PNAS
www.pnas.org

Supplementary Information for
Three genomes in the algal genus *Volvox* reveal the fate of a haploid sex-determining region after a transition to homothallism

Kayoko Yamamoto, Takashi Hamaji, Hiroko Kawai-Toyooka, Ryo Matsuzaki, Fumio Takahashi, Yoshiki Nishimura, Masanobu Kawachi, Hideki Noguchi, Yohei Minakuchi, James G. Umen, Atsushi Toyoda and Hisayoshi Nozaki

Hisayoshi Nozaki
Email: nozaki@bs.s.u-tokyo.ac.jp

This PDF file includes:

> Supplementary text 1-5
> Figures S1 to S11
> Tables S1 to S7
> SI References 1-27

Supplementary Information Text 1

Bridging between Contig011 and Contig058 of *Volvox reticuliferus* female genotype.
Genomic DNA of *V. reticuliferus* female culture strain NIES-3785 was prepared according to the method of Miller *et al.* (1). Specific primers (VrF_058_F1_788.3k and VrF_011_R1_2.7k; *SI Appendix*, Table S4) were designed for terminal regions of Contigs 011 and 058 to fill in a gap between the two contigs. PCR was performed using the total DNA and the two specific primers sets with KOD FX Neo (TOYOBO, Osaka, Japan) to amplify DNA fragments of ca 5 kbp long. PCR schedule was 2 min at 94°C, followed by 5 cycles of 10 s at 98°C and 60 s at 74°C, 5 cycles of 10 s at 98°C and 60 s at 72°C, 5 cycles of 10 s at 98°C and 60 s at 70°C, 40 cycles of 10 s at 98°C and 60 s at 68°C, 420 s at 68°C. For determining the nucleotide sequences, direct sequencing of the PCR-amplified fragments was carried out by cycle sequencing reactions with BigDye™ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems™, Thermo Fisher Scientific, Waltham, Massachusetts, USA) using specific primers (*SI Appendix*, Table S4). The determined sequence from all sets was 598 bp (Acc. No. LC582539) and demonstrated a bridging between Contig011 (368-1 positions) and Contig058 (790402-790173 positions) without gap between the contigs. Although the 598 bp sequence alternatively suggested another bridging between Contig011 (368-1 positions) and Contig013 (2289448-2289677 positions) without gap between the contigs, this bridging would result in discrepancy in one (*CGL55~TH/10*, Fig. 3A) of the two peripheral regions of *MT* which the large male Contig010 completely harbors. Thus, the bridging between Contig011 and Contig058 is more probable than that between Contig011 and Contig013.

Supplementary Information Text 2

Bridging between Contig018 and Contig132 of *Volvox africanus*. Genomic DNA of *V. africanus* culture strain NIES-3780 was prepared according to the method of Miller *et al.* (1). Specific primers (VxAfr-SF1 and VxAfr-SR3; *SI Appendix*, Table S5) were designed for terminal regions of Contig018 and Contig132 to fill in a gap between the two contigs. PCR was performed using the total DNA and the two specific primers (VxAfr-SF1 and VxAfr-SR1/VxAfr-SR3) with Tks Gflex™ DNA Polymerase (Takara Bio., Osaka, Japan) to amplify DNA fragments of ca 5 kbp long. PCR schedule was 1 min at 94°C, followed by 40 cycles of 10 s at 98°C and 180 s at 68°C. For determining the nucleotide sequences, direct sequencing of the PCR amplified fragments was carried out by cycle sequencing reactions with BigDye™ Terminator v3.1 Cycle Sequencing Kit using internal primers (*SI Appendix*, Table S5). A sequence of 5148 bp (Acc. No. LC582540) was determined just internal to the primer pair (VxAfr-SF1 and VxAfr-SR3) to bridge Contig018 and Contig132. Based on the blastn search by GenomeMatcher 3.0 (2), 949-5148 positions of the 5148 sequence were best-fitted with the anterior 4198 bp (1-4198 positions) of Contig132 whereas the remaining 948 bp (1-948 positions) of the 5148 bp correspond to 1532837-1533784 positions of Contig018 (1534657 bp). Thus, the most posterior 873 bp (1533785-1534657 positions) of Contig018 were removed for bridging between Contig018 and Cobtig132.

Supplementary Information Text 3

Identification of *FUS1* sequence of *Volvox africanus*. The polyadenylated mRNAs were directly isolated from *V. africanus* cells using Dynabeads Oligo(dT)$_{25}$ (Thermo Fisher Scientific), reverse transcribed with Superscript III reverse transcriptase (Thermo Fisher Scientific), and subjected to PCR (2 min at 94°C, followed by 30 cycles of 10 sec at 98°C, 30 sec at 65°C and 30 sec at 68°C) with two specific primers (FUS1_E2_F1: 5′-CCGTTCGCATTTACAGTGCAGCTC-3′; FUS1_E4_R1: 5′-GACAACCGACGCAGCTGGAGAAA-3′) and KOD FX Neo (TOYOBO, Osaka, Japan). To extend the cDNA sequences, 5' RACE and 3' RACE were performed using the GeneRacer kit (Thermo Fisher Scientific) and three specific primers (FUS1_E2_R1: 5'-GCTGTATGGACGCCTCTGATTGGT-3'; FUS1_E2_R2: 5'-TGCGGAACCTGGCAAGAGTTTACA-3'; FUS1_E4_F1: 5'-TGCGTCGGTTGTCCACATGTTAAG-3'). The PCR products were directly sequenced, or first cloned into the pCR4Blunt-TOPO vector (Thermo Fisher Scientific). For determining the nucleotide sequences, direct sequencing of the PCR-amplified fragments was carried out by cycle sequencing reactions with BigDye™ Terminator v3.1 Cycle Sequencing Kit. Based on the cDNA sequence and genome sequence,

only sequence harboring exon2-exon4 of *FUS1* was recognized (*SI Appendix*, Fig. S4). No RNA-seq data were obtained regarding *V. africanus FUS1*.

Supplementary Information Text 4

RNA-seq data mapping. For RNA-seq data, total RNA was extracted from asexual and sexually induced culture for each culture strain (*SI Appendix*, Table S2) using RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. Contaminating DNA was removed using RNase-Free DNase I (Takara Bio Inc., Shiga, Japan). The RNA-seq libraries were constructed using a TruSeq Stranded mRNA Library Prep (Illumina) and were sequenced on the Illumina HiSeq 2500 instruments. The Illumina reads from each sample were mapped to the reference genome with HISAT v2-2.1.0 (3), setting the parameters of "-q --phred33 -p 16 --rna-strandness RF". The mapped reads were then assembled into transcripts by StringTie v1.3.4d (3) with "--rf" option. Lastly, the candidate coding regions in transcript sequences were predicted by TransDecoder v5.2.0 (https://github.com/TransDecoder/TransDecoder/wiki).

