Recurrent neo-sex chromosome evolution in kiwifruit

Recurrent neo-sex chromosome evolution in kiwifruit

Takashi Akagi[1,2,*,§], Erika Varkonyi-Gasic[3,*], Kenta Shirasawa[4], Andrew Catanach[3], Isabelle M. Henry[5], Daniel Mertten[3], Paul Datson[3], Kanae Masuda[1], Naoko Fujita[1], Eriko Kuwada[1], Koichiro Ushijima[1], Kenji Beppu[6], Andrew C. Allan[3,7], Deborah Charlesworth[8], Ikuo Kataoka[6]

1. Graduate School of Environmental and Life Science, Okayama University, Okayama, Japan
2. Japan Science and Technology Agency (JST), PRESTO, Kawaguchi-shi, Saitama 332-0012, Japan
3. The New Zealand Institute for Plant and Food Research Limited (PFR), Private Bag 92169, Auckland 1142, New Zealand
4. Kazusa DNA Research Institute, Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
5. Department of Plant Biology and Genome Center, University of California Davis, Davis, CA, USA
6. Faculty of Agriculture, Kagawa University, Miki, Kagawa 761-0795, Japan
7. School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand
8. Institute of Ecology and Evolution, University of Edinburgh, Charlotte Auerbach Road, Edinburgh EH9 2FL, United Kingdom

*These authors contributed equally

§ Corresponding author

Email: takashia@okayama-u.ac.jp, TEL: +81-86-251-8337

Abstract

Sex chromosome evolution is thought to be tightly associated with the acquisition and maintenance of sexual dimorphisms. Plant sex chromosomes have evolved independently in many lineages, and can provide a powerful comparative framework to study this. We assembled and annotated genome sequences of three kiwifruit species (genus *Actinidia*) and uncovered recurrent sex chromosome turnovers in multiple lineages. Specifically, we observed structural evolution of the neo-Y chromosomes, which was driven via rapid bursts of transposable element insertions. Surprisingly, sexual dimorphisms were conserved in the different species studied, despite the fact that the partially sex-linked genes differ between them. Using gene-editing in kiwifruit, we demonstrated that one of the two Y chromosome-encoded sex determining genes, *Shy Girl*, shows pleiotropic effects that can explain the conserved sexual dimorphisms. These plant sex chromosomes therefore maintain sexual dimorphisms through the conservation of a single gene, without a process involving interactions between separate sex-determining genes and genes for sexually dimorphic traits.

MAIN TEXT

Chromosomal sex determination is common in many species (reviewed by Bachtrog et al.[1]). Dioecious plants, unlike many gonochoristic animals, evolved independently in different lineages from functional hermaphrodite ancestors (reviewed by Ming et al.[2], Henry et al.[3]). Two *de novo* evolutionary paths to genetic sex determination have been characterized. One path results in two distinct loci, one male-determining and one female-suppressing factor, located in separate, but closely linked sex chromosome regions[4, 5, 6, 7, 8]. The other path results in a system with a single sex-determining gene, as observed in persimmon[9, 10]. Single gene control can also emerge through “turnovers”, involving either movement of a pre-existing sex-determining gene, or its replacement by a new sex-determining gene, which takes over the control of male versus female development, and creates a sex-determining locus in a different genomic location. Examples of both types in animal species are reviewed in Vicoso[11] and Pan et al.[12]. Instances of sex chromosome turnover have also been detected in plants,

including in the family Salicaceae[13, 14]), and in the genus *Fragaria* (strawberries), in which a 13-kb transposable "cassette" has moved a currently unidentified female-determining gene(s) to different genomic locations[15].

The two-locus sex determining system creates selection for close linkage between the two genes[16], potentially leading to recombination suppression across large chromosome regions, which subsequently undergo the degeneration and rearrangements characteristic of sex chromosomes in many species[17, 18]. In a male heterogametic (or XY) system, this creates a male-specific Y-linked region (MSY). Extended MSY regions can also evolve if mutations that benefit only one sex (sexually antagonistic, or SA) establish polymorphisms in the genome region closely linked to the sex-determining region. SA polymorphisms can also favor turnovers through the establishment of a new sex-determining gene nearby, again favoring closer linkage[19]. Here, we investigated recurrent sex chromosome turnovers and the evolution of sexual dimorphism in kiwifruit species (the genus *Actinidia*).

In kiwifruit, sex (gender) is determined by two Y-linked genes, *Shy Girl* and *Friendly Boy*[5]. The neofunctionalized *Shy Girl* suppresses gynoecium development, and was established before the largely dioecious genus *Actinidia* radiated, about 20 million years ago[4, 20]. Across a wide variety of *Actinidia* species, males produce more flowers than females, after a shorter juvenile phase (Extended Data Figure S1, Testolin et al.[21]). There is no obvious lineage-specific sexual dimorphism beside those present across the genus. Conservation of sexual dimorphisms across species might suggest establishment before the divergence of *Actinidia* species, perhaps after a non-recombining MSY evolved to preserve advantageous combinations of sex determining factors and SA genes.

