AGRICULTURE AND NATURAL RESOURCES

Journal homepage: http://anres.kasetsart.org

Research article

Investigation of sex-linked DNA regions in *Cycas* species using amplified fragment length polymorphism markers

Chatuporn Kuleung

Department of Genetics, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand

Article Info

Article history:
Received 19 December 2019
Revised 14 February 2020
Accepted 4 March 2020
Available online 30 October 2020

Keywords:
AFLP,
Cycad,
Cycas,
Sex identification,
Sex-linked marker

Abstract

Cycads are dioecious plant species in which the male and female organisms can be visually identified only by their reproductive organs or cones. However, these plants must grow for several years before producing their first cones. DNA markers constitute an efficient tool that facilitates sex differentiation of cycads at an early age. Therefore, this research used amplified fragment length polymorphism (AFLP) markers to scan the male and female genomes of four *Cycas* species (*C. chamaoensis*, *C. clivicola*, *C. edentata* and *C. siamensis*) for polymorphic regions and their later utilization as specific markers. The DNA fingerprints of the male and female groups of each species were compared using 64 combinations of AFLP primers. Three primer pairs did not yield any DNA bands. Sixty-one primer pairs yielded 6–58 DNA bands, with an average of 30 bands per primer pair. In total, 302 polymorphic bands were obtained with 6.06%, 3.62%, 3.54% and 2.92% from *C. edentata*, *C. siamensis*, *C. chamaoensis* and *C. clivicola*, respectively. Next, 54 primer pairs designed from sequences of 141 selected polymorphic bands were tested on pooled and individual DNA. Though 35 primer pairs yielded a single amplified product, none could differentiate between male and female plants.

Introduction

Donaldson (2003) described cycads and their uses and management issues. Cycads are the most primitive extant seed-producing species. This makes them valuable for evolutionary studies of plant developmental systems. With their fascinating form, cycads are popular in landscape gardening. Habitat loss, retarded growth and smuggling have caused rapid declines in their natural populations, threatening the existence of these species. Management for conservation and utilization requires information on their population structure, including sex distribution. However, sex identification of the plants can only be done morphologically using its cone, restricting the observation to adult plants. Genomic variation linked to sex can be used as a marker to assist in sex identification of a plant at any developmental stage.

Cycas species, like other cycads, are dioecious with all species having an equal chromosome number (2n = 22) and nearly all of them have indistinguishable sex chromosomes (Sangduen et al., 2009). However, their sex determination is assumed to be based on an XX/XY system as is the case with *C. revoluta*, which has distinct sex chromosomes (Segawa et al., 1971). Different numbers and locations of satellites between male and female karyotypes are also observed in *C. pectinata*, *C. clivicola*, *C. edentata*, *C. elephantipes*, *C. siamensis* and *C. transachana*, but the sex chromosomes of these species are not recognized (Abraham and Mathew, 1962; Sangduen et al., 2009). Using the random amplified polymorphic DNA (RAPD) technique, Gangopadhyay et al. (2007) investigated genomic variation between male and female plants of *C. circinalis* and detected two sex-linked

E-mail address: fscictp@ku.ac.th (C. Kuleung)

https://doi.org/10.34044/j.anres.2020.54.5.14

bands that can be used as sex-specific markers, facilitating sex identification in this species.

In species like cycads with limited sequence information, randomly amplified techniques can be used to simultaneously amplify multiple regions of genomes, increasing the likelihood of finding polymorphic regions associated with sex (Gangopadhyay et al., 2007; Mwase et al., 2007). Amplified fragment length polymorphism (AFLP) is one technique used to search for a genomic region associated with a trait; in turn, the DNA sequence retrieved from these regions can be used as a sequence-characterized amplified region (SCAR) marker and sex-specific markers derived from the AFLP technique have been reported in many non-model organisms such as *Uapaca kirkiana* Muell. Årg. (Mwase et al., 2007), *Eucommia ulmoides* Oliv. (Wang et al., 2011) and *Piper betle* L. (Goyat et al., 2019).

The objective of this research was to detect regions that varied between male and female genomes in four *Cycas* species using the AFLP technique, for the purpose of later using the sequences derived from these regions as sex-specific markers.

Materials and Methods

Leaf samples of male (M) and female (F) plants of *C. chamaoensis* (3 M, 5 F), *C. clivicola* (5 M, 5 F), *C. edentata* (2 M, 3 F) and *C. siamensis* (3 M, 3 F) were collected from Nong Nooch Tropical Botanical Garden, Chonburi, Thailand. Genomic DNA was extracted using an Invisorb Spin Plant Mini Kit (Invitek; Germany) according to the manufacturer's instructions. From each species, equal volumes of 50 ng/μL DNA samples from individuals of the same sex were pooled together. The pooled DNA was digested using FastDigest *Eco*RI and FastDigest *Tru*1I (Thermo Fisher Scientific Inc.; USA) and then ligated to adapters. Adapter sequences were used and the AFLP procedure was followed (Vos et al., 1995). Next the ligated DNA was pre-selectively amplified with *Eco*RI and *Tru*1I primers having one selective base (E-A: 5'-GACTGCGTACCAATTCA-3' and T-C: 5'-GATGAGTCCTGAGTAAC-3', respectively). The polymerase chain reaction (polymerase chain reaction) products were then used as templates for selective amplification using 64 combinations of AFLP primers with three selective bases (Table 1). The amplified products were separated in 6% denaturing polyacrylamide gel and stained with silver nitrate (Benbouza et al., 2006). Some of the polymorphic bands with clear bands were cut and cloned with pGem-T easy vector following the manufacturer's instructions (Promega Corporation; USA). Three colonies from the same AFLP band were selected for colony PCR. The clone with the expected insert size was sent for sequencing by Macrogen Inc. (Korea).

The sequences that had high-quality base calls, a size larger than 100 bp excluding AFLP primers, and non-redundancy were subjected for primer design using Primer3 (Koressaar and Remm, 2007; Untergasser et al., 2012). All primer pairs were tested on DNA pools and those showing polymorphism between male and female groups were then further tested individually to verify their use for sex identification.

