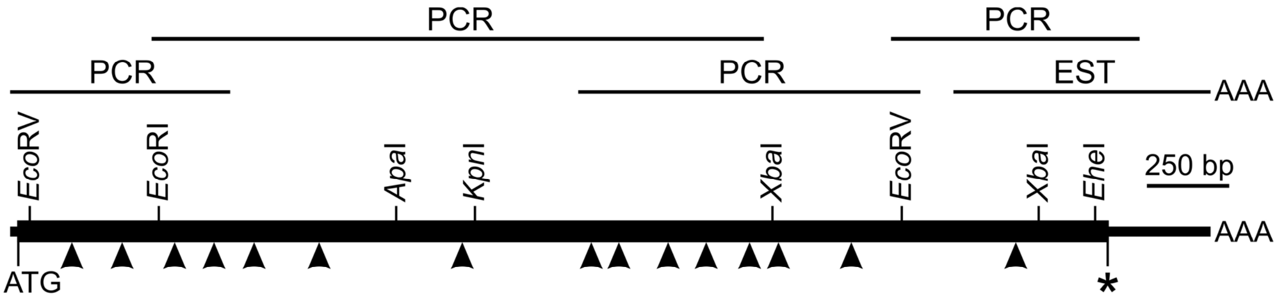
Supplemental Figure 1. Physical Map of the *Volvox RBR1* cDNA (Splice Variant 1) and Overlapping Fragments Used to Assemble the Complete Sequence.

Fragments obtained by RT-PCR or from an EST library are indicated. ATG, translation start site; asterisk (*), translation stop site; thick horizontal bar, open reading frame; thinner horizontal bar, 5'- and 3'-UTRs; AAA, poly A tail; arrowheads, positions of introns.



Supplemental Figure 2. Trimmed Alignment of 24 RB-Related Proteins from Green Algae, Moss, and Higher Plants and Animals.

Alignment of protein sequences was done using the MUltiple Sequence Comparison by Log-Expectation program (MUSCLE) (Edgar, 2004). Minor manual optimization of alignments, trimming, and management of multi-aligned data was done with BioEdit v7.0.9 (Hall, 1999). Source organisms: Vc RBR1, *Volvox carteri*; Ci MAT3, *Chlamydomonas incerta*; Cr MAT3, *Chlamydomonas reinhardtii*; Ot RBR, *Ostreococcus tauri*; Pp RBR, *Physcomitrella patens*; Gg RB1, *Gallus gallus*; Mm RB1, *Mus musculus*; Hs RB1 (RB), *Homo sapiens*; Hs RB2 (p130), *Homo sapiens*; Hs RBL1 (p107), *Homo sapiens*; Dm RBF, *Drosophila melanogaster*; Pt RBR, *Populus tremula* x *Populus tremuloides*; Ee RBR, *Euphorbia esula*; At RBR, *Arabidopsis thaliana*; Nb RBR, *Nicotiana benthamiana*; Nt RBR, *Nicotiana tabacum*; Sb RBR, *Scutellaria baicalensis*; Ps RBR, *Pisum sativum*; Chr RBR, *Chenopodium rubrum*; Zm RBR1, *Zea mays*; Zm RBR2, *Zea mays*; Zm RBR2b, *Zea mays*; Zm RBR3, *Zea mays*; Cn RBR, *Cocos nucifera*. The corresponding GenBank accession numbers are given in the Methods section.