Supplementary Information Text 5

Molecular phylogenetic analyses. Homologous protein sequences of fully sex-linked genes in *V. reticuliferus* and *V. africanus* (*SI Appendix*, Table S6) were retrieved from database of other volvocine algae by BLASTP (cutoff maximum E-value: 5e−2) on NCBI. *Chlamydomonas reinhardtii* protein sequences were treated as outgroups. When such an outgroup sequence was not retrieved, homologous sequences were extracted by BLASTP (cutoff maximum E-value: 1e−1) from *Chlamydomonas reinhardtii* v5.5 genome data in phytozome v12.1 (https://phytozome.jgi.doe.gov/pz/portal.html#). Phylogenetic analyses were performed using MUSCLE (4)-aligned full-length protein sequences of fully sex-linked genes (Fig. 3; *SI Appendix*, Figs. S6-S8). The maximum likelihood method was subjected to each alignment with complete deletion option and bootstrap analysis based on 1000 replications by MEGA X (5) using the best-fitted model selected by MEGA X (*SI Appendix*, Table S7). In addition, Bayesian inference for the respective alignments was carried out using MrBayes 3.2.7a (6) with the best-fitted model selected by ModelTest-NG 0.1.6 (7) (*SI Appendix*, Table S7), with 1,000,000 generations of Markov chain Monte Carlo iterations (discarding the first 25% as burn-in). Convergence of the analysis was confirmed by the average standard deviation of split frequencies below 0.01.
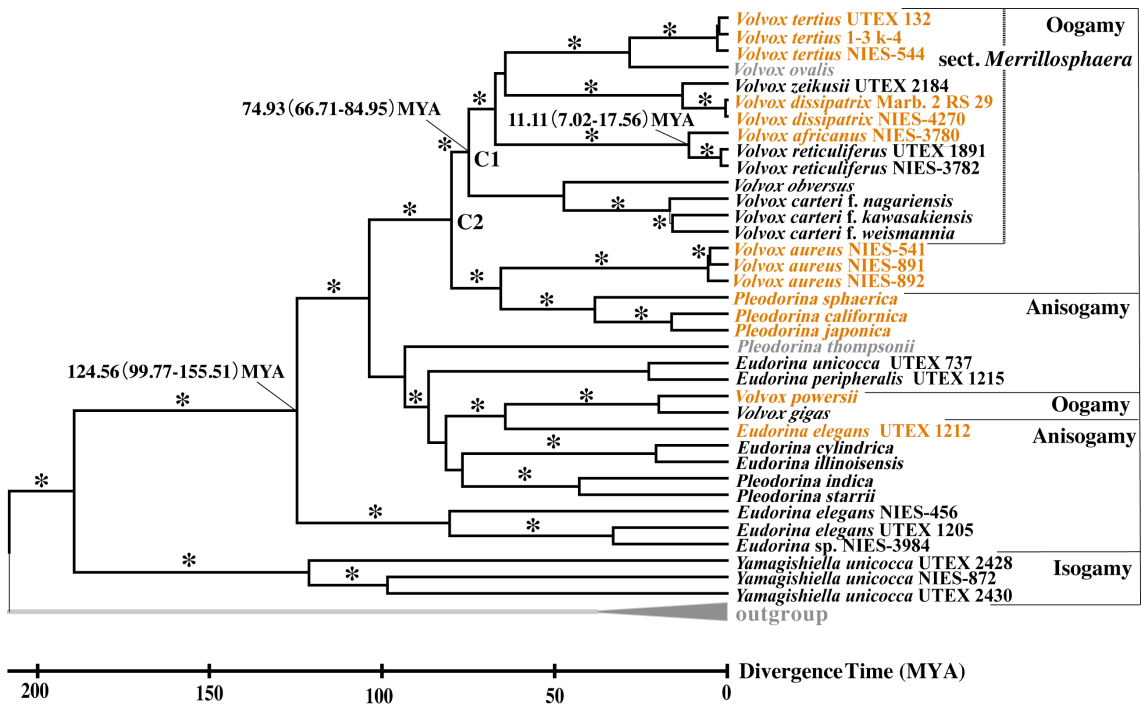
Fig. S1. Timetree analysis of advanced members of the colonial volvocine algae. Note that the divergence time between heterothallic *Volvox reticuliferu*s and homothallic *V. africanus* is "11.11 (7.02-17.56) MYA". Species names in orange or black indicate homothallic or heterothallic sexuality, respectively. Tree topology was inferred by maximum likelihood (ML) analyses of 6021 base pairs of five chloroplast genes (8) [TreeBase (9, 10) ID 26647 plus *Eudorina* sp. NIES-3984 (Acc. No. MH267732)] with a model selected by MEGA X (5). Asterisks on branches indicate 80% or more bootstrap values (based on 1000 replicates) by the ML analyses. A timetree was inferred by applying the RelTime method (11, 12) to the ML tree whose branch lengths were calculated using the ML method and the General Time Reversible substitution model (13). The timetree was computed using two calibration constraints (C1 (65-90 MYA) and C2 (50-90 MYA)) based on TimeTree: the Time Scale of Life < http://www.timetree.org/> (14, 15)). The Tao *et al.* (16) method was used to set minimum and maximum time boundaries on nodes for which calibration densities were provided. The estimated log likelihood value of the tree is -38392.49. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.8168)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 57.55% sites). This analysis involved 46 nucleotide sequences. Codon positions included were 1st+2nd+3rd. There were a total of 6021 positions in the final dataset. Evolutionary analyses were conducted in MEGA X.
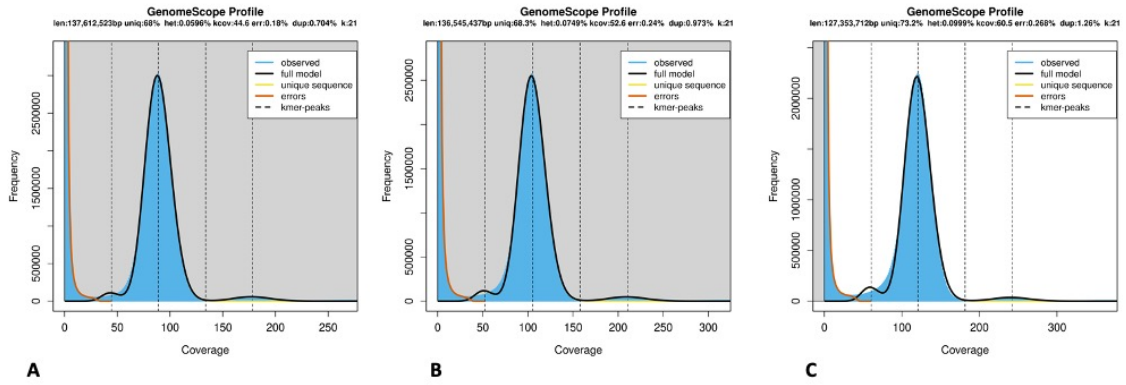
Fig. S2. GenomeScope profiles (17) showing *k-mer* frequencies in three genomes of *Volvox* (*SI Appendix*, Table S2). A. *V. reticuliferus* female. B. *V. reticuliferus* male. C. *V. africanus*.