However, our chromosome-scale whole-genome assemblies based on PacBio HiFi reads and optical mapping in male accessions of three *Actinidia* species, *A. arguta*, *A. polygama*, and *A. rufa* (Extended Figure S2, Supplementary Data S1-2), and comparison with the male *A. chinensis* genome (Tahir et al.[22]), suggest a different evolutionary history. None of the species has an extensive MSY

(approx. 1.1-2.7Mb), and the genomic locations of the MSYs differ. The two lineages seem to have undergone independent sex chromosome turnovers (Figure 1, Extended Data Figure S2). The species' phylogeny (Figure 1a, see also Methods), and, as described in detail later, silent site divergence (*dS*) values between X- and Y-allelic genes in the pseudoautosomal regions, or "PARs", suggest that Chr 25 was the most likely ancestral state for both turnover events. Remarkably, both turnovers translocated the two Y-linked sex-determining factors as a unit, to the end of Chr 4 in *A. rufa*, and the middle of Chr 3 in *A. arguta* (Figure 1a). These locations are supported by genetic mapping (Extended Data Figure S3). Analyses of genomic read mapping (Figure 1b and d) and synteny between the X and Y chromosomes (Figure 1c and e) in *A. rufa* and *A. arguta* showed that neither these neo-sex determining regions, nor the original Chr 25 region (Extended Data Figure S4), have any X-linked counterpart sequences, indicating that each translocation created a hemizygous neo-Y-linked region. However, the MSY regions of the four *Actinidia* species (including the new MSYs) exhibit no clear synteny with each other, and appear highly rearranged, although short fragments (< 50kb) are more frequently conserved between Chr. 25 of *A. chinensis* and *A. polygama* than in the other inter-chromosomal comparisons (Extended Data Figure 5). Importantly, these MSYs share only 3 genes (Figure 1g, Supplementary Data S3): the two known *A. chinensis* sex-determining genes, *Friendly Boy* and *Shy Girl*, as well as *FLOWERING LOCUS-T* (the FT-like gene described in Akagi et al. 2018). The relative physical locations of these three genes are not conserved, even between the two species with the putatively ancestral MSYs on Chr25, or between these and the new MSYs (Extended Data Figure S6, Figure 4). Rearrangements, including inversions and transposable element (TE) insertions, which are discussed further below, may explain the evolution of the MSYs, which are all in regions that are often rich in repetitive sequences that can lead to rearrangements[23,24]: the *A. rufa* new MSY has remained telomeric, and the *A. arguta* new MSY is near the Chr3 centromere.

To further understand the timing and order of events, we estimated silent site divergence (*dS*) values between sequences. The time back to the common ancestor of the genus *Actinidia* corresponds to an average *dS* of 0.102 ± 0.006, based on the same 3,002 genes as those used in

the phylogeny in Figure 1a. Because the new MSYs on *A. rufa* Chr 4 and *A. arguta* Chr 3 arose by translocation events that created hemizygous Y-linked copies, we cannot estimate Y-X sequence divergence in these regions. We therefore used nearby sequences from our X and Y assemblies of each species (in 500 kb windows from the Y-linked *Shy Girl* gene); these "Y-X *dS*" values are for genes in the PARs, created when the male-determining genes translocated into the region, which would have hindered pairing in the flanking regions[25]. Supporting the view that the ancestral sex-linked region was on the *A. chinensis* Chr 25, divergence values are highest in this species (Fig. 1f). Although the values differ considerably between windows, they are similar to the values above that reflect the origin of the genus *Actinidia*. Elevated *dS* values in the PARs of the *A. chinensis* probably reflect linkage disequilibrium with the MSY, but possibly also the presence of two chromosomal inversions adjacent to the MSY (Extended Data Figure S4a-b), which would suppress recombination and Y-linked sequences to diverge from their X-linked alleles. The estimated Y-X *dS* in windows surrounding both of the new MSYs consistently indicate lower divergence times, supporting the hypothesis that these translocations occurred more recently.

The sizes of the new MSY regions in *A. rufa* and *A. arguta* range from *ca.* 1.5 to 2.7 Mb (Figure 1b-e), slightly longer than their probable progenitor MSYs in *A. chinensis* and *A. polygama* (ca. 1.1-1.5 Mb) (Extended Data Figure S4). This can be explained because hemizygosity of the Y-linked regions created by the translocations prevents recombination in the inserted region, allowing repeats to accumulate, and structural changes to occur without causing deleterious effects such as chromosome breaks during crossing over[23]. Both the *A. rufa* and *A. arguta* new MSYs are highly enriched with diverse TEs (Figure 2a-b, Supplementary Data S4). Specifically, the *A. rufa* new MSY has accumulated LTR-unknown, non-LTR *CACTA*, *hAT*, and *Helitron*-like TEs, while the new MSY in *A. arguta* is occupied mainly by *Gypsy*-like TEs (Figure 2a-b, $p < 1e^{-10}$ with Fisher's exact test against the genome-wide TE distribution). Although *Gypsy*-like TEs are abundant in all four *Actinidia* MSYs (Figure 2b), the sequence types (defined as clusters with >90% identity) within the MSYs are species-specific (Figure 2d), suggesting that they accumulated independently within these regions, and could

126 have contributed to the increased sizes of the two new MSYs. The corresponding neo-X regions
127 exhibit no recent TE accumulation, and nor do the Chr 3 (pericentromeric) or Chr 4 (peritelomeric)
128 regions, apart from the MSYs themselves (Extended Data Figure S7). A complete lack of recombination
129 in these sex-linked regions is not required, as TEs accumulate even in recombining regions, though
130 low recombination rates correlate with greater accumulation, other things being equal.

