OPEN ACCESS Freely available online

PLOS ONE

Linkage Mapping Identifies the Sex Determining Region as a Single Locus in the Pennate Diatom *Seminavis robusta*

Ives Vanstechelman[1,2,3], Koen Sabbe[1], Wim Vyverman[1], Pieter Vanormelingen[1], Marnik Vuylsteke[2,3]*

1 Laboratory of Protistology and Aquatic Ecology, Department of Biology, Ghent University, Gent, Belgium, **2** VIB Department of Plant Systems Biology, Gent, Belgium, **3** Department of Plant Biotechnology and Bioinformatics, Ghent University, Gent, Belgium

Abstract

The pennate diatom *Seminavis robusta*, characterized by an archetypical diatom life cycle including a heterothallic mating system, is emerging as a model system for studying the molecular regulation of the diatom cell and life cycle. One of its main advantages compared with other diatom model systems is that sexual crosses can be made routinely, offering unprecedented possibilities for forward genetics. To date, nothing is known about the genetic basis of sex determination in diatoms. Here, we report on the construction of mating type-specific linkage maps for *S. robusta*, and use them to identify a single locus sex determination system in this diatom. We identified 13 mating type plus and 15 mating type minus linkage groups obtained from the analysis of 463 AFLP markers segregating in a full-sib family, covering 963.7 and 972.2 cM, respectively. Five linkage group pairs could be identified as putative homologues. The mating type phenotype mapped as a monogenic trait, disclosing the mating type plus as the heterogametic sex. This study provides the first evidence for a genetic sex determining mechanism in a diatom.

Citation: Vanstechelman I, Sabbe K, Vyverman W, Vanormelingen P, Vuylsteke M (2013) Linkage Mapping Identifies the Sex Determining Region as a Single Locus in the Pennate Diatom *Seminavis robusta*. PLoS ONE 8(3): e60132. doi:10.1371/journal.pone.0060132

Editor: Tongming Yin, Nanjing Forestry University, China

Received November 23, 2012; **Accepted** February 21, 2013; **Published** March 20, 2013

Copyright: © 2013 Vanstechelman et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: PV is a postdoctoral research fellow with the Research Foundation – Flanders (FWO). The authors acknowledge Belspo for funding the Belgian Co-ordinated Culture Collections of Micro-organisms (BCCM) and Ghent University Research Funding (BOFGOA 01G01911 and BOF BOF09/24J/103). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: marnik.vuylsteke@psb.vib-ugent.be

Introduction

Diatoms (Bacillariophyceae) belong to the Stramenopila, which comprise several microalgal groups dominating primary production in aquatic environments [1]. The diatoms are one of the most diverse and productive groups of algae, with an estimated 200,000 species responsible for almost 20% of global primary production [2]. They are also promising from a biotechnological point of view, and hold great potential for the production of high-value bioproducts such as lipids, pigments and biofuels [3]. The available genomic resources for diatoms have grown rapidly over the past few years [4,5]. In addition, tools for reverse genetics have been developed [6]. However, sexual reproduction, a key feature of most diatom life cycles [7], has never been demonstrated for the most commonly used model diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum* [3]. This prevents the use of forward genetics to link phenotype to genotype, including the use of mutagenesis and QTL mapping [8].

Life cycles, including sexual reproduction, have been studied in detail for only a minute fraction of known diatom species, but these represent most principal diatom lineages. The diatom life cycle comprises two main stages: a prolonged vegetative stage lasting months to years, which is diploid, and a short sexual stage lasting hours to days [3,7,9]. During the vegetative stage of the life cycle, a gradual reduction in cell size takes place, caused by physical constraints imposed by their silica cell wall. This cell wall comprises two parts (or thecae), one of which (the epitheca) is slightly larger than, and overlaps, the other (the hypotheca). During mitosis, each daughter cell inherits one maternal valve (which becomes the new epitheca) and de novo synthesizes a smaller hypotheca. Below a species-specific size threshold (the sexual size threshold, SST) cells become capable of sexual reproduction. Restoration of the maximal cell size generally occurs through sexual reproduction. The zygote matures into an auxospore, which expands to on average two to three times the size of the parental cells. After reaching its maximum size, a new so-called initial cell is formed inside the auxospore envelope, initiating a new round of vegetative multiplication. Because of its crucial role in cell size restoration, sexual reproduction is an obligatory stage in the life cycle of most diatoms.

The raphid pennate diatom *S. robusta* has recently been advocated as a model organism to study diatom biology, and in particular life cycle regulation [3]. Like most pennate diatoms [7,10–12], *S. robusta* has a heterothallic mating system with two mating types (MT^+ and MT^-) [13,14]. After cell pairing, each of the two cells forms two morphologically and behaviorally identical gametes. Subsequent zygote formation and auxospore expansion finally result in two initial cells with a cell length of 64–73 μm. Sexual reproduction is easily induced in cultured strains with a cell size below the SST by adding a suitable mating partner which allows sex to be reliably controlled in mating experiments. However, as in some other pennate diatoms (e.g. *T. tabulata* and *F. delicatissima*) [12,15], intraclonal reproduction (homothally) can