Fig. S3. *Volvox reticuliferus* dotplot between male (vertical) and female (horizontal) specific regions [male SDR (*MTM*, blue) and female SDR (*MTF*, red), respectively] and parts of pseudo autosomal regions (gray). Green and blue dots indicate forward and reverse alignments, respectively. For details of *MTM* and *MTF*, see Fig. 3.
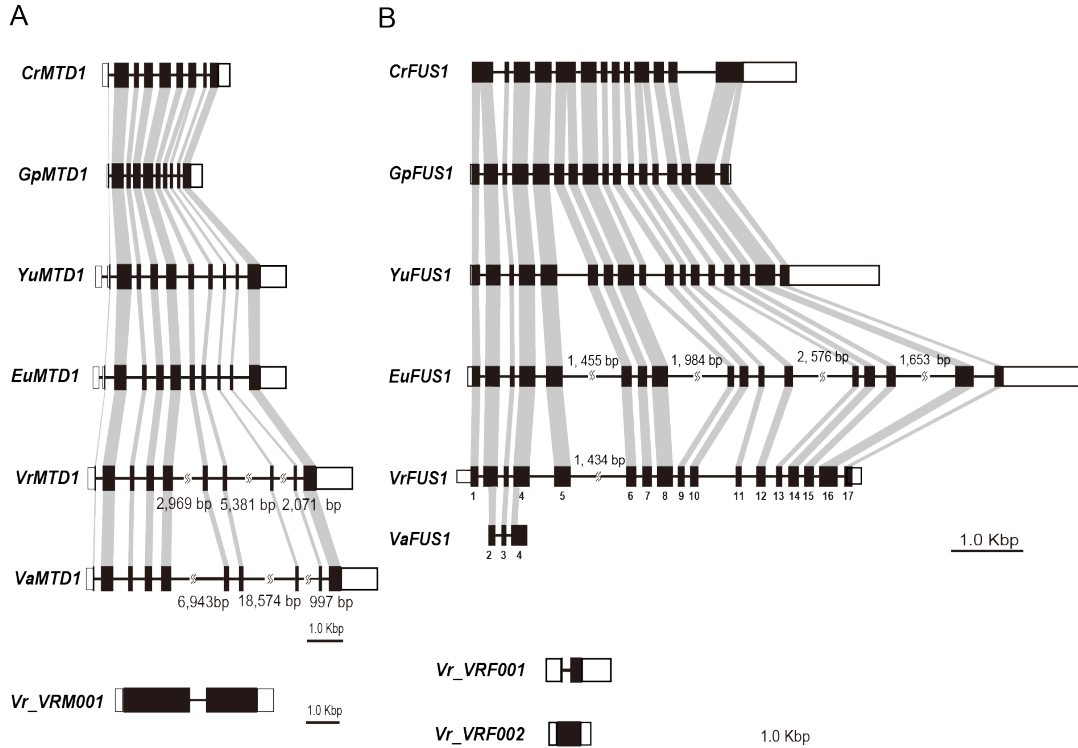
Fig. S4. The exon-intron structures of male- and female-specific genes of *Volvox reticuliferus* and their homologs in other volvocine species (Fig. 3). *Cr*, *Gp*, *Yu*, *Eu*, *Vr* and *Va* at the prefixes of gene names represent *Chlamydomonas reinhardtii*, *Gonium pectorale*, *Yamagishiella unicocca*, *Eudorina* sp., *V. reticuliferus* and *V. africanus*, respectively. Filled and open boxes represent coding and non-coding exon sequences, respectively. Lines between boxes represent introns. Gray boxes link homologous coding sequences. A. Two male-specific genes: *MTD1* and *VRM001*. For *MID* homologs, see Yamamoto *et al.* (18). B. Three female-specific genes: *FUS1*, *VRF001* and *VRF002*. Numbers below boxes for *VrFUS1* and *VaFUS1* indicate exon numbers.
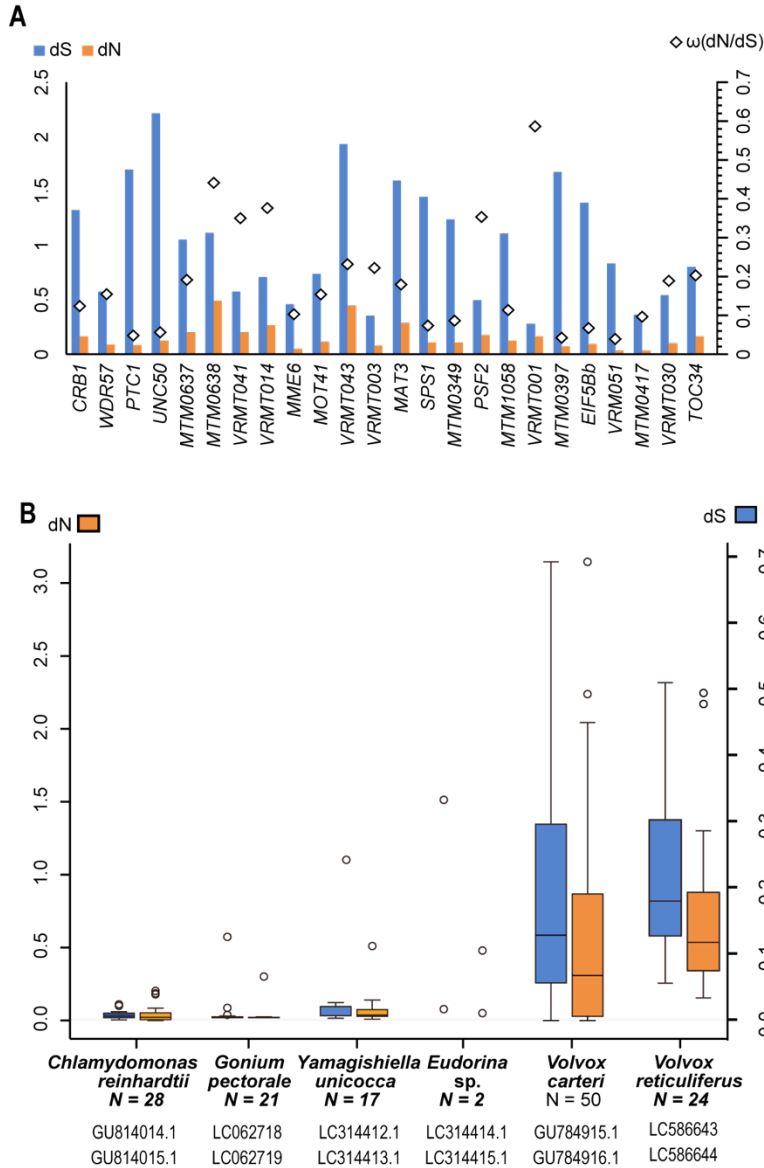
Fig. S5. Molecular evolutionary analyses of gametologs in *Volvox reticuliferus*. A. dN/dS ratios of gametologs in rearranged regions or sex-determining regions (SDRs) of *V. reticuliferus*. There are no prominently dimorphic gametologs under positive selection between sexes (dN/dS > 1). B. Box-whisker plots comparing the distributions of synonymous (dS, blue/left) and non-synonymous (dN, orange/right) substitution values for gametolog pairs found in SDRs of volvocine algal haploid UV chromosomes. Open dots are outliers from interquartile ranges except for those of *Eudorina* sp. which indicate two gametologs.
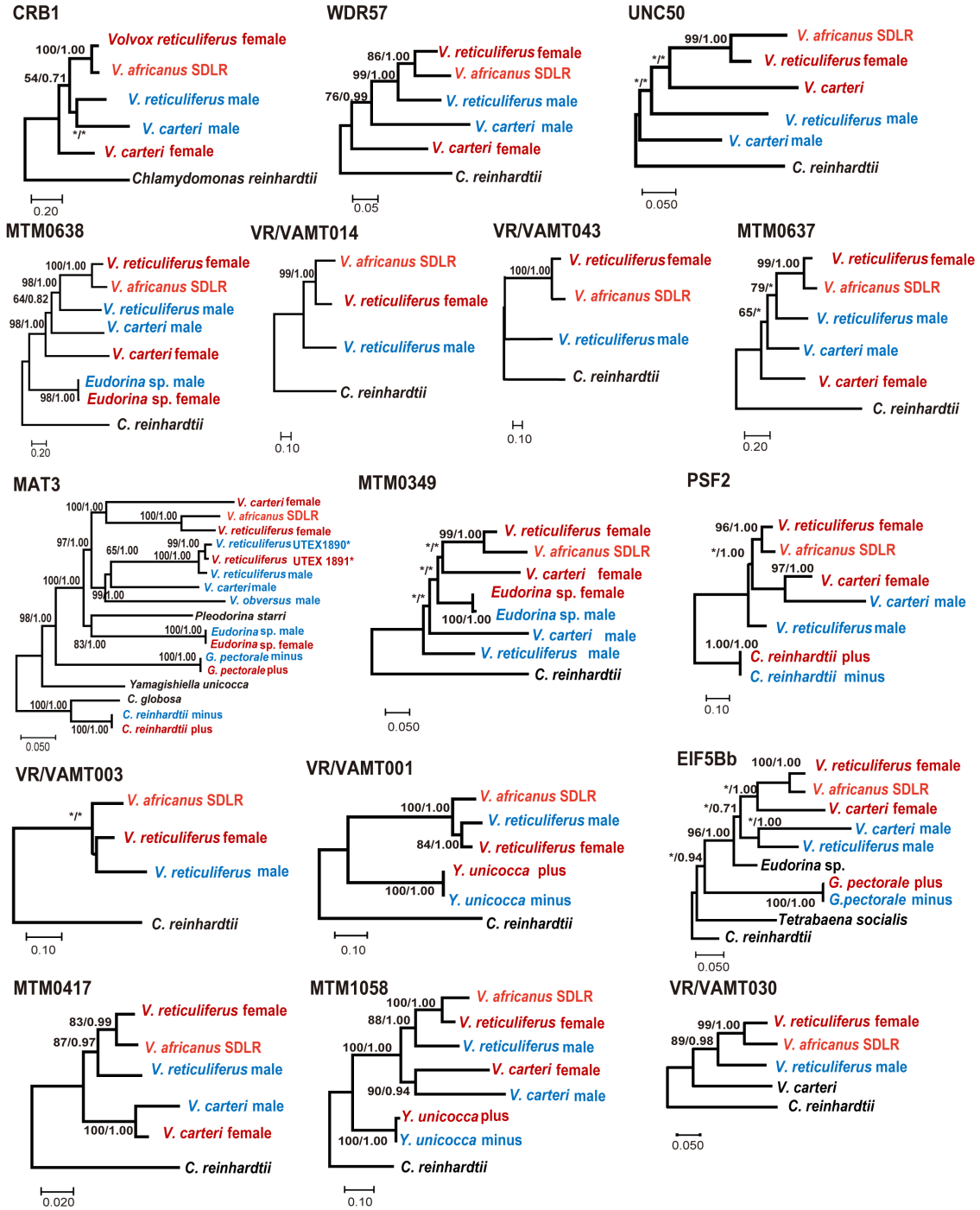
Fig. S6. Maximum likelihood (ML) phylogeny of 16 homologs of gametologs harboring in *Volvox africanus* SDLR (Fig. 3). Red and blue represent homologs of gametologs from female SDR (SDLR) and male SDR, respectively. Note that 14 *V. africanus* homologs represent *V. reticuliferus* female SDR-related whereas male- or female-relation of the other two (*VAMT003* and *VAMT001*) are ambiguous. Numbers in left and right sides at branches indicate bootstrap values of ML analysis and posterior probabilities of Bayesian inference, respectively. For the other four gametolog homologs in *Volvox africanus* SDLR, see Fig. 3 and Fig. S7 in *SI Appendix*.