```
                   10        20        30        40        50        60        70        80        90       100
           ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Ot RBR     TPISQAMASASWLQDIVCASSDLETLHKKVDVLIFFFFLNRILRAESQVNIVALLQSSRFTKSLLACCMEVIVATYTLTFPATTHLLGIHPFDLTTIIEP
Pp RBR     TPVSITMTTAKWLRTVIAPLSAEPSMKVRANVILYYKVLGAMCTAESHRHL-PFLPNERFHRCMLACSAELVLASHTMTFPAVLEPAGITAFDLSKVIEG
Zm RBR1    TPVTSAMTTAKWLREVISSLPDKPSVTERVSIVLYYRVLEAICRAELQNNLTPLLSNERFHRCLIACSADLVLATHIMMFPAVLESTGLTAFDLSKIIEN
Zm RBR2b   TPITSAMTTAKWLREVISSLPEKPSVTERVSIVLYYRVLEAICRAELQNNLTPLLSNERFHRCLIACSAELVLATHIMMFPAVLESTGLTSFDLSKIIEN
Zm RBR2    TPITSAMTTAKWLREVISSLPEKPSVTERVSIVLYYRVLEAICRAELQNNLTPLLSNERFHRCLIACSADVVLATHIMMFPAVLESAGLTSFDLSKIIEN
Zm RBR3    TPVSTAMTTAKWLRNTVSPLPSRPSIAHRAGIILYYRVLESMCRAESQINLTSLLSNERFHRCMIACSAELVLATHTMMFPAVLEKTGITAFDLSKVIEG
Chr RBR    TPVSTAMTTAKWLRTYISPLPSKPSVVRRAHIILYYRVLETMCTAEAQLNLTSLLTNERFHRCMLACSAELVLATHTMLFPAVLEKTGITAFDLSKVIES
At RBR     TPVSTAMTTAKWLRTVISPLLPKPSVTRRAHIILYYRVLEAMCKAEAQINLNSLLTNERFHRCMLACSAELVLATHTMLFPAVLERTGITAFDLSKVIES
Nt RBR     TPVSTAMTTARWLRTVIAPLQPKPSVIRRAQIILYYRVLQTMCTAESQINLTSLLTNERFHRCMLACSAELVLATHTMLFPAVLERTGITAFDLSKVIES
Nb RBR     TPVSTAMTTARWLRTVIAPLQPKPSVIRRAQIILYYRVLQTMCTAESQINLTSLLTNERFHRCMLACSAELVLATHTMLFPAVLERTGITAFDLSKVIES
Sb RBR     TPVTTAMTTAKWLRTVIAPLPSKPSVTRRAQIILYFRVLQALCTAESQINLTSLLTNERFHRCMLACSAELVLATHTMLFPTVLERTGITAFDLSKVIES
Ps RBR     TPVSTAMTTAKWLRTVISPLPSKPSVIRRAQIILYYRVLATMCRAEAQINLTSLLTNERFHRCMLACSAELVLATHTMLFPAVLERTGITAFDLSKVIES
Ee RBR     TPVTTAMTTAKWLRTVISPLPSKPSVTRRAQIILSYRVLEAMCTAEAQVNLTSLLTNERFHRCMLSCSAELVLATHTMLFPAVLERTGITAFDLSKVIES
Pt RBR     TPVSTAMTTAKWLRTIISPLPSKPSVIRRAQIILYYRVLESMCTAEAQINLTSLLTNERFHRCMLACSAELVVATYTMLFPAVLERTGITAFDLSKVIES
Cn RBR     TPVSTAMTTAKWLRNIISPLPSRPSVTRRASIILYYRVLETMCRAESHINLTSLLSNDRFHRCMLACSAELVLATHTMMFPAVLEKTGITAFDLSKVIES
Vc RBR1    TPVSEIMSTSAWLRGITASLAAEPSLNQRVQELVYYLSLERILQAEEKNGVTSLLSAGKFHRALVACSAEVVTACYSCTFPKVLDALHIKAFDLAKMIQC
Ci MAT3    TPISEAMGASAWLRGVTSSMAAEPSLMKRVRDLLYCHSLETIMLNEERTAALALLGSAKFQRGLIACCIEVVAACYSCAFPKVVDALRIKAFDMASIIGT
Cr MAT3    TPISEAMGASAWLRGVTSNMVAEPSLMKRVRDLLYCHSLDTILQNEQRTAALALLGSAKFQRGLIACCIEVVAACYSCAFPNVLDALRIKAFDMANIIGT