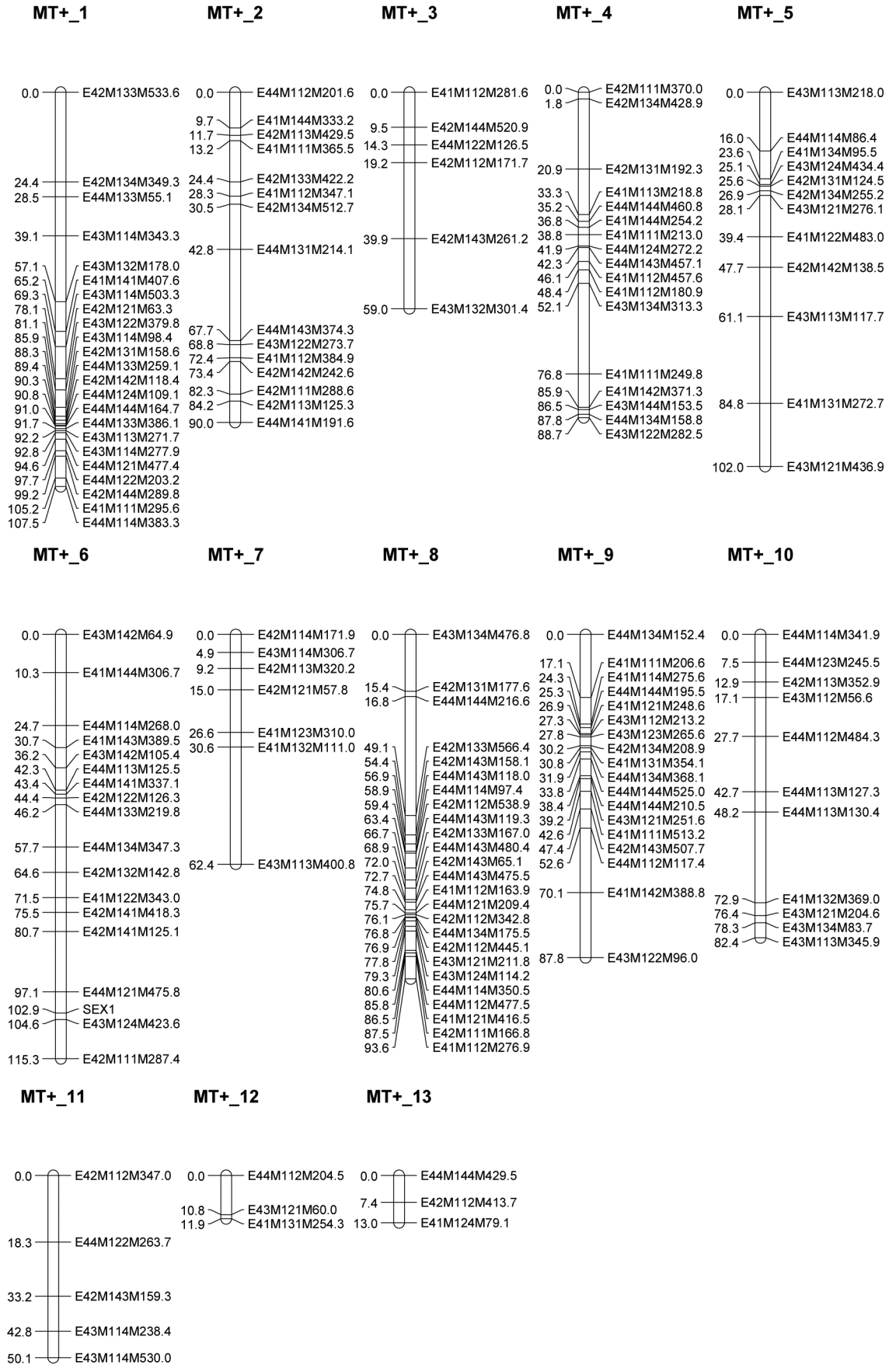
MT+_1
0.0 E42M133M533.6
24.4 E42M134M349.3
28.5 E44M133M55.1
39.1 E43M114M343.3
57.1 E43M132M178.0
65.2 E41M141M407.6
69.3 E43M114M503.3
78.1 E42M121M63.3
81.1 E43M122M379.8
85.9 E43M114M98.4
88.3 E42M131M158.6
89.4 E44M133M259.1
90.3 E42M142M118.4
90.8 E44M124M109.1
91.0 E44M144M164.7
91.7 E44M133M386.1
92.2 E43M113M271.7
92.8 E43M114M277.9
94.6 E44M121M477.4
97.7 E44M122M203.2
99.2 E42M144M289.8
105.2 E41M111M295.6
107.5 E44M114M383.3
MT+_2
0.0 E44M112M201.6
9.7 E41M144M333.2
11.7 E42M113M429.5
13.2 E41M111M365.5
24.4 E42M133M422.2
28.3 E41M112M347.1
30.5 E42M134M512.7
42.8 E44M131M214.1
67.7 E44M143M374.3
68.8 E43M122M273.7
72.4 E41M112M384.9
73.4 E42M142M242.6
82.3 E42M111M288.6
84.2 E42M113M125.3
90.0 E44M141M191.6
MT+_3
0.0 E41M112M281.6
9.5 E42M144M520.9
14.3 E44M122M126.5
19.2 E42M112M171.7
39.9 E42M143M261.2
59.0 E43M132M301.4
MT+_4
0.0 E42M111M370.0
1.8 E42M134M428.9
20.9 E42M131M192.3
33.3 E41M113M218.8
35.2 E44M144M460.8
36.8 E41M144M254.2
38.8 E41M111M213.0
41.9 E44M124M272.2
42.3 E44M143M457.1
46.1 E41M112M457.6
48.4 E41M112M180.9
52.1 E43M134M313.3
76.8 E41M111M249.8
85.9 E41M142M371.3
86.5 E43M144M153.5
87.8 E44M134M158.8
88.7 E43M122M282.5
MT+_5
0.0 E43M113M218.0
16.0 E44M114M86.4
23.6 E41M134M95.5
25.1 E43M124M434.4
25.6 E42M131M124.5
26.9 E42M134M255.2
28.1 E43M121M276.1
39.4 E41M122M483.0
47.7 E42M142M138.5
61.1 E43M113M117.7
84.8 E41M131M272.7
102.0 E43M121M436.9
MT+_6
0.0 E43M142M64.9
10.3 E41M144M306.7
24.7 E44M114M268.0
30.7 E41M143M389.5
36.2 E43M142M105.4
42.3 E44M113M125.5
43.4 E44M141M337.1
44.4 E42M122M126.3
46.2 E44M133M219.8
57.7 E44M134M347.3
64.6 E42M132M142.8
71.5 E41M122M343.0
75.5 E42M141M418.3
80.7 E42M141M125.1
97.1 E44M121M475.8
102.9 SEX1
104.6 E43M124M423.6
115.3 E42M111M287.4
MT+_7
0.0 E42M114M171.9
4.9 E43M114M306.7
9.2 E42M113M320.2
15.0 E42M121M57.8
26.6 E41M123M310.0
30.6 E41M132M111.0
62.4 E43M113M400.8
MT+_8
0.0 E43M134M476.8
15.4 E42M131M177.6
16.8 E44M144M216.6
49.1 E42M133M566.4
54.4 E42M143M158.1
56.9 E44M143M118.0
58.9 E44M114M97.4
59.4 E42M112M538.9
63.4 E44M143M119.3
66.7 E42M133M167.0
68.9 E44M143M480.4
72.0 E42M143M65.1
72.7 E44M143M475.5
74.8 E41M112M163.9
75.7 E44M121M209.4
76.1 E42M112M342.8
76.8 E44M134M175.5
76.9 E42M112M445.1
77.8 E43M121M211.8
79.3 E43M124M114.2
80.6 E44M114M350.5
85.8 E44M112M477.5
86.5 E41M121M416.5
87.5 E42M111M166.8
93.6 E41M112M276.9
MT+_9
0.0 E44M134M152.4
17.1 E41M111M206.6
24.3 E41M114M275.6
25.3 E44M144M195.5
26.9 E41M121M248.6
27.3 E43M112M213.2
27.8 E43M123M265.6
30.2 E42M134M208.9
30.8 E41M131M354.1
31.9 E44M134M368.1
33.8 E44M144M525.0
38.4 E44M144M210.5
39.2 E43M121M251.6
42.6 E41M111M513.2
47.4 E42M143M507.7
52.6 E44M112M117.4
70.1 E41M142M388.8
87.8 E43M122M96.0
MT+_10
0.0 E44M114M341.9
7.5 E44M123M245.5
12.9 E42M113M352.9
17.1 E43M112M56.6
27.7 E44M112M484.3
42.7 E44M113M127.3
48.2 E44M113M130.4
72.9 E41M132M369.0
76.4 E43M121M204.6
78.3 E43M134M83.7
82.4 E43M113M345.9
MT+_11
0.0 E42M112M347.0
18.3 E44M122M263.7
33.2 E42M143M159.3
42.8 E43M114M238.4
50.1 E43M114M530.0
MT+_12
0.0 E44M112M204.5
10.8 E43M121M60.0
11.9 E41M131M254.3
MT+_13
0.0 E44M144M429.5
7.4 E42M112M413.7
13.0 E41M124M79.1

Figure 1. MT^+ linkage groups of *S. robusta* containing markers originating from parental strain 96A.
doi:10.1371/journal.pone.0060132.g001

sporadically occur (in *S. robusta* in MT^+). Additional advantages of *S. robusta* are that the cells are reasonably large (SST ~50 μm apical cell length) and can grow on surfaces allowing easy non-intrusive monitoring of cell and life cycle stages using inverted microscopy.