Results and Discussion

The DNA profiles of all four *Cycas* species generated using the AFLP technique showed high similarity between male and female groups. Of the AFLP-primer combinations, 64 yielded 1,893, 1,882, 1,864 and 1,850 total bands for *C. chamaoensis*, *C. clivicola*, *C. edentata* and *C. siamensis*, respectively (Table 1). On average, about 30 bands per primer combination were produced, with the highest numbers of bands from the E-ACC/T-CAA and E-AAG/T-CGC primer pairs and the lowest from the E-AGT/T-CAC primer pair. Polymorphism between male and female groups was found at frequencies of 6.06%, 3.62%, 3.54% and 2.92% in *C. edentata*, *C. siamensis*, *C. chamaoensis* and *C. clivicola*, respectively. Out of the 302 polymorphic bands from all species, 141 clear bands were randomly selected for further cloning. From 91 AFLP bands, 263 colonies could be reamplified with the expected insert size and were subjected to sequencing, resulting in 57 non-redundant sequences larger than 100 bp being obtained and 54 SCAR markers were subsequently designed (supplementary file). Testing these markers on male and female DNA pools showed that 35 markers yielded a single amplified product, but none yielded a band associated with sex. Therefore, none of the SCAR markers were tested further on individuals.

A comparison of the sequences to those found in the NCBI database, using the BLASTn tool, revealed that two sequences showed substantial similarities to known genes. One sequence, T-CGA/E-AGC_063_1_C.chamaoensis_F, was matched to 26S ribosomal RNA genes found in many conifer and cycad species (identity 91–92% and e-value < $3e^{-80}$). Another sequence, T-CGG/E-AGG_173_2_C. chamaoensis_F, aligned with transcripts of predicted 40S ribosomal protein S16, present in many species in the Phaseoleae tribe and Rosaceae family (identity 80–88% and e-value < $1e^{-34}$). Most sequences showed no homology to sequences in the database, suggesting that AFLP fragments were primarily produced from variable non-coding sequences. Similar results have been reported in *Ficus fulva* (Parrish et al., 2004) and *Eucommia ulmoides* Oiv. (Wang et al., 2011). The AFLP technique has proven useful for detecting genomic variation in low-divergence samples. Even given high genetic similarity among the samples tested in this study, several polymorphic bands could be detected. Unfortunately, no specific marker developed from these AFLP bands could be used to differentiate sex in the tested *Cycas* species. Loss of polymorphism could be caused by the removal of polymorphic ends from the AFLP-derived sequences. The remaining sequences may not contain variation; alternatively, they may contain variations such as single nucleotide polymorphisms (SNPs), which cannot be differentiated using agarose gel electrophoresis. The presence of SNPs that create another recognition site within the restriction fragments can produce polymorphisms within homologous sequences; however, SCAR markers do not have this effect (Zhang and Stommel, 2001).

Table 1 Total numbers of amplified fragment length polymorphism-generated bands and polymorphic bands between male and female DNA pools

Primer combination	*C. chamaoensis*		*C. clivicola*		*C. edentata*		*C. siamensis*	
	Total bands	Polymorphic bands	Total bands	Polymorphic bands	Total bands	Polymorphic bands	Total bands	Polymorphic bands
E-AAC/T-CAA	34	0	35	0	28	0	33	0
E-AAC/T-CAC	0	0	0	0	0	0	0	0
E-AAC/T-CAG	38	1	33	0	43	1	34	1
E-AAC/T-CAT	12	0	12	0	10	0	13	0
E-AAC/T-CGA	26	1	28	0	36	0	23	0
E-AAC/T-CGC	45	4	44	0	41	2	45	4
E-AAC/T-CGG	36	1	39	5	37	0	35	2
E-AAC/T-CGT	24	0	27	1	30	1	25	1
E-AAG/T-CAA	36	0	36	0	37	0	35	0
E-AAG/T-CAC	0	0	0	0	0	0	0	0
E-AAG/T-CAG	26	2	23	0	10	0	24	1
E-AAG/T-CAT	15	0	15	0	17	0	15	0
E-AAG/T-CGA	28	0	26	0	32	4	25	0
E-AAG/T-CGC	51	1	55	2	55	6	58	4
E-AAG/T-CGG	21	1	20	1	21	0	23	2
E-AAG/T-CGT	36	0	37	1	33	0	37	2
E-ACA/T-CAA	39	0	36	0	41	0	38	0
E-ACA/T-CAC	0	0	0	0	0	0	0	0
E-ACA/T-CAG	36	0	35	0	39	2	35	0
E-ACA/T-CAT	12	0	12	0	13	0	12	0
E-ACA/T-CGA	31	0	28	0	21	0	25	0
E-ACA/T-CGC	42	3	41	1	41	5	38	3
E-ACA/T-CGG	31	2	31	1	24	2	29	0
E-ACA/T-CGT	41	4	48	3	42	2	43	3
E-ACC/T-CAA	55	2	55	2	54	1	50	1
E-ACC/T-CAC	51	1	48	1	52	1	47	3
E-ACC/T-CAG	31	1	29	0	30	2	31	3
E-ACC/T-CAT	24	0	23	0	25	0	24	0
E-ACC/T-CGA	39	1	37	0	39	0	37	0
E-ACC/T-CGC	34	4	39	2	32	2	32	3
E-ACC/T-CGG	27	0	29	0	33	3	28	2
E-ACC/T-CGT	43	2	44	0	34	2	42	0
E-AGA/T-CAA	45	2	40	0	38	1	42	0
E-AGA/T-CAC	31	0	32	0	36	1	34	0
E-AGA/T-CAG	33	0	26	0	32	1	29	3
E-AGA/T-CAT	32	0	32	0	30	1	32	0
E-AGA/T-CGA	22	1	19	1	22	0	20	2
E-AGA/T-CGC	36	1	31	1	29	3	30	0
E-AGA/T-CGG	36	2	41	1	28	0	35	1
E-AGA/T-CGT	34	0	36	0	36	4	38	0
E-AGC/T-CAA	41	1	40	3	35	7	40	1
E-AGC/T-CAC	39	1	38	2	32	8	38	1
E-AGC/T-CAG	31	3	31	1	33	4	31	0
E-AGC/T-CAT	27	0	25	0	33	0	24	0
E-AGC/T-CGA	19	1	25	0	25	4	20	2
E-AGC/T-CGC	24	2	29	2	29	3	27	1
E-AGC/T-CGG	18	0	15	0	23	1	15	1
E-AGC/T-CGT	38	0	39	0	38	7	39	2
E-AGT/T-CAA	12	0	12	0	6	0	12	0
E-AGT/T-CAA	37	2	37	1	41	1	37	0
E-AGT/T-CAC	6	0	6	0	7	0	7	0
E-AGT/T-CAC	32	0	31	0	46	3	35	0
E-AGT/T-CAG	45	1	43	1	45	3	45	2
E-AGT/T-CAG	41	0	36	1	37	0	39	1
E-AGT/T-CAT	15	0	15	0	20	0	15	0
E-AGT/T-CAT	21	0	20	0	20	0	21	0
E-AGT/T-CGA	28	4	26	1	20	1	20	0
E-AGT/T-CGA	20	3	16	1	14	0	16	0
E-AGT/T-CGC	31	3	31	1	27	5	35	4
E-AGT/T-CGC	32	4	35	10	26	5	34	3
E-AGT/T-CGG	11	0	10	1	9	1	10	0
E-AGT/T-CGG	24	4	28	2	24	3	25	1
E-AGT/T-CGT	39	1	40	2	41	4	41	3
E-AGT/T-CGT	29	0	32	3	32	6	28	4
Total	1893	67	1882	55	1864	113	1850	67