131 The conserved sexual dimorphisms in kiwifruit, combined with widely variable MSY locations, are
132 inconsistent with the presence of SA mutations at loci closely linked to an ancient male-determining
133 locus. They could instead reflect pleiotropic effects of the genes shared by the MSYs[26]. We tested
134 this possibility by disrupting the Y-encoded gynoecium suppressing gene, *Shy Girl*, by CRISPR-Cas9
135 gene-editing. This not only converted flowers from male to hermaphrodite, but also resulted in loss of
136 most sexual dimorphisms (Figure 3). Male *A. chinensis* (cv. Bruce) plants produce clustered flowers
137 (Figure 3a) and, compared to females, more flowers per inflorescence, more floral nodes per shoot,
138 and more flowering shoots (i.e. a shorter juvenile phase). The gene-edited (*sygl*) lines exhibited
139 substantially fewer flowers per inflorescence (Figure 3f, p = 0.0031), floral nodes (Figure 3g), and
140 adult shoots than the wild-type (WT) male cultivar (Figure 3h, p = 0.0082), thus becoming similar to
141 female cultivars (Extended Data Figure S8). Transcriptome analyses in developing buds showed that
142 genes involved in response to stimulus/stress are highly enriched among genes that are differentially
143 expressed between wild-type males and gene-edited *sygl* lines (Extended Data Figure S9,
144 Supplementary Data S5). Interestingly, these genes include orthologs of an autosomal sex
145 determining gene in persimmon, *MeGI*[9], which pleiotropically affects flower numbers per inflorescence
146 and the length of the juvenile phase[26], resembling the sexual dimorphisms in kiwifruit. This gene family
147 also includes the *Vrs1* gene [27] involved in the transition from two-rowed to six-rowed spikelet barley,
148 and several genes involved in axillary bud initiation in *Arabidopsis thaliana*, *HB21/40/53*[28] (Extended
149 Data Fig. S10). The demonstration of pleiotropic effects shows that the evolution of sexual dimorphism
150 need not involve SA genes separate from the primary sex-determining factor(s), or SA polymorphism
151 that selects for suppressed recombination in the sex-determining chromosome region involved.

That sexual dimorphisms can be explained by the pleiotropic effects of a sex determining factor was originally suggested by Darwin[29], who used the term “compensation” (today’s terminology is “trade-offs”, to stress the idea of conflict between the sexes, and that an increase in a trait in one sex can be at the expense of the same trait in the other sex). Establishment of higher functioning males by a femaleness-suppressing mutation in a cosexual ancestor involves a trade-off, by definition; a trade-off is therefore guaranteed in the two-gene model[16]. Our evidence in kiwifruit supports this, as *Shy Girl* simultaneously establishes maleness and male-specific sexual dimorphisms that are likely to have sexually antagonistic effects, as males of many dioecious plants show larger flower numbers per inflorescence, and a reduced juvenile phase, compared with females[30]. We cannot currently exclude contributions from the other two conserved genes in the MSYs, *Friendly Boy* and *FT*-like, although they have no known effects on kiwifruit sexual dimorphisms[4, 5, 26].

Overall, our results suggest that, as proposed in Figure 4, lineage-specific translocations events resulted in the retention of only three genes within the new MSYs, including the two sex determining genes, preserving their pleiotropic effects creating sexual dimorphisms, while the translocated regions accumulated independent inversions, and also TEs, expanding the new MSYs. If the translocations were favored because they created close linkage between the sex-determining genes and SA polymorphisms in autosomal regions, and resulted in suppressed recombination, the SA gene(s) involved should be located within the new MSYs. The presence of only three common genes within the MSYs, argues against this scenario, although we cannot definitively exclude the possibility that the new MSYs in the *A. arguta* Chr 3 or *A. rufa* Chr 4 might include SA genes involving further, currently uncharacterized, lineage-specific sexual dimorphisms in these species. It is also unlikely that pseudo-autosomal SA polymorphisms are present in the nearby neo-sex chromosome regions (creating selection to suppress recombination, without a response to the selection having yet occurred), because the *A. rufa* and *A. arguta* neo-sex chromosomes did not exhibit elevated Y-X divergence values, as would be expected if such polymorphisms were present[31](Fig. 1f). The slightly expanded non-recombining new MSY regions might be a direct effect of the translocations into the *A.*

178 *rufa* and *A. arguta* regions of cassettes carrying the *SyGI/FrBy* and *YFT* genes, hindering pairing in
179 meiosis. Although we do not currently know whether the translocations in *Actinidia* species were
180 mediated by TEs, as in *Fragaria* (see above), active TEs may have caused the inversions near the
181 new MSYs. Taken together, our observations suggest that maintenance of sexual dimorphism and
182 recombination arrest are independent in kiwifruit neo-sex chromosome evolution. In other words, the
183 pleiotropy of one sex-determining gene, controlling sexual dimorphisms in this system, probably
184 allows rapid Y chromosome turnovers to any genome location, not just regions where such movement
185 would establish linkage to a polymorphic SA gene. This can explain the absence of recombination
186 suppression near these species' sex-determining genes.
187
188
189

METHODS

Plant materials

The three *Actinidia* species used for whole genome assembly, *A. rufa* (sel. Nakamura-B, male), *A. polygama* (sel. Hokkaido, male), and *A. arguta* (var. *hypoleuca*) (sel. Ubaishi) were originally sampled at Shimanto, Kochi, Japan (N33.00, E132.56), Setana, Hokkaido, Japan (N42.13, E139.49), and Futtsu, Chiba, Japan (N35.18, E139.51), respectively. For observation of sexual polymorphisms in *Actinidia* accessions, 3 *A. deliciosa*, 2 *A. chinensis*, 4 *A. rufa*, 2 *A. polygama*, and 5 *A. arguta*, germplasm collections maintained in Kagawa University, Japan (N34.28, E134.13) were assessed in 2019 and 2021. In native conditions, inflorescences were assessed for 4 *A. rufa*, and 6 *A. polygama* at approx. 100-800m altitude in the Konan-Koga area of Shiga prefecture (N35.00, E136.05), and 11 *A. polygama*, at approx. 800-1200m in the Iya area of Tokushima prefecture (N33.53, E133.49).
Four families were used for genetic mapping of sex-associated variants: *A. chinensis* x *A. rufa* (CR, 109 females, 94 males), *A. chinensis* X *A. arguta* var. *melanandra* (CM, 43 females, 32 males), *A. chinensis* X *A. arguta* var. *purpurea* (CP, 54 females, 59 males), and a family of 16 *A. chinensis* X *A. valvata* (CV, 7 females and 8 males).