9

MTM0397
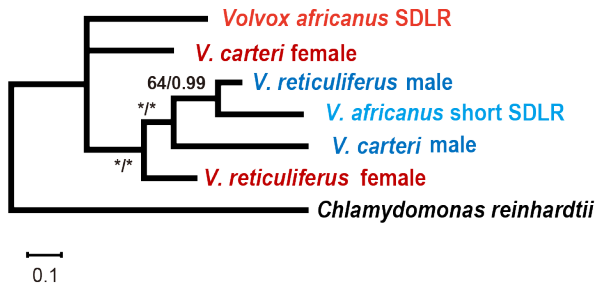
VR/VAMT051
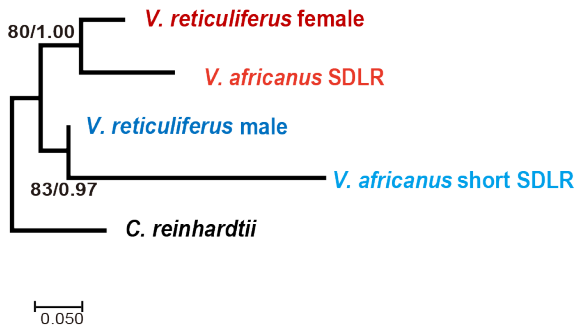
Fig. S7. Maximum likelihood (ML) phylogeny of two homologs of gametologs harboring in both SDLR and short SDLR of *Volvox africanus* (Fig. 3). Red and blue represent homologs of gametologs from female SDR (SDLR) and male SDR (short SDLR), respectively. Numbers in left and right sides at branches indicate bootstrap values of ML analysis and posterior probabilities of Bayesian inference, respectively. For the other gametolog homologs in *Volvox africanus* SDLR and short SDLR, see Fig. 3 and Fig. S6 in *SI Appendix*.

MOT41

TOC34
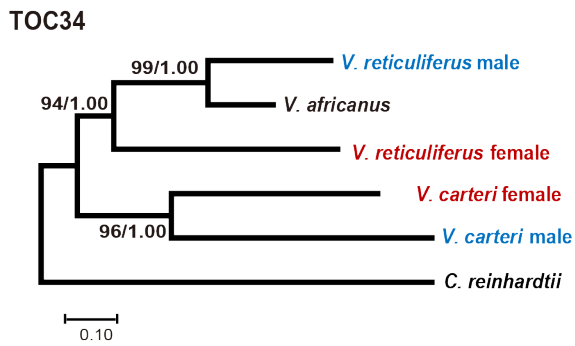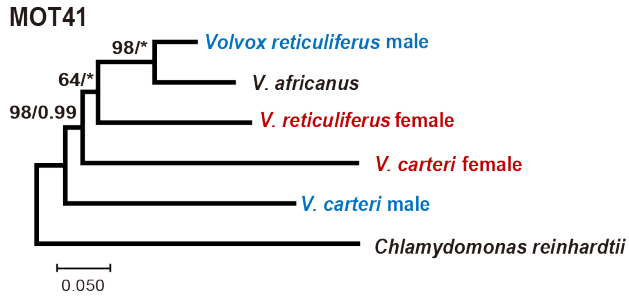
Fig. S8. Maximum likelihood (ML) phylogeny of and two gametologs of *Volvox reticuliferus*. Red and blue color in gene names represent gametologs from female and male SDR, respectively. Note that two *V. africanus* homologs are found within autosome-like regions (outside SDLR and short SDLR) (Fig. 4A). Numbers in left and right sides at branches indicate bootstrap values of ML analysis and posterior probabilities of Bayesian inference, respectively.

