S5).

Table 1 Results from the t-test or Wilcoxon rank sum test for significant correlation between the sex of individuals and the IRD for plants of stages 1 and 2. 'W' represents the sum of the rankings of values from one of the two groups in the comparison in Wilcoxon rank sum test; 't' represents the ratio of the difference between the means of the two groups in t-test; 'df' indicates degrees of freedom of the test

Stage	Section of the leaf	df	t	p-value	W
1	Distal part of the leaf petiole	110	7.8486	$2.979e^{-12}$	
1	Median part of the leaf petiole	117	7.3342	$3.161e^{-11}$	
1	Proximal part of the leaf petiole			$9.722e^{-13}$	2360
2	Distal part of the leaf petiole	267.25	2.6672	0.008115	
2	Median part of the leaf petiole	271.68	2.2602	0.0246	
2	Proximal part of the leaf petiole	270.37	2.2263	0.02682	

In April of 2024, we conducted a replicate analysis on a different batch of 40 plants (stage 0) that were gathered from La Floricola greenhouse. Based on molecular analysis, we identified 22 males and 18 females and measured their IRD. The results were consistent with those obtained in the prior analysis, as illustrated in Figure S6 and Table S3.

3.4 Discussion

We identified a vegetative morphological trait that reliably discriminates between the sexes of *C. revoluta* at the seedling stage. In a different way from our work, Kaviani et al. (2014) measured only the interval between leaflets number 29 and 30 of the leaf rachis of adult leaves in off-shoots and found a non-overlapping distribution between the two sexes (Table 2 in Kaviani et al. 2014). In the three plant stages that we examined, we never found any single interval between leaflets that showed a non-overlapping distribution of the IRD (Table S1). However, when we averaged the intervals between leaflets only at the distal part of the leaf petiole (i.e. corresponding to the upper 2-3 intervals at stage 0), we found that, at the juvenile stage 0, the trait has a clear distinction between male and female *C. revoluta* young individuals with no overlapping values between sexes. This suggests that sexual dimorphism in the leaf is already present very early in plant life. However, only a specific section of the leaf (i.e. the distal part) has to be examined for accurate sex determination of *C. revoluta* seedlings. We achieved highly concordant results in two different batches of plants with distinct genetic backgrounds and growth conditions, indicating that this sex-specific trait is not affected by the environment or plant source.

We found that IRD is greater in female individuals but with some exceptions in stage 2 (Table S1). The male leaflets are located almost opposite each other on the central rachis, while the female leaflets are alternate. A complex genetic and molecular network controls the formation of the compound *Cycas* leaf. Cycad leaves resemble fern leaves due to their prolonged apical growth, circinate tips, and leaflets (Stevenson 1990). The marginal position of the leaflets' veins indicates that the leaflets develop exclusively through marginal meristem activity, like in *Ginkgo* (Boyce 2005; Boyce and Knoll 2002). Gene expression of C1KNOX transcription factors (TFs) in *Zamia integrifolia* (Cycadales) was found at the sites of leaflet initiation in older leaf primordia (Bharathan et al. 2002). This may suggest that C1KNOX TFs in Cycads can play a role in the prolonged apical growth of leaves and the formation of leaflets by establishing new meristems (Romanova et al. 2023). KNOX factors also play a crucial role in the formation of folioles in angiosperms (Hasson et al. 2010). Leaflet formation requires the maintenance of an undifferentiated environment, determined by KNOX factors, and a local accumulation of auxin (Hasson et al. 2010). More precisely, peaks of auxin response contribute to the outgrowth of leaflet primordia, whereas repression of the auxin response is required to inhibit the outgrowth of the regions between leaflets (Barkoulas et al. 2008; Koenig et al. 2009). It could be that plant sex determines a different response to auxin in the early stages of leaf development, contributing to the development of the observed morphological trait dimorphism. It is interesting to note that while the relationship between plant hormones and sex determination is not yet fully understood in gymnosperms, it has been well documented in flowering plants (i.e. angiosperms) (Song et al. 2013; Yamasaki et al. 2003). Genetic and physiological aspects could therefore explain the presence of sex dimorphism in IRD. Sex dimorphism could evolve as a consequence that males and females have different reproductive roles. Alternatively, a secondary sexual trait could have no direct adaptive significance or could even be counter-adaptive if it is a pleiotropic by-product of physiological or genetic aspects of the mechanism of primary sex determination (Lloyd and Webb 1977). We have no evidence to support an adaptive or non-adaptive significance for this leaf trait. Further studies are also needed to understand why the observed morphological dimorphism is not sustained as the plant grows.

When available, morphological traits as tools for sex differentiation are preferable over molecular identification of sex, particularly in nurseries where thousands of young plants are grown every year. In the next step, the identified phenotypic differences between the sexes could be used to develop a rapid screening tool. For instance, Li and Tang (2017) developed a low-cost system for the three-dimensional (3D) reconstruction of indoor plants and the characterization of their morphological traits such as collar height, length, and leaf area. Alternatively, leaf phenotyping and measurement can be performed directly in the field, using fast and low-cost LiDAR-based systems, such as those introduced by Panjvani et al. (2019) for the measurement of leaf traits (length, width, and area). Both approaches potentially overcome the need to scan the cycad leaves on a flat surface as we did. With our study, we have established the basis for sex identification in *C. revoluta* at an early stage of plant development by employing a sexually dimorphic (measurable) morphological trait. The subsequent step involves creating a faster and more cost-effective identification tool, ideally also in the form of an app for mobile devices based on image recognition and measurement. This functionality can be beneficial not only for setting an efficient and cost-effective system to assist large-scale horticultural companies that produce thousands of plants annually but also for enthusiasts in the field.

Author Contribution

YDL, DC, and AW conceived the project and designed experiments; YDL and SC wrote the manuscript; AC performed greenhouse plant measurements; YDL performed molecular analyses and analyzed data; DC administered the project; all the authors read and approved the manuscript contents.

3.5 Supplementary Information

Title: Early sex identification by leaflet distance in plantlets of *Cycas revoluta*

Journal name: Euphytica

Authors: Ylenia De Luca[1], Salvatore Cozzolino[1], Antonia Cristaudo[2], Alex Widmer[3], Donata Cafasso[1]

Authors' affiliation:

[1] Department of Biology, University of Naples Federico II, Complesso Universitario MSA, 80126 Naples, Italy

[2] Department of Biological, Geological and Environmental Sciences, University of Catania, 58 Empedocle St., 95128 Catania, Italy

[3] Department of Environmental Systems Science, ETH Zurich, 16, Universitätstrasse St., 8092 Zurich, Switzerland

Corresponding author: Salvatore Cozzolino, Department of Biology, University of Naples Federico II, Complesso Universitario MSA, 80126 Naples, Italy; E-mail: cozzolin@unina.it; Orcid-ID: 0000-0002-3176-8130

Author Contribution Statement

YDL, DC and AW conceived the project and designed experiments; YDL and SC wrote the manuscript; AC performed greenhouse plant measurements; YDL performed PCR analyses and analyzed data; DC administered the project; all the authors read and approved the manuscript contents.

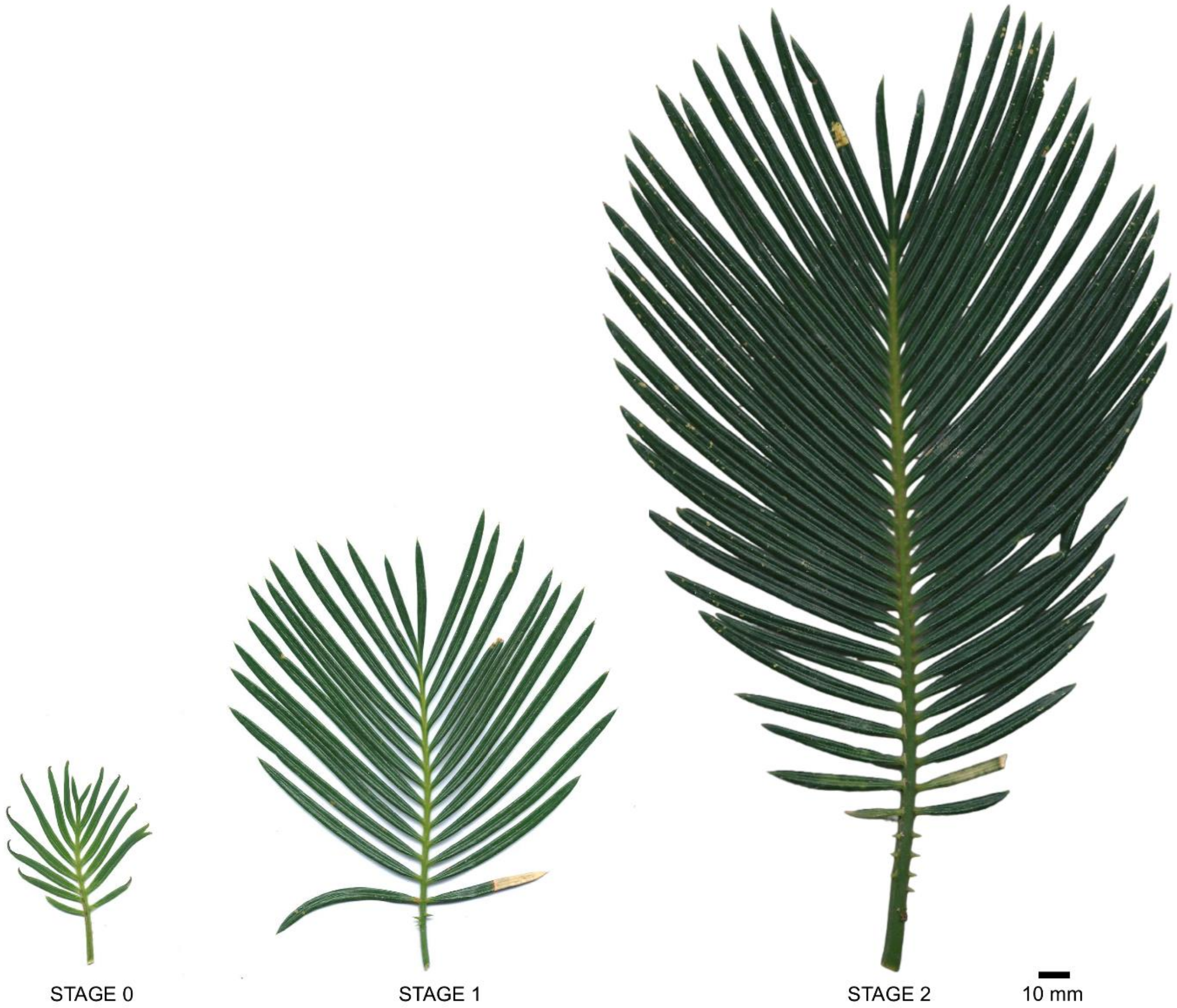

Fig. S1 Growth stages of *Cycas revoluta* seedlings. Stage 0 corresponds to plantlets of 12 months; stage 1 corresponds to plantlets of 24 months and stage 2 corresponds to plantlets of 36 months after seed germination. These plantlets were all grown in a greenhouse Piante Faro in Sicily

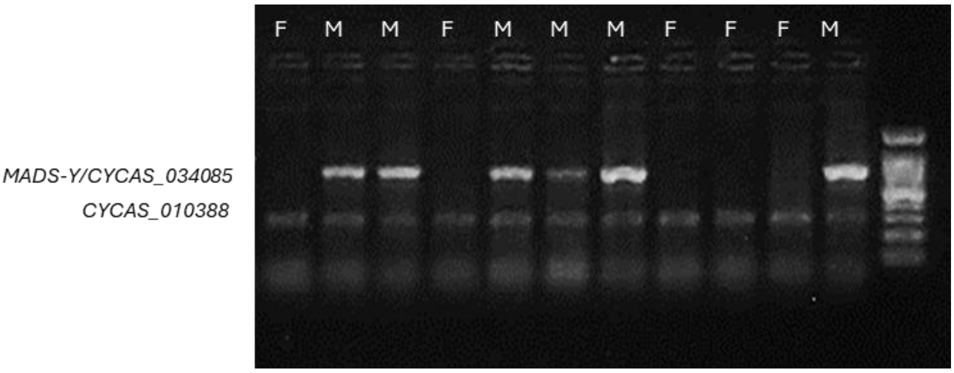

Fig. S2 PCR amplifications of the male-specific locus MADS-Y/CYCAS_034085 and locus CYCAS_010388 (control). The latter amplifies both in male and female plants. A 100 bp molecular ladder is in the rightmost lane.