Genome assembly and gene/repeat annotations

Genomic DNA for sequencing was extracted from young kiwifruit leaves using the Qiagen Genome-tip G20 or G100, following the standard protocol. The genome DNA was sheared into ~20 kb DNA fragments with a g-tube machine (Covaris, Woburn, MA, USA) and a HiFi SMRTbell library was constructed with a SMRTbell Express Template Prep Kit 2.0 (PacBio, Menlo Park, CA, USA). The library DNA was fractionated using the BluePippin (Sage Science, Beverly, MA, USA) to eliminate fragments <20 kb and sequenced with a single SMRT cell 8 M on the Sequel II system (PacBio). The sequence reads were converted into HiFi reads with the CCS pipelines (PacBio, https://ccs.how) and assembled with Hifiasm[31]. Contig sequences were aligned to the chromosome sequences of *A. chinensis* cv. RED5[32] as a reference, and scaffolded with RaGoo[33] to build pseudomolecule

sequences for *A. rufa*, *A. polygama*, and *A. arguta*.

Potential protein-coding genes were predicted with the Maker pipeline[34], using Illumina short reads (SR50) mRNA-seq data from various *Actinidia* species, as well as peptide sequences predicted from the genomes of *A. chinensis* (Hong Yang v3[35]; RED5[32]), *A. eriantha* (White[36]) and *A. rufa* (*A. rufa* Fuchu_1.0, NCBI accession number PRJDB8483). Short gene sequences (<300 bp), genes in repeat sequences, and genes with annotation edit distance (AED) >0.5 were removed to facilitate the selection of high-confidence genes.

Repetitive sequences in the assemblies were identified with RepeatMasker[37] using repeat sequences registered in Repbase[38] and de novo repeat libraries built with RepeatModeler Kirkpatrick[39]. Repeat elements were classified into nine types with RepeatMasker: short interspersed nuclear elements (SINEs), long interspersed nuclear elements (LINEs), long terminal repeat (LTR) elements, DNA elements, small RNA, satellites, simple repeats, low complexity repeats, and unclassified.

Genetic mapping of the sex determining regions

Three mapping families, CM, CP, and CV were genotyped using capture-sequencing (Rapid Genomics) using ~10,000 baits. The CR family was genotyped using flex-sequencing (Rapid Genomics) using ~3,000 amplicons. Raw data were delivered as paired fastq files. Sequencing data qualities were checked using FastQC and reads were trimmed along with removal of adaptors and index sequences using the bbduk.sh script of BBMap/38.33, with the following parameters: trimpolyg = 50, trimpolyc = 50, k = 13, trimq = 15, minlength = 50. Trimmed reads were aligned to the respective genome reference sequences using bwa-mem v.0.7.17. Output files were sorted, readgroups were added, and bam file indexes were generated, using samtools v.1.12. Variants were called using freebayes 1.1.0 in parallel mode with the following parameters: --report-genotype-likelihood-max --min-mapping-quality 10, --genotype-qualities, --use-mapping-quality, --no-mnps, --no-complex. The resulting VCF files were filtered using bcftools to exclude variants with Phred-scaled quality scores below 51.

To identify variant sites linked to sex in our families VCF files were converted into genotypes using vcfR v.1.8.0 in R v3.5.0 (usually 0/0, 0/1 or 1/1). Variants were selected if all individuals of one sex were heterozygous, while all individuals of the other sex were homozygous. Fisher's exact test (R v3.5.0) was used to test associations between sex and genotypes. Whole genome plots of –log10P values for sites in different genomic locations were created using Circlize[40] in R v3.5.0.

Synteny analyses

Chromosome-scale synteny analysis was performed with MCScanX[41] and collinearity was visualized using SynVisio (https://synvisio.github.io/#/). All-versus-all BLASTP analyses were performed between the protein sequences from four *Actinidia* species, with an e-value cut-off of $<1e^{-50}$. Syntenic blocks were constructed using MCScanX, with BLASTP and gff files, after preprocessing to be suitable for MCScanX. Large scale synteny was also evaluated with D-genies[42] for dot-plot visualization (http://dgenies.toulouse.inra.fr/). Two whole genome sequences (fasta) files were aligned with Minimap2 with the D-genies default conditions.

Local synteny, especially in the sex determining and adjacent regions, was defined based on physical orders of homologous gene pairs, with Strudel[43]. All-versus-all BLASTP analyses were preformed among the protein sequences in the genomic regions of interest with e-value $<1e^{-100}$. The resultant files were processed to be fit Strudel format for visualization.

Characterization of the genomic context around the sex determining region and Y-chromosome

For the construction of Illumina genomic libraries, we used approximately 0.5 µg of genomic DNA of *A. chinensis* cvs. Soyu (male) and Rainbow Red (female), *A. rufa* sels. Nakamura-B (male) and Fuchu (female), *A. polygama* sels. Hokkaido (male) and Wakayama (female), and *A. arguta* (var. *hypoleuca*) sels. Ubaishi (male) and Nagao (female). The libraries were constructed using the KAPA HyperPlus Library Preparation Kit (KAPA Biosystems), following Akagi et al.[5], and sequenced using Illumina's

HiSeq4000 (100-bp or 150-bp paired-end reads, x20-30 for the averaged coverage). All Illumina sequencing was conducted at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, and the raw reads were processed using custom Python scripts developed in the Comai laboratory and available online (http://comailab.genomecenter.ucdavis.edu/index.php/), as previously described[5]. The preprocessed reads were aligned to the whole genomic scaffolds of each species, with Burrows-Wheeler Aligner (BWA)-mem[44], allowing up to 12% mismatches. The mapped reads were visualized with Integrative Genomics Viewer (IGV)[45].

de novo transposable element annotations and enrichment test

Transposable elements in the genomes of the four *Actinidia* species, *A. chinensis*, *A. rufa*. *A.polygama*, and *A. arguta*, were analysed with the Extensive *de-novo* TE Annotator (EDTA) pipeline, that integrates structure- and homology-based approaches for TE identification; including LTRharvest (v1.5.10)[46], LTR_FINDER_parallel (v1.0)[47], LTR_retriever (v2.6)[48], Generic Repeat Finder (v1.0)[49], TIR-Learner (v1.23)[50], MITE-Hunter (v1.0)[51], and HelitronScanner (v1.0)[52] with extra basic and advanced filters. The genomic distributions of the detected TEs were visualized with DNAfeature (https://github.com/Edinburgh-Genome-Foundry/DnaFeaturesViewer). Clustering of TEs within the *Gypsy* family amongst the *Actinidia* species was conducted with cd-hit-est in the Cd-hit (Cluster Database at High Identity with Tolerance[53]), with the -c 0.9 (>90% sequence identity) option.