Gg RB1     TPVRAAMNTIQQLMMILNSATDKPSILKRVES-LYYRVMESMLKSEEERNFSKLLNDNIFHTSLLACALEIVMATYDLSFPWILNVFDLKAFDFYKVIES
Hs RB1     TPVRTVMNTIQQLMMILNSASDQPSILKRVKDILYYRVMESMLKSEEERNFSKLLNDNIFHMSLLACALEVVMATYDLSFPWILNVLNLKAFDFYKVIES
Mm RB1     TPVRTVMNTIQQLMVILNSASDQPSILKRVKDVLYYRVMESMLKSEEERNFSKLLNDNIFHMSLLACALEVVMATYDLSFPWILNVLNLKAFDFYKVIES
Dm RBF     EPVRNATNNVKQLSAF--GRITEPTLLSIIEEIFFFYLLDQILQAEIRNDLKRLLVQKIFNITLMACCVELVLEAYELKFPWVLDCFSISAFEFQKIIEI
Hs RBL1    TPVASATQSVSRLQSIVAGLKNAPSIMKILKGILYYKILETVMVQETRRDMSVLLEQDIFHRSLMACCLEIVLFAYPRTFPWIIEVLNLQPFYFYKVIEV
Hs RB2     TPVSTATHSLSRLHTMLTGLRNAPSIANRLKEMLYYKVLESVIEQEQKRDLSGILEQDAFHRSLLACCLEVVTFSYPGNFPFITEIFDVPLYHFYKVIEV
```

```
                   110       120       130       140       150       160       170       180       190       200
            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Ot RBR      FVRADMDLPREIVRHFNSLEEKTLERLAWCKGSALFDFLVLQIFFAKVMQLAARRLGDLASRLTRDVYALIEHVIYEQTNLVYNRHVDQIILASVYGVCK
Pp RBR      FVRHEETLPRELKRHLNSIEERLLESMAWEKGSSMFNCLPINVFFQKVLKLAAVRIRNLCERLVERVNRALHHALNHETGLFFNRHIDQIILCCVYGICK
Zm RBR1     FVRHEETLPRELKRHLNSLEEQLLESMAWEKGSSLYNSLPIHIFFSKILKLAAIRIRNLCERVTERVYNVFKQILEQQTTLFFNRHIDQLILCCLYGVAK
Zm RBR2b    FVRHEETLPRELKRHLNSLEEQILESMAWEKGSSLYNSLPIQIFFSKILKLAAIRIRNLCERITERVYNVFKQILDQQTTLFFNRHMHQLILCCLYGVAK
Zm RBR2     FVRHEETLPRELKRHLNSLEEQILESMAWEKGSSLYNSLPIQIFFSKILKLAAIRIRNLCERITERVYNVFKQILDQQTTLFFNRHIDQLILCCLYGVAK
Zm RBR3     FVRHEDTLPRELKRHLNSLEERLLESMAWEKGSSMYNSLPIGVFFSKISKLAAIRIRSLCERLLERVYSLVQQILSQQTGLFFNRHIDQIILCSIYGVAK
Chr RBR     FIRHEDSLPRELRRHLNSLEERLLESMAWEKGSSMDNSLPINIFFSKIVKLAAVRINGMIERLRENVYCRFHIILNQRTSLFFNRHIDQIILCAFYGVAK
At RBR      FIRHEDSLPRELRRHLNSLEERLLESMVWEKGSSMYNSLPINIFFTKINKLAAVRINGMVERLRESVYCFFQHVLAQRTSLLFSRHIDQIILCCFYGVAK
Nt RBR      FIRHEESLPRELRRHLNSLEERLLESMVWEKGSSMYNSLPINVFFGKIVKLAAVRINGMIERLRETVYCLFQKILSQRTSLFFNRHIDQIILCSFYGVAK
Nb RBR      FIRHEESLPRELRRHLNSLEERLLESMVWEKGSSMYNSLPINVFFGKIVKLAAVRINGMIERLRETVYCLFQKILSQRTSLFFNRHIDQIILCSFYGVAK
Sb RBR      FIRHEESLPRELRRHLNSLEERLLESMVWEKGSSMYNSLPVNVFFSKIVKLAAVRINGMVERLRESVYCLFQKILSQRTTLFFNRHIDQILLCCFYGVSK
Ps RBR      FIRYEESLPRELRRHLNSLEERLLESLVWEKGSSMYNSLPISVFFSKIVKLGAVRISGMVERLRENVYCLFQRILNQWTSLFFNRHIDQIILCCFYGVAK