Table S1 Dimension of IRD (mm) for any single interval between leaflets in stages 0,1 and 2. The minimum and maximum values for each interval are shown. "M" male individuals and "F" female individuals. These values correspond to the raw measures without log transformation for the statistical analysis

Leaflets interval	**IRD Stage 0**		**IRD Stage1**		**IRD Stage 2**	
1-2	M(1.22-2.83)	F(2.63-5.45)	M(1.52-2.93)	F(2.91-7.66)	M(1.56-4.58)	F(1.95.4.34)
3-4	M(1.79-4.11)	F(2.25-5.73)	M(2.21-4.11)	F(2.38-6.96)	M(1.85-3.62)	F(1.70-3.81)
5-6	M(1.33-3.41)	F(2.66-4.47)	M(2.08-4.11)	F(2.10-6.44)	M(1.49-3.49)	F(1.88-4.01)
7-8	M(1.17-3.76)	F(2.04-4.85)	M(1.70-4.16)	F(1.87-6.68)	M(1.30-4.51)	F(1.96-4.30)
9-10	M(1.53-4.10)	F(2.03-3.96)	M(1.50-4.40)	F(2.14-5.41)	M(1.90-4.31)	F(1.95-5.01)
11-12	M(1.64-3.46)	F(2.06-3.94)	M(1.53-3.40)	F(1.70-6.16)	M(1.97-4.46)	F(1.82-4.54)
13-14	M(1.32-3.62)	F(1.71-3.09)	M(1.21-3.34)	F(2.24-5.22)	M(1.62-5.03)	F(1.85-5.24)
15-16	M(1.53-2.95)	F(1.93-3.96)	M(1.46-3.38)	F(2.11-6.04)	M(1.41-3.62)	F(1.66-3.84)
17-18	M(1.51-2.70)	F(2.63-3.71)	M(1.69-3.08)	F(1.89-5.56)	M(1.69-3.62)	F(1.85-3.61)
19-20	M(1.20-2.81)	F(2.28-3.40)	M(1.60-2.56)	F(2.24-6.28)	M(1.80-4.72)	F(1.86-3.94)
21-22			M(1.62-2.95)	F(2.85-6.04)	M(1.60-4.00)	F(1.64-4.00)
23-24			M(1.71-2.34)	F(2.19-6.39)	M(1.51-4.71)	F(1.80-3.76)
25-26			M(1.78-2.71)	F(2.60-6.02)	M(1.68-5.32)	F(1.86-3.66)
27-28					M(1.93-4.31)	F(1.95-3.98)
29-30					M(1.53-3.47)	F(1.90-3.71)
31-32					M(2.06-4.40)	F(1.99-3.41)
33-34					M(1.80-4.72)	F(1.67-3.90)
35-36					M(1.94-4.08)	F(2.12-4.35)
37-38					M(1.85-4.72)	F(1.95-4.69)
39-40					M(1.78-6.07)	F(2.12-3.86)
41-42					M(1.51-5.46)	F(1.80-4.73)
43-44					M(1.37-4.88)	F(2.16-4.48)
45-46					M(1.77-6.57)	F(2.19-4.64)
47-48					M(1.81-6.72)	F(1.82-4.96)
49-50					M(2.08-6.10)	F(2.30-4.76)
51-52					M(1.80-5.48)	F(3.39-4.40)
53-54					M(2.02-4.82)	F(2.47-4.30)
55-56					M(2.51-3.88)	F(2.44-4.69)
57-58					M(2.72-3.89)	F(2.47-4.54)
59-60					M(3.25-4.28)	F(3.29-3.95)
61-62					M(3.25-4.28)	F(3.06-3.78)

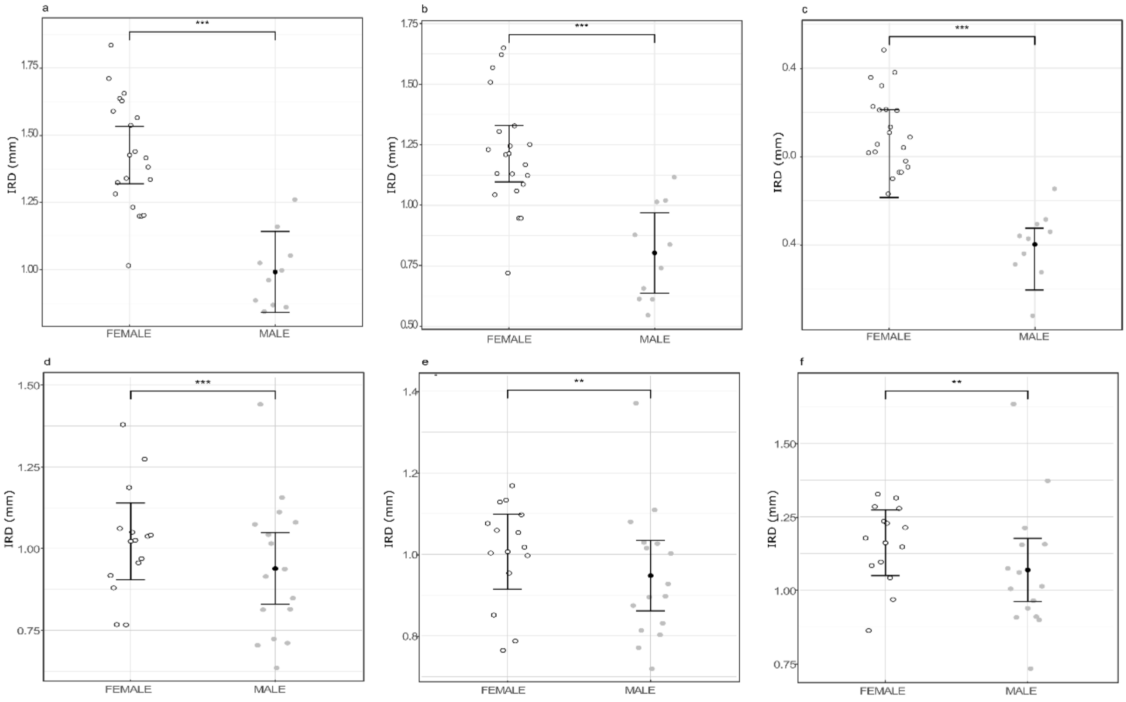

Fig. S3 Dimension of leaf IRD (as the medium of all measurements for each leaf section with error bar, i.e. 95% confidence interval of the mean) of female and male plants for stages 1 (top) and 2 (bottom) in the three leaf sections. (a,d) distal part of the leaf petiole, (b,e) median part of the leaf petiole, and (c,f) proximal part of the leaf petiole. White dots correspond to female individuals and grey dots represent male individuals. *** (p-value <0.01), ** (p-value <0.05)

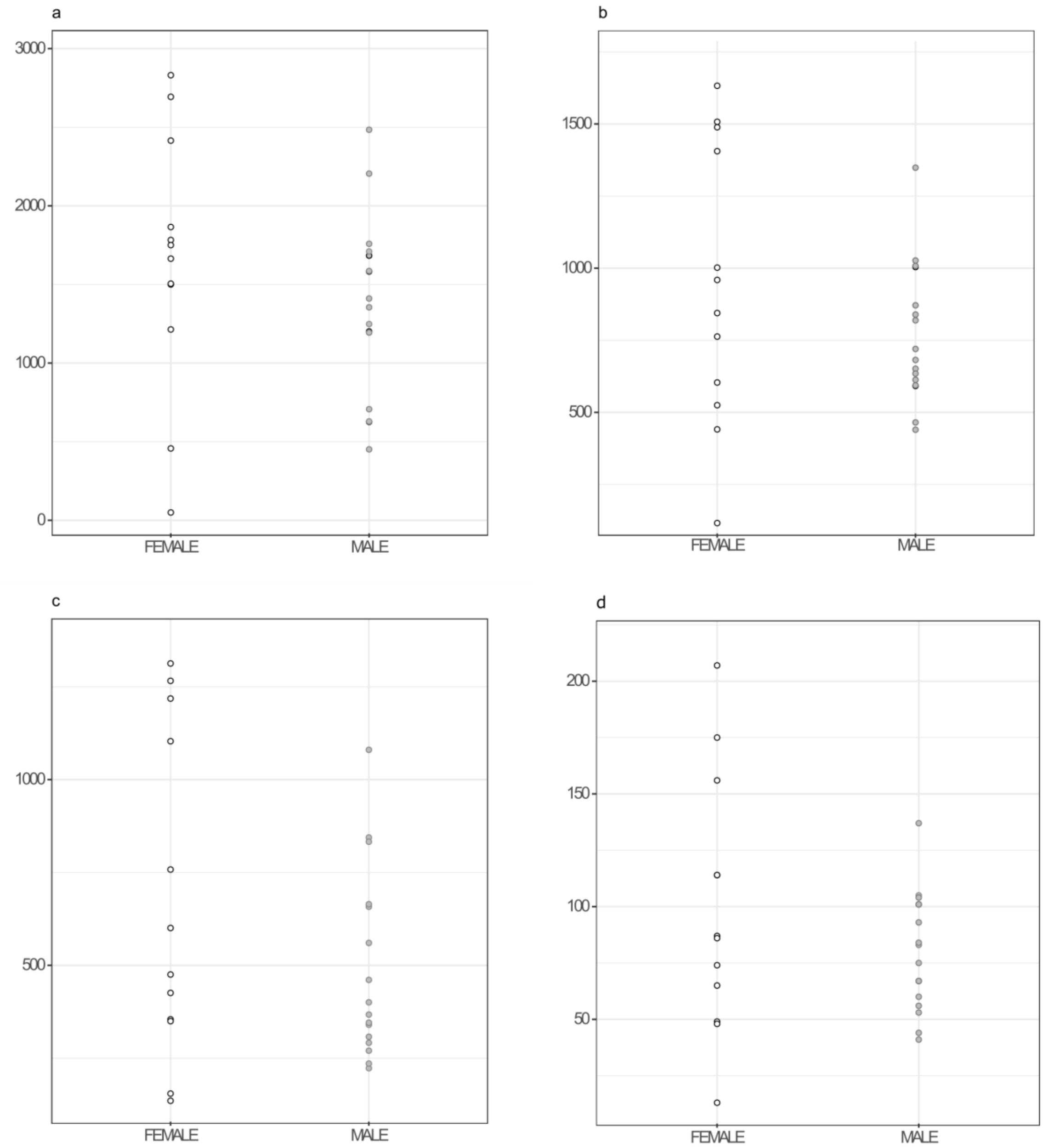

Fig. S4 Dimension of leaf traits a) area (mm^2), b) perimeter (mm), c) leaf complexity ($perimeter^2/area$), d) number of indents in female and male plants at stage 0

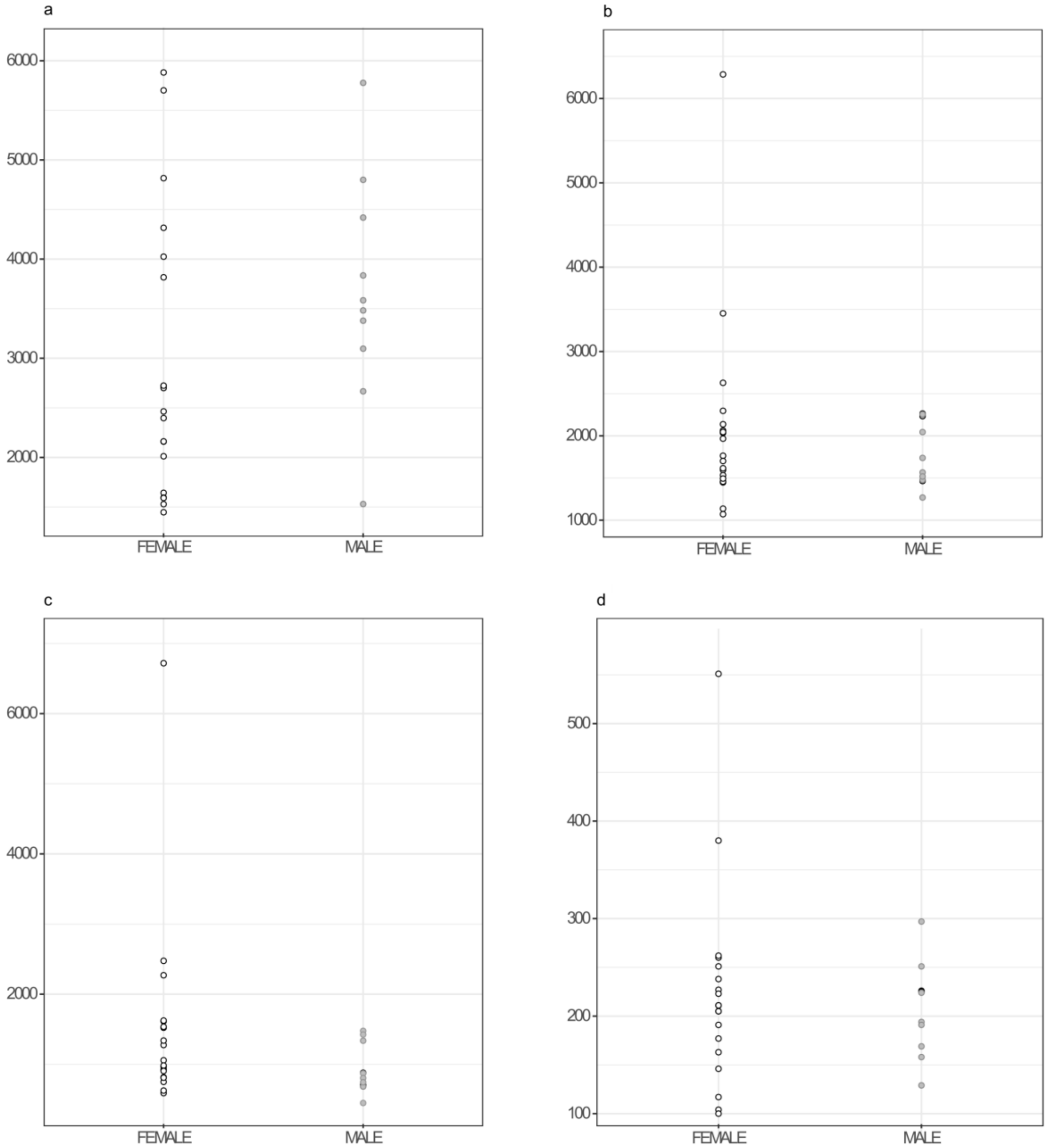

Fig. S5 Dimension of leaf traits a) area (mm^2), b) perimeter (mm), c) leaf complexity ($perimeter^2$/area), d) number of indents in female and male plants at stage 1