To date, genetic maps exist for only a small number of species outside the opisthokont and Plantae lineages. Within the stramenopile lineage, linkage maps have been reported for the brown alga *Ectocarpus* siliculosus [16] and several oomycete species [17–20], while a preliminary linkage map has been published for the kelp Laminaria japonica [21].

Experimental evidence suggests that mating type determination in heterothallic pennate diatoms is genetic [7,15], but how mating type determination is achieved is unknown. Here, we report on the construction of a sex-specific linkage map for *S. robusta* based on AFLP markers [22,23] and its use to identify the sex determining region in this diatom. We present AFLP markers co-segregating with the mating type phenotype, the genetic structure of the mating type locus and the identification of MT^+ as the heterogametic sex in *S. robusta*.

Materials and Methods

Production of a F_1 mapping population

The experimental strains were selected from a collection of cryopreserved strains of *S. robusta* publicly available in the BCCM/DCG diatom collection (http://bccm.belspo.be/about/dcg.php). A full-sib (FS) family, containing 116 individual F_1 progeny, was produced from a cross between strains H73A (DCG 0123) and 96A (DCG 0128), having the MT^- and MT^+ mating type, respectively. Their average cell size at the time of crossing was 37.1±0.4 μm and 26.3±0.3 μm respectively. This difference in cell size between the parental strains allowed to distinguish hetero- from homothallic F_1 auxospores and initial cells after mating. Cultures were inoculated for dark synchronization [24] by growing them for 3 days at 18°C in a 12:12 h light:dark regime with cool-white fluorescent lamps at approximately 80 μmol photons m^{-2} s^{-1}. At the end of day 3, the dark period was extended for another 18 h, after which cells were mixed by transferring a quarter of the suspended cells of both mating types to a new flask which was then replenished with an equal amount of fresh Guillard's F/2 medium. After another 9 h of darkness, light was switched on again, allowing the mixed cultures to progress synchronously through the different sexual stages. Two days later, 200 F_1 auxospores and initial cells resulting from heterothallic reproduction (derived from two parental gametangia of unequal size) were picked and transferred to 96-well culture plates containing 0.2 ml F/2 medium.

Phenotyping: determining the mating type of the F_1 progeny

After four months of weekly re-inoculating the F_1 cultures in 24-well culture plates (2 ml), cells reached the SST (~50 μm). The mating type of each F_1 progeny was then determined by backcrossing to the parental strains as well as to two other strains (85A and 85B) of known mating type. The mating type of 116 F_1 progeny could be determined unambiguously, yielding a mapping population of 57 MT^+ and 59 MT^- strains. These strains have been deposited and cryopreserved in the BCCM/DCG culture collection.

DNA extraction of *S. robusta* cultures

For DNA extraction, cultures were grown in 100 ml culture flasks and harvested in exponential phase. Most culture medium was removed and cells were scraped from the surface (cell scrapers; MLS) in 10 ml of remaining culture medium and centrifuged in 15 ml falcon tubes for 15 minutes at 1000 g (4K15(SIGMA)). After removal of the supernatant, the cells were transferred to a 2 ml Eppendorf tube and centrifuged (15 minutes; 1000 g). Eppendorf tubes with cell pellets were frozen at −20°C.

For DNA extraction, 0.5 g Zirconia/Silica Beads 0.1 mm diameter (BioSpec Products), 0.5 ml TE buffer (10 mM TrisHCl (pH 7.6), 1 mM EDTA (pH 8.0)) and 0.5 ml buffered phenol were added to each sample. The samples were bead-beaten (3×; 85 s; 30 Hz). Each time, the samples were cooled on ice. The tubes were centrifuged (Centrifuge 5415R(EPPENDORF); 4°C; 5 m; 10,000 rpm). The water phase containing the DNA was then transferred to a new 2 ml Eppendorf tube and 0.5 ml PCI (phenol:chlorophorm:isoamyl alcohol = 25:24:1v/v) was added. The mixture was centrifuged (10,000 rpm; 5 m; 4°C). The water phase was transferred to a new Eppendorf, 0.5 ml PCI was added again and the mixture was centrifuged. The water phase was transferred to a 1.5 ml Eppendorf. 50 μl 3 M NaAc (pH = ±5), 1 ml 96% ethanol (−20°C) and 2 μl glycogen (−20°C) were added. The samples were incubated over night at −20°C. The next day, the Eppendorfs were centrifuged (13,000 rpm; 30 m; 4°C). The liquid phase was removed and the DNA pellet became visible. 1 ml 70% ethanol was added and the sample was centrifuged (13,000 rpm; 5 m; 4°C). Ethanol was removed and the sample was again centrifuged (13,000 rpm; 5 m; 4°C). The remaining ethanol was removed with a 200 μl pipet and the pellet was dried for at least 20 minutes. Preheated (55°C; 50 μl) TE buffer (pH 8.0) was added to the pellets and the samples were incubated (20 m; 55°C). DNA of the mapping population was stored at −20°C.

Segregation and linkage analysis

The 116 offspring were analyzed with 54 *Eco*RI+2/*Mse*I+3 AFLP primer combinations (PC's; Table S1) as described in [23]. Detection of the AFLP fragments was made possible by fluorescent labeling of the *Eco*RI+2 primer in the final selective amplification reaction, and AFLP images were generated using LI-COR automated DNA sequencers. AFLP markers were scored on the basis of relative fragment intensities, using the image analysis software AFLPQuantar*Pro* (http://www.keygene-products.com). AFLP markers heterozygous in MT^+ parent (96A) and homozygous in the MT^-, and expected to segregate 1:1 in the F_1 generation, were termed MT^+ specific markers. Markers heterozygous in the MT^- parent (H73A) and homozygous in the MT^+ were termed MT^- specific markers. Markers heterozygous in both parents, and expected to segregate 1:2:1 in the F_1 generation were termed "biparental markers". Whenever feasible, biparental markers were scored co-dominantly (i.e., following a 1:2:1 segregation pattern). They were scored dominantly (i.e., conforming to a 3:1 segregation pattern) when heterozygosity could not reliably be discriminated from homozygosity for the present AFLP marker allele. Each AFLP marker was identified by a specific code referring to the corresponding PC and the estimated molecular size of the fragment in nucleotides as estimated by AFLP-Quantar*Pro*. Linkage analysis and segregation distortion tests were performed using the software package JoinMap 4.0 [25]. The