In addition, the cleavage properties of the restriction enzymes used for DNA digestion can affect the presence or absence of an AFLP band. *Eco*RI, a typical rare-cutter enzyme used in the AFLP approach included in this study, is CpG methylation-sensitive meaning that the enzyme does not cut some methylated recognition sites but does cut unmethylated sites (Pray, 2008). Therefore, differential methylation results in different banding patterns for identical sequences. Consequently, a SCAR marker designed from this sequence will yield a monomorphic pattern for both sexes. In *Cycas*, Kanchanaketu et al. (2007) reported variable methylation patterns in CCGG sequences between male and female groups, indicating that differential methylation between sexes actually occurs and could be another explanation for AFLP polymorphic loss.

The discovery rate of a sex-specific marker located via genome scanning with multilocus markers varies regardless of the technique used, as shown in the review by Milewicz and Sawicki (2013). For example, one male-specific band and one female-specific band were identified from 10 RAPD primers in *C. clivicola* (Gangopadhyay et al., 2007), one male-specific marker was obtained from 300 RAPD primers in *Phoenix dactylifera* L. (Al-Qurainy et al., 2018) and one male-specific marker was screened from 64 AFLP primer combinations in *Eucommia ulmoides* Oliv. (Wang et al., 2011). Based on *in-silico* restriction fragment analysis, Caballero et al. (2013) recommended using an enzyme pair with numerous GC content restriction sites combined with high-GC content-selective nucleotides of AFLP primers to increase the probability of AFLP fragments linking to a gene.

Development of sex-linked markers in plants tends to be more successful in species with heteromorphic sex chromosomes such as *Cannabis sativa* (Sakamoto et al., 1995), *Silene latifolia* (Zhang et al., 1998), and *Rumex acetosa* (Moriotti et al., 2009; Gangopadhyay et al., 2007). However, heteromorphic sex chromosomes are uncommon in dioecious plants (Charlesworth, 2002). The difficulty in developing a sex-linked marker indicates that DNA regions involved in sex differentiation are probably very small. Thus, a sequence-level investigation is required. Alternatively, sex differentiation may possibly be regulated by other mechanisms apart from nucleotide differences.

Epigenetic sex-determination mechanisms regulate differential gene expression, thereby influencing sex differentiation, with methylation known to have epigenetic effects and DNA methylation bias in Y chromosomes has been observed in the males of many dioecious species. For example, in papaya (*Carica papaya*), Zhang et al. (2008) found a stronger 5mC immunofluorescence signal at the heterochromatic knobs on the male-specific region of the Y chromosome (MSY), compared to its X counterpart. Bräutigam et al. (2017) detected the distinct methylation pattern of a gene in a sex-determining region, revealing highly methylated promoter and intron 1 of PbRR9 in the male balsam poplar (*Populus balsamifera*). Similarly, Lorenzo et al. (2018) reported that the promoters of sex-linked stratum-I genes of males were highly methylated compared to those of females, with higher levels in Y alleles than in X alleles. For cycads, the different methylation-sensitive amplification polymorphism patterns observed in male and female groups of *Cycas* and *Zamia* presented evidence of differential methylation, suggesting that cycad sex is possibly controlled by epigenetic factors (Kanchanaketu et al., 2007).

The accumulation of repetitive sequences plays an important role in shaping sex chromosomes. Insertion of transposable elements may lead to methylation of the elements and neighbor sequences via genomic defense mechanisms. For example, Na et al. (2014), by analyzing repetitive sequences in hermaphrodite-specific regions of the papaya Y chromosome (HSY), found that repetitive content in HSY was greater than in its X counterpart and that most of the repetitive sequences were retrotransposons. BLAST analysis of the repeats identified 21 potential sex-specific repeats, with 20 repeats specific to HSY and 1 to the X counterpart. Many other studies have reported additional types of repetitive sequences associated with sex or sex chromosomes, such as higher numbers of repeated AC and AAC microsatellite motifs that occur in males of *Rumex acetosa*, as well as the RAYSI satellite DNA family found on the Y1 and Y2 chromosomes of *R. acetosa* (Kejnovský et al., 2013; Hobza et al., 2017).

Sex-linked markers may simplify sex identification in *Cycas* species. However, the negative results of this research indicated that other new techniques are required to discover sex-linked markers for these species. Currently, next-generation sequencing (NGS) technology provides detailed sequence information at affordable prices and detects DNA modifications, such as methylation. Large quantities of NGS data could allow numerous types of genomic variation associated with sex, including SNP, differential methylation and repetitive sequences, to be identified in *Cycas*.