Table S2 Results from the t-test or the Wilcoxon rank sum test for significant correlation between the sex of individuals in plants of stages 0 and 1. ‘W’ represents the sum of the rankings of values from one of the two groups in the comparison in Wilcoxon rank sum test; ‘t’ represents the ratio of the difference between the means of the two groups in t-test; ‘df’ indicates degrees of freedom of the test

Leaf Trait	Stage	df	t	p-value	W
Leaf Complexity	0	26	0.81436	0.4228	
Leaf Complexity	1			0.0623	136
Perimeter	0	26	1.2348	0.2280	
Perimeter	1			0.7352	103
Area	0	26	1.0681	0.2953	
Area	1			0.1378	62
Number of indents	0	26	1.2161	0.2349	
Number of indents	1	27	0.11196	0.9117	

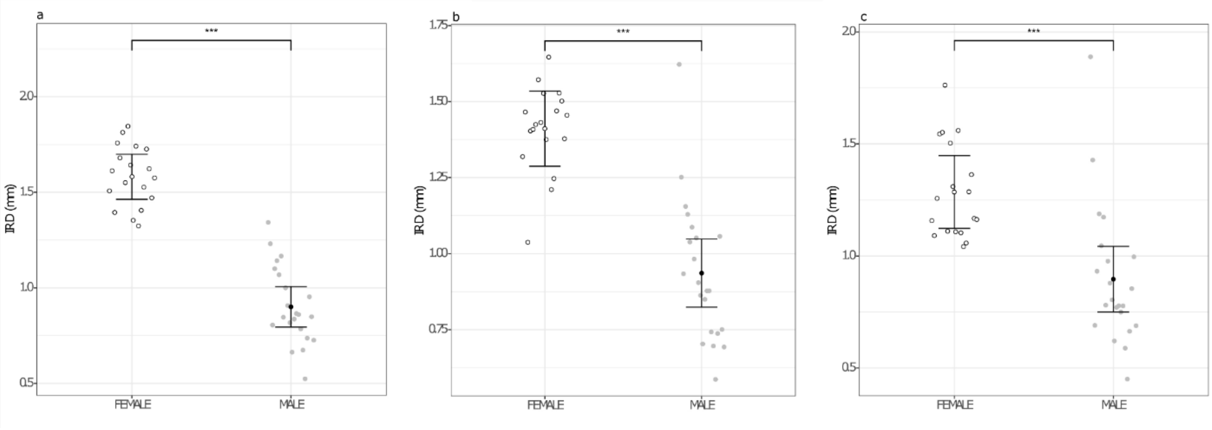

Fig. S6 Dimension of leaf IRD (as the medium of all measurements for each section with error bar, i.e. the 95% confidence interval of the mean) in female and male plants of the stage 0 collected at greenhouse La Floricola (Ispica, Sicily, Italy) in April 2024. White dots correspond to female individuals and grey dots represent male individuals. a) distal part of the leaf petiole, b) median part of the leaf petiole, and c) proximal part of the leaf petiole. *** (p-value <0.01)

Table S3 Results from the t-test for significant correlation between the sex of individuals and the IRD for plants of stage 0 collected at the greenhouse La Floricola (Ispica, Sicily, Italy) in April 2024. ‘t’ represents the ratio of the difference between the means of the two groups in t-test; ‘df’ indicates the degrees of freedom of the test

Stage	**Section of the leaf**	**df**	**t**	**p-value**
0	Distal part of the leaf petiole	150	20.802	$2.2e^{-16}$
0	Median part of the leaf petiole	165	9.3166	$2.2e^{-16}$
0	Proximal part of the leaf petiole	156	6.131	$6.867e^{-13}$

3.6 References

Ashman TL (2009) Sniffing out patterns of sexual dimorphism in floral scent. Funct Ecol 23:852–862. https://doi.org/10.1111/j.1365-2435.2009.01590.x

Barrett SCH, Hough J (2012) Sexual dimorphism in flowering plants. J Exp Bot 64:67–70. https://doi.org/10.1093/jxb/ers308

Barkoulas M, Hay A, Kougioumoutzi E, Tsiantis M (2008) A developmental framework for dissected leaf formation in the *Arabidopsis* relative *Cardamine hirsuta*. Nat Genet 40:1136–1141. https://doi.org/10.1038/ng.189

Bates D, Mächler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67:1–48. https://doi.org/10.18637/jss.v067.i01

Bertrand PV (1988) Review of akaike information criterion statistics. J R Stat Soc Ser A Stat Soc 151:567–568. https://doi.org/10.2307/2983028

Bharathan G, Goliber TE, Moore C, Kessler S, Pham T, Sinha NR (2002) Homologies in leaf form inferred from KNOXI gene expression during development. Science 296:1858–1860. https://www.doi.org/10.1126/science.1070343

Boyce CK (2005) Patterns of segregation and convergence in the evolution of fern and seed plant leaf morphologies. Paleobiology 31:117–140. https://doi.org/10.1666/0094-8373(2005)031<0117:posaci>2.0.co;2

Boyce CK, Knoll AH (2002) Evolution of developmental potential and the multiple independent origins of leaves in Paleozoic vascular plants. Paleobiology 28:70–100. https://doi.org/10.1666/0094-8373(2002)028

Bylesjö M, Segura, V, Soolanayakanahally RY, Rae AM, Trygg J, Gustafsson P, Jansson S, Street NR (2008) LAMINA: a tool for rapid quantification of leaf size and shape parameters. BMC Plant Biol 8. https://doi.org/10.1186/1471-2229-8-82

Chang SS, Chan YL, Wu ML, Deng JF, Chiu TF, Chen JC, Wang FL, Tseng CP (2004) Acute *Cycas* seed poisoning in Taiwan. J Toxicol Clin Toxicol 42:49–54. https://doi.org/10.1081/CLT-120028744

Correns C (1929) Bestimmung, Vererbung und Verteilung des Geschlechts bei den höheren Pflanzen. Z.Ver-erbungslehre 49:168. https://doi.org/10.1007/BF01847575

Cornelissen T, Stiling P (2005) Sex-biased herbivory: a meta-analysis of the effects of gender on plant-herbivore interactions. Oikos 111:488–500. https://doi.org/10.1111/j.1600-0706.2005.14075.x

Dawson TE, Geber MA (1999) Sexual dimorphism in physiology and morphology. In: Geber MA, Dawson TE, Delph LF (ed) Gender and sexual dimorphism in flowering plants. Springer, Berlin, Heidelberg, pp 175–215

Delph LF (1999) Sexual dimorphism in life history. In: Geber MA, Dawson TE, Delph LF (ed) Gender and sexual dimorphism in flowering plants. Springer, Berlin, Heidelberg, pp 149–173

Delph LF, Galloway LF, Stanton ML (1996) Sexual dimorphism in flower size. Am Nat 148:299–320. https://www.doi.org/10.1086/285926

Delph LF, Gehring JL, Arntz MA, Levri M, Frey FM (2005) Genetic correlations with floral display lead to sexual dimorphism in the cost of reproduction. Am Nat 166:S31–S41. https://doi.org/10.1086/444597

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Echenard V, Lefort F, Calmin G, Perroulaz R, Belhahri L (2008) A new and improved automated technology for early sex determination of *Ginkgo biloba*. Arboric Urban For 34:300–307. https://www.doi.org/10.48044/jauf.2008.041

Eckhart VM (1999) Sexual dimorphism in flowers and inflorescences. In: Geber MA, Dawson TE, Delph LF (ed) Gender and sexual dimorphism in flowering plants. Springer, Berlin, Heidelberg, pp 123–148

Freeman DC, Wachocki BA, Stender MJ, Goldschlag DE, Michaels HJ (1994) Seed size and sex ratio in spinach: application of the Trivers-Willard hypothesis to plants. Écoscience 1:56–63. https://doi.org/10.1080/11956860.1994.11682228

Garcia MB, Antor RJ (1995) Age and size structure in populations of a long-lived dioecious geophyte: *Borderea pyrenaica* (Dioscoreaceae). Int J Plant Sci 156:236–243. https://doi.org/10.1086/297246

Hasson A, Blein T, Laufs P (2010) Leaving the meristem behind: the genetic and molecular control of leaf patterning and morphogenesis. C R Biol 333:350–360. https://doi.org/10.1016/j.crvi.2010.01.013

Heikrujam M, Sharma K, Prasad M, Agrawal V (2015) Review on different mechanisms of sex determination and sex-linked molecular markers in dioecious crops: a current update. Euphytica 201:161–194. https://doi.org/10.1007/s10681-014-1293-z

Inoti SK, Chamshama SAO, Thagana WM, Lulandala LLL, Dodson R (2015) Sex determination of young nursery Jojoba (*Simmondsia chinensis L.*) plants using morphological traits in semi-arid areas of Voi, Kenya. Biol Agric Health Sci 5:113–123

Käfer J, Marais GAB, Pannell JR (2017) On the rarity of dioecy in flowering plants. Mol Ecol 26:1225–1241. https://doi.org/10.1111/mec.14020

Kaviani B, Mahtabi H, Saedi Mehrvarz S, Vali M, Ghaziani F (2014) Identification of male and female *Cycas revoluta* thunb. before maturity using morphological and anatomical features. Sci Papers Ser B Hortic 58:327–336

Koenig D, Bayer E, Kang J, Kuhlemeier C, Sinha N (2009) Auxin patterns *Solanum lycopersicum* leaf morphogenesis. Development 136:2997–3006. https://doi.org/10.1242/dev.033811

Lenth R, Singmann H, Love J, Buerkner P, Herve M (2019) Emmeans: estimated marginal means, aka least-squares means (Version 1.3.4). Emmeans Estim Marg Means Aka Least-Sq Means https://CRAN. R-project. org/package= emmeans

Li J, Tang L (2017) Developing a low-cost 3D plant morphological traits characterization system. Comput Electron Agric 143:1–13. https://doi.org/10.1016/j.compag.2017.09.025

Liu Y, Wang S, Li L, Yang T, Dong S, Wei T, Wu S et al (2022) The Cycas genome and the early evolution of seed plants. Nat Plants 8:389–401. https://doi.org/10.1038/s41477-022-01129-7

Lloyd DG, Webb CJ (1977) Secondary sex characters in plants. Bot Rev 43:177–216. https://doi.org/10.1007/BF02860717

Mank JE (2022) Are plant and animal sex chromosomes really all that different? Phil Trans R Soc B 377:20210218. https://doi.org/10.1098/rstb.2021.0218

Newell SJ (1985) Intrapopulational variation in leaflet morphology of *Zamia pumila L*. in relation to microenvironment and sex. Am J Bot 72:217–221. https://doi.org/10.1002/j.1537-2197.1985.tb08286.x

Niklas KJ, Marler TE (2008) Sex and population differences in the allometry of an endangered cycad species, *Cycas micronesica* (Cycadales). Int J Plant Sci 169:659–665. https://doi.org/10.1086/533606

Nishida K, Kobayashi A, Nagahama T (1956) Studies on cycasin, a new toxic glycoside, of *Cycas revoluta* thunb.: part VI. Polarography of cycasin. Bull Chem Soc Jpn 20:122–126. https://doi.org/10.1080/03758397.1956.10857320

Ornduff R (1996) Gender performance in a cultivated cohort of the Cycad *Zamina Integrifolia* (Zamiaceae). Am J Bot 83:1006–1015. https://doi.org/10.2307/2445989

Panjvani K, Dinh AV, Wahid KA (2019) LiDARPheno – a low-cost LiDAR-based 3D scanning system for leaf morphological trait extraction. Front Plant Sci 10:147. https://doi.org/10.3389/fpls.2019.00147