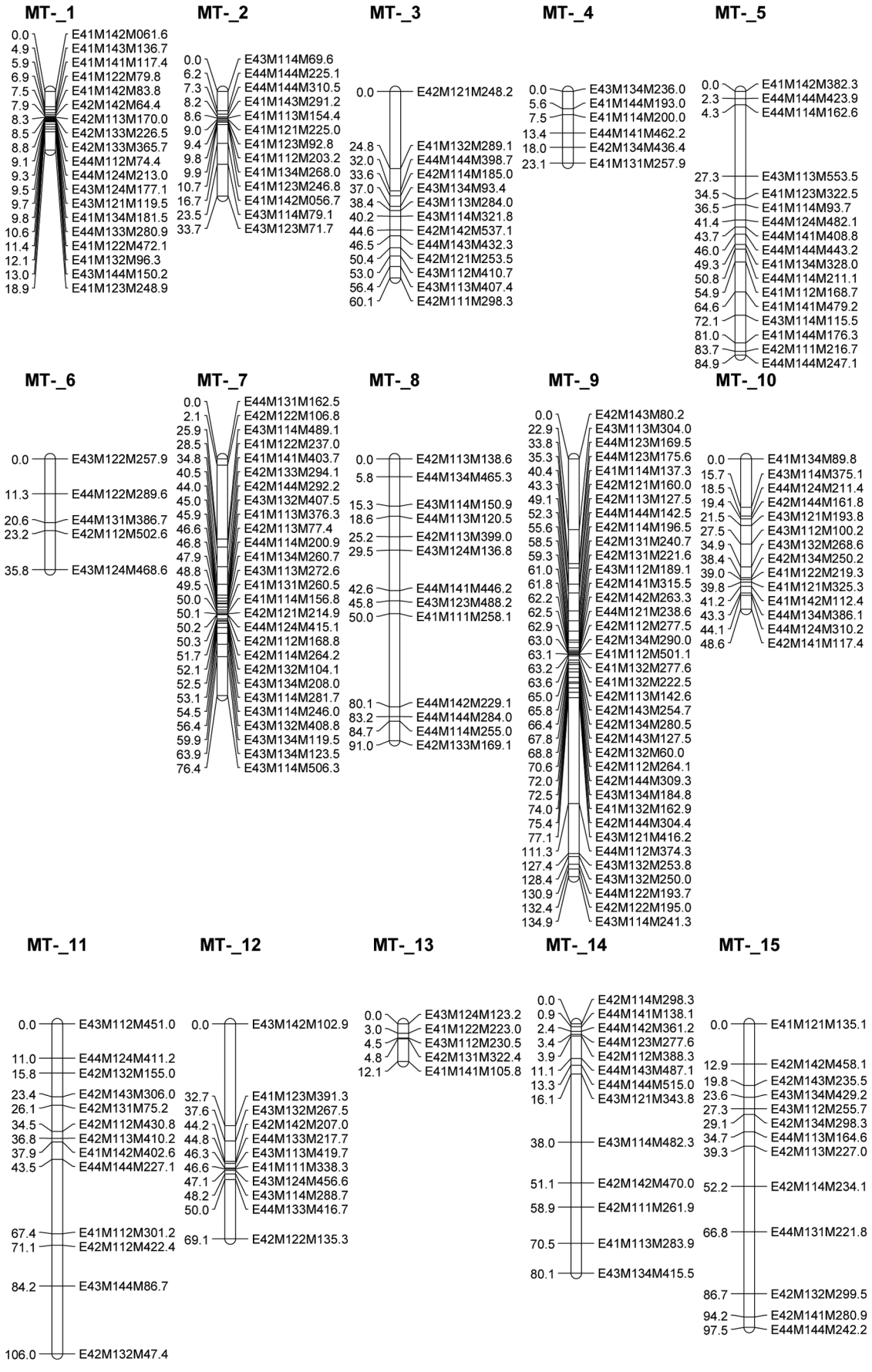
MT-_1
0.0 E41M142M061.6
4.9 E41M143M136.7
5.9 E41M141M117.4
6.9 E41M122M79.8
7.5 E41M142M83.8
7.9 E42M142M64.4
8.3 E42M113M170.0
8.5 E42M133M226.5
8.8 E42M133M365.7
9.1 E44M112M74.4
9.3 E44M124M213.0
9.5 E43M124M177.1
9.7 E43M121M119.5
9.8 E41M134M181.5
10.6 E44M133M280.9
11.4 E41M122M472.1
12.1 E41M132M96.3
13.0 E43M144M150.2
18.9 E41M123M248.9
MT-_2
0.0 E43M114M69.6
6.2 E44M144M225.1
7.3 E44M144M310.5
8.2 E41M143M291.2
8.6 E41M113M154.4
9.0 E41M121M225.0
9.4 E41M123M92.8
9.8 E41M112M203.2
9.9 E41M134M268.0
10.7 E41M123M246.8
16.7 E41M142M056.7
23.5 E43M114M79.1
33.7 E43M123M71.7
MT-_3
0.0 E42M121M248.2
24.8 E41M132M289.1
32.0 E44M144M398.7
33.6 E42M114M185.0
37.0 E43M134M93.4
38.4 E43M113M284.0
40.2 E43M114M321.8
44.6 E42M142M537.1
46.5 E44M143M432.3
50.4 E42M121M253.5
53.0 E43M112M410.7
56.4 E43M113M407.4
60.1 E42M111M298.3
MT-_4
0.0 E43M134M236.0
5.6 E41M144M193.0
7.5 E41M114M200.0
13.4 E44M141M462.2
18.0 E42M134M436.4
23.1 E41M131M257.9
MT-_5
0.0 E41M142M382.3
2.3 E44M144M423.9
4.3 E44M114M162.6
27.3 E43M113M553.5
34.5 E41M123M322.5
36.5 E41M114M93.7
41.4 E44M124M482.1
43.7 E44M141M408.8
46.0 E44M144M443.2
49.3 E41M134M328.0
50.8 E44M114M211.1
54.9 E41M112M168.7
64.6 E41M141M479.2
72.1 E43M114M115.5
81.0 E41M144M176.3
83.7 E42M111M216.7
84.9 E44M144M247.1
MT-_6
0.0 E43M122M257.9
11.3 E44M122M289.6
20.6 E44M131M386.7
23.2 E42M112M502.6
35.8 E43M124M468.6
MT-_7
0.0 E44M131M162.5
2.1 E42M122M106.8
25.9 E43M114M489.1
28.5 E41M122M237.0
34.8 E41M141M403.7
40.5 E42M133M294.1
44.0 E42M144M292.2
45.0 E43M132M407.5
45.9 E41M113M376.3
46.6 E42M113M77.4
46.8 E44M114M200.9
47.9 E41M134M260.7
48.8 E43M113M272.6
49.5 E41M131M260.5
50.0 E41M114M156.8
50.1 E42M121M214.9
50.2 E44M124M415.1
50.3 E42M112M168.8
51.7 E42M114M264.2
52.1 E42M132M104.1
52.5 E43M134M208.0
53.1 E43M114M281.7
54.5 E43M114M246.0
56.4 E43M132M408.8
59.9 E43M134M119.5
63.9 E43M134M123.5
76.4 E43M114M506.3
MT-_8
0.0 E42M113M138.6
5.8 E44M134M465.3
15.3 E43M114M150.9
18.6 E44M113M120.5
25.2 E42M113M399.0
29.5 E43M124M136.8
42.6 E44M141M446.2
45.8 E43M123M488.2
50.0 E41M111M258.1
80.1 E44M142M229.1
83.2 E44M144M284.0
84.7 E44M114M255.0
91.0 E42M133M169.1
MT-_9
0.0 E42M143M80.2
22.9 E43M113M304.0
33.8 E44M123M169.5
35.3 E44M123M175.6
40.4 E41M114M137.3
43.3 E42M121M160.0
49.1 E42M113M127.5
52.3 E44M144M142.5
55.6 E42M114M196.5
58.5 E42M131M240.7
59.3 E42M131M221.6
61.0 E43M112M189.1
61.8 E42M141M315.5
62.2 E42M142M263.3
62.5 E44M121M238.6
62.9 E42M112M277.5
63.0 E42M134M290.0
63.1 E41M112M501.1
63.2 E41M132M277.6
63.6 E41M132M222.5
65.0 E42M113M142.6
65.8 E42M143M254.7
66.4 E42M134M280.5
67.8 E42M143M127.5
68.8 E42M132M60.0
70.6 E42M112M264.1
72.0 E42M144M309.3
72.5 E43M134M184.8
74.0 E41M132M162.9
75.4 E42M144M304.4
77.1 E43M121M416.2
111.3 E44M112M374.3
127.4 E43M132M253.8
128.4 E43M132M250.0
130.9 E44M122M193.7
132.4 E42M122M195.0
134.9 E43M114M241.3
MT-_10
0.0 E41M134M89.8
15.7 E43M114M375.1
18.5 E44M124M211.4
19.4 E42M144M161.8
21.5 E43M121M193.8
27.5 E43M112M100.2
34.9 E43M132M268.6
38.4 E42M134M250.2
39.0 E41M122M219.3
39.8 E41M121M325.3
41.2 E41M142M112.4
43.3 E44M134M386.1
44.1 E44M124M310.2
48.6 E42M141M117.4
MT-_11
0.0 E43M112M451.0
11.0 E44M124M411.2
15.8 E42M132M155.0
23.4 E42M143M306.0
26.1 E42M131M75.2
34.5 E42M112M430.8
36.8 E42M113M410.2
37.9 E41M142M402.6
43.5 E44M144M227.1
67.4 E41M112M301.2
71.1 E42M112M422.4
84.2 E43M144M86.7
106.0 E42M132M47.4
MT-_12
0.0 E43M142M102.9
32.7 E41M123M391.3
37.6 E43M132M267.5
44.2 E42M142M207.0
44.8 E44M133M217.7
46.3 E43M113M419.7
46.6 E41M111M338.3
47.1 E43M124M456.6
48.2 E43M114M288.7
50.0 E44M133M416.7
69.1 E42M122M135.3
MT-_13
0.0 E43M124M123.2
3.0 E41M122M223.0
4.5 E43M112M230.5
4.8 E42M131M322.4
12.1 E41M141M105.8
MT-_14
0.0 E42M114M298.3
0.9 E44M141M138.1
2.4 E44M142M361.2
3.4 E44M123M277.6
3.9 E42M112M388.3
11.1 E44M143M487.1
13.3 E44M144M515.0
16.1 E43M121M343.8
38.0 E43M114M482.3
51.1 E42M142M470.0
58.9 E42M111M261.9
70.5 E41M113M283.9
80.1 E43M134M415.5
MT-_15
0.0 E41M121M135.1
12.9 E42M142M458.1
19.8 E42M143M235.5
23.6 E43M134M429.2
27.3 E43M112M255.7
29.1 E42M134M298.3
34.7 E44M113M164.6
39.3 E42M113M227.0
52.2 E42M114M234.1
66.8 E44M131M221.8
86.7 E42M132M299.5
94.2 E42M141M280.9
97.5 E44M144M242.2