Conflict of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

The authors thank Mr. Anders J. Lindstrom and the Nong Nooch Tropical Botanical Garden in Chonburi, Thailand for providing the cycad samples used in this research. Associate Professor Mingkwan Mingmuang provided useful advice. This research was co-funded by the Thailand Research Fund, the Commission on Higher Education, and Kasetsart University, Bangkok, Thailand (MRG5480233). In addition, the project was supported in part by the Faculty of Science, Kasetsart University, under the Science Research Fund program.

References

Abraham, A., Mathew, P.M. 1962. Cytological studies in the cycads: Sex chromosomes in *Cycas*. Ann. Bot. 26: 261−266. doi.org/10.1093/oxfordjournals.aob.a083792

Al-Qurainy, F., Al-Ameri, A.A., Khan, S., Nadeem, M., Gaafar, A.Z., Tarroum, M. 2018. SCAR marker for gender identification in date palm (*Phoenix dactylifera* L.) at the seedling stage. Int. J. Genomics. 2018: 3035406. doi.org/10.1155/2018/3035406

Benbouza, H., Jacquemin, J.M., Baudoin, J.P., Mergeai, G. 2006. Optimization of a reliable, fast, cheap and sensitive silver staining method to detect SSR markers in polyacrylamide gels. Biotechnol. Agron. Soc. Environ. 10: 77−81.

Bräutigam, K., Soolanayakanahally, R., Champigny, M., Mansfield, S., Douglas, C., Campbell, M.M., Cronk, Q. 2017. Sexual epigenetics: Gender-specific methylation of a gene in the sex determining region of *Populus balsamifera*. Sci. Rep. 7: 45388. doi: 10.1038/ srep45388

Caballero, A., García-Pereira, M.J., Quesada, H. 2013. Genomic distribution of AFLP markers relative to gene locations for different eukaryotic species. BMC Genomics 14: 528. doi:10.1186/1471-2164-14-528

Charlesworth, D. 2002. Plant sex determination and sex chromosomes. Heredity 88: 94–101.

Donaldson, J.S. 2003. Cycads. Status Survey and Conservation Action Plan. IUCN/SSC Cycad Specialist Group. IUCN, Gland, Switzerland and Cambridge, UK.

Gangopadhyay, G., Roy, S.K., Ghose, K., Poddar, R., Bandyopadhyay, T., Basu, D., Mukherjee, K.K. 2007. Sex detection of *Carica papaya* and *Cycas circinalis* in pre- flowering stage by ISSR and RAPD. Curr. Sci. 92: 524–526.

Goyat, S., Grewal, A., Singh, D., Katiyar, R.S., Tewari, S.K., Nainwal, R.C., Hima Bindu, K. 2019. Sex-linked AFLP marker identification in dioecious Betelvine (*Piper betle* L.). J. Hortic. Sci. Biotechnol. 94: 422–427. doi.org/ 10.1080/14620316.2018.1538707

Hobza, R., Cegan, R., Jesionek, W., Kejnovsky, E., Vyskot, B., Kubat, Z. 2017. Impact of repetitive elements on the Y chromosome formation in plants. Genes 8: 302. doi: 10.3390/genes8110302.

Kanchanaketu, T., Hongtrakul, V., Sangduen, N. 2007. Analysis of sex determination in some cycads using methylation-sensitive amplification polymorphism (MSAP). Kasetsart J. (Nat. Sci.). 41: 641–650.

Kejnovský, E., Michalovova, M., Steflova, P., et al. 2013. Expansion of microsatellites on evolutionary young Y chromosome. PloS One 8: e45519. doi.org/10.1371/journal.pone. 0045519

Koressaar, T., Remm, M. 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics. 23: 1289–1291. doi.org/10.1093/ bioinformatics/btm091

Lorenzo, J.L.R., Hobza, R., Vyskot, B. 2018. DNA methylation and genetic degeneration of the Y chromosome in the dioecious plant *Silene latifolia*. BMC Genomics 19: 540. doi:10.1186/s12864-018-4936-y

Milewicz, M., Sawicki, J. 2013. Sex-linked markers in dioecious plants. POJ. 6: 144–149.

Moriotti, B., Manzano, S., Kejnovský, E., Vyskot, B., Jamilena, M. 2009. Accumulation of Y-specific satellite DNAs during the evolution of *Rumex acetosa* sex chromosomes. Mol. Genet. Genomics. 281: 249–259. doi: 10.1007/s00438-008-0405-7.

Mwase, W.F., Erik-Lid, S., BjØrnstad, Å., Stedje, B., Kwapata, M.B., Bokosi, J.M. 2007. Application of amplified fragment length polymorphism (AFLPs) for detection of sex-specific markers in dioecious *Uapaca kirkiana* Muell. Årg. Afr. J. Biotechnol. 6: 137–142.

Na, J.K., Wang, J., Ming, R. 2014. Accumulation of interspersed and sex-specific repeats in the non-recombining region of papaya sex chromosomes. BMC Genomics 15: 335. doi: 10.1186/1471-2164-15-335.

Parrish, T.L., Koelewijn, H.P., van Dijk, P.J. 2004. Identification of a male-specific AFLP marker in a functionally dioecious fig, *Ficus fulva* Reinw. ex Bl. (Moraceae). Sex. Plant. Reprod. 17: 17–22. doi: 10.1007/s00497-004-0208-x

Pray, L. 2008. Restriction enzymes. Nature Education 1: 38.

Sakamoto, K., Shimomura, K., Komeda, Y., Kamada, H., Satoh, S. 1995. A male-associated DNA sequence in a dioecious plant, *Cannabis sativa* L. Plant Cell Physiol. 36: 1549–1554.

Sangduen, N., Toahsakul, M., Hongtrakul, V. 2009. Comparative karyomorphological study between male and female plants of some *Cycas* and *Zamia* species. Kasetsart J. (Nat. Sci.). 43: 476–485.