Primack RB (1985) Longevity of individual flowers. Annu Rev Ecol Evol Syst 16:15–37. https://doi.org/10.1146/annurev.es.16.110185.000311

R Core Team (2023) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

Renner SS (2014) The relative and absolute frequencies of angiosperm sexual systems: dioecy, monoecy, gynodioecy, and an updated online database. Am J Bot 101:1588–1596. https://doi.org/10.3732/ajb.1400196

Renner SS, Müller NA (2021) Plant sex chromosomes defy evolutionary models of expanding recombination suppression and genetic degeneration. Nat Plants 7:392–402. https://doi.org/10.1038/s41477-021-00884-3

Renner SS, Ricklefs RE (1995) Dioecy and its correlates in the flowering plants. Am J Bot 82:596–606. https://doi.org/10.2307/2445418

Romanova MA, Domashkina VV, Maksimova AI, Pawlowski K, Voitsekhovskaja OV (2023) All together now: cellular and molecular aspects of leaf development in lycophytes, ferns, and seed plants. Front Ecol Evol 11:1097115. https://doi.org/10.3389/fevo.2023.1097115

Rourke JP (1989) The inflorescence morphology and systematics of *Aulax* (Proteaceae). S Afr J Bot 53:464–480. https://doi.org/10.1016/S0254-6299(16)31381-3

Segalla R, Pinheiro F, Baronio GJ, Morellato LPC (2021) Male-biased effective sex ratio across populations of the threatened *Zamia boliviana* (Zamiaceae). Plant Ecol 222:587–602. https://doi.org/10.1007/s11258-021-01127-3

Song Y, Ma K, Ci D, Chen Q, Tian J, Zhang D (2013) Sexual dimorphic floral development in dioecious plants revealed by transcriptome, phytohormone, and DNA methylation analysis in *Populus tomentosa*. Plant Mol Biol 83:559–576. https://doi.org/10.1007/s11103-013-0108-2

Stehlik I, Barrett SCH (2005) Mechanisms governing sex-ratio variation in dioecious *Rumex nivalis*. Evolution 59:814–825. https://doi.org/10.1111/j.0014-3820.2005.tb01755.x

Stevenson DW (1990) Morphology and systematics of the Cycadales. Mem NY Bot Gard 57:8–55

Thomas SC, LaFrankie JV (1993) Sex, size and interyear variation in flowering among dioecious trees of the Malayan Rain Forest. Ecology 74:1529–1537. https://doi.org/10.2307/1940080

van Beek TA, Bombardelli E, Morazzoni P, Peterlongo F (1998) *Ginkgo biloba L.* Fitoterapia 69:195–244

Wada K, Haga M (1997) Food poisoning by *Ginkgo biloba* seeds. In: Hori T, Ridge RW, Tulecke W, Del Tredici P, Trémouillaux-Guiller J, Tobe H (ed) Ginkgo Biloba a global treasure. Springer, Tokyo, pp 309–321

Yamasaki S, Fujii N, Takahashi H (2003) Characterization of ethylene effects on sex determination in cucumber plants. Sex Plant Reprod 16:103–111. https://doi.org/10.1007/s00497-003-0183-7

Yakimowski SB, Glaettli M, Barrett SCH (2011) Floral dimorphism in plant populations with combined versus separate sexes. Ann Bot 108:765–776. https://doi.org/10.1093/aob/mcr025

Zhou Z, Zheng S (2003) The missing link in *Ginkgo* evolution. Nature 423:821–822. https://doi.org/10.1038/423821a

4 Is ddRAD an alternative approach to WGS to identify sex-specific molecular markers in plants with large genomes?

Ylenia De Luca · Niklaus Zemp · Salvatore Cozzolino ·Antonia Cristaudo · Alex Widmer · Donata Cafasso

Y. De Luca · S. Cozzolino (✉) · D. Cafasso
Department of Biology, University of Naples Federico II, Complesso Universitario MSA, 80126 Naples, Italy
Y. De Luca
e-mail: ylenia.deluca@unina.it
S. Cozzolino
e-mail: cozzolin@unina.it
D. Cafasso
e-mail: donata.cafasso@unina.it

A. Cristaudo
Department of Biological, Geological and Environmental Sciences, University of Catania, 58 Empedocle St., 95128 Catania, Italy
e-mail: acristau@unict.it

N. Zemp · A. Widmer
Department of Environmental Systems Science, ETH Zurich, 16, Universitätstrasse St., 8092 Zurich, Switzerland
A. Widmer
e-mail: alex.widmer@usys.ethz.ch
N. Zemp
e-mail: niklaus.zemp@usys.ethz.ch

Acknowledgments: We are very grateful to Piante Faro (Giarre, Italy) for providing us with the greenhouse structures and plants.

Funding: YDL was funded by PON Dottorato Industriale «Ricerca e innovazione 2014-2020» – Asse "Capitale Umano" (DOT1318787). SC and DC were funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4 - NextGenerationEU; Award Number: Project code CN_00000033, "National Biodiversity Future Center - NBFC."

Abstract

Some dioecious plants have sex chromosomes for sex determination, varying between XY and ZW systems. Recent advancements in sequencing techniques have significantly bolstered the identification of genomic regions responsible for sex determination in dioecious species. In this context, the application of representative reductive genomics analysis (ddRAD) as an alternative to whole genome sequencing (WGS) has proven successful in some model plants. In this study, we tested the potential of ddRAD and a dedicated bioinformatic pipeline to identify sex-specific markers in *Cycas revoluta*, a non-model plant with a very large genome. A region responsible for sex determination has already been identified by sequencing the full genome of the related *Cycas panzhihuaensis*. By comparing the results of our dedicated ddRAD pipeline in *Cycas revoluta* with those from WGS of *Cycas panzhihuaensis*, we confirmed the potential of our approach to identify sex-specific genomic regions. This work demonstrates the efficacy of ddRAD-seq as an alternative method for the rapid development of sex-linked molecular markers in non-model plants with large genomes.

Keywords ddRAD-sequencing · sex-linked molecular marker · *Cycas revoluta* · next-generation sequencing ·plant sex determination

4.1 Introduction

Many dioecious plants have a genetic system for determining sex, often involving sex chromosomes. The sex chromosomes are different from autosome chromosomes because they prevent meiotic recombination around the genes responsible for sex determination creating a region on the chromosome tied specifically to one sex. In certain species, the sex of an individual is determined by the chromosomes they inherit. When males determine the sex, this is known as heterogametic male (genotype XY for males and XX for females). On the other hand, when females determine the sex, this is known as heterogametic females (genotype ZZ for males and ZW for females) (Scharmann et al. 2019). The unique regions on these chromosomes that determine sex (such as the male-specific region of the Y, or MSY, and the female-specific region of the W) have distinct characteristics. For instance, the MSY can have sequences that are not present on the X chromosome, making them specific to males. The evolution of X and Y chromosomes is believed to originate from a pair of autosome chromosomes, and the acquisition of male-specific sequences can be explained by various mechanisms, such as new mutations, translocations, or the expansion of certain sequences due to the absence of recombination (Tennessen et al. 2018). Over time, the MSY may undergo genetic degeneration, losing functional genes that were initially shared with the X chromosome (Bachtrog, 2013). Sex chromosomes in plants have evolved independently many times, likely resulting in different ages and degrees of degeneration. A few sex chromosomes have changed enough in size or structure to be visually distinguishable under a microscope (heteromorphic sex chromosomes). Examples of plants with widely differentiated sex chromosomes are *Silene latifolia* (Kejnovsky and Vyskot, 2010), *Rumex acetosa* (Kihara and Ono, 1923), and *Coccinia grandis* (Sousa et al. 2013). However, most dioecious plants only have subtler differences that require more advanced genetic methods for the detection of sex (homomorphic sex chromosomes).

Progress in the identification of genomic regions responsible for sex determination in dioecious species has increased dramatically due to novel sequencing techniques and the employment of reverse genetic methods. Recently, candidates for sex determination have been discovered in several plant species (Hobza et al. 2018). In kiwifruit (*Actinidia* spp.), Y-encoded genes *SyGI* and *FrBy* act independently as the suppressor of feminization (S*uF*) and the maintenance of male (Akagi et al. 2019). In persimmon (*Diospyros* spp.), a Y-encoded pseudogene *OGI* is sufficient for the expression of androecium and repression of gynoecium through encoding a small RNA and targeting its autosomal counterpart *MeGI* (Akagi et al. 2014). Two Y-encoded factors, named *SOFF* and *aspTDF*, have been discovered in garden asparagus (*A. officinalis*) for the expression of maleness and repression of female development (Harkess et al. 2017).

Understanding these sex determination systems and finding genetic markers linked to sex is crucial for molecular sex identification, especially for non-flowering plants, in areas like agriculture, breeding, and conservation. This is particularly important for cultivated plants that need to determine their gender well in advance of developing reproductive organs and have a long generation time. *Phoenix dactylifera L.* is a long-living, dioecious evergreen fruit tree (Dhawan et al. 2013). The date palm generally reaches reproductive age between 5 and 8 years after sowing (Aberlenc-Bertossi et al. 2011; Bendiab et al. 1993). Only females produce fruit and only a few male plants are needed for pollination. Thus, early screening out of male plants is essential for efficient date palm orchard establishment. Early sex determination is also relevant for some ornamental trees. For instance, *Ginkgo biloba* trees are dioecious, and only males are preferred in public gardens, but usually do not mature sexually until 20 to 30 years of age (van Beek, 1998). Similarly, male Cycads are preferred in landscape gardening due to toxicity of female plant seeds, but plants reach sexual maturity only after ca. 15 -20 years (Segalla et al. 2021).

Next-generation sequencing techniques (NGS) have now significantly increased knowledge about sex-linked genes (reviewed by Muyle et al. 2017). However, they require prior knowledge of heterogamety, controlled breeding, or whole genome sequencing, which may represent the best approach to find sex-specific markers (Andrews et al. 2016). However, large, and complex genomes, such as those of several Gymnosperms, Cycads included, present cost problems for current NGS approaches (Alexeyenko et al. 2014). Further problems concern the generation of a sufficient number of reads to account for the overrepresentation of highly repeated elements typical of these large genomes (Catchen et al. 2017). Accordingly, the size of an organism's genome greatly affects the cost of sequencing its genome, which in turn affects the number of organisms for which genomic data are already available (Andrews et al. 2016). An alternative class of strategies to WGS explores the use of population polymorphism to infer sex-specific markers (reviewed by Muyle et al. 2017), even without pedigrees. These strategies can potentially enable the discovery of sex-linked regions with cheaper reduced-representation sequencing (RRS) methods, such as RAD-seq and ddRAD-seq (Baird et al. 2008; Peterson et al. 2012). These techniques are extremely advantageous in long-living species, with large genomes or that are not easily cultivable (Clugston et al. 2019). Several studies have successfully used RADseq to determine plant sex (Gamble et al. 2015; Jeffries et al. 2018; Scharmann et al. 2019; Morgan et al. 2020; Matsumura et al. 2014). Undoubtedly, RADseq appears to be sufficiently versatile and robust to characterize genome-wide polymorphisms in non-model species that do not have a reference genome and allow reliable identification of loci associated with sex determination (Feron et al. 2021; Palmer et al. 2019; Peterson et al. 2012). RAD-seq/ddRAD-seq has been successfully used to identify sex-linked DNA sequences in dioecious plants, e.g. in *Excoecaria agallocha* (Zhou et al. 2018) and *S. latifolia* (Qiu et al. 2016) but their application in species with large genomes has never been tested. A recent study has identified a sex-specific region in Cycads (Liu et al. 2022) based on WGS of *Cycas panzhihuaensis* and a genome-wide association study (GWAS). Considering that the ddRAD-seq approach is faster because it focuses on specific genomic regions instead of sequencing and assembling a whole genome and is cheaper because it may require less sequencing data than a genome-wide analysis such as GWAS-seq (Clugston et al. 2019), in this study we tested the use of a representative reductive genomics analysis (ddRAD) and a dedicated bioinformatic pipeline to identify sex-specific markers in *Cycas revoluta*. We chose this plant as a study system because it is a long-lived dioecious plant, with a very large genome (20 G), in which a sex-specific region has been already identified in a close relative species (Liu et al. 2022). Therefore, we aim by comparing the results of our dedicated ddRAD pipeline in *Cycas revoluta* with those from WGS of *Cycas panzhihuaensis*, we aim to confirm the reliability and reproducibility of our method as a viable alternative to whole genome sequencing for the rapid development of sex-linked molecular markers in non-model plant species with large genomes.

4.2 Materials and methods

4.2.1 Sampling, ddRAD-seq, and genotyping

We identified potential sex-linked genetic markers using the approach of Scharmann et al. (2019). We sampled young leaves of adult *Cycas revoluta* plants of known sex, 24 male individuals and 24 female individuals. We extracted DNA with the Kingfisher method (Macherey-Nagel) and prepared sequencing libraries following the ddRAD-seq protocol (Peterson et al. 2012) using *EcoRI* and *TaqI* enzymes. 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 an average length of approximately 96 bases), read mapping, genotype calling, and quality filtering followed a modified dDocent pipeline (Puritz et al. 2014), and the code is deposited at https://github.com/mscharmann.