Figure 2. MT$^-$ linkage groups of *S. robusta* containing markers originating from parental strain H73A.
doi:10.1371/journal.pone.0060132.g002

appropriate mapping population type was set to option CP, a population resulting from a cross between two heterogeneously heterozygous and homozygous diploid parents, linkage phases originally unknown. Because for population type CP, the segregation type (SEG) might vary across the loci, a code indicating the segregation type has to be given. The SEG was set to <nn×np> for the MT$^+$ markers, <lm×ll> for the MT$^-$ markers and <hk×hk> for the biparental markers. The two characters left and right of the "×" in these codes correspond to the AFLP marker alleles of the first and second parent, respectively. Each different AFLP marker allele is represented by a different character. We first ran through a fairly wide range of logarithm of the odds (LOD) thresholds, from 2.0 to 14.0, to obtain a proper view of what might be the best grouping. In general, we decided to use the grouping obtained with a LOD score of 6.0. In a few cases, the grouping obtained at a LOD score of 8.0 and 14.0 was used. Only linkage groups containing at least three markers were considered for map construction. Maps were constructed in three rounds, each producing a linkage map. In this map-building procedure, each map was calculated by using the pairwise data of loci present on the map, with default settings (recombination frequency (REC) <0.4; LOD threshold >1). Once the well-fitting markers (causing a change in goodness of fit smaller than the threshold = 5) were positioned on the map (after two rounds), the remaining markers were forced onto the map by setting the jump threshold to zero. When the markers in the third map caused a jump in goodness of fit larger than an arbitrary threshold of 10, the second map was selected as the final map, otherwise the third map. Single markers with a segregation ratio in discordance with the flanking markers (i.e., markers showing strong segregation distortion flanked by a number of non-distorted markers) were discarded, and the map construction was repeated. A marker order was not forced on any linkage group during map construction. Recombination frequencies were converted to Kosambi centiMorgans (cM) prior to the map estimation. Linkage groups were constructed using the MT$^+$ and MT$^-$ markers. Biparental markers were included in both the MT$^+$ and MT$^-$ linkage maps. This way, homologous linkage groups were identified on the presence of identical biparental markers. Editing the linkage groups was done with the software MapChart [26].

Table 1. Statistics for the integrated MT$^+$ and MT$^-$ linkage maps of *S. robusta*.

	MT$^+$	MT$^-$
No. of linkage groups	13	15
No. of markers per linkage group		
Min	3	5
Max	25	37
Median	12	13
Mean	12.5	14.7
Total	163	221
Size of linkage groups (cM)		
Min	11.9	12.1
Max	115.3	134.9
Median	87.8	69.1
Mean	74.1	64.8
Total	963.7	972.2
Intermarker distance (cM)		
Min	0.1	0
Max	32.3	34.2
Median	4.1	2.3
Mean	6.4	4.7

doi:10.1371/journal.pone.0060132.t001

Mapping the MT locus

The mapping of the mating type phenotype was done in two ways: 1) by including the phenotype as a single marker segregating as a MT$^+$ or MT$^-$ specific marker, indicated as SEX1 and SEX2, respectively, and 2) by QTL analysis using mixed models as implemented in the QTL menu in GenStat 14 [27]. Because too few homologous linkage groups are identified, QTL analysis was done for the two mating type specific linkage maps separately. In a preliminary search for QTL, we tested the association of individual marker loci with mating type every 5 cM along the genome, using the commonly known simple interval mapping (SIM) procedure. In a second step, we tested for QTL at particular positions after correcting for QTL elsewhere in the genome, as were identified in the preliminary analysis. This procedure is commonly known as composite interval mapping (CIM). The genome-wide type I error rate was set to $\alpha = 0.05$. The *P* values calculated assume normally distributed errors, when a binomial distribution is more appropriate in the case of mating type. A previous study [28] has shown that applying the mixed model to binary traits is robust and do not result in an excess of low *P* values (i.e. false discoveries) as long as the minor allele frequency of the response variable and the markers is not too low.