Fig. S9. Comparison of homologous genes in expanded sex-determining regions (SDRs) of *Volvox carteri* and *V. reticuliferus*. Red, blue and gray regions represent female SDR (*MTF*), male SDR (*MTM*) and pseudo autosomal regions, respectively. Note two male-specific genes (backed blue) in *MTM* and 17 gametologs in *MTM* and *MTF* for each species. Only homologous genes between *V. carteri* and *V. reticuliferus* are shown. Based on the present study (Fig. 3) and Ferris *et al.* (19).
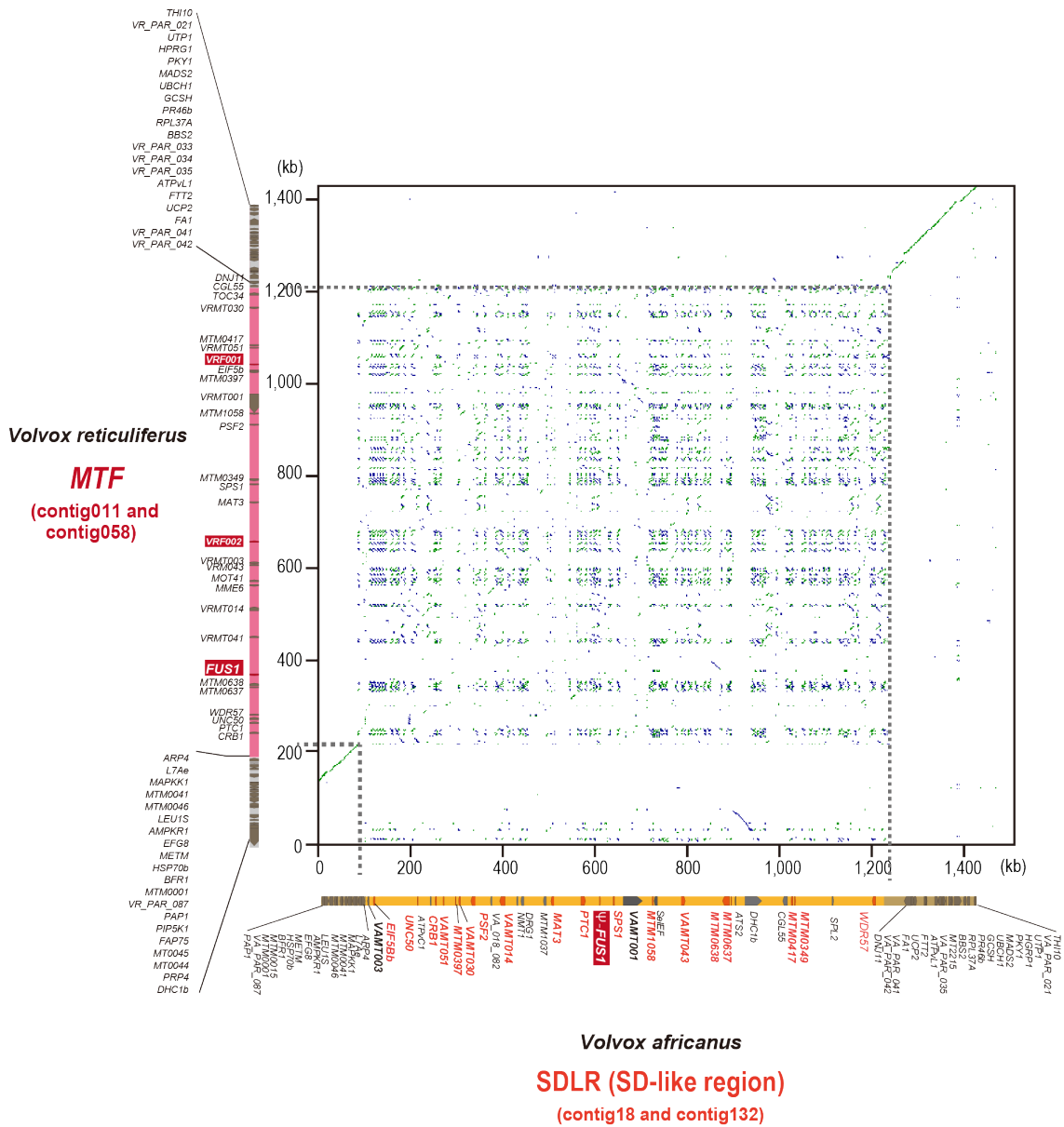
Fig. S10. Dotplot between *Volvox reticuliferus* fully sex-linked region (vertical, female SDR or *MTF*) and its homologous region of *Volvox africanus* (horizontal, SDLR) and parts of flanking or pseudo autosomal regions (gray). Green and blue dots indicate forward and reverse alignments, respectively. For details of *MTF* and SDLR, see Fig. 3.

13

Volvox africanus
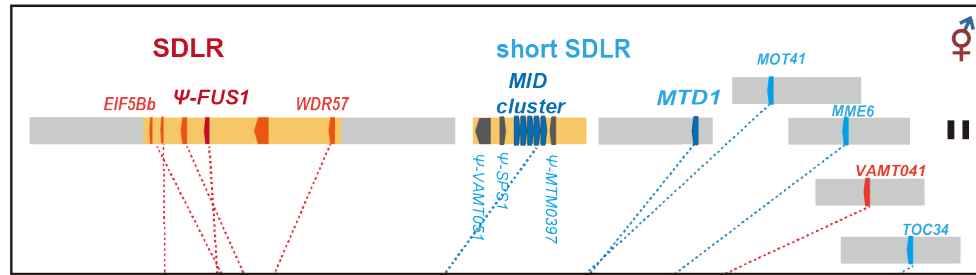
Volvox reticuliferus
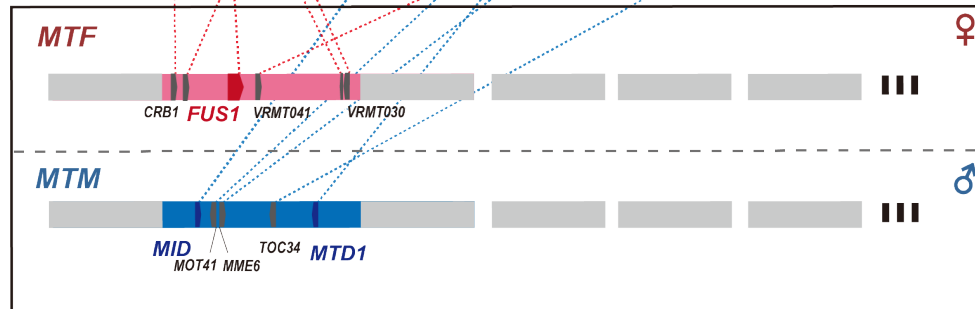
Fig. S11. Diagrammatic comparison between sex-determining region (SDR)-like sequences (SDLR and short SDLR) of homothallic *V. africanus* (upper panel) and female SDR (*MTF*) and male SDR (*MTM*) of heterothallic *Volvox reticuliferus* (lower panel). Only four homologs of female-related gametologs in SDLR are shown. Based on the present study.