We define sex-specific contig as contigs where sequencing reads can only be aligned from one sex, indicating a presence-absence polymorphism. However, this may not always reflect true presence-absence, as absence could be influenced by methodological factors specific to each dataset. Imperfect detection can arise from various factors such as the number and relatedness of individuals, genome characteristics, library preparation, sequencing depth, and bioinformatics processing. Presence/absence statistics (samtools, Li et al. 2009) for all individuals and reference contigs were then subjected to the *privacy rarefaction* procedure of Scharmann et al. (2019) (https://github.com/mscharmann/privacy-rarefaction). This algorithm separates real biological genomic presence/absence from stochasticity, yielding a list of contigs that, with high confidence, occur in one of the two sexes exclusively (i.e. with bootstrap support).

4.2.2 Privacy rarefaction procedure

We quantify sex-specificity by assessing the deviation of observed sex-specific contigs from zero and qualitatively evaluate each locus to retain the most reliable candidates for further analysis. To ensure unbiased comparisons among datasets with varying total numbers of individuals and sex ratios, we employ a privacy rarefaction procedure. This procedure subsamples male and female individuals at a 1:1 sex ratio, allowing for comprehensive assessments across different dataset compositions. We bootstrap n random sets of males and females without replacement, ranging from subsample size 1 to the maximum possible subsample size determined by the minimum of N males and N females. The subsample size, also referred to as 'stringency', reflects the confidence level in the results, with higher subsample sizes indicating greater confidence. In quantitative tests, we count sex-specific loci for each combination of individuals, generating distributions of observed male- and female-specific loci counts through bootstrap resampling. These distributions are then separately compared to a null distribution obtained from permutations of the sexes, estimating the number of loci that might erroneously appear sex-specific if individuals were interchangeable. For qualitative assessment, we again bootstrap n random sets of males and females without replacement for each possible subsample size. Subsequently, for each locus, we tally the number of bootstrapped male-female comparisons in which it exhibits sex-specificity. True sex-specific loci are expected to appear more frequently in such comparisons compared to false positives occurring randomly. The bootstrap support value for sex-specificity is calculated by dividing the count by the number of bootstraps.

4.2.3 BLAST (Basic Local Alignment Search Tool) analysis and alignments

All top four hypothetical contigs were analyzed and compared with the genomic data obtained by Liu et al. 2022. BLAST analyses were conducted on each of the four identified contigs, using both the genome of a female individual (sample details: seed endosperm tissue) of *Cycas panzhihuaensis* available on NCBI (GCA_023213395.1), and the Y chromosome of *Cycas panzhihuaensis* available on https://db.cngb.org/codeplot/datasets/public_dataset?id=PwRftGHfPs5q. Multi-alignments were made using the MEGA software (Kumar et al. 2008) between the contigs identified using our ddRAD approach, the regions on the Y chromosome and chromosome 8 (i.e. the X chromosome according to Liu et al. 2022) that had the highest percentage of identical positions from the BLAST analysis.

4.2.4 Development of a molecular marker

The top one of the four hypothetical contigs, namely dDocent_ass_91454 male-specific contig, was chosen for PCR validation. Based on the alignment made among this contig, X, and Y chromosomes, PCR primers were designed in the region where SNPs occurred between the X and Y chromosomes, i.e. in the regions of divergence (Fig. 2). PCR primers were designed in Primer3web 4.0.0 (Untergasser et al. 2012) and tested according to the protocol described in Text S1.

4.2.5 Exploring sex ratio in the F1 generation of *Cycas revoluta*

Once a marker allowing us to discriminate the sex of *Cycas revoluta* was developed, we applied it to study the sex ratio in an F1 generation of *Cycas revoluta*. To do this, we manually crossed a male adult plant with a female adult plant, obtaining after two years a progeny of 180 seedlings, growing at the greenhouse Piante Faro. We extracted DNA from all 180 individuals following the protocol described by Doyle and Doyle, (1990). To determine the sex of all individuals of the F1 generation, the identified male-specific contig dDocent_ass_91454 was amplified by PCR using primer pair listed in Table S1- Supplementary materials. As a control, the male-specific MADS-Y/CYCAS_034085 region was amplified with a sex-specific primer pair as described in (Liu et al. 2022),

4.3 Results

We first searched for sex-specific contigs in *Cycas revoluta* using privacy rarefaction according to Scharmann et al. (2019). When small numbers of individuals of each sex were analyzed, the procedure yielded similar numbers of male- and female-specific candidates. 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. 1). Hence, these curves correctly diagnosed an XY-system for *Cycas revoluta* and rejected a ZW-system. The non-redundant candidate male-specific contigs (cd-hit-est at similarity 1.00) were ranked according to the cumulative support they received (sum of the passed stringency levels overall alternative reference assemblies).

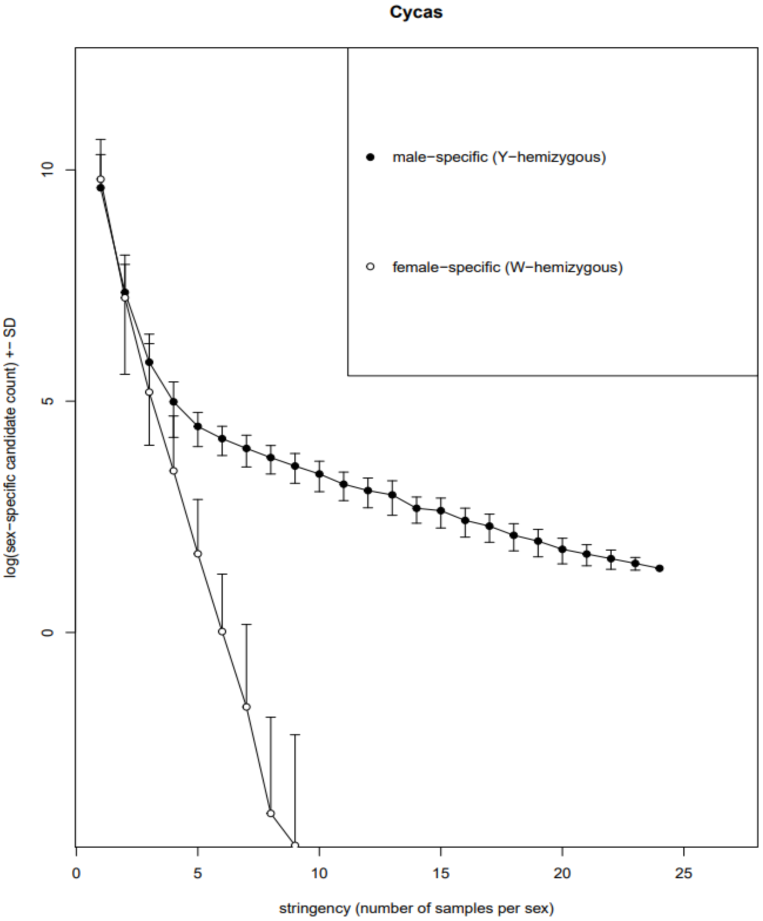

Fig. 1 Evidence for male-specific loci and XY sex-determination systems in *Cycas revoluta* (privacy rarefaction curves). Shown are counts of sex-specific contigs (y-axis) as a function of the number of individuals of each sex sampled to score sex specificity (x-axis, stringency). Sex-specific contigs are defined as those to which sequencing reads from only one sex can be aligned. Dots represent averages, and whiskers the 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. Male-specific candidates were found in *Cycas revoluta* up to the maximum possible stringency (the minimum number of male individuals and female individuals).

After the *privacy rarefaction* analysis, the four contigs identified were searched via BLATn analysis on both the Y chromosome and chromosome 8 (i.e., the X chromosome) of *Cycas panzhihuaensis* assembled by Liu et al. (2022). The results indicate (see Table 1) that all four contigs are present on both the X and Y chromosomes.

Specifically, the contig dDocent_ass_91454 shows a 98.6% identical position percentage on scaffold 3 of the Y chromosome and 96.1% on the X chromosome.

The dDocent_Ass_Contig_80549 from 1 to 150 bp shows 98.0% identical positions on scaffold 1 of the Y chromosome and 92,8% on X-chromosome, and from position 161 to 305 bp shows 96,5% identical positions on scaffold 1 of the Y chromosome and 92,2% on X-chromosome.

The dDocent_Ass_Contig_262742 from 161 to 305 bp displays 100% and 96.5% identity on scaffold 4 and scaffold 1 respectively of the Y-chromosome, and 100% on X-chromosome. From 1 to 150 bp shows 92.6% identical positions on scaffold 1 of the Y chromosome, 82,8% on scaffold 4 of the Y-chromosome and 94% on the X-chromosome.

Lastly, the dDocent_Ass_Contig_295273 from 161 to 305 bp exhibits 96.5% identical positions on scaffold 1 and 6 of the Y chromosome and 97% on the X-chromosome. From 1 to 150 bp shows 93,2% identical positions on scaffold 1 of the Y-chromosome and 95,9% on the X-chromosome.

PCR validation

For developing a sex-specific molecular marker we designed the PCR primers by aligning dDocent_ass_91454 contig, and the X and Y chromosomes of *Cycas panzhihuaensis*. Specifically, we positioned the forward primer in the way that its last nucleotide (T) corresponds to a nucleotide difference between the male and female sex chromosomes (T and C, respectively). As a consequence, on the X chromosome, the forward primer is unable to effectively bind to the annealing site, failing the amplification (Fig. 2) and selectively working as a male-specific molecular marker, i.e. enabling the amplification of a 280 bp fragment only in male individuals (Fig. S1).

```
Species/Abbrv         ******  **  ********
1. chm8 genome      AATTCTTCGTGAGGCACCCTCCACTCCAACAGCAGGCCACGTGGGAATTGATAAAACCTATCACTAAATCTGTA
2. scaffold3 Ychrom AATTCTTCATGAAGCACCCTCTACTCCAATAGCAGGCCACGTGGGAATTGATAAAACCTATCACCAAATCTGTA
3. contig 91454     AATTCTTCACGAAGCACCCTCTACTCCAATAACAGGCCACGTGGGAATTGATAAAACCTATCACCAAATCTGTA
4. primer F         --TTCTTCACGAAGCACCCTCT-----------------------------------------------------
```

Fig. 2 ClustalW alignment through the MEGA software between dDocent_ass_91454 contig (the third sequence in the alignment), chromosome 8 (i.e., X-chromosome, the first sequence in the alignment), scaffold 3 Y-chromosome (the second sequence in the alignment), and Forward primer (indicated as primer F in the alignment).

The designed molecular marker was then used to discriminate sex in an F1 generation of *Cycas revoluta* seedlings. The same individuals were also subjected to amplification using the primers designed by Liu et al. (2022) for the sex-specific MADS-Y region, as control. Results show that by using both the primers designed by Liu et al. (2022) and those designed from the dDocent_ass_91454 contig we were able to gather identical results by identifying the same 77 female individuals and 103 male individuals in the F1 generation of *Cycas revoluta* seedlings.

ddrad sex hypotethical contig	**Subject**	**Percentage of identical positions**	**Alignment length**	**Number of mismatches**	**Number of gap openings**	**Star of alignment in query**	**End of alignment in query**	**Star of alignment in subject**	**End of alignment in subject**	**Expect value**	**Bit score**
dDocent_ass_91454	hic_scaffold_3_y_chrome	98.592	284	4	0	1	284	569683	569400	3.40e-142	503
dDocent_ass_91454	chromosome_8_genome	96.127	284	11	0	1	284	49376178	49375895	3.51e-128	464
dDocent_Ass_Contig_80549	hic_scaffold_1_y_chrome	98.000	150	3	0	1	150	384991	385140	2.44e-69	261
dDocent_Ass_Contig_80549	hic_scaffold_1_y_chrome	96.552	145	5	0	161	305	384895	384751	3.18e-63	241
dDocent_Ass_Contig_80549	chromosome 8_genome	92.857	140	10	0	1	140	795507741	795507880	9.15e-50	204
dDocent_Ass_Contig_80549	chromosome 8_genome	92.254	142	11	0	164	305	693158365	693158506	3.29e-49	202

dDocent_Ass_Contig_262742	hic_scaffold_4_y_chrome	100.000	145	0	0	161	305	3873509	3873365	1.46e-71	268
dDocent_Ass_Contig_262742	hic_scaffold_4_y_chrome	82.895	152	15	8	1	148	3873562	3873706	9.30e-29	126
dDocent_Ass_Contig_262742	hic_scaffold_1_y_chrome	96.528	144	5	0	161	304	14047616	14047759	1.15e-62	239
dDocent_Ass_Contig_262742	hic_scaffold_1_y_chrome	92.667	150	11	0	1	150	14047563	14047414	5.37e-56	217
dDocent_Ass_Contig_262742	chromosome_8_genome	94.000	150	9	0	1	150	291249269	291249418	5.43e-57	228
dDocent_Ass_Contig_262742	chromosome_8_genome	100.000	145	0	0	161	305	608855675	608855531	3.20e-69	268
dDocent_Ass_Contig_295273	hic_scaffold_6	96.552	145	5	0	161	305	2043934	2044078	3.18e-63	241

dDocent_Ass_Contig_295273	hic_scaffold_6	90.476	147	14	0	4	150	467146	467292	2.51e-49	195
ddocent_ass_contig_295273	hic_scaffold_1	96.552	145	5	0	161	305	7905671	7905815	3.18e-63	241
ddocent_ass_contig_295273	hic_scaffold_1	93.243	148	10	0	1	148	7905641	7905494	1.49e-56	219
ddocent_ass_contig_295273	chromosome_8_genome	97.241	145	4	0	161	305	267449071	267448927	1.50e-62	246
ddocent_ass_contig_295273	chromosome_8_genome	95.973	149	6	0	2	150	541814852	541814704	1.94e-61	243