Segawa, M., Kishi, S., Tatuno, S. 1971. Sex chromosomes of *Cycas revoluta*. Jap. J. Genet. 46: 33–39. doi.org/10.1266/jjg.46.33

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen. S.G. 2012. Primer3-new capabilities and interfaces. Nucleic Acids Res. 40: e115. doi.org/10.1093/nar/gks596

Vos, P., Hogers, R., Bleeker, M., et al. 1995. AFLP: A new technique for DNA fingerprinting. Nucleic Acids Res. 23: 4407–4414. doi: 10.1093/ nar/23.21.4407

Wang, D.W., Li, Y., Li, Z.Q. 2011. Identification of a male-specific amplified fragment length polymorphism (AFLP) and a sequence characterized amplified region (SCAR) marker in *Eucommia ulmoides* Oliv. Int. J. Mol. Sci. 12: 857–864. doi: 10.3390/ijms12010857

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. doi: 10.1101/ gr.078808.108

Zhang, Y., Stommel, J.R. 2001. Development of SCAR and CAPS markers linked to the beta gene in tomato. Crop Sci. 41: 1602–1608.

Zhang, Y.H., Stilio, V.S., Rehman, F., Avery, A., Mulcahy, D., Kesseli, R. 1998. Y chromosome specific markers and the evolution of dioecy in the genus *Silene*. Genome. 41: 141–147. doi.org/10.1139/g97-094

Supplementary

Amplified fragment length polymorphism-derived DNA sequences, where sequence-characterized amplified region primer sequences are underlined

>T-CAA/E-ACC_017_3_C.clivicola_M

CCTCATCTAGGTCTAGCCTAGTCTCCAAATACCAAAGGACAAGTTTCTCACCTCAGTCTGCCACTTGATACACCTTACAGCCAATGGTTTCTTCTAATATCTGCAGAATTCATCTCTTCATCCATAGCTGCTTTCTACGTTTGT

>T-CAA/E-AGC_019_2_C.edentata_F

GCCAAGCACGGGAACCCGCAGGACGGAATACAGAACGCACGGGAGTCGAACCCGGACGGGCGAGGTCGGGGCGAAACGCGGAGCAGTTCACGTGACGGGGCGT

>T-CAA/E-AGC_022_2_C.chamaoensis_F

GCTAGACTGGGACATTATGCATAATAGAAGTCGACTAGACTATAGAAAGAAATGCTACAAGTCCCACCCCATGAAAACATGATTTGCTCTCAAAGCATGGGTAAATTGAGGT

>T-CAA/E-AGG_024_1_C.edentata_M

GGTATATCAATGGTTATGGTAAAGGGATAGTTCCATTAGCGTTCCCTATCACCAGATGGAGGGAGATCTTTAGGATCAAAAGAGATGACAAGTTTACGACTTAGGGTGATACCCAAGACATGGACATACGT

>T-CAC/E-AGC_011_5_C.edentata_F

GCTTGAAATATAAGAAAGGAGTTATCAACAAAATATCCATTGACAAGCTACTCGCAAGACTCCCCTAGGAGAGTAATGCCTCTGACAAGATCTTGGGAAATTGTGTCCTAGAGAAGATAACAAAGGGCATCGATTGTTAGCACATATGGGGATGGT

>T-CAC/E-AGC_013_2_C.edentata_F
GCACCAGATTCGAAGCCAATTATCAAATACAGT
CAACCATTCCAAGTAGGATGGGGGGTGTATATATACTG
TAACTCTCTTTATTGGAATCAAGAGCTAATGGAAAACATTCT
GCTGGT

>T-CAG/E-AAG_034_1_C.siamensis_M
AGGTTCATCCATATGGAGGAGCCTAGAGTCCATG
AGAAAAATGATAACGATGGGAGGCCCAATGGAGCTTGG
GAGATGGAAAAGATACGGACTTTTGGGAGGACACTTGA
CATGGGCATAGACAATTGAACAAATGACCTACTTTCGC
CTCCTTGCAACCTGCCT

>T-CAG/E-AAG_034_2_C.siamensis_M
TCAACTACTTTGATGTAGCCCATCCATTTGAGAG
AGGAATGAAATCCCTCTTATCACTCAAATAAGGTTACG
ATGAAATGAATCGTGATTTTGTAAGATGATTACATGACC
TCCTTCGCTCTACTAAGGCCACGCTGCCCAAGGAGCTA
AAGAGATACCTTTTC

>T-CAG/E-AAG_035_1_C.chamaoensis_F
AGAGGAAGCTGGACCGAGTGGAGAGAGTAATAATGAG
GAGTTGTCACTTGAAGAAATACCTACAGCACCAAGTCCTA
TACCAATATGGAGAGGGAAGATAATAGAGGTTTCCAGAGT
GGCGGTGCAAAACTTTTCTGATGACAAAAGAACCAGAA
GCCAACATGTTAGTCAGACTTTGCT

>T-CAG/E-ACA_036_1_C.edentata_M
CAAAACTGCATCTACTTCGCTTATGGGCATGGCAAACA
TACAAGTTGTGAACTTGTAATCATTGATCTAGAGAGTCATG
TCTTTGCAGTATCCTCCACACAGCATTGTTCTCCCATTTGCT

>T-CAG/E-ACA_036_2_C.edentata_M
CATCAACAATAACAACAAGAACAACAACCTCAACAAC
AAAGGCAACATCAACATCAACAACAATGCAAAGAACGAC
CACCTGGGCAATGTGGTCAAGTTTCGAATGTAGTAGAAAA
GTCACCT

>T-CAG/E-ACC_029_1_C.siamensis_F
CCATCATTAGGTGCTCCTAAGCAATGAGGTTCTACTAC
TAAAGGGACAACATTTCCTAGCTATAAAGAATGGTTAGG
AGTGGTACCTATTAGGTAAGTACATGTATACATCCTCAC
CTCTTGCCCTATTCACTTTGATGGCT

>T-CAG/E-ACC_029_2_C.siamensis_F
CCATCTCCAGGCTCTAGTGGTGTGTCAAATAGCTTCC
TTGTGAAATAGATCTCCCAAGATGCCACAAGACCTTCAT
CTACCTTCTCTTTGTTTTCAGTCATGCTCTTTGAGCTATCC
TTTGAAGTAAGATTTCCCCGAACCACT

>T-CAG/E-ACC_030_1_C.edentata_F
CCCATACTGGTAGTCCATCTACCAGAACCTCCTCTCTC
ATCGCCCAGGGATAGAGCCGAATTGAAATCTCCCGCCAA
CAGCCAAAGTTGTTGGTAGTCCCTATGAACCTCCAGTAT
CCTATCCCAGAGTTTCTTCCTGGCCT