Table S1. Four types of mating systems in the green algal genus *Volvox*[*].

| Basic system	heterothallic mating system[**]	homothallic mating system[***]		
Sexual spheroids produced[****]	male and female spheroids by different genotypes	male and female spheroids by a single genotype	male and bisexual spheroids by a single genotype	bisexual spheroids
Volvox species	*V. carteri*, *V. reticuliferus*	*V. aureus*	*V. africanus*	*V. ferrisii*
References	Starr (20), Nozaki *et al.* (21)	Darden (22)	Nozaki *et al.* (21)	Isaka *et al.* (23)

[*]Based on Starr (24).

[**]In the heterothallic (with distinct male and female genotypes, associated with a mating-type system that prevents fusion of gametes of the same sex) mating system, male or female is determined by the two complementary genotypes and a single clonal culture produces only one type of gametes, eggs in female spheroids or sperm in male spheroids. Thus, male or female is determined genetically in the heterothallic species or clonal culture strains. The heterothallic species perform only outcrossing.

[***]In the homothallic (bisexual, with the ability to self-fertilize) mating system, gametes of both sexes (eggs and sperm) are produced within a single clonal culture (the same genotype). Thus, the homothallic species have possibility to form zygotes between gametes from identical genotype or a single clonal culture (selfing).

[****]*Volvox* sexual spheroids (individuals) may be "unisexual" (male or female) or "bisexual" (production of both sperm and eggs) depending upon a species or lineage. Thus, the homothallic *Volvox* mating system is further classified into three subtypes based on production of the types of sexual spheroids (24): production of both male and female sexual spheroids, production of both male and bisexual spheroids and production of only bisexual spheroids in a single clonal culture (genotype). Male, female and bisexual spheroids produce only male gametes (sperm), only female gametes (eggs) and both sperm and eggs, respectively.

Table S2. Comparison of details of whole nuclear genomes and SDR or SDLR/short SDLR of three *Volvox* culture strains.

	Species	*V. reticuliferus*		*V. africanus*
	Mating type/ sex (culture strain)	Female (NIES-3786)	Male (NIES-3785)	Homothallic (NIES-3780)
Whole genome	Total length (bp) [Total number of gaps/non-ATGC nucleotides (bp)]	133,065,142 [0/0]	133,961,728 [0/0]	129,328,469 [0/0]
	Number of contigs	200	230	448
	Min (bp)	18,733	21,643	10,030
	Max (bp)	4,940,096	5,524,336	6,701,515
	contig N50	1,907,605	1,866,906	1,356,898
	%GC	54	54	53
	Number of genes	13,860	14,050	12,903
	Gene density (genes/Mb)	104.2	104.9	99.7
	Repeats (%)*	26.82	26.82	25.85
	BUSCO scores**	C: 97.7% S: 95.9% D: 1.8% F: 0.3% M:2.0%	C: 94.9% S: 93.0% D: 1.9% F: 0.4% M:4.7%	C: 98.1% S: 96.4% D: 1.6% F: 0.5% M: 1.4%
	Estimated genome size (Mbp)***	137.6	136.5	127.4
SDR or SDLR/ short SDLR	Size (Mbp)	1.01	0.98	1.02/0.20
	%GC	51	51	51/50
	Number of genes	28	28	30/4
	Gene density (genes/Mb)	27.64	28.82	29.44/20.00
	Repeats (%)*	70.15	70.20	74.83/84.83

*Repetitive sequences identified using RepeatMasker-open-4-0-9 (http://www.repeatmasker.org) with Dfam3.0, generated libraries of repeats by RepeatScout (http://bix.ucsd.edu/repeatscout/).
**BUSCO v4.0.4 (25) scores calculated based on chlorophyta_odb10 (1,519 BUSCO). C: complete; S: complete and single-copy; D: complete and duplicated; F: fragmented; M: missing.
***Based on *k-mer* profiles by GenomeScope (17).

Table S3. Comparison of whole genome and SDR or SDLR/short SDLR properties of the volvocine algae (Fig. 2)*.

Species	Mating type/sex	Whole genome				
		Size/total length (Mb)	%GC	Number of genes	Gene density (genes/Mb)	DDBJ/ENA/ GenBank accessions no.
V. reticuliferus	Female	133	54	13860	104.2	BNCP01000001- BNCP01000200
	Male	134	54	14050	104.8	BNCQ01000001- BNCQ01000230
V. africanus	Homothallic	127	53	13716	108.1	BNCO01000001- BNCO01000448
V. carteri	Female	131.1	56.1	14958	114	GCA_000143455.1
	Male	N.d.	N.d.	N.d.**	N.d.	N.d.
Eudorina sp.	Female	184	61	N.d.	N.d.	GCA_003117195.1
	Male	168.6	61.3	N.d.	N.d.	GCA_003117095.1
Y. unicocca	Plus	134.2	61.1	N.d.	N.d.	GCA_003116995.1
	Minus	140.8	60.8	N.d.	N.d.	GCA_003117035.1
G. pectorale	Plus	149	64.5	17990	121	GCA_001584585.1
	Minus	N.d.	N.d.	N.d.	N.d.	N.d.
C. reinhardtii	Plus	111.1	64.1	17732	159.6	GCA_000002595.2
	Minus	N.d.	N.d.	N.d.	N.d.	N.d.

Table S3. Continued.

Species	Mating type/sex	SDR or SDLR/short SDLR				
		Size (Mb)	%GC	Number of genes	Gene density (genes/Mb)	DDBJ/ENA/GenBank accessions no.
V. reticuliferus	Female	1.01	51	28(25***)	27.7	LC586643
	Male	0.98	51	28(25)	28.6	LC586644
V. africanus	Homo-thallic	1.02	51/50	30/4	29.4/20	LC586641/ LC586642
V. carteri	Female	1.51	52	55(50)	39	GU784915.1
	Male	1.13	53	60(50)	54	GU784916.1

Eudorina sp.	Female	0.09	53.9	3(2)	33.3	LC314414.1
	Male	0.007	51.4	3(2)	428	LC314415.1
Y. unicocca	Plus	0.268	60.1	18(17)	67.2	LC314412.1
	Minus	0.165	60.3	18(17)	109	LC314413.1
G. pectorale	Plus	0.366	59.7	24(21)	58	LC062718
	Minus	0.499	61	24(21)	46	LC062719
C. reinhardtii	Plus	0.31	60	35(22)	109	GU814014.1
	Minus	0.204	61	25(22)	118	GU814015.1

*References: *V. reticuliferus* and *V. africanus* (the present study); *C. reinhardtii* (19); *G. pectorale* (26); *Y. unicocca* and *Eudorina* sp. (27); *V. carteri* (19).
**Not determined.
***The number of gametologs in parentheses.

Table S4. Primers used for amplification and sequencing of the DNA fragments filling in a gap between Contigs 011 and 058 of *Volvox reticuliferus* female genotype.