Table 1 Results of BLASTn analysis between the 4 hypothetical ddRAD contigs (Query) and the Subject Y and X chromosome (i.e. chromosome 8) of the genome assembled by Liu et al. 20222

4.4 Discussion

In this study, we employed a reduced representation genomic analysis, i.e. ddRAD sequencing and a dedicated pipeline to identify sex-specific markers in *Cycas revoluta*. The privacy rarefaction resulted in a divergence in counts of male and female specific candidates with increasing stringency, which is characteristic of organisms with sex chromosomes (Scharmann et al. 2019). More precisely, only male-specific candidate markers were obtained with high confidence. This result is characteristic of a male-heterogametic system, and hence the existence of Y-chromosomes (Scharmann et al. 2019), although the Y-specific sequences were overall extremely rare.

ddRADseq protocols (Double-digest RADseq) (Peterson et al. 2012) are particularly suitable for detecting regions associated with sex chromosomes in large genomes. These protocols can be easily adapted to target different numbers of loci in each species by varying the restriction enzymes used and/or the fragment size selection window (Hime et al. 2019). When the non-recombinant portion of the sex chromosomes is small and there is a low probability of finding a suitable marker to associate with it, it can be relatively straightforward to use a restriction enzyme that cuts more frequently, thereby increasing the likelihood of resolving small differences between the sexes.

The approach provides a powerful and flexible alternative to previous methods such as RAPD and AFLP since RAD-seq sequencing data can be directly transformed into simple PCR assays, avoiding additional cloning steps for primer design (Guillemin et al. 2012; Griffiths and Orr, 1999).

Additionally, sequenced genomes are not required for conducting RAD-seq (Baird et al. 2008), although their availability can enhance the value and interpretative power of the results. In our case, we compared the results obtained with ddRAD on *Cycas revoluta* with those derived from whole-genome sequencing conducted by Liu et al. (2022), who identified a unique region of the Y chromosome in the related species *Cycas panzhihuaensis*. This allowed us to compare the results and gain a better interpretation of our findings. However, unlike Liu et al. (2022), who detected a sex-linked genomic region exclusive of the Y chromosome (i.e. absent in the X chromosome), the contigs selected in our study aligned with both the X and Y chromosomes even if in regions of genetic divergence between sexes. Indeed, the presence of an SNP allowed us to develop a male-specific molecular marker representing a reliable tool for distinguishing male from female plants in *Cycas revoluta*.

According to Liu et al (2022) the assembled MSY had 32 Mb different in length from the corresponding region on the X chromosome, which agrees with the heteromorphy of the Cycas sex chromosomes. As the haploid Genome size of Cycas is approx. 10.5 Gb, the size difference between the two sexes is less than 0,3%. Accordingly, a higher marker density and higher coverage than those gathered in our pilot study would have been necessary to map and detect a ddRAD marker in this exclusive Y- specific chromosome region. We argue the combination of multiple enzyme combinations (only EcoRI /TaqI, in our study) may increase the chance of detecting specific sex-linked chromosome regions in large genomes with small length differences between sexual chromosomes.

Validation of our sex-linked marker on the F1 *Cycas revoluta* progeny of unknown sex further strengthened its utility and confirmed its specificity for the Y chromosome. It is noteworthy that in some plant species, mechanisms can influence the proportion between males and females, with different implications for sex ratio distortion. Our results showed a ratio of approximately 1:1 between males and females, confirming the XY sex-determination system and the absence of ecological or physiological factors influencing sex distribution.

In conclusion, the reduced representation genomic approach of ddRAD sequencing allowed us to develop a male-specific sex molecular marker in *Cycas revoluta*, which proved to be rapid, economic, easily amplifiable, and reliable. These markers, which have successfully passed the validation stage, may eventually be submitted for patenting, and a low-cost rapid screening kit may be developed for greenhouses, allowing significant rationalization in terms of both production and labor costs for plant breeding that requires a selection of plant material based on sex.

4.5 Supplementary materials

Text S1. Sex-linked marker design for *Cycas revoluta*

DNA extraction from adult *Cycas revoluta* plants

Leaves of adult Cycas revoluta plants that had reached sexual maturity were used for DNA extraction following the protocol (Doyle and Doyle, 1990). Specifically, 29 male individuals and 32 female individuals were used.

Candidate male-specific contigs obtained from privacy rarefaction, the top dDocent_ass_91454 was used for the development of the molecular marker:

dDocent_ass_91454

AATTCTTCACGAAGCACCCTCTACTCCAATAACAGGCCACGTGGGAATTGATAAAACCT
ATCACCAAATCTGTAATACTTATTTTTGTAATGGAATGAAAGCAAACATTCACAAGTTCAT
CTCAAAATGCGACATATGTCAAAGAAACAGGAACAAAAGCATAAAATTACTAGGACTAC
TACAACCGCTACATATACCTGAACGGAAATGGGAGGAAATCTTGATGGATTTCATAATAG
GACTCCCCACATCAAAAAACAAGGATGCAATCATTGTCGTGGTCG

dDocent_Ass_Contig_80549

ATTTTTGAGGCCAAAGGGCATTGCCTTGTAGGCAAAGGTGTCCCATAGGGTGGTGAAGG
CTGTCTTATGGCAGTCTTTAGGAGAGATTTAAATCTGGTTATAGCCAGAGAAGCCATCCA
TTAAGGACATGACTTCAAACCTGACGATTCGNNNNNNNNNNGCATGTCATAGCCCATCA
CCCTTTAGACTTGTGAGCCATCTTCACACGATGCCACAAGTACCACGTAAGTTTAAACCC
ACTCAAATGTGTGTTTCTCGTCACTTCTGGCAATCTCCTTAGGTTCCTTATGTCTATGCAA
GGAATT

dDocent_Ass_Contig_262742

ACTTACATCAGGTTTCTGACCTGTCCAAGCTTCTTCTGGAGTGACGTTCTTCAGTGTCCT
GCTTGGACTTCTGTTTATCAAGTAGCATGCTGTGCTAACTGCTTCTACCCAGTATGCGTTT
GATATTTTTGTCCCATACATCATGCTTCGNNNNNNNNNNGGTCCAAATTAGATGCAAAAA
GCATAGAATGTGTCTTAATAGGATATAGCAAAAATTCAAAGGCATATAAGTGCATAGATCC

AAGAACAAATAGAATCTATGTGAGTAGAGATGTTGTCTTTAATGAAGATAGCTCAGATGA
ATT

dDocent_Ass_Contig_295273

TGCTGCTTGTATCTCTGTTGAGATTTATGAAGCTGCTCTTCAACCTGTTGATGGATTTTCC
TGATTCTTTCCAAGAATCTTTGTGTCTTTTCATCCCCTTCTTCCTCCTTTCCATTTCTTACA
AAAAGTTGAGCAGAAAATTCTAAATCGNNNNNNNNNNTTAAAGAAGGAGATAGAGTTT
GGTTACACCTCAGTAAGAAAAGATTGCAAGGTGAAGGCAAAAAGCTAAATCCTTTACG
ATATGGTCCTTTTACAATTTTGAAGAAAATTGGGGAAAATGCTTTCCAGCTTGAACTTTT
AGAATT

‘N’ represents an unknown number of bases in paired-end contigs.

For the dDocent_ass_91454 contig, we designed the primers using Primer3web 4.0.0 (Untergasser et al. 2012) (Table S1) for PCR amplification.

Table S1: Sequence of primers designed for each hypothetical contig and the expected amplified fragment size.

Contig	**Primer sequences (5’-3’)**	**Size (bp)**
dDocent_ass_91454	F: TTCTTCACGAAGCACCCTCT R: CTCCCATTTCCGTTCAGGTA	280

Sequencing and PCR amplification

All PCR reactions were conducted in an Applied Biosystems 2720 thermal cycler (Thermo Fisher, US). The reaction is performed in 25 µl volumes containing 2.5 mM MgCl2, 250 µM of each dNTP, 0.5 units of Taq DNA polymerase (Promega, Wisconsin, USA), 1x Taq buffer (Promega), 0.5 µM of each of the primers (Table S1), and 1 µl of template DNA (c. 5-20 ng/µl). After initial denaturation for 2 min at 95°C, 30 cycles are run with denaturation for 30s at 95°C, annealing for 30s at 55°C, and extension for 40s at 72°C, followed by a final extension step of 5 min at 72°C. From an initial amplification reaction, it emerged that the contig dDocent_Ass_Contig_80549 amplified in both sexes, so it was used as an internal control for each amplification reaction, as suggested by Hobza and Widmer (2008). The common target is used as an internal control of each reaction. It ensures that poor template quality or other technical issues are recognised as such, instead of confusion with the true absence of the sex-specific target, i.e. the expectation for females. The two regions can also be amplified in separate reactions.

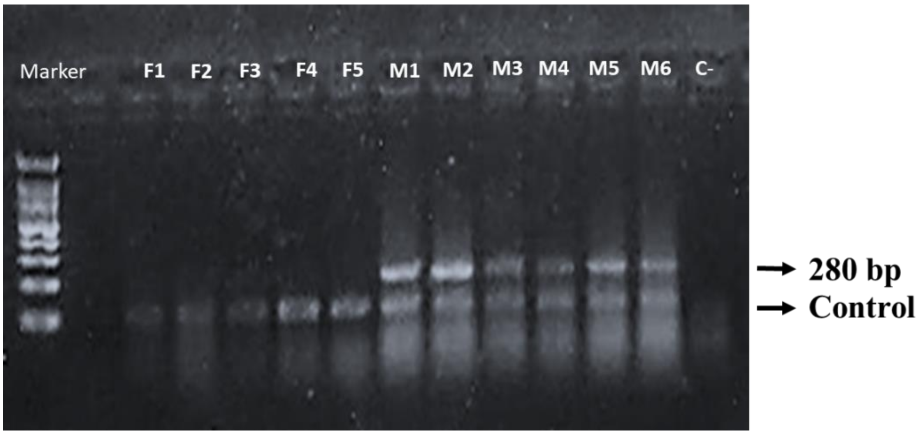

Fig. S1 PCR amplifications of the male-specific contig dDocent_ass_91454 (280 bp) and dDocent_Ass_Contig_80549 (control). The latter amplifies both in male and female plants. A 100 bp molecular ladder is in the leftmost lane

References

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Hobza R, Widmer A (2008) Efficient molecular sexing in dioecious *Silene latifolia* and *S. dioica* and paternity analysis in F1 hybrids. Mol Ecol Resou 8:1274–1276. https://doi.org/10.1111/j.1755-0998.2008.02344.x

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3—new capabilities and interfaces. Nucleic Acids Res 40:e115–e115. https://doi.org/10.1093/nar/gks596

4.6 References

Aberlenc-Bertossi F, Daher A, Chabrillange N, Tregear JW, Mohamed N (2011) Sex determination in date palm: new perspectives on an old theme. In Plant and animal genomes XIX conference W 519:15–19

Akagi T, Pilkington SM, Varkonyi-Gasic E, Henry IM, Sugano SS, Sonoda M, Tao R (2019) Two Y-chromosome-encoded genes determine sex in kiwifruit. Nat plants 5:801–809. https://doi.org/10.1038/s41477-019-0489-6

Alexeyenko A, Nystedt B, Vezzi F, Sherwood E, Ye R, Knudsen B, et al (2014) Efficient de novo assembly of large and complex genomes by massively parallel sequencing of Fosmid pools. BMC Genomics 15:1–10. https://doi.org/10.1186/1471-2164-15-439

Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA (2016) Harnessing the power of RADseq for ecological and evolutionary genomics. Nat Rev Genet 17:81–92. https://doi.org/10.1038/nrg.2015.28

Bachtrog D (2013) Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. Nat Rev Genet 14:113–124. https://doi.org/10.1038/nrg3366

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, et al (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PloS One 3:e3376. https://doi.org/10.1371/journal.pone.0003376

Bendiab K, Baaziz M, Brakez Z, Sedra H (1993) Correlation of isoenzyme polymorphism and Bayoud-disease resistance in date palm cultivars and progeny. Euphytica 65: 23–32

Catchen JM, Hohenlohe PA, Bernatchez L, Funk WC, Andrews KR, Allendorf FW (2017) Unbroken: RADseq remains a powerful tool for understanding the genetics of adaptation in natural populations. Mol Ecol Resou 17:362–365.