Results

Segregation analysis and linkage mapping

A total of 54 *EcoRI*+2/*MseI*+3 AFLP PCs, generating on average 8.6 markers, resulted in a total of 463 AFLP fragments segregating in the 116 F_1 progeny. In total, 162 MT$^+$ and 221 MT$^-$ markers were used for the construction of 13 MT$^+$ and 15 MT$^-$ -specific linkage groups, covering 963.7 cM and 972.2 cM respectively (Figure 1 and 2). Of those marker loci, 28% displayed significantly distorted segregation ratios at the $\alpha = 0.05$ significance level. As segregation distortion is a normal phenomenon in wide crosses, these markers were not a priori excluded, but evaluated after map construction. Although some larger genomic regions did not reveal any markers (e.g. 32.3 cM in the MT$^+$_8 linkage group and 34.2 cM in linkage group MT$^-$_9), the median inter-marker distances were relatively low (4.1 and 2.3 cM for the MT$^+$ and the MT$^-$ linkage maps respectively) (Table 1).

Sixty-four biparental markers were mapped to the MT$^+$ and MT$^-$ map separately and 20 of those markers showed cosegregation with MT$^+$ and MT$^-$ specific linkage groups. Five homologous maps were identified based on the presence of one or more identical biparental markers (Figure 3). No significant difference was observed between the intervals of the 20 biparental markers in the MT$^+$ and MT$^-$ linkage groups (paired *t*-test, $t = 0.79$; $P = 0.22$), suggesting that recombination frequencies do not differ much between both mating types in *S. robusta*.

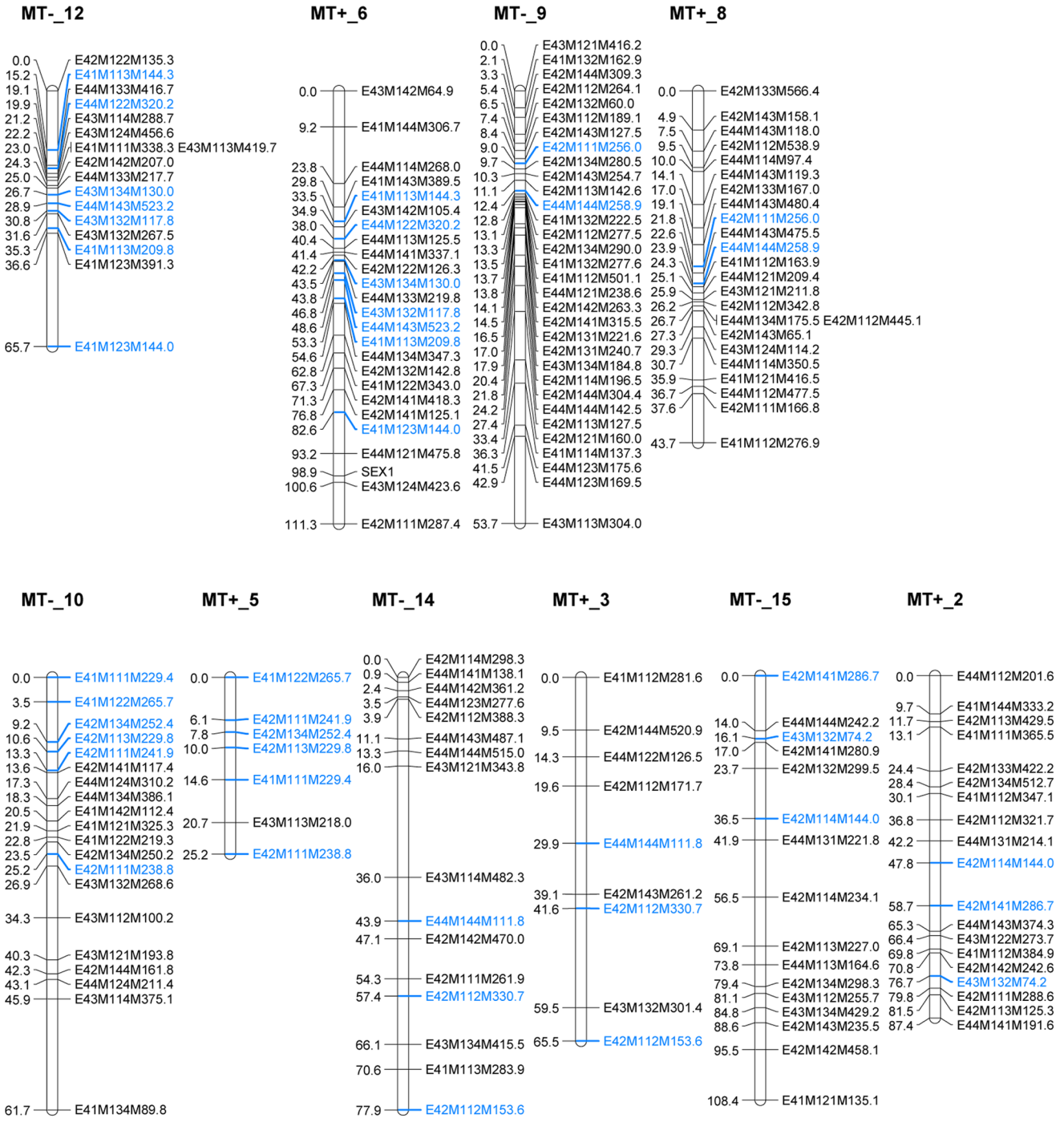

Figure 3. Homologous linkage groups of *S. robusta*. Common biparental markers between the MT^+ and the MT^- linkage groups are shown in blue.
doi:10.1371/journal.pone.0060132.g003

Mapping the mating type locus

The mating type phenotype was mapped as a single marker segregating as either a MT^+ (SEX1) or MT^- specific marker (SEX2). SEX1 could be assigned to the MT^+_6 linkage group, including 18 markers and spanning 115.3 cM. SEX1 is flanked by E43M124_M423.6 at 1.7 cM and E44M121_M475.8 at 5.8 cM distance. In contrast, SEX2 did not cosegregate with any of the markers of the MT^- linkage groups (highest LOD = 2).