>T-CAG/E-ACC_031_1_C.chamaoensis_F
CCAAATAATGAAGAATAGATTACCAACATTTGATTGAT
CCAAGATTAGCCCAAAATCATGTTGAACTCAAGATCAAA
GTCAATGATGGTGACTTTTACTATAATAGTCTTGCCT

>T-CAG/E-ACC_031_2
CCATTATATACCTTACAAGAACATTTGGATCCCTGCTTTT
CTGGGTTTATAACTTATCATGTTATAGGACCACACCCTTTCA
ACGGAGACAGTTGATCAATGAAGTCGTATCCT

>T-CAG/E-ACC_032_1_C.siamensis_F
CCCAATTTTGCAAGGATAATGGTATCAAACGATAGGTTT
TTGCCCCTTACACACCTTAGCAAAATGGAATTGTTGAGAG
AATGAATTGGATCTTGGTTAGTAAAGGCAAAGCCATGCTT
GCT

>T-CAG/E-ACC_032_2_C.siamensis_F
CCAAGGCCTAGGTTACAAATCTTCAAACCCAAGCATG
CCCGCCCTTCCATTTCTAATGTCTATTTATCACATGCCCTTT
GCCATTCTTGATGGCCTTTAGCATTTACTTATCTGCT

>T-CAG/E-AGA_045_2_C.edentata_M
GAAGGGAAAACGACCTATAGGAAACTATAGCAAGGAAC
AAGATGGCAAAGATGAAAAAGAAACTACATTCATTGTTGC
ACTTACAGTAAGCCTTGAAGACGAGGAGGTGTGGTACATT
GATTCCAGAGCATCAAGACATATGATTAGGCAAGAGGATT
GGTCCAAATCCATAAAACACTACACCTCCAACAAGTCTGT
TGAACTAGGGGACAAGCATCCT

>T-CAG/E-AGA_047_1_C.siamensis_F
GACTCGAAGGCCACAATTGTGGTGTGAGTTGGTTTGTT
GGATGTCATCTGAAGGCGATTGGCAATCTGTAATAGTGATT
GAAGCCAATAGAGCACGTTCTGATGGTGATTCGATGCT

>T-CAG/E-AGC_041_1_C.chamaoensis_F
GCTTTGTAAATCTGCACTGTGCTGCTAAATTTCACTGCT
GTTATGAACTCAACTTTGTAAATCTACACTGTGCTGCTAAA
TTTCACTGTTGTTATGAACTCAGCTTTGTAAATCTGAACTG
TGTTGCTAAATTTCACTGCTGTTATGAACTCAGCCTGGTAA
TTCTGCCT

>T-CAG/E-AGC_041_2_C.chamaoensis_F
GCAGGAATGGCAGATGATACCAGAAGACGACAAGAGA
CGCCAGTTAGTCCGTCACATAGAGTTGCAGAAATTCGTGA
CAGTGGGAGTGTTTGGAAATGCAGAAAAATTCATGCAAA
GAATACTTCCGCTACTAGCTAAGTACCCGCAAATGGTAC
AAATAACTGAAGACCT

>T-CAG/E-AGT_049_2_C.edentata_F
GTAGCAGATTGGTTGGCAAACTGGGCAGTGTCAAAT
AAAAGAACAAAAGTTTAGAAAGAAGAATACCTAGAGGA
GTTATTGAGACTGGTCAGGGTTGATGCACATATGAGGGA
GGTAGTCTGACAAGAACGAGGGATCTAACCTTGGGCTA
GACCATGGGCTTGTCACGGTTATCCTTGCTAAGCGGTTT
GGCCTTGATGTACAAGGTTTGGGGACCCT

>T-CAG/E-AGT_050_1_C.edentata_F
GTGATAACTGGATATGCAAGTTTGAATAGATTGTCTCAA
CAGATGGAGACTGAAGAGTTTGTGCCCAACACAAACTAAC
TACTATAGACTGGCATGCCAAATGTGAGAGCATAGAGGGTG
CGTGCAAAGTGCTTGACAATATGCT

>T-CGA/E-AAC_055_1_C.chamaoensis_F
ACGGTTTCATTCTATAAACGGGCGCATTTCTATAAAC
GGGCGCAGTTTCAGAAATTTCCAGGACTGGACTCACGTCA
CGGGATGTGAGTATAATCAAAGCACTCCACCAAAAAACAT
CTGTCCCTAATAAATCTAATATGATTAGAGCATTAGCAGCA
GCAAAAGGGGCCAGTCTTTATCTGAGAATACCTCTGCCTC
GATCTAAATTATCAATTTTTGAACTATC

>T-CGA/E-AAG_056_1_C.edentata_F

TCAACTACTTCGATGTAGCTTGTCCATTCGAAAACG
AAATAAAATCTCTTCTATCACTCAAACAAGGTTATGATG
AAATGAATCATGACTTTGTGTGATGATTCCATGACCTTCTT
CACTCCACTAAGGCCACATTGCCCAAGGAGTTGAAGA
GAAACCTTTTT

>T-CGA/E-AAG_056_2_C.edentata_F

AGGTGTTTATACCCAATTTTGTAGTATGTAGATATGTA
TTAGGTTAGTTTATTTCTCATCTTTTTTTTAGTGTGTGTCTT
GGTTGATTTTATTCTATTGGTGATATATTGCGATCTTGCCCA
TCTCTTGTGCCACACGTC

>T-CGA/E-AAG_057_1_C.edentata_F

AGGGGTATGCACCATTTTTCAAATAGATATTTTTGCAG
AATCCTCCACAGGGTGCTGTCTTCCAGTTGGCAATTGAG
ACTTGGAAGGTTTTCAACAGTTCCACCTTTATTCCTAGA
TTTTATGCTGTTGATTTGTC