Clugston JA, Kenicer GJ, Milne R, Overcast I, Wilson TC, Nagalingum NS (2019) RADseq as a valuable tool for plants with large genomes—A case study in cycads. Mol Ecol Resou 19:1610–1622. https://doi.org/10.1111/1755-0998.13085

Dhawan C, Kharb P, Sharma R, Uppal S, Aggarwal RK (2013) Development of male-specific SCAR marker in date palm (*Phoenix dactylifera L.*). Tree Genet Genomes 9:1143–1150. https://doi.org/10.1007/s11295-013-0617-9

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Feron R, Pan Q, Wen M, Imarazene B, Jouanno E, Anderson J, et al (2021) RADSex: A computational workflow to study sex determination using restriction site-associated DNA sequencing data. Mol Ecol Resou 21:1715–1731. https://doi.org/10.1111/1755-0998.13360

Gamble T, Coryell J, Ezaz T, Lynch J, Scantlebury DP, Zarkower D (2015) Restriction site-associated DNA sequencing (RAD-seq) reveals an extraordinary number of transitions among gecko sex-determining systems. Mol Biol Evol 32:1296–1309. https://doi.org/10.1093/molbev/msv023

Griffiths R, Orr K (1999) The use of amplified fragment length polymorphism (AFLP) in the isolation of sex-specific markers. Mol Ecol 8:671–674. https://doi.org/10.1046/j.1365-294x.1999.00578.x

Guillemin ML, Huanel OR, Martínez EA (2012) Characterization of genetic markers linked to sex determination in the haploid-diploid red alga *Gracilaria chilensis*[1]. J Phycol 48: 365–372.

Harkess A, Zhou J, Xu C et al (2017) The asparagus genome sheds light on the origin and evolution of a young Y chromosome. Nat Commun 8:1279. https://doi.org/10.1038/s41467-017-01064-8

Hime PM, Briggler JT, Reece JS, Weisrock DW (2019) Genomic data reveal conserved female heterogamety in giant salamanders with gigantic nuclear genomes. G3: Genes, Genomes, Genetics 9:3467–3476. https://doi.org/10.1534/g3.119.400556

Hobza R, Hudzieczek V, Kubat Z, Cegan R, Vyskot B, Kejnovsky E, Janousek B (2018) Sex and the flower–developmental aspects of sex chromosome evolution. Ann Bot 122:1085–1101. https://doi.org/10.1093/aob/mcy130

Hobza R, Widmer A (2008) Efficient molecular sexing in dioecious *Silene latifolia* and *S. dioica* and paternity analysis in F1 hybrids. Mol Ecol Resou 8:1274–1276. https://doi.org/10.1111/j.1755-0998.2008.02344.x

Jeffries DL, Lavanchy G, Sermier R, et al (2018) A rapid rate of sex-chromosome turnover and non-random transitions in true frogs. Nat Commun 9:4088. https://doi.org/10.1038/s41467-018-06517-2

Kejnovsky E, Vyskot B (2010) *Silene latifolia*: the classical model to study heteromorphic sex chromosomes. Cytogenet Genome Res 129:250–262. https://doi.org/10.1159/000314285

Kihara H and Ono T (1923) Cytological studies on Rumex L. I. Chromosomes of *Rumex acetosa L*. Bot Mag (Tokio) 37:84–90

Kumar S, Nei M, Dudley J, Tamura K (2008) MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinf 9:299–306. https://doi.org/10.1093/bib/bbn017

Lang PL, Weiß CL, Kersten S, Latorre SM, Nagel S, Nickel B, Burbano HA (2020) Hybridization ddRAD-sequencing for population genomics of nonmodel plants using highly degraded historical specimen DNA. Mol Ecol Resour 20:1228–1247. https://doi.org/10.1111/1755-0998.13168

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liu Y, Wang S, Li L, Yang T, Dong S, Wei T, Wu S et al (2022) The Cycas genome and the early evolution of seed plants. Nat Plants 8:389–401. https://doi.org/10.1038/s41477-022-01129-7

Matsumura H, Miyagi N, Taniai N, Fukushima M, Tarora K, Shudo A, Urasaki N (2014) Mapping of the gynoecy in bitter gourd (*Momordica charantia*) using RAD-seq analysis. PloS One 9:e87138. https://doi.org/10.1371/journal.pone.0087138

Morgan EJ, Kaiser-Bunbury CN, Edwards PJ, Scharmann M, Widmer A, Fleischer-Dogley F, Kettle CJ (2020) Identification of sex-linked markers in the sexually cryptic coco de mer: are males and females produced in equal proportions? AoB Plants 12:plz079. https://doi.org/10.1093/aobpla/plz079

Muyle A, Shearn R, Marais GA (2017) The evolution of sex chromosomes and dosage compensation in plants. Genome Biol Evol 9:627–645. https://doi.org/10.1093/gbe/evw282

Palmer DH, Rogers TF, Dean R, Wright AE (2019) How to identify sex chromosomes and their turnover. Mol Ecol 28:4709–4724. https://doi.org/10.1111/mec.15245

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PloS one, 7:e37135. https://doi.org/10.1371/journal.pone.0037135

Puritz JB, Hollenbeck CM, Gold JR (2014) dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. PeerJ 2:e431. https://doi.org/10.7717/peerj.431

Qiu S, Bergero R, Guirao-Rico S, Campos JL, Cezard T, Gharbi K, Charlesworth D (2016) RAD mapping reveals an evolving, polymorphic and fuzzy boundary of a plant pseudoautosomal region. Mol Ecol 25:414–430. https://doi.org/10.1111/mec.13297

Scharmann M, Grafe TU, Metali F, Widmer A (2019) Sex is determined by XY chromosomes across the radiation of dioecious Nepenthes pitcher plants. Evol Lett 3:586–597. https://doi.org/10.1002/evl3.142

Segalla R, Pinheiro F, Baronio GJ, Morellato LPC (2021) Male-biased effective sex ratio across populations of the threatened *Zamia boliviana* (Zamiaceae). Plant Ecol 222:587–602. https://doi.org/10.1007/s11258-021-01127-3

Sinclair JP, Emlen J, Freeman DC (2012) Biased sex ratios in plants: theory and trends. Bot Rev 78:63–86. https://doi.org/10.1007/s12229-011-9065-0

Sousa A, Fuchs J, Renner SS (2013) Molecular cytogenetics (FISH, GISH) of *Coccinia grandis*: a ca. 3 myr-old species of Cucurbitaceae with the largest Y/autosome divergence in flowering plants. Cytogenet Genome Res 139:107–118. https://doi.org/10.1159/000345370

Stehlik I, Friedman J, Barrett SC (2008) Environmental influence on primary sex ratio in a dioecious plant. Proc Natl Acad Sci 105:10847–10852. https://doi.org/10.1073/pnas.0801964105

Tennessen JA, Wei N, Straub SC, Govindarajulu R, Liston A, Ashman TL (2018) Repeated translocation of a gene cassette drives sex-chromosome turnover in strawberries. PLoS Biol 16: e2006062. https://doi.org/10.1371/journal.pbio.2006062

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3—new capabilities and interfaces. Nucleic Acids Res 40:e115-e115. https://doi.org/10.1093/nar/gks596

van Beek TA, Bombardelli E, Morazzoni P, Peterlongo F (1998) *Ginkgo biloba L*. Fitoterapia 69:195–244

Zhou Y, Wu W, Ning Z, Zhou R (2018) Identification and characterization of sex-specific markers in the milky mangrove *Excoecaria agallocha* using double digest restriction site-associated DNA sequencing. Aquat Bot 144:54–60. https://doi.org/10.1016/j.aquabot.2017.11.004

5 Conclusion

In conclusion, our study addresses the urgent need for early sex identification in dioecious plants, focusing specifically on *C. revoluta*, a popular ornamental species with toxic seeds produced by female plants. By examining morphological traits in seedlings at different growth stages and correlating them with molecular sex determination, we aimed to identify a reliable method for early sex identification in *C. revoluta*. Our results reveal that a specific morphological trait, the distance between leaflets on the leaf rachis (IRD), exhibits sexual dimorphism in *C. revoluta* seedlings. Particularly, we found that at the earliest stage of plant development (stage 0), the IRD in the distal part of the leaf petiole provides a clear distinction between male and female individuals, with no overlapping values between sexes. This indicates that sexual dimorphism in leaf morphology is established very early in the life cycle of *C. revoluta*. The identification of this sexually dimorphic trait opens up possibilities for the development of rapid and cost-effective sex identification tools for use in nurseries and greenhouses. Leveraging advancements in image recognition and measurement technologies, such as 3D reconstruction and LiDAR-based systems, we can optimize the process of sex identification in large-scale cultivation operations. Furthermore, the development of mobile applications based on image recognition can empower enthusiasts and professionals alike to accurately determine the sex of *C. revoluta* plants in the field. Our study underscores the importance of integrating morphological and molecular approaches in plant sex identification research. While molecular markers offer precision, morphological traits provide practical and accessible tools for early sex identification, especially in contexts where molecular analyses may be impractical or cost-prohibitive. Moving forward, further research is needed to understand the genetic and physiological mechanisms underlying the observed morphological dimorphism in *C. revoluta*. Additionally, efforts to refine and optimize sex identification tools based on morphological traits will increase their utility and accessibility in horticultural and conservation contexts.

Overall, our study contributes to the development of efficient and sustainable practices for the cultivation and management of dioecious plant species, with implications for commercial horticulture. Additionally, in this study, we applied reduced representation genomic analysis using ddRAD sequencing coupled with a dedicated bioinformatics pipeline to identify sex-specific markers in *C. revoluta*, a non-model plant species with a large genome. Our goal was to ascertain whether this approach could be a viable alternative to whole-genome sequencing for identifying sex-linked molecular markers in dioecious plants, particularly those with complex genomes like *C. revoluta*.

Through the privacy rarefaction procedure, we identified male-specific contigs, indicative of an XY sex-determination system. This method allowed us to distinguish between male and female individuals with increasing stringency, highlighting the power of ddRAD sequencing in detecting sex-linked genomic regions. Furthermore, we conducted BLASTn analysis and alignments, confirming the presence of the identified contigs on both the X and Y chromosomes of the related species *Cycas panzhihuaensis*. Despite being present on both chromosomes, an SNP within one contig allowed us to develop a male-specific molecular marker. Validation of the developed molecular marker on an F1 generation of *C. revoluta* seedlings confirmed its specificity for the Y chromosome, further strengthening its reliability for sex determination in this species. The nearly equal sex ratio observed in the F1 generation indicated the absence of ecological or physiological factors influencing sex distribution, supporting the XY sex-determination system in *C. revoluta*.

Overall, our study highlights the efficacy of ddRAD sequencing as an alternative method for the rapid development of sex-linked molecular markers in non-model plant species with large genomes. The developed male-specific molecular marker presents a valuable tool for distinguishing between male and female *C. revoluta* plants, with potential applications in agriculture, breeding, and conservation efforts.

6 References

Abraham A, Mathew P (1962) Cytological studies in the Cycads: sex chromosomes in Cycas. Ann Bot 26:261–66. https://doi.org/10.1093/oxfordjournals.aob.a083792

Ahmadi H, Bringhurst R (1991) Genetics of sex expression in Fragaria species. Am Jou Bot. 78:504–514.