Primer name	Sequence (5'-3')	Nucleotide positions
VrF_058_F1_788.3k	ATCACGACGAAGGGTATCATCATTAGGC	788289-788316 in Contig058
VrF_058_F2_790.1k	TGCCTAGCTTTCTTTGCCGCAATAGAC	790135-790161 in Contig058
VrF_011_R2_0.2k	CATTAGGCTGCTTTAGTTCGCGTATGTG	209-183 in Contig011
VrF_011_R1_2.7k	GGTTGCATGCACTCCCATTTCTTCAG	2751-2726 in Contig011

Table S5. Primers used for amplification and sequencing of the DNA fragments filling in a gap between Contigs 018 and 132 of *Volvox africanus*.

Primer name	Sequence (5'-3')	Nucleotide positions
VxAfr-SF1	GCCTGAGCTCGGTGATACTGACTTCGGGAG	1532807-1532836 in Contig018
VxAfr-SR3	AACGTGCTGGGGTGAATCATACCCTTGGCG	4228-4199 in Contig132
VxAFr-In1	TCGGAGGGACAGATGGTTTTGCTGAGCACT	1123-1152*
VxAfr-In2	GGGATAACTCACAATATTCGACTCGCGTCG	1398-1427*
VxAfr-In3	GGCCTTGATTGAGCAGACGAGGAGGCTGAG	2372-2401*
VxAfr-In4	CGTGAACGGATACTTGCTGTATCTGACCGG	2408-2437*
VxAfr-In5	GTGCGCAGCGTCGTAAGCCGGATCCTGTGC**	3436-3407*
VxAfr-In6	GGTGCGCACGACCGTTGTCGCAGCGGTAGG	3493-3464*
VxAfr-In7	GTCCGGATCCAGCAGTAAGGCCCACTTAGT	181-152*
VxAfr-In8	GTTTAACGGTGTGGATGTAGCAGGCGTGCC	4222-4251*
VxAfr-In9	TCTTCGCTTATGCAAGGGGACTCTGTTGGG	4714-4743*
VxAfr-In10	GGTGGTTCAGGTCGACCTTGGTGGTTGTTT	4399-4428*
VxAfr-In11	GGCGTGCCGGTTCCGTTGACGGGGCCGGCA	4244-4273*
VxAfr-SF2	ACTAAGTGGGCCTTACTGCTGGATCCGGAC	152-181*
VxAfr-SR1	TGCCGGCCCCGTCAACGGAACCGGCACGCC	4273-4244*
VxAfr-SR2	GCAAACGACACGAAGGGTACCTTGCCCAAC	4767-4738*
VxAfrR-3	ATCTGGCGAACAACAGAAAGACTTAGC	1230-1204*
VxAfrR-4	AAATTCGCTCGGACAGTATTCACAAAG**	1594-1568*

*Nucleotide positions in the bridging sequence (5148 bp: Acc. No. LC582540) between VxAF-SF1 (Contig018) and VxAF-SPR3 (Contig132).
** Not completely identical to the bridging sequence.

Table S6. List of genes in SDR and pseudo autosomal regions of *Volvox reticuliferu*s (Fig. 3) and their orthologs in *V. africanus* and *V. carteri*. Genes in SDR or R domain are shown in bold.

V. reticuliferus gene	*V. africanus* ortholog		*V. carteri* ortholog
	Location	Name	
THI10	autosome-like	THI10	THI10
VR_PAR_021	autosome-like	VA_PAR_021	-
UTP1	autosome-like	UTP1	**UTP1**
HGRP1	autosome-like	HGRP1	**HGRP1**
PKY1	autosome-like	PKY1	PKY1
MADS2	autosome-like	mads2	MADS2
UBCH1	autosome-like	UBCH1	UBCH1
GCSH	autosome-like	GCSH	GCSH
PR46b	autosome-like	PR46b	PR46b
RPL37A	autosome-like	RPL37A	RPL37a
BBS2	autosome-like	BBS2	BBS2
VR_PAR_033	-	-	VOLCADRAFT_99150
VR_PAR_034	-	-	VOLCADRAFT_99151
VR_PAR_035	-	-	-
ATPvL1	autosome-like	ATPvL1	**ATPvL1**
FTT2	autosome-like	FTT2	**FTT2**
UCP2	autosome-like	UCP2	**UCP2**
FA1	autosome-like	FA1	**FA1**
VR_PAR_041	autosome-like	VA_PAR_041	-
VR_PAR_042	autosome-like	VA_PAR_042	VOLCADRAFT_88639
DNJ11	autosome-like	DNJ11	dnj11
CGL55	SDLR	CGL55	CGL55
VRMT001	SDLR	VAMT001	-
VRMT003	SDLR	VAMT003	-
MTM1058	SDLR	MTM1058	**MTM1058**
MTM0349	SDLR	MTM0349	**MTM0349**
MTM0417	SDLR	MTM0417	**MTM0417**
MTD1*	autosome-like	MTD1	**ψ-MTD1***
VRM001*	-	-	-
VRMT014	SDLR	VAMT014	-

WDR57	SDLR	WDR57	**WDR57**
VRMT030	SDLR	VAMT030	VOLCADRAFT_89775
TOC34	autosome-like	TOC34	**TOC34**
SPS1	SDLR/short SDLR	SPS1/ψ-SPS1	**SPS1**
EIF5Bb	SDLR	EIF5Bb	**EIF5Bb**
MAT3	SDLR	MAT3	**MAT3**
CRB1	SDLR	CRB1	**CRB1**
VRMT041	autosome-like	VAMT041	-
VRMT043	SDLR	VAMT043	-
MTM0397	SDLR/short SDLR	MTM0397/ ψ-MTM0397	**MTM0397**
MME6	autosome-like	MME6	**MME6**
MOT41	autosome-like	MOT41	**MOT41**
VRMT051	SDLR/short SDLR	VAMT051	-
UNC50	SDLR	UNC50	**UNC50**
MID*	short SDLR	MID	**MID***
MTM0637	SDLR	MTM0637	**MTM0637**
MTM0638	SDLR	MTM0638	**MTM0638**
PTC1	SDLR	PTC1	**PTC1**
PSF2	SDLR	PSF2	**PSF2**
ARP4	autosome-like	ARP4	**ARP4**
L7Aef	autosome-like	L7Ae	**L7Ae**
MAPKK1	autosome-like	MAPKK1	**MAPKK1**
MTM0041	autosome-like	MTM0041	**MTM0041**
MTM0046	autosome-like	MTM0046	**MTM0046**
LEU1S	autosome-like	LEU1S	**LEU1S**
AMPKR1	autosome-like	AMPKR1	**AMPKR1**
EFG8	autosome-like	EFG8	**EFG8**
METM	autosome-like	METM	**METM**
HSP70b	autosome-like	HSP70b	**HSP70**
BFR1	autosome-like	BFR1	**BFR1**
MTM0001	autosome-like	MTM0001	**MTM0001**
VR_PAR_087	autosome-like	VA_PAR_087	VPS53

PAP1	autosome-like	PAP1	PAP1
PIP5K1	-	-	PIP5K1
FAP75	-	-	FAP75
MT0045	autosome-like	Va.22190.1	MT0045
MT0044	autosome-like	MT0044	MT0044
PRP4	autosome-like	PRP4	PRP4
DHC1b	SDLR	DHC1b	**DHC1b**
VRF001**	-	-	-
VRF002**	-	-	-
FUS1**	SDLR	ψ-FUS1	-

*Male-specific gene.
**Female-specific gene.