>T-CGA/E-AAG_058_2_C.edentata_F

AGCAGAAAGAAGGAATTGACTAAGACGTTTACACCA
GTCTCTAGAATGACAACAGTCTGTTTACTCATTTCCTTAG
CAACTGCGTTAGATTAGGAGTTACACCAGATGGTC

>T-CGA/E-AGA_065_2_C.siamensis_M

GAGGAGTCAGGTAAAGACCACGGGGAGTGATTGA
TTTTTCCCCATTTGGGTAGGAAGAGAGGACTTAGTTCTCT
GGTTCATGTCTTGTGACTTCCATGATAGAAGACATATAG
AAAACACTCAAACATGACCCATATATGCAAAGCCCTTTAT
CCTTTTAGAGTAACAGTTGAACCTCACCTATGGTGAACAA
CAATGTGGCTGGAATATCCTCCTTTAGTAACTGTGCCATC
TACATCGATGTC

>T-CGA/E-AGC_060_1_C.siamensis_F

GCACACATACATAGCAGAAATCAACCTGGCTCTGATA
CCATGTTAGAATGAGGGATATAAGCTAGCCTCACAACAT
AAGTCACGTGCAAAAAAAAGCAAGAGGTGATGATACG
CCAATAAGACTAAACTCTCTAATTCTCATTGAAAAAATT
CTTGATGATACAAGTACATATTTATAAATTTGCATCATAGC
TAAACATAGGGGTTGGTTTCAACGAACTCCAACTTCCC
TTTATTTTTGGTGTGCTCAAACTACCTC

>T-CGA/E-AGC_063_1_C.chamaoensis_F

GCTTCACAATGCATAGGAAGAGCCAACATCAAAGGAT
CAAAAAGTAATGTCGCTATGAACACTTGGCTGCCACAAG
CCAGTTATACCTATGGTAACTTTTCTAACACGTCTAGCCTC
AAATTCTGTAAATATAAAGGATTGATAGGCCATGCTTTCAT
GATTTGTATTCATACTGAAAATCAAAATCAAATGAGCTTTT
ACCCTTTTGTTCCACATGAGATTTCTGTTCTC

>T-CGA/E-AGG_074_1_C.chamaoensis_F

GGCATAATTTAGAGACCATGATATGCCTTGAAGAGCAA
TCTTATTGGCCGACATAACATGCCAATACTATTGGCCCGTT
ATTCGGCAAGGAGCAAACTCGACAACATTATGCATTACA
TTTCAATTACAATATGGTACAACTCATGATACATAGATAC
AAAGGGTACATTTACAAATAAGATGTC

>T-CGA/E-AGG_075_2_C.chamaoensis_F

GGGATGAGTTGTGTAGGGTCCCAACCACGTAAGCAC
TACGGATATAAGCGCTGTAGAGTCATAGTCTATCGCGAA
TTTTCCGAGGCAACGGAAATCACGTGGTAAATTTGAAGG
CTATATACCCATCCTATCGCGACGCAAGTGGTTTCATCGG
TCTCCCAGTACCTGCGCGCCCATCCGATC

>T-CGA/E-AGG_076_1_C.chamaoensis_F

GGTACAAAATCAAGTGCCTTCGAATAGACAATAGT
GGGGAGTACACATCCAATGAATTTGAAGACTTTCTCTG
CCACCATGGGATCAAACATTAGACTACCATCGTAGGTGC
CCCGTAACAGAATGGAGTC

>T-CGA/E-AGT_071_1_C.clivicola_M

GTCGGAAGGATGGCCTTGGTTCCATAAACCAACTGAA
ATGGTGTTTGTCCCATTGATCTTTTATATGTCGTTCTGAAGG
TCCATAATACTGATAGTAGTTTCTCGGCCCAATCTGACCTT
CCTACATCACAAATCTTCTGCAAGCAATTCTC

>T-CGA/E-AGT_072_1_C.chamaoensis_M

GTCGGTGCCTGATGTTCGGCGATCAAGGTACAAGATTT
ATGCTATGCTCAACATTCAGCATCTATATGGCCTTAGAATTT
GTGAAAAGAATAATATTTTTTTGTATTGGTTGTTCACGGCTC

>T-CGC/E-AAC_078_1_C.siamensis_M

ACAATCAAAAATACTAAAACCCACTAGTTCTGGGCT
AAAAGGTATCTTCAGCCTGGAAAAATTTTGGAAGTTGC
AATTCTAGATGGCCCAGAATGATTTTTAGGCCCATGGAA
TGGGCACTTAGGAGCTCCAAAAACCCTAATGCCCCCAAA
TAAACCCATTAGCCATGACAATCAAAGGGCAATAAAA
ATCTAATAAAAGGC

>T-CGC/E-AAC_079_2

ACCAATAAGCCATTGCCCAAGCTGGTCATTCATTGTGG
AGCCACAAACACTTACTTATCACCTTGCCATGCCTATATG
CAACCGTAACCTATGTTGGATATTGTGACAACAATATTTT
AGC

>T-CGC/E-AAC_080_1_C.chamaoensis_F

ACAACTCGTATTGTTATAGCTTGGAAAATGCACCTG
TATGGTGCATGAGAGAAATATGTATGCTTGTAACACGTT
GCAATATAGCCGACACGTCACGAGCAATTAGTACTAG
TTTTGC

>T-CGC/E-AAG_094_1_C.edentata_F

AGGGGGAAGAAGATAAAATATATAAACTCAAAAAGG
AACTTTATGATTTGAAGCAAGACCTAGGGCCTACCATTC
TTGCATTACTCAGGACTCATCGTCAGGGAGCTCCATCTA
GATATCGTCTAATAAAGGCTTATCTAACCTCACCAGGGC
ACAAATTCTGGC

>T-CGC/E-AAG_094_2_C.edentata_F

AGGCTATGCAAAAGGTGATAGCGATGGCAGATGAGG
ATGAATAGGAGAAAGTGATAGTAGAAGGAGACTCCTTG
AATGTGATTAGGAGTATAAAGGGTGAGTTGGAGGTGGATT
GGGGCAACTAGGATATGGTTAGGGATGCTTAGGATACATA
TAGGGATGTTCAAATCCAGTATACATATAGGGAAGGAAA
TGCTACAGCAGATTGGCTTATAAATTGGGCTGC