Barrett SCH (2002) The evolution of plant sexual diversity. Nat Rev Gen 3:274–84. https://doi.org/10.1038/nrg776

Bell CD, Soltis DE, Soltis PS (2010) The age and diversification of the angiosperms re-visited. Am J Bot 97:1296–303. https://doi.org/10.3732/ajb.0900346

Bensen RJ, Johal GS, Crane VC, Tossberg JT, Schnable PS, Meeley RB, Briggs SP (1995) Cloning and characterization of the maize An1 gene. Plant Cell 7:75–84. https://doi.org/10.1105/tpc.7.1.75

Blocka-Wandas M, Sliwinska E, Grabowska-Joachimiak A, Musial K, Joachimiak JA (2007) Male gametophyte development and two different DNA classes of pollen grains in *Rumex acetosa L*., a plant with an XX/XY1Y2 sex chromosome system and a female-biased sex ratio. Sex Plant Repr 20:171–80. https://doi.org/10.1007/s00497-007-0053-9

Byers RE, Baker LR, Sell HM, Herner RC, Dilley DR (1972) Ethylene: a natural regulator of sex expression in cucumber. Proc Natl Acad Sci USA 69:717–20

Charlesworth B, Charlesworth D (2000) The degeneration of Y chromosomes. Philos Trans R Soc Lond B 355:1563–72. https://doi.org/10.1098/rstb.2000.0717

Charlesworth D (2015) Plant contributions to our understanding of sex chromosome evolution. New Phytol 208:52–65. https://doi.org/10.1111/nph.13497

Chen RY, Song WQ, Li XL (1987) Study on the sex chromosomes of *Ginkgo biloba*. Proc Sino-Jpn Symp Plant Chromosome Res 381–86

Cuñado N, Navajas-Perez R, de la Herran R, Ruiz Rejon C, Ruiz Rejon M, et al (2007) The evolution of sex chromosomes in the genus Rumex (Polygonaceae): identification of a new species with heteromorphic sex chromosomes. Chrom Res 15:825–33. https://doi.org/10.1007/s10577-007-1166-6

Danilova T, Karlov G (2006) Application of inter simple sequence repeat (ISSR) polymorphism for detection of sex-specific molecular markers in hop (*Humulus lupulus L.*). Euphytica 151:15–21. https://doi.org/10.1007/s10681-005-9020-4

Delph LF, Arntz AM, Scotti-Saintagne C, Scotti I (2010) The genomic architecture of sexual dimorphism in the dioecious plant *Silene latifolia*. Evolution 64:2873–86. https://doi.org/10.1111/j.1558-5646.2010.01048.x

DeLong A, Calderon-Urrea A, Dellaporta SL (1993) Sex determination gene TASSELSEED2 of maize encodes a short-chain alcohol dehydrogenase required for stage-specific floral organ abortion. Cell 74:757–768. https://doi.org/10.1016/0092-8674(93)90522-R

Fraser LG, Tsang GK, Datson PM, De Silva HN, Harvey CF, et al (2009) A gene-rich linkage map in the dioecious species *Actinidia chinensis* (kiwifruit) reveals putative X/Y sex-determining chromosomes. BMC Genom 10:102. https://doi.org/10.1186/1471-2164-10-102

Freeman DC, Doust JL, El-Keblawy A, Miglia KJ, McArthur ED (1997) Sexual specialization and inbreeding avoidance in the evolution of dioecy. Bot Rev. 63:65–92. https://doi.org/10.1007/BF02857918

Goldberg MT, Spigler RB, Ashman TL (2010) Comparative genetic mapping points to different sex chromosomes in sibling species of wild strawberry (Fragaria). *Genetics* 186:1425–1433. https://doi.org/10.1534/genetics.110.122911

Grabowska-Joachimiak A, Sliwinska E, Pigula M, Skomra U, Joachimiak AJ (2006) Genome size in *Humulus lupulus L.* and *H. japonicus Siebold & Zucc.* (Cannabaceae). Acta Soc Bot Pol 75:207–14

Grabowska-Joachimiak A, Mosiolekb M, Lecha A, Goralski G (2011) C-banding/DAPI and in situ hybridization reflect karyotype structure and sex chromosome differentiation in *Humulus japonicus Siebold & Zucc*. Cytogenet Genome Res 132:203–211

Hizume M, Shiraishi H, Tanaka A (1988) A cytological study of *Podocarpus macrophyllus* with special reference to sex chromosomes. Jpn J Genet 63:413–23. https://doi.org/10.1266/jjg.63.413

Howell EC, Armstrong SJ, Filatov DA (2009) Evolution of neo-sex chromosomes in *Silene diclinis*. Genetics 182:1109–15. https://doi.org/10.1534/genetics.109.103580

Jaarola M, Martin RH, Ashley T (1998) Direct evidence for suppression of recombination within two pericentric inversions in humans: a new sperm-fish technique. Am J Hum Gene. 63:218–24

Janousek B, Mrackova M (2010) Sex chromosomes and sex determination pathway dynamics in plant and animal models. Biol. J Linn Soc 100:737–52. https://doi.org/10.1111/j.1095-8312.2010.01470.x

Jakse J, Stajner N, Kozjak P, Cerenak A, Javornik B (2008) Trinucleotide microsatellite repeat is tightly linked to male sex in hop (*Humulus lupulus L.*). Mol Breed 21:139–48. https://doi.org/10.1007/s11032-007-9114-x

Karlov GI, Danilova TV, Horlemann C, Weber G (2003) Molecular cytogenetics in hop (*Humulus lupulus L.*) and identification of sex chromosomes by DAPI-banding. Euphytica 132:185–90. https://doi.org/10.1023/A:1024646818324

Kenigsbuch D, Cohen Y (1990) The inheritance of gynoecy in muskmelon. Genome 33:317–20

Lan T, Chen RY, Li XL, Dong FP, Qi YC, Song WQ (2008) Microdissection and painting of the W chromosome in *Ginkgo biloba* showed different labelling patterns. Bot Studies 49:33–37.

Lee CL (1954) Sex chromosomes in *Ginkgo biloba*. Am J Bot 41:545–49. https://doi.org/10.2307/2438713

Lewis D (1942) The evolution of sex in flowering plants. Bio Rev Cambridge Philos Soc 17:46–67

Loptien H (1979) Identification of the sex chromosome pair in asparagus (*Asparagus officinalis L.*). Z Pflanzenz 82:162–73

Maloisel L, Rossignol JL (1998) Suppression of crossing-over by DNA methylation in Ascobolus. Genes Dev 12:1381–89. https://doi.org/10.1101/gad.12.9.1381

Marks M (1973) A reconsideration of the genetic mechanism for sex determination in *Asparagus officinalis*. In Proc. EUCARPIA Meeting on Asparagus (Asparagus officinalis L.) 123–28. Wageningen, Netherlands: EUCARPIA

Ming R, Yu Q, Moore PH (2007) Sex determination in papaya. In Semin Cell Biol 18:401–408. Academic Press. https://doi.org/10.1016/j.semcdb.2006.11.013

Ming R, Bendahmane A, Renner S (2011) Sex chromosomes in land plants. Ann Rev Plan Biol. 62:485–514. https://doi.org/10.1146/annurev-arplant-042110-103914

Moore RC, Kozyreva O, Lebel-Hardenack S, Siroky J, Hobza R, et al (2003) Genetic and functional analysis of DD44, a sex-linked gene from the dioecious plant *Silene latifolia*, provides clues to early events in sex chromosome evolution. Genetics 163:321–34. https://doi.org/10.1093/genetics/163.1.321

Nicolas M, Marais G, Hykelova V, Janousek B, Laporte V, et al (2005) A gradual process of recombination restriction in the evolutionary history of the sex chromosomes in dioecious plants. PLoS Biol 3:47–56. https://doi.org/10.1371/journal.pbio.0030004

Papadopoulou E, Little HA, Hammar SA, Grumet R (2005) Effect of modified endogenous ethylene production on sex expression, bisexual flower development and fruit production in melon (Cucumis melo L.). Sex Plant Reprod 18:131–42. https://doi.org/10.1007/s00497-005-0006-0

Polley E, Seigner E, Ganal MW (1997) Identification of sex in hop (*Humulus lupulus*) using molecular markers. Genome 40:357–61. https://doi.org/10.1139/g97-048

Pollock EG (1957) The sex chromosomes of maidenhair tree. J Hered 48:290–94. https://doi.org/10.1093/oxfordjournals.jhered.a106747

Poole CF, Grimball PC (1939) Inheritance of new sex forms in *Cucumis melo L*. J Hered 30:21–25

Reamon-Buttner SM, Schondelmaier J, Jung C (1998) AFLP markers tightly linked to the sex locus in *Asparagus officinalis L*. Mol Breed 4:91–98. https://doi.org/10.1023/A:1009650221460

Renner SS, Ricklefs RE (1995) Dioecy and its correlates in the flowering plants. Am J Bot 596–606. https://doi.org/10.1002/j.1537-2197.1995.tb11504.x

Rick CM, Hanna GC (1943) Determination of sex in *Aspiragus officinalis L*. Am J Bot 30:711–7l4

Ross MT, Grafham DV, Coffey AJ, Scherer S, McLay K, Muzny D, Joseph SS (2005) The DNA sequence of the human X chromosome. Nature 434:325–337. https://doi.org/10.1038/nature03440

Sakai AK, Weller SG (1999) Gender and sexual dimorphism in flowering plants: a review of terminology, biogeographic patterns, ecological correlates, and phylogenetic approaches. In Gender and Sexual Dimorphism in Flowering Plants, ed. MA Geber, TE Dawson, LF Delph, pp. 1–31. Berlin: Springer

Sakamoto K, Shimomura K, Komeda H, Satoh S (1995) A male-associated DNA sequence in a dioecious plant, *Cannabis sativa L*. Plant Cell Physiol 36:1549–54

Seefelder S, Ehrmaier H, Schweizer G, Seigner E (2000) Male and female genetic linkage map of hops, *Humulus lupulus*. Plant Breed 119:249–55. https://doi.org/10.1046/j.1439-0523.2000.00469.x

Segawa M, Kishi S, Tatuno S (1971) Sex chromosomes of *Cycas revoluta*. Jpn J Genet 46:33–39. https://doi.org/10.1266/JJG.46.33

Siljak-Yakovlev S, Benmalek S, Cerbah M, Coba de la Pena T, Bounaga N, et al (1996) Chromosomal sex determination and heterochromatin structure in date palm. Sex Plant Reprod 9:127–32. https://doi.org/10.1007/BF02221391

Smith CA, Roeszler KN, Ohnesorg T et al (2009) The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. Nature 461:267–271. https://doi.org/10.1038/nature08298

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Page DC (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423:825–837. https://doi.org/10.1038/nature01722

Tanurdzic M, Banks JA (2004) Sex-determining mechanisms in land plants. Plant Cell 16:S61 S71. https://doi.org/10.1105/tpc.016667

Telgmann-Rauber A, Jamsari A, Kinney MS, Pires JC, Jung C (2007) Genetic and physical maps around the sex-determining M-locus of the dioecious plant asparagus. Mole Genetic Genomic 278:221–234. https://doi.org/10.1007/s00438-007-0235-z

Trebitsh T, Staub JE, O'Neill SD (1997) Identification of a 1-aminocyclopropane-1-carboxylic acid synthase gene linked to the female (F) locus that enhances female sex expression in cucumber. Plant Physiol 113:987–95. https://doi.org/10.1104/pp.113.3.987

Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Rokhsar D (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313:1596–1604. https://doi.org/10.1126/science.1128691

Volz SM, Renner SS (2008) Hybridization, polyploidy, and evolutionary transitions between monoecy and dioecy in Bryonia (Cucurbitaceae). Am J Bot 95:1297–306. https://doi.org/10.3732/ajb.0800187

Wellmer F, Riechmann JL, Alves-Ferreira M, Meyerowitz EM (2004) Genome-wide analysis of spatial gene expression in Arabidopsis flowers. Plant Cell 15:1314–26. https://doi.org/10.1105/tpc.021741

Westergaard M (1964) Ojving winge. Biogr Memoirs Fellows R Soc 10:357–69

Yamasaki S, Fujii N, Matsuura S, Mizusawa H, Takahashi H (2001) The M locus and ethylene-controlled sex determination in Andromonoecious cucumber plants. Plant Cell Physiol 42:608–19. https://doi.org/10.1093/pcp/pce076

Yampolsky C, Yampolsky H (1922) Distribution of sex forms in the Phanerogamic flora. Bibl Genet 3:1–62.

Yin T, DiFazio SP, Gunter LE, Zhang X, Sewell MM, et al (2008) Genome structure and emerging evidence of an incipient sex chromosome in Populus. Genome Res 18:422–30. https://doi.org/10.1101/gr.7076308

Younis RAA, Ismail OM, Soliman SS (2008) Identification of sex-specific DNA markers for date palm (*Phoenix dactylifera L.*) using RAPD and ISSR techniques. Res J Agric Biol Sci 4:278–84.

Yu Q, Navajas-Perez R, Tong E, Robertson J, Moore PH, et al (2008) Recent origin of dioecious and ´ gynodioecious Y chromosomes in papaya. Trop Plant Biol 1:49–57. https://doi.org/10.1007/s12042-007-9005-7

Yu Q, Hou S, Feltus FA, Jones MR, Murray JE, Veatch O, Ming R (2008) Low X/Y divergence in four pairs of papaya sex-linked genes. Plant Journ 53:124–132. https://doi.org/10.1111/j.1365-313X.2007.03329.x

Zhang X, Feng B, Zhang Q, Zhang D, Altman N, Ma H (2005) Genome-wide expression profiling and identification of gene activities during early flower development in Arabidopsis. Plant Mol Biol 58:401–19. https://doi.org/10.1007/s11103-005-5434-6

Zhang W, Wang X, Yu Q, Ming R, Jiang J (2008) DNA methylation and heterochromatinization in the male-specific region of the primitive Y chromosome of papaya. Genome Res 18:1938–1943. https://doi.org/10.1101/gr.078808.108