A QTL analysis of the mating type phenotype confirmed the monogenic nature of mating type and the identification of MT^+ as the heterogametic sex in *S. robusta* (Figure 4). The genome-wide significance threshold ($P = 0.05$) for detection of QTL co-segregating with mating type was calculated as $-\log_{10}(P) = 3.21$. The mating type phenotype mapped significantly ($-\log_{10}(P) = 198.29$; E43M124M423.6) to a single locus located on the MT^+_6 linkage group (Figure 4a). In contrast, no significant

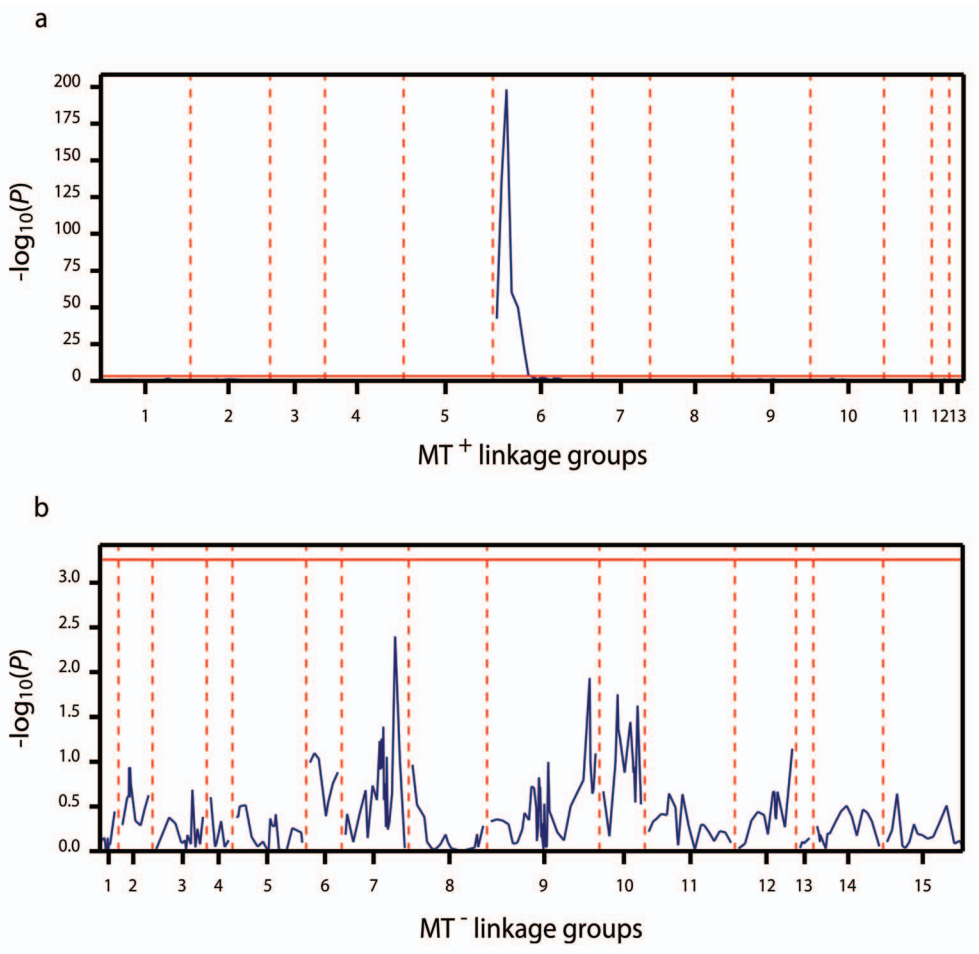

Figure 4. QTL mapping of the mating type phenotype in 116 *S. robusta* F_1 progeny. QTL analysis was done for the two mating type specific linkage maps separately. Linkage scores (plotted as $-\log_{10}(P)$) for MT^+ (a) and MT^- (b) markers are shown according to genome position. The linkage analysis indicates that a single locus on the MT^+_6 linkage group determines the mating type in *S. robusta*.
doi:10.1371/journal.pone.0060132.g004

association was identified between the mating phenotype and MT^- specific markers. (Figure 4b).

The two homologous linkage groups MT^-_12 and MT^+_6 (Figure 3) contain six biparental markers widely spread across the two linkage groups. These identify a relatively large region of recombination between the two sex homologous chromosomes.

Discussion

In this study we exploit the high multiplex ratio of AFLP technology to construct the first linkage maps for a diatom species. We applied these maps to demonstrate that sex determination in the heterothallic pennate species *S. robusta* is genetic, and identify the sex determining region as a single locus.

Segregation and linkage analysis of the 463 AFLP markers scored for 116 individuals of an F_1 mapping population resulted in 13 MT^+ and 15 MT^- specific linkage groups. The use of biparental markers, segregating in a 1:2:1 mode and scored co-dominantly, allowed the detection of five putative homologous linkage groups, including those carrying the mating type locus.

The mating type phenotype cosegregates with markers of a MT^+ specific linkage group, identifying MT^+ as the heterogametic sex and MT^- as the homogametic sex.

Unlike 'classical' sex chromosomes (like the XY chromosomes of mammals and WZ chromosomes of birds), which have only a small recombining pseudo-autosomal region (PAR) or do not recombine at all (e.g. the *Drosophila* Y chromosome), the homologous sex linkage groups in *S. robusta* appear to have a relatively large region of recombination. It is therefore likely that its X and Y chromosomes are cytologically indistinguishable (or non-heteromorphic, cf. Bergero and Charlesworth [29]). Sex chromosomes with extensive recombining PAR regions (in which both homologue chromosomes carry the same gene content) and small non-recombining regions, have been hypothesized to represent recently evolved sex chromosome systems. However, few estimates are available for the age of non-recombining regions (e.g. <2 MY ago in the papaya plant *Carica papaya*), and some data suggest that such regions may also be much older (e.g. in some bird and snake species) [29]. Studies on divergence times in the non-recombining regions of the sex chromosomes in *S. robusta* and other diatoms can contribute to our understanding of the evolution of sex chromosomes, as the diatoms, having an extensive fossil record, have a relatively well time-calibrated evolutionary record. Heterothally is to date known only in the pennate diatoms, which in the fossil record appeared in the Late Cretaceous (about 75 MY ago) [30]. All studied centric diatoms, from whom the pennate diatoms evolved, appear to be homothallic (i.e. they have no mating type

differentiation). In some cases, clones are predominantly male or female, but homothally is still present [7]. While more studies on mating systems in centric diatoms are still needed to confirm that homothally is the rule in centric diatoms, all available evidence so far suggests that in diatoms the evolution of heterogametic sex determination associated with mating type differentiation is a relatively recent event, coinciding with the transition from homothally to heterothally.

The first diatom linkage maps presented in this study will constitute an important resource for future genetic analyses in *S. robusta*. Linkage maps provide important insights into genome organization and can be used for genetic studies of traits of interest [31]. A particular advantage is that each individual progeny of the F_1 mapping population can be clonally propagated and they are maintained as cryopreserved strains. A panel of immortalized F_1 individuals has a number of advantages for genetic mapping identical to those of Recombinant Inbred Lines (RILs), often used in plant or rodent genetics: one needs to genotype each progeny only once and can phenotype multiple individuals from each clonal culture to reduce individual, environmental and measurement variability for multiple traits.

The linkage maps will also provide important information about the *S. robusta* genome sequence which is under construction (A. Bones & T. Brembu, pers. comm.). The completion of the genome sequence will also be the opportunity to further progress on the linkage mapping, as it will represent a source of single nucleotide polymorphisms (SNPs) and insertion/deletion (INDEL) markers for mapping. This will provide a framework to solve the position and order of scaffolds during assembly [31]. This information can be used to construct pseudo-chromosomes by concatenating adjacent supercontigs, and to carry out broad analyses of genome composition.