Gene-editing in kiwifruit

Gene editing, plant transformation and genotyping of *SyGI*-edited lines were performed as previously described[20]. Briefly, 4 *SyGI*-specific guides (and two guides targeting *CEN4*) were included in a polycistronic tRNA - sgRNA cassette placed downstream of the Arabidopsis U6 - 26 promoter were cloned into the binary vector pDE-KRS, introduced into *Agrobacterium tumefaciens* strain EHA105 by electroporation and transformed into *A. chinensis* Planch. var. *chinensis* ‘Bruce’ according to Wang et al.[54]. To identify *SyGI* gene-edited lines, genomic DNA (DNeasy Plant Mini Kit, Qiagen) was amplified

294 (iProof High - Fidelity DNA Polymerase, Bio - Rad) using *SyGI* gene-specific oligonucleotide primers,
295 cloned and sequenced. Two lines with small deletions (-10 nt) and one with a small insertion (+1 nt)
296 in the first exon of *SyGI*, 85 nt downstream from the translation start site (and no mutations in the
297 *CEN4* locus) were identified.
298
299 **Characterization of sexual dimorphisms**
300 *SyGI*-edited and non-edited male lines were grown at ambient conditions in the containment
301 greenhouse at Plant & Food Research, Auckland, New Zealand. After exposure to naturally
302 shortening days in autumn, the plants were manually defoliated and transferred to the cold room (7°C,
303 dark) for 8 weeks, then transferred to the glasshouse in spring. Plants were pruned to maintain a
304 vertically grown trunk and up to three horizontally grown branches of comparable lengths. The first
305 flowering (maturity) was observed after 2 years. Phenotyping for flower morphology, self-fertility,
306 number of floral shoots, and ratio of floral shoots to total shoots, number of floral nodes per shoot and
307 number of flowers per inflorescence was performed.
308
309 **Transcriptome analyses in inflorescence primordia**
310 For the transcriptome analyses, samples from a *SyGI*-edited and non-edited original male (WT) line
311 were collected in the third year. The sampling was performed 0, 4, 10 and 14 days after transferring
312 from the cold room into spring conditions. The minimum of 10 embryonic shoots at day 0 and day 4
313 and the minimum of five inflorescences at day 10 and day 14 were sampled from individual branches
314 (biological replicates) at each time point. All sampling was performed at midday. Dissection of
315 embryonic shoots and inflorescences was performed using the Leica MZ FLIII (Leica Microsystems)
316 stereomicroscope and representative stages were photographed using the Olympus DP74 camera
317 (Olympus Optical Co Ltd). Total RNA was isolated using the Spectrum Plant Total RNA Kit (Sigma-
318 Aldrich). Extracted total RNA were processed in preparation for Illumina Sequencing, according to a
319 previous report[55]. In brief, mRNA was purified using the Dynabeads mRNA purification kit (Life

Technologies). Next, cDNA was synthesized via KAPA RNA HyperPrep kit (Roche), followed by a DNA cleanup step with AMPure XP beads (Beckman Coulter; AMPure:reaction, 0.8:1). The constructed libraries were sequenced on Illumina's HiSeq 4000 sequencer (50-bp single-end reads). All Illumina sequencing was conducted at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley. Raw sequencing reads were processed using Python scripts (https://github.com/Comai-Lab/allprep/blob/master/allprep-13.py) for preprocessing and demultiplexing of sequencing data. The mRNA-seq reads were aligned to the reference coding sequences (CDSs) of cv. RED5[32], using the default parameters of the Burrows-Wheeler aligner (BWA) (version 0.7.15)[56] (http://bio-bwa.sourceforge.net/). Read counts per CDS were generated from the aligned SAM files using R script. PCA of mRNA expression was conducted using the genes with RPKM > 1, using prcomp in R. Differential expression between the WT male and the *sygI* hermaphrodite (herm.) flowers was analyzed using DESeq2 package[57], using four replicates per sample type. The DEGs were filtered according to RPKM and FDR values (RPKM > 1, FDR < 0.01). Clustering of the DEGs was conducted using the hclust function in R, and visualized via a heatmap using the ComplexHeatmap package. Putative functions of each gene were determined with a BLASTX search of the TAIR10 database (https://www.arabidopsis.org/index.jsp). A GO enrichment analysis was performed on the genes in the three networks using agriGO (http://bioinfo.cau.edu.cn/agriGO).

Phylogenetic analysis

For the evolutionary topology of the genome-wide structure of *Actinidia* species, we used 327 genes defined by OrthoMCL as conserved single genes in rice (*Oryza sativa*, monocot), *Arabidopsis thaliana* (eurosid II), grape (*Vitis vinifera*, rosid), tomato (*Solanum lycopersicum*, euasterid), and three *Actinidia* (the order Ericale in asteroid) genomes. Their sequences were aligned with MAFFT ver. 7 with the L-INS-i model, followed by manual pruning using SeaView. The alignments were concatenated, and all sites, including missing and gap data, were used to construct phylogenic trees with the maximum likelihood method (ML) in Mega v. 5[58], using the WAG+I+G model (# of discrete gamma categorization

= 3). 1,000 bootstrap replications, calculated using the nearest neighbour interchange technique, are shown on the branches as 1/10 of the calculated values.

The orthologs of HB21/40/53-like homeobox (HD-ZIP1) genes were extracted using BLASTp (<1e-10) in Phytozome (JGI release version 12.0, https://phytozome.jgi.doe.gov/pz/portal.html). An *Arabidopsis* ortholog of the ATHB5 was used as the outgroup gene. Alignment analyses on amino acid sequences were conducted using MAFFT ver. 7 with the L-INS-i model, followed by manual revision using SeaView ver. 4. The evolutionary topology was examined using the maximum likelihood method (ML) by MEGAX[59] with WAG+G model and 1,000 replications for bootstraps. All sites, including missing and gap data, were used to estimate phylogenic trees, and the nearest neighbour interchange technique was used. Bootstraps were shown on the branches as 1/10 of the calculated values.

Accession numbers

All Illumina and three *Actinidia* genomes sequencing data have been deposited in the DDBJ database: Short Read Archives (SRA) database (BioProject ID PRJDB13958, Run ID DRR396285-DRR396297 for Illumina sequencing reads; BRYE01000001-BRYE01000030 for *A. arguta* genome; BRYF01000001 - BRYF01000030 for *A. polygama* genome; BRYG01000001-BRYG01000029 for *A. rufa* genome).

REFERENCES

1. Bachtrog, D. et al. Sex determination: why so many ways of doing it? *PLOS Biol.* https://doi.org/10.1371/journal.pbio.1001899 (2014).
2. Ming, R., Bendahmane, A. & Renner, S. S. Sex chromosomes in land plants. *Annu. Rev. Plant Biol.* **62**, 485-514 (2011).
3. Henry, I. M., Akagi, T., Tao, R. & Comai, L. One hundred ways to invent the sexes: theoretical and observed paths to dioecy in plants. *Annu. Rev. Plant Biol.* **69**, 553-575 (2018).
4. Akagi, T. et al. A Y-encoded suppressor of feminization arose via lineage-specific duplication of a cytokinin response regulator in kiwifruit. *Plant Cell*. **30**, 780-795 (2018).
5. Akagi, T. et al. Two Y-chromosome-encoded genes determine sex in kiwifruit. *Nat. Plants* **5**, 801-809 (2019).
6. Harkess, A. et al. The asparagus genome sheds light on the origin and evolution of a young Y chromosome. *Nat. Commun.* **8**, 1-10 (2017).
7. Harkess, A. et al. Sex determination by two Y-linked genes in garden asparagus. *Plant Cell* **32**, 1790-1796 (2020).
8. Massonnet, M. et al. The genetic basis of sex determination in grapes. *Nat. Commun.* **11**, 1-12 (2020).
9. Akagi, T., Henry, I. M., Tao, R. & Comai, L. A Y-chromosome–encoded small RNA acts as a sex determinant in persimmons. *Science* **346**, 646-650 (2014).
10. Yang, H. W., Akagi, T., Kawakatsu, T. & Tao, R. Gene networks orchestrated by *MeGI*: a single‐factor mechanism underlying sex determination in persimmon. *Plant J.* **98**, 97-111 (2019).
11. Vicoso, B. Molecular and evolutionary dynamics of animal sex-chromosome turnover. *Nat. Ecol. Evol.* **3**, 1632-1641 (2019).
12. Pan, Q. et al. Evolution of master sex determiners: TGF-β signalling pathways at regulatory crossroads. *Philos. Trans. R. Soc. B* **376**, 20200091 (2021).
13. Müller, N. A. et al. A single gene underlies the dynamic evolution of poplar sex determination. *Nat. Plants* **6**, 630-637 (2020).
14. Hallingbäck, H. R., Pucholt, P., Ingvarsson, P. K., Rönnberg-Wästljung, A. C. & Berlin, S. Genome-wide association mapping uncovers sex-associated copy number variation markers and female hemizygous regions on the W chromosome in *Salix viminalis*. *BMC Genom.* **22**, 1-14 (2021).
15. Tennessen, J. A. et al. Repeated translocation of a gene cassette drives sex-chromosome turnover in strawberries. *PLOS Biol.* **16**, e2006062 (2018).
16. Charlesworth, B. & Charlesworth, D. A model for the evolution of dioecy and gynodioecy. *Am. Nat.* **112**, 975-997 (1978).
17. Charlesworth, D., Charlesworth, B. & Marais, G. Steps in the evolution of heteromorphic sex chromosomes. *Heredity* **95**, 118-128 (2005).
18. Bergero, R. & Charlesworth, D. The evolution of restricted recombination in sex chromosomes.

Trends Ecol. Evol. **24**, 94-102 (2009).
19. Van Doorn, G. & Kirkpatrick, M. Turnover of sex chromosomes induced by sexual conflict. *Nature* **449**, 909-912 (2007).
20. Varkonyi-Gasic, E. et al. *Shy Girl*, a kiwifruit suppressor of feminization, restricts gynoecium development via regulation of cytokinin metabolism and signalling. *New Phytol.* **230**, 1461-1475 (2021).
21. Testolin, R., Cipriani, G. & Costa, G. Sex segregation ratio and gender expression in the genus *Actinidia*. *Sex. Plant Reprod.* **8**, 129-132 (1995).
22. Tahir, J. et al. First chromosome-scale assembly and deep floral-bud transcriptome of a male kiwifruit. *Front. Genet.* https://www.frontiersin.org/articles/10.3389/fgene.2022.852161 (2022).
23. Charlesworth, B., Sniegowski, P. & Wright, S. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* **371**, 215-220 (1994).
24. Kapun, M. & Flatt, T. The adaptive significance of chromosomal inversion polymorphisms in Drosophila melanogaster. *Mol. Ecol.* **28**, 1263-1282 (2019).
25. Charlesworth, D. Young sex chromosomes in plants and animals. *New Phytol.* **224**, 1095-1107 (2019).
26. Akagi, T. & Charlesworth, D. Pleiotropic effects of sex-determining genes in the evolution of dioecy in two plant species. *Proc. R. Soc. B.* http://doi.org/10.1098/rspb.2019.1805 (2019).
27. Komatsuda, T. et al. Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 1424-1429 (2007).
28. González-Grandío et al. Abscisic acid signaling is controlled by a *BRANCHED1/HD-ZIP I* cascade in *Arabidopsis* axillary buds. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E245-E254 (2017).
29. Darwin, C. R. *The Different Forms of Flowers on Plants of the Same Species*. London: John Murray (1877).
30. Barrett, S. C. & Hough, J. Sexual dimorphism in flowering plants. *J. Exp. Bot.* **64**, 67-82 (2013).
31. Cheng, H., Concepcion, G. T., Feng, X., Zhang, H. & Li, H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170-175 (2021).
32. Pilkington, S. M. et al. A manually annotated *Actinidia chinensis* var. *chinensis* (kiwifruit) genome highlights the challenges associated with draft genomes and gene prediction in plants. *BMC Genom.* **19**, 1-19 (2018).
33. Alonge, M. et al. RaGOO: fast and accurate reference-guided scaffolding of draft genomes. *Genome Biol.* **20**, 1-17 (2019).
34. Cantarel, B. L. et al. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* **18**, 188-196 (2008).
35. Wu, H. et al. A high-quality *Actinidia chinensis* (kiwifruit) genome. *Hort. Res.* **6**, 117 (2019).
36. Tang, W. et al. Chromosome-scale genome assembly of kiwifruit *Actinidia eriantha* with single-molecule sequencing and chromatin interaction mapping. *GigaScience* **8**, giz027 (2019).
37. Smit, A. F. A., Hubley, R. & Green, P. *RepeatMasker Open*-4.0 http://www.repeatmasker.org

(2013–2015).
38. Jurka, J. et al. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.* **110**, 462-467 (2005).
39. Smit, A. F. A. & Hubley, R. *RepeatModeler Open* version 1.0. http://www.repeatmasker.org (2008–2015).
40. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. *circlize* implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811-2812 (2014).
41. Wang, Y. et al. *MCScanX*: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **40**, e49 (2012).
42. Cabanettes, F. & Klopp, C. D-GENIES: dot plot large genomes in an interactive, efficient and simple way. *PeerJ*, https://doi.org/10.7717/peerj.4958 (2018).
43. Bayer, M. et al. Comparative visualization of genetic and physical maps with Strudel. *Bioinformatics* **27**, 1307-1308 (2011).
44. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
45. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24-26 (2011).
46. Ellinghaus, D., Kurtz, S. & Willhoeft, U. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinform.* **9**, 1-14 (2008).
47. Ou, S. & Jiang, N. LTR_FINDER_parallel: parallelization of LTR_FINDER enabling rapid identification of long terminal repeat retrotransposons. *Mob. DNA* 10, 1-3 (2019).
48. Ou, S. & Jiang, N. LTR_retriever: a highly accurate and sensitive program for identification of long terminal repeat retrotransposons. *Plant Physiol.* **176**, 1410-1422 (2018).
49. Shi, J. & Liang, C. Generic repeat finder: a high-sensitivity tool for genome-wide de novo repeat detection. *Plant Physiol.* **180**, 1803-1815 (2019).
50. Su, W., Gu, X. & Peterson, T. TIR-Learner, a new ensemble method for TIR transposable element annotation, provides evidence for abundant new transposable elements in the maize genome. *Mol. plant* **12**, 447-460 (2019).
51. Han, Y. & Wessler, S. R. MITE-Hunter: a program for discovering miniature inverted-repeat transposable elements from genomic sequences. *Nucleic Acids Res.* **38**, e199 (2010).
52. Xiong, W., He, L., Lai, J., Dooner, H. K. & Du, C. HelitronScanner uncovers a large overlooked cache of *Helitron* transposons in many plant genomes. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 10263-10268 (2014).
53. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658-1659 (2006).
54. Wang, T., Atkinson, R. & Janssen, B. The choice of Agrobacterium strain for transformation of kiwifruit. *ISHS Acta Horticulturae* **753**, 227-232 (2007).
55. Masuda, K. et al. Reinvention of hermaphroditism via activation of a *RADIALIS*-like gene in hexaploid persimmon. *Nat. Plants* **8**, 217-224 (2022).

56. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25** 1754-60 (2009)
57. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1-21 (2014).
58. Tamura, K.et al. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.*, **28**, 2731-2739 (2011).
59. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547-1549 (2018).

SUPPLEMENTARY INFORMATION

Extended Figure S1-S10

Supplementary Data S1-S5

ACKNOWLEDGEMENT

We thank Dr. Luca Comai (Dept. Plant Biol and Genome Center University of California Davis, USA) for discussion and comments for this study, and Kenichi Miyata (Board of Education of Tokushima prefecture, Japan) for sampling of native *Actinidia* germplasm collections. This work was supported by PRESTO from Japan Science and Technology Agency (JST) [JPMJPR20Q1] and Grant-in-Aid for Transformative Research Areas (A) from JSPS [22H05172 and 22H05173] to T.A.

AUTHOR CONTRIBUTION

T.A., D.C. and K.I. conceived the study. T.A., E.V., A.C., I.M.H., and D.C. designed experiments. T.A., E.V., A.C., K.S., K.M., N.F., and E.K. conducted the experiments. T.A., E.V., A.C., K.S., P.D., K.M., and N.F. analyzed the data. T.A., E.V., K.S., I.M.H., P.D., K.U., K.B., A.C.A, and I.K. contributed to plant resources and facilities. T.A., E.V., A.C., I.M.H., D.C. and I.K. drafted the manuscript. All authors approved the manuscript.

AUTHOR INFORMATION

Affiliations

1. Graduate School of Environmental and Life Science, Okayama University, Okayama, Japan

2 Japan Science and Technology Agency (JST), PRESTO, Kawaguchi-shi, Saitama 332-0012, Japan

3 The New Zealand Institute for Plant and Food Research Limited (PFR), Private Bag 92169, Auckland 1142, New Zealand

4 Kazusa DNA Research Institute, Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan

5. Department of Plant Biology and Genome Center, University of California Davis, Davis, CA, USA

514 6. Faculty of Agriculture, Kagawa University, Miki, Kagawa 761-0795, Japan

515 7. School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland 1142, New
516 Zealand

517 8. Institute of Ecology and Evolution, University of Edinburgh, Charlotte Auerbach Road, Edinburgh
518 EH9 2FL, United Kingdom

519 **Competing interests**

520 The authors declare no competing interests.

521 **Corresponding author**

522 Correspondence to Takashi Akagi (takashia@okayama-u.ac.jp)

523 ORCHID ID: 0000-0001-9993-8880

524

525 **FIGURE LEGENDS**

526 **Figure 1. Movement of both sex-determining factors, *Shy Girl* and *Friendly Boy*, and formation**
527 **of new MSYs around them.**

528 **a**, In *A. chinensis* and *A. polygama*, the sex-determining genes *SyGI* and *FrBy* are located near one
529 end of Chr 25 (or Y-chromosome), as indicated by the dark blue symbols. In *A. arguta* and *A. rufa*,
530 however, these genes are on Y-haplotype region of Chr 3 and Chr 4, respectively; these chromosomes
531 have no other major syntenic relationship with any Chr 25 region (see also Extended Data Figure S2).
532 The phylogeny was estimated based on analysis of 3,002 non-repetitive genes in the species (see
533 Methods). **b-c**, Close-up of the Y-haplotype region of Chr 4 in *A. rufa*. **b**, The ~1.5 Mb putative new
534 MSY region exhibited male-specific coverage by genomic reads. SLR: regions showing sex-linkage
535 in genetic maps of F1 populations (Extended Data Figure S3), S: *Shy Girl*, F: *Friendly Boy* **c**, synteny
536 between Y- and X-haplotypes. Green bands indicate predicted genes. **d-e**, Close-up views of the Chr
537 3 Y-haplotype region in *A. arguta*, formatted as in parts **b** and **c**. A ~2.7Mb region show new MSY
538 characteristics. **f**, Silent divergence (*dS*) values between alleles in 500 kb windows of the PAR
539 adjoining the X-Y or neo-Y regions. The *dS* values in *A. chinensis* and *A. polygama* (Chr 25) were
540 comparable to those associated with the *Actinidia* origin (and nearly equal to values since the origin
541 of *Shy Girl*: Akagi et al. 2018). Values in *A. rufa* (Chr 4), and *A. arguta* (Chr 3) are lower, consistent

with the hypothesis that Chr 25 is the oldest sex-linked region in the genus, and with the evolutionary topology of these 4 species. **g**, Venn diagram of genes shared by the putative MSYs of all 4 *Actinidia* species shown in part **a**. Apart from *Shy Girl*, and *Friendly Boy*, only one other *Y*-linked gene, *FT* (*YFT*) was found in all four species.

Figure 2. Rapid evolution of new MSYs via independent transposable elements (TEs) insertions.

a, TE density in the *A. rufa* and *A. arguta* MSYs, compared with the adjacent partially sex-linked regions (PARs). **b**, Results of Fisher's exact tests for the enrichment of TEs in the MSY regions in all 4 *Actinidia* species. The *A. rufa* MSY is enriched for several TE families (including LTR-unknown, *CACTA*, *hAT*, and *helitron* families). LTR-*Gypsy* may be enriched in the MSYs of all four species. **c**, Venn diagram showing the orthologous relationships of *Gypsy* TEs (>90% sequence homology) in the MSYs in 4 *Actinidia* species. Most *Gypsy* TEs are species-specific and none was common to all 4 species. **d**, Phylogenetic analysis of *Gypsy* in *A. rufa* (blue) and *A. arguta* (brown) MSYs. These elements exhibited frequent lineage-specific bursts (indicated by blue and brown bands) in independent clades.

Figure 3. Pleiotropic functions of *Shy Girl* explain the conserved sexual dimorphisms.

a, WT male flower (cv. Bruce). **b**, Hermaphrodite flower of the gene-edited *sygI* male line. An: anther, Ov: ovary, Sg: stigma. Self-pollination of the *sygI* line (**c**), and cross-pollination of Hort16A (female) x the *sygI* line (**d**), suggested functional male and female functions in the *sygI* line. Se: fertile seeds. **e**, The WT male exhibited clusters of flowers in most shoots. Fl-sh: floral shoot (orange arrows). **f**, Hermaphrodite *sygI* line exhibited female-like sexual dimorphisms, represented by drastic reduction in flower numbers and fewer floral shoots (or late-flowering). Vg-sh: vegetative shoot (turquoise arrows). **g-j**, flower numbers per inflorescence in the WT male (**g-h**) and *sygI* line (**i**), were statistically different (**j**) (p = 0.0031, Student's t-test). **k-l**, floral/vegetative node numbers per shoot (**k**), and adult

568 (floral) shoots ratio per individual (**I**) were also different between the WT and *sygI* lines.

569

570 **Figure 4. Model for recurrent neo-sex chromosome evolution in the genus *Actinidia*.**

571 Independent TE-mediated movement and shuffling of the sex determining genes. Sexual

572 dimorphisms conserved amongst the *Actinidia* species, are controlled by *Shy Girl* and independent of

573 the evolution of male-specific regions and recombination arrest in the Y-chromosomes.

574