Ee RBR      FIRHEESLPRELRRHLNSLEERLLESMVWEKGSSMYNSLPISIFFCKINKLAAVRINGMIERLRENVYRLFQQILSHRTSLFFNRHIDQIILCCFYGVAK
Pt RBR      FIRHEESLPRELRRHLNSLEERLLDSMVWEKGSSLYNSLPINVFFTKINKLAAVRINGMIEKLRENVYRLFQLILSHQTSLFFNRHIDQIILCCFYGVAK
Cn RBR      FVRHEETLPRELKRHLNSLEERLLESMAWERGSSMYNSLPINIFFNKIVKLAAIRIRSLCERLLEWVYCLIQQILSQRTALFFNRHIDQIILCSFYGVAK
Vc RBR1     FVRSITTLPRDLKRHLFLVEEKILESLAWEADSPLYSLIKLNHFCCKVLKLAAFRLAVLCDNFNTKVYQTIEHALFFQTHLFYNRHIDQIILSALYGFCK
Ci MAT3     FVKSIATLPRELKRHLFLIEEKILECLAWQPGSSLYHLIPLFEFLRKVLKLTSFRLALLCENVSSKVYETIEHALYKQTHLFYNRHIDQVMLSTLYGYCK
Cr MAT3     FVKSIATLPRELKRHLFLIEEKILECLAWEPGSSLYHLIPLFEFLRKVLKLTSFRLALLCENFNSKVYEAIEHALYKQTHLFYNRHIDQIMLSTLYGYCK
Gg RB1      FIKVEPSLTRDMIKHLERCEHRIMESLAWQSDSPLFDLIQLSLFYKKVFRLAYLRLHTLFFRLEPLIWTLFQHTLQNESELMRDRHLDQIMMCSMYGICK
Hs RB1      FIKAEGNLTREMIKHLERCEHRIMESLAWLSDSPLFDLIQLSLFYKKVYRLAYLRLNTLCERLEHIIWTLFQHTLQNEYELMRDRHLDQIMMCSMYGICK
Mm RB1      FIKVEANLTREMIKHLERCEHRIMESLAWLSDSPLFDLIQLALFYKKVYRLAYLRLNTLCARLEHIIWTLFQHTLQNEYELMRDRHLDQIMMCSMYGICK
Dm RBF      VVRHEGCLNRSLIKHLNSIEETCLERLAWARNSTVWEMIPLQIFLRKVYLLGWLRIQKLCSELPESIWHIFEHSITHETELMKDRHLDQNIMCAIYIYIR
Hs RBL1     VIRSEEGLSRDMVKHLNSIEEQILESLAWSHDSALWEALALALFYRKVYHLASVRLRDLCLKLRRKIWTCFEFTLVHCPDLMKDRHLDQLLLCAFYIMAK
Hs RB2      FIRAEDGLCREVVKHLNQIEEQILDHLAWKPESPLWEKIPLSLFFRKVYHLAAVRLRDLCAKLRKKIWTCFEFSIIQCPELMMDRHLDQLLMCAIYVMAK
```

```
                   210       220
            ....|....|....|....|
Ot RBR      DIISFYNKVFVSRVRTFLLA
Pp RBR      DIIRFYNEVFVPSTKSFLLQ
Zm RBR1     GIITFYNEVFVPAAKPFLVS
Zm RBR2b    DIITFYNEVFVPAAKPFLVS
Zm RBR2     DIITFYNEVFVPAAKPFLVS
Zm RBR3     DIITFYNEVFIPAVKSLLVE
Chr RBR     DIITFYNEVFIPSVKPLLGE
At RBR      DIITFYNEIFIPAVKPLLVE
Nt RBR      DIITFYNEMFIPSVKPLLVE
Nb RBR      DIITFYNEMFIPSVKPLLVE
Sb RBR      DIITFYNEIFVPAVKPLLVE
Ps RBR      DIISFYNEVFIPSVKPLLVE
Ee RBR      DIITFYNEIFIPAAKPLLVE
Pt RBR      DIITFYNEIFIPAAKPLLVD
Cn RBR      DIITFYNEVFIPSVKPLLVE
Vc RBR1     DIIAFYNQVFVPCMKSFLLR
Ci MAT3     DIIAFYNAVFVPAMKSFLLK
Cr MAT3     DIIGFYNAVFVPAMRNFLLK
Gg RB1      SIIVFYNLVFMQKLKTNILQ
Hs RB1      SIIVFYNSVFMQRLKTNILQ
Mm RB1      SIIVFYNSVFMQRLKTNILQ
Dm RBF      DIIHFYNHTYVPLMRQFVID
Hs RBL1     DLIKFYNTIYVGRVKSFALK
Hs RB2      DLIQFYNNIYIKQIKTFAMK
```

Supplemental Figure 3. Nucleotide Sequence Alignment of *Volvox RBR1* and *Chlamydomonas MAT3* Coding Sequences.

BLASTn result (basic local alignment search tool, nucleotide) using the National Center for Biotechnology Information (NCBI) databases. Numbers on both sides of the sequence alignment correspond to nucleotide numbering of *Volvox RBR1* and *Chlamydomonas MAT3* coding sequences.

```
Identities: 637 of 962 nucleotides (66%), gaps: 45 of 962 nucleotides (4%).

RBR1  160   GACTACGCGGGCGAGCAGGATGAAGAGGTCAAACATCTTTGTACACAGGCGAGCCTTCTT  219
            || || |||||||| ||||| || || || || ||    || |  ||||| ||||| ||
MAT3  151   GAGTATGCGGGCGACCAGGACGAGGACGTGAAGCAGGCGTGCAAGCAGGCCAGCCTGCTC  210

RBR1  220   TTCGTGGCGCGCGTACTTTTCGCGGCGCAGGCGCA-----AGGA-GGCAG---AAC--AG  268
            ||||||||||| || || | ||||||||||  | |     |||| |||||   |||  ||
MAT3  211   TTCGTGGCGCGAGTTCTCTCCGCGGCGCAGAAGAAGGAGAAGGATGGCAGCGAAACGCAG  270

RBR1  269   TATCC-AACTTTGG-CTGTCCAATATCACGGATTGTTGCTGCTGCGGGCATTAATCTTCT  326
              | | ||||| || || |||| | || || || || ||||  || ||||| || || ||
MAT3  271   CCTTCGAACTTGGGGCTTTCCATT-TCTCGCATCGTGGCTGTGGCCGGCATCAACCTGCT  329

RBR1  327   TGATTTTTTCCGCGAGGTGAATGTCGTGGTCAGCAAGCTTTCTTCATACTTTGAAGCGCG  386
             || || || ||||||||| | ||||| || ||| |||| ||  | || || ||  ||||
MAT3  330   GGACTTCTTTCGCGAGGTGCACGTCGTCGTGAGCGAGCTGTCGGCTTATTTCGAGTCGCG  389

RBR1  387   TGGGGGTAGTTCGCGATTGTTTGCGCAGCAGGCTCAGCTGAAGGAGAATTCGGAGACTGT  446
             ||  | || ||      |||    |||||||| | |||||||||||| ||||| || ||
MAT3  390   CGGCAGCAGCTCAAAGCAGTTCAGCCAGCAGGCGCTGCTGAAGGAGAACTCGGAAACGGT  449

RBR1  447   TGTAGTCATGGGTTTGCTCGCCAAAAAGTACAAGGACAATTTCGGCATGCTTCTGCATCA  506
             || || |||||  | || ||||| || ||||||||||| |||  ||   | ||||| ||
MAT3  450   GGTCGTGATGGGACTCCTGGCCAAGAAATACAAGGACAACTTCAACACCTTCCTGCACCA  509

RBR1  507   GTTAGACTTCTTTAAGCAGATAGTGCTTAAGCTTGGGTGGAACGCATTCCTTGTGCTACG  566
            | | |||||||  |||||| | || ||    || || |||| ||| || || ||||| ||
MAT3  510   GCTGGACTTCTACAAGCAGGTCGTCCTGCGCCTGGGCTGGAGCGCTTTTCTCGTGCTGCG  569

RBR1  567   GATCAAGCTTTTGAGTTCCTTTCCCGATGTCGTCAGTTGCGTTGAACTTCTACCGTGCGT  626
              |||||||  ||||  | |||||||| || || || || || || || || ||| |
MAT3  570   AGTCAAGCTCCTGAGCGCGTTTCCCGACGTTGTGAGCTGTGTGGAGCTGCTGCCCTGCAT  629

RBR1  627   TTTCGCCATCCTCGTTAGTCACGCCCCGCAGCTTCCAGACTGCCTTAGCCATGTCAGTGG  686
             ||||||||||| |  || || ||||  ||||| |||||||| |  ||  |||   |
MAT3  630   CTTCGCCATCCTTGCCAGCCATGCGCCGCGTCTTCCGGACTGCCTCACGCACATCACGCG  689

RBR1  687   GGCAGAATACCGTTCAGCACTTCTC----AAATCACTTTCAGAGATATGCAAGGCAGACT  742
            |   ||| ||||        |||||    | | || |  | || |  |||||||| ||||
MAT3  690   G---GAAGACCGGGGGAAGTTTCTCCTGAAGAGCA-TGGCGGACACGTGCAAGGCGGACT  745

RBR1  743   ACAGCAGGGTGCAGGCGCGTATTCCCAGTGTCGAAGCTCTGCTCTCTCATGTCCTGACGT  802
            || |  | ||||||||||| || || || || || || || ||| | || ||  | ||
MAT3  746   ACGGGCGCGTGCAGGCGCGGATGCCAAGCGTGGAGGCGCTACTCACGCAAGTGTTCACTA  805

RBR1  803   CTGCTGTTCCAGAATGGCGGGAGGCGGCGTCCCATGCCCAATTGGAATCAAACTTAGCAA  862
             ||| || || || |||||    ||||   |  | ||       ||         ||
MAT3  806   GTGCCGTGCCGGAGTGGCGCACAGCGGTAGCGGAGGC-------GA---------AGGCG  849

RBR1  863   AAAGTGCAACCGCA-GCGGAAAAGTGCAC-ATTTCCG--GGGGTTGTGGACCTTGTATCT  918
            || | ||   || | |||     | || |   | |||  || |  |||||||| ||  ||
MAT3  850   AAGGCGCTGTCGGACGCG---TCGGGCCCTGCTGCCGTTGGAGGCGTGGACCTGGTTGCT  906

RBR1  919   AGTCGTTTGCTGGAGGGCCTTGTCACAGATGGCAATCGCATGCAGCGCGCACTTTCTGCG  978
               |   ||||||||||||| || || ||     | |||||| | ||||   |  | ||
MAT3  907   TCGCCGGTGCTGGAGGGCCTGGTGACGGACACGGACCGCATGAACCGCGTCGTGGCCGCT  966

RBR1  979   TTGGAGACGGAATACGCGCAGCATTACACACGCGGCGGCAGCGAACTCGATGAAAGAGAG  1038
             |||||   || || | |||||| ||| | ||||||| ||  || || || ||  | |||
MAT3  967   CTGGAGCGCGAGTATGAGCAGCACTACTCGCGCGGCGCCACGGAGCTGGACGAGCGCGAG  1026

RBR1  1039  TTCTTGTTCACAGACTTCACAAAATTTGCATCGCCCCGTTTTTCACCACGGCACATGCAG  1098
            ||| ||| ||| |||||||| || ||||| ||||| || || || ||  ||||||||||
MAT3  1027  TTCCTGTCCACGGACTTCACCAAGTTTGCGTCGCCGCGCTTCTCGCCGGGGCACATGCAC  1086

RBR1  1099  TC  1100
            ||
MAT3  1087  TC  1088
```

```
Identities: 251 of 365 nucleotides (68%), gaps: 6 of 365 nucleotides (1%).

RBR1    1558   GAGCGGCGTATTGAGGTCACTAAGGTATATTACCTTAGCCTGGAGCGCATCTTGCAGGCC  1617
               |||||||||||||| | ||| ||| | || | ||  ||||||||   |||| |||||  |
MAT3    1516   GAGCGGCGTATTGAAGCCACCAAGCTGTACTGCCACAGCCTGGACACCATCCTGCAGAAC  1575

RBR1    1618   GAGGAGAAGAACAACGGTGTGGCAGGTGTTACT-TCCCTCCTTTCAGCTGGCAAGTTCCA  1676
               ||  ||  | ||||  | |  || || |   ||  | || |||    | | |||||||||
MAT3    1576   GAACAG-CGCACAAACGGGCTGCCGGCGGCGCTGGCGCTGCTTGGCTCGGCCAAGTTCCA  1634

RBR1    1677   TCGCGCCTTAGTTGCGTGTTCTGC--CGAGGTGGTGACTGCATGCTATCGCATGGTTTCT  1734
                |||| | |  |||| ||  ||||  |||||| ||| | || ||||| |||||||| ||
MAT3    1635   GCGCGGCCTCATTGCCTG--CTGCATCGAGGTCGTGGCGGCCTGCTACCGCATGGTCTCC  1692

RBR1    1735   TGTACGTTCCCAAAAGTGTTGGATGCATTACATATCAAGGCGTTTGACTTGGCCAAGATG  1794
               ||  | || || || ||| |||| ||  | |  |||||||| || ||| ||||||| ||
MAT3    1693   TGCGCCTTTCCCAATGTGCTGGACGCGCTGCGCATCAAGGCCTTCGACATGGCCAACATC  1752

RBR1    1795   ATACAATGCTTCGTCAGGAGTATTACAACGTTGCCCCGGGATCTCAAACGGCACCTTTTC  1854
               ||      |||||||| ||| ||  | ||| |||| || || ||||| || ||| | |||
MAT3    1753   ATCGGCACCTTCGTCAAGAGCATCGCCACGCTGCCGCGTGAGCTCAAGCGCCACTTGTTC  1812

RBR1    1855   CTCGTTGAAGAAAAAATCCTCGAATCCTTGGCGTGGGAGGCGGATTCTCCACTCTACAGC  1914
               ||  | || || ||||| || |   |||| |||||| |||  ||  ||||||||  |
MAT3    1813   CTGATCGAGGAGAAGATCCTGGAGTGTCTGGCTTGGGAGCCGGGCTCGTCACTCTACCAC  1872

RBR1    1915   CTCAT  1919
               || ||
MAT3    1873   CTGAT  1877
```

```
Identities: 248 of 347 nucleotides (71%), gaps: 0 of 347 nucleotides (0%).

RBR1    2230   AAGGTACTGAAGCTTGCCGCCTTCAGGCTTGCGGTCTTATGCGATAACTTTGACTTCAGT  2289