Despite the rapidly growing amount of diatom genomic information [32], almost nothing is known about the regulation of the unique diatom life cycle. This is changing fast with the introduction of new model diatoms in which the life cycle and sexual reproduction can be reliably manipulated experimentally, including *S. robusta* [3,14]. Further characterization of the mating type locus will prove crucial for our understanding of regulation of the diatom life cycle. The identification of the *S. robusta* genomic region carrying the mating type locus in this study provides a starting point for further fine-mapping of the locus and the identification of the gene(s) and sequence polymorphism(s) underlying mating type dimorphism in *S. robusta*. In turn, this will pave the road for understanding the mechanisms underlying mating system switches to alternative reproductive modes (homothally, paedogamy, and apomixis), which are regularly observed among closely related pennate diatoms [12,15,33], and more broadly, the evolution of the MT locus in diatoms following the evolution of the pennate lineage from a homothallic centric ancestor.

Supporting Information

Table S1 List of primer combinations used for AFLP analysis; E: *Eco*RI primer with two selective bases; M: *Mse*I primer with three selective bases, selective bases: 1,2,3,4 corresponds to A, C, G, T.
(DOCX)

Table S2 AFLP markers underlying the single QTL detected for the MT phenotype.
(DOCX)

Author Contributions

Conceived and designed the experiments: IV KS WV PV MV. Performed the experiments: IV PV. Analyzed the data: IV MV. Contributed reagents/materials/analysis tools: PV MV. Wrote the paper: IV KS WV PV MV.

References

1. Granum E, Raven JA, Leegood RC (2005) How do marine diatoms fix 10 billion tonnes of inorganic carbon per year? Canadian Journal of Botany-Revue Canadienne De Botanique 83: 898–908.
2. Mann DG (1999) The species concept in diatoms. Phycologia 38: 437–495.
3. Chepurnov VA, Mann DG, von Dassow P, Vanormelingen P, Gillard J, et al. (2008) In search of new tractable diatoms for experimental biology. BioEssays 30: 692–702.
4. Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, et al. (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. Science 306: 79–86.
5. Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, et al. (2008) The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. Nature 456: 239–244.
6. De Riso V, Raniello R, Maumus F, Rogato A, Bowler C, et al. (2009) Gene silencing in the marine diatom *Phaeodactylum tricornutum*. Nucleic Acids Research 37: e96.
7. Chepurnov VA, Mann DG, Sabbe K, Vyverman W (2004) Experimental studies on sexual reproduction in diatoms. In: Jeon KW, editor. International Review of Cytology. San Diego: Elsevier Academic Press. 91–154.
8. Tierney MB, Lamour KH (2005) An introduction to reverse genetic tools for investigating gene function. The Plant Health Instructor.
9. Lewis WM Jr (1984) The diatom sex clock and its evolutionary significance. American Naturalist 123: 73–80.
10. Mann DG, Chepurnov VA, Droop SJM (1999) Sexuality, incompatibility, size variation, and preferential polyandry in natural populations and clones of *Sellaphora pupula* (Bacillariophyceae). Journal of Phycology 35: 152–170.
11. Chepurnov VA, Mann DG (2004) Auxosporulation of *Licmophora communis* (Bacillariophyta) and a review of mating systems and sexual reproduction in araphid pennate diatoms. Phycological Research 52: 1–12.
12. Roshchin AM (1994) Zhiznennye Tsikly Diatomovykh Vodoroslej. Kiev, Ukraine: Naukova Dumka.
13. Chepurnov VA, Mann DG, Vyverman W, Sabbe K, Danielidis DB (2002) Sexual reproduction, mating system, and protoplast dynamics of *Seminavis* (Bacillariophyceae). Journal of Phycology 38: 1004–1019.
14. Gillard J, Frenkel J, Devos V, Sabbe K, Paul C, et al. (2012) Metabolomics enabled structure elucidation of the first diatom sex pheromone. Angewandte Chemie-International Edition in press.
15. Davidovich NA, Kaczmarska I, Ehrman JM (2010) Heterothallic and homothallic sexual reproduction in *Tabularia fasciculata* (Bacillariophyta). Fottea 10: 251–266.
16. Heesch S, Cho GY, Peters AF, Le Corguillé G, Falentin C, et al. (2010) A sequence-tagged genetic map for the brown alga Ectocarpus siliculosus provides large-scale assembly of the genome sequence. New Phytologist 188: 42–51.
17. Van der Lee T, De Witte I, Drenth A, Alfonso C, Govers F (1997) AFLP linkage map of the oomycete *Phytophthora infestans*. Fungal Genetics and Biology 21: 278–291.
18. van der Lee T, Testa A, Robold A, van 't Klooster J, Govers F (2004) High-density genetic linkage maps of *Phytophthora infestans* reveal trisomic progeny and chromosomal rearrangements. Genetics 167: 1643–1661.
19. May KJ, Whisson SC, Zwart RS, Searle IR, Irwin JAG, et al. (2002) Inheritance and mapping of 11 avirulence genes in *Phytophthora sojae*. Fungal Genetics and Biology 37: 1–12.
20. Sicard D, Legg E, Brown S, Babu NK, Ochoa O, et al. (2003) A genetic map of the lettuce downy mildew pathogen, *Bremia lactucae*, constructed from molecular markers and avirulence genes. Fungal Genetics and Biology 39: 16–30.
21. Yang G, Sun Y, Shi Y, Zhang L, Guo S, et al. (2009) Construction and characterization of a tentative amplified fragment length polymorphism-simple sequence repeat linkage map of *Laminaria* (Laminariales, Phaeophyta). Journal of Phycology 45: 873–878.
22. Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, et al. (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23: 4407–4414.
23. Vuylsteke M, Peleman JD, van Eijk MJT (2007) AFLP technology for DNA fingerprinting. Nature Protocols 2: 1387–1398.
24. Gillard J, Devos V, Huysman MJJ, De Veylder L, D'Hondt S, et al. (2008) Physiological and transcriptomic evidence for a close coupling between chloroplast ontogeny and cell cycle progression in the pennate diatom *Seminavis robusta*. Plant Physiology 148: 1394–1411.
25. Van Ooijen JW (2006) JoinMap® 4, Software for the calculation of genetic linkage maps in experimental populations. Wageningen: Kyazma B.V.

26. Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. Journal of Heredity 93: 77–78.
27. VSN International (2011) GenStat *for Windows* 14th Edition. Hemel Hempstead, UK: VSN International.
28. Cockram J, White J, Zuluaga DL, Smith D, Comadran J, et al. (2010) Genome-wide association mapping to candidate polymorphism resolution in the unsequenced barley genome. Proc Natl Acad Sci U S A 107: 21611–21616.
29. Bergero R, Charlesworth D (2009) The evolution of restricted recombination in sex chromosomes. Trends in Ecology & Evolution 24: 94–102.
30. Sims PA, Mann DG, Medlin LK (2006) Evolution of the diatoms: insights from fossil, biological and molecular data. Phycologia 45: 361–402.
31. Foulongne-Oriol M (2012) Genetic linkage mapping in fungi: current state, applications, and future trends. Applied Microbiology and Biotechnology 95: 891–904.
32. Lommer M, Specht M, Roy AS, Kraemer L, Andreson R, et al. (2012) Genome and low-iron response of an oceanic diatom adapted to chronic iron limitation. Genome Biol 13: R66.
33. Vanormelingen P, Chepurnov VA, Mann DG, Sabbe K, Vyverman W (2008) Genetic divergence and reproductive barriers among morphologically heterogeneous sympatric clones of Eunotia bilunaris sensu lato (Bacillariophyta). Protist 159: 73–90.