>T-CGC/E-ACC_086_1_C.chamaoensis_F

CCCATTGAATATCAATATGTAGTGATATAGTCACCCA
TTTTGATGAACCTACTACCTGTGCATCCATGATATATACTA
GGTTCACCTAGCATATAATATTTGCAGTCGTCCACATCCAA
GCTGTAAATGAAAAGTCACTAATTCACAAGTTACATTTGTA
CGGTGTATTCTCTGCCCTGCTTGGGTAATAATGGC

>T-CGC/E-ACC_090_2_C.clivicola_M

CCCCTCAACCTTTTAGAACTCCTAATAAGAACTACCCA
TAAGTTGAGGTATGTGATATAAATGTTAGAGAAATTTAGTG
AAGAAACTATGGTATTTCATCCCAAAGAATACTAGC

>T-CGC/E-AGC_114_1_C.clivicola_M

GCCCTACCACTGTGCCTATAGTTTGGACCTCAGTAATG
TGCCAGAACCCAACGCATTCACACCCAAATAAGCCAGTGC
AGCAGAGATCTATAAGTAGAAAAGGCTCAATCAACATCCA
AAATGTAGCCAACCTTGTTGTATTGCTATAAGAACACAGTA
GGTAATGC

>T-CGG/E-AAC_148_1_C.clivicola_F

ACAGTCTGATCTCTCCCTACATGCCTCCAAAGTTGGG
TGGCACTCCACTGTGTAAGGGAATGGCTCTAGTTCCTCA
TGAGAGCCAGTGAAAATGAATCCATTTTCTATTTGATTT
CCTTATATAGGCGAGGGCCCCAATGGGCACTTCTTCCA
TTTTTGGCCTCACGACAGCC

>T-CGG/E-AAC_150_1_C.clivicola_F

ACATAATATATGATAGTATGGCCTAAAAGTGGAGT
AGAGGAAGTTCATGCAAGGAAAAATCTAGCATCATATAA
TATGAATGGGTTTACCCAAGGTTAGGGAGCAAGCAAGAG
TTGGTAATGTTGAGAAGATAAGGTCCCAACC

>T-CGG/E-AAC_150_2_C.clivicola_F

GGCATATATGCATCAGTCTGGTCAGGAGTAGTCTAGG
CCAGGAACCTGTTCTCAGTCTTTCATTCATCTATAGGGA
TTGATACAACAGGCAAGAATATGTGAGGCAGTGGCAAGA
CTCATGTTCAATATTTACAGTCGTGAATCAATCTCCTGATT
CATCTGTAGGT

>T-CGG/E-AAG_161_2_C.clivicola_M

GAGGTCAGGTAATTAGTAGTCTGGTTTAGAAAGCTAC
TTCCTTGATCCGACATTATCTCTAATGGGGATCCAAAACAA
GCAATTATCACCTCATAGAGGAATTTGGTTGTACTTCTCGG
ATCACATTGTCTCATTGGCTTGGCCTTTATCCATTTTGTCA
TGAAATTCGTGGCC

>T-CGG/E-AGG_173_2_C.chamaoensis_F

GGAAAAAAGGAAAACCTAACGGTAAGATTTCTGGAATC
TAGCACGGGCACCGCGACCGCCGAACTTCTTGGGCTCG
CAGCGCCTGGGATCAGCCACGAGTAAAGTTCGATCGTAT
CTGACCAAGATGTCTTTGATCTCCTTCTTGGACTGCTCG
TCGACGTATTTCTGGTAGAAAGCCACCAGAGCCTTAGC
GATGCTCTGCCGTATGGCGTAAATCTGTGAGGTGTGGCC
GCCCCC

>T-CGG/E-AGG_176_1_C.siamensis_M

GGGGTTGAAGAAACATTCCGAAGACTCAAAGAAGC
AGTGGTTACTACACTAGTATTAGCCCTACTTGATTTTTCT
GAACCTTTTATTTTAGAATGCGATACATCTAGACTCGAAA
TGGGAATGGTATTACTACAGAAAGAGCATCCAATAGTTT
ACGAATGTTGTTTTCCTACGATTTCCCC

>T-CGT/E-AAC_182_1_C.edentata_M

ACATGGACCTTGTAGTATGTAGGCTATTTATACTCTA
GTGGTGAACTATCATAGGATTTCCTTAGTTGGATGGGA
TTGTAGAGTTTCGCTACATAAGATGAAGGCCATGCC
TAGAGTTGACAGTAAGGGTCAACTGTGTAGATTAGC
TAACAC

>T-CGT/E-ACA_197_1_C.chamaoensis_F

CACAAAGGGAAGGCATTGACTACACTAAAATCTTTA
CTACTATAACCAAGTGGCCTACAATCAGAACTCTTTTAG
CCTTGGTTACTCAACAGGATTAGAAAATCTACTAACTCGAC

>T-CGT/E-ACA_197_2_C.chamaoensis_F

CATGCCTTATTGTGATCATCACTCTCATCAACCCTA
GCCCCTCTTGATAAACAACCTTTTTCACATCACTCATT
GGCATTCTTGATCATATCTATGCCACACTCTATACTAAGGAC

>T-CGT/E-AGC_202_1_C.edentata_F

GCCCTACCACTGTACCTATAGTTTGGACCTCAGTAA
TGTGCCAGAACCCAACGCATTCACACCCAAATAAGCCA
GTGCAGCATAGATCTATAAGTAGAAAAGGCTCAATCA
ACATCCAAAATGTAGCCAACCTTGTTGTATTGCTATAA
GAACCCAGTAGGTAATAC

>T-CGT/E-AGC_202_2_C.edentata_F

GCTTGCAATCTGGACTCAAAATGATCCTCTCTGTG
CAGCCACATGAAATGTCAGACTTGGATGACTTTACGAT
AAATCTGGTCTCTTCTACTCTTCTGGCAGACTCGGCTA
CTAGTACAGACAGCAGGACAAAATTGGGATCACCAAA
AAACCTACTCTGATACCAC

>T-CGT/E-AGG_228_1_C.edentata_M

GGACAGGAGCTTGCGAGATTATCGGCGCATTTACA
GAGGCTTGGAGGTTTTTCACAGTTTCTCTTGGGGTTG
CAGTTCCTGTTTTGGTTTTACATTATTTTGGCTTTAC