Table S7. Models used for the maximum likelihood analyses (ML) and Bayesian inference (BI) in molecular phylogenetic analyses of homologs of fully sex-linked genes (male- and female-specific genes and gametologs) of *Volvox reticuliferus* and *V. africanus* (Figs. 3, 4; *SI Appendix*, Figs. S6-S8).

Gene name	ML_model*	BI_model**	Figure
CRB1	JTT +G	Dayhoff +I+G	Figure S6
WDR57	JTT+G	WAG+G	Figure S6
PTC1	JTT+G	WAG+G	Figure 3B
UNC50	JTT+G	Cprev +G	Figure S6
MTM0637	JTT+G+I	Dayhoff+G	Figure S6
MTM0638	JTT+G+I+F	WAG+I+G	Figure S6
VR/VAMT041	JTT+G	Dayhoff+G	Figure 4B
VR/VAMT014	JTT+G	Vt +G	Figure S6
MME6	JTT+G	WAG+I+G	Figure 4B
MOT41	JTT+G	WAG+G	Figure S8
VR/VAMT043	JTT+G	WAG+G	Figure S6
VR/VAMT003	LG+G ^f	Dayhoff+G	Figure S6
MAT3	JTT+G	Dayhoff+I+G	Figure S6
SPS1	JTT+G	Dayhoff+G	Figure 3B
MTM0349	JTT+G	WAG+G	Figure S6
PSF2	JTT+G	Dayhoff+I+G	Figure S6
MTM1058	JTT+G	Dayhoff+G	Figure S6
VR/VAMT001	JTT+G+F	WAG+G	Figure S6
MTM0397	LG+G	WAG+G	Figure S7
eif5Bb	JTT+I	Dayhoff+I	Figure S6
VR/VAMT051	JTT	Vt	Figure S7
MTM0417	JTT+G	Dayhoff+G	Figure S6
VR/VAMT030	LG+G	Vt+G	Figure S6
TOC34	JTT+G	WAG+G	Figure S8
MID	JTT+G	WAG+G	Figure 3B
MTD1	JTT+G	WAG+G	Figure 4B
FUS1	JTT+G	WAG+G	Figure 3B

*The best-fitted model selected by MEGA X (5).
**The best-fitted model was selected by Modeltest-NG 0.1.6 (7).

SI References

1. S. M. Miller, R. Schmitt, D. L. Kirk, *Jordan*, an active *Volvox* transposable element similar to higher plant transposons. *Plant Cell* **5**, 1125–1138 (1993).
2. Y. Ohtsubo, W. Ikeda-Ohtsubo, Y. Nagata, M. Tsuda, GenomeMatcher: A graphical user interface for DNA sequence comparison. *BMC Bioinformatics* **9**, 376 (2008).
3. M. Pertea, D. Kim, G. M. Pertea, J. T. Leek, S. L. Salzberg, Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols.* **11**, 1650–1667 (2016).
4. R. C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
5. S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
6. F. Ronquist *et al.*, MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol*. **61**, 539–542 (2012).
7. D. Darriba *et al.*, ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Mol. Biol. Evol.* **37**, 281–294 (2020).
8. H. Nozaki *et al.,* Morphology, reproduction and taxonomy of *Volvox dissipatrix* (Chlorophyceae) from Thailand, with a description of *Volvox zeikusii* sp. nov. *Phycologi*a **58**, 192–199 (2019).
9. R. A. Vos, H. Lapp, W. H. Piel, V. Tannen, TreeBASE2: Rise of the machines. *Nat. Prec.* https://doi.org/10.1038/npre.2010.4600.1. Deposited 30 June 2010.
10. R. A. Vos *et al.*, NeXML: rich, extensible, and verifiable representation of comparative data and metadata. *Syst. Biol.* **61**, 675–689 (2012).
11. K. Tamura *et al.*, Estimating divergence times in large molecular phylogenies. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 19333–19338 (2012).
12. K. Tamura, Q. Tao, S. Kumar, Theoretical foundation of the RelTime method for estimating divergence times from variable evolutionary rates. *Mol. Biol. Evol.* **35**, 1770–1782 (2018).
13. M. Nei, S. Kumar S, *Molecular Evolution and Phylogenetics* (Oxford University Press, 2000).
14. M. D. Herron, J. D. Hackett, F. O. Aylward, R. E. Michod, Triassic origin and early radiation of multicellular volvocine algae. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 3254–3258 (2009).
15. M. Liss, D. L. Kirk, K. Beyser, S. Fabry, Intron sequences provide a tool for high-resolution phylogenetic analysis of volvocine algae. *Curr. Genet.* **31**, 214–227 (1997).
16. Q. Tao, K. Tamura, B. Mello, S. Kumar, Reliable confidence intervals for RelTime estimates of evolutionary divergence times. *Mol. Biol. Evol.* **37**, 280–290 (2020).
17. G. W. Vurture *et al.*, GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics* **33**, 2202–2204 (2017).
18. K. Yamamoto *et al.*, Molecular evolutionary analysis of a gender-limited MID ortholog from the homothallic species *Volvox africanus* with male and monoecious spheroids. *PLoS ONE* **12**, e0180313 (2017).
19. P. Ferris *et al.*, Evolution of an expanded sex-determining locus in *Volvox*. *Science* **328**, 351–354 (2010).
20. R. C. Starr, Structure, reproduction and differentiation in *Volvox carteri* f. *nagariensis* Iyengar, strains HK9 & 10. *Arch. Protistenkd.* **111**, 204–222 (1969).
21. H. Nozaki, R. Matsuzaki, K. Yamamoto, M. Kawachi, F. Takahashi, Delineating a new heterothallic species of *Volvox* (Volvocaceae, Chlorophyceae) using new strains of '*Volvox africanus*'. *PLoS ONE* **10**, e0142632 (2015).
22. W. H. Darden Jr., Sexual differentiation in *Volvox aureus*. *J Protozool*. **13**, 239–55 (1966).
23. N. Isaka, H. Kawai-Toyooka, R. Matsuzaki, T. Nakada, H. Nozaki, Description of two new monoecious species of *Volvox* sect. *Volvox* (Volvocaceae, Chlorophyceae), based on comparative morphology and molecular phylogeny of cultured material. *J Phycol.* **48**, 759–767 (2012).

24. R. C. Starr, "Sexual reproduction in *Volvox africanus*" in *Contribution in Phycology*, B. C. Parker, R. M. Brown Jr., Eds. (Allen Press, 1971), pp. 59–66.
25. F. A. Simão, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatic*s **31**, 3210–3212 (2015).
26. T. Hamaji *et al.*, Sequence of the *Gonium pectorale* mating locus reveals a complex and dynamic history of changes in volvocine algal mating haplotypes. *G3 (Bethesda)* **6**, 1179–1189 (2016).
27. T. Hamaji *et al.*, Anisogamy evolved with a reduced sex-determining region in volvocine green algae. *Commun. Biol.* **1**, 17 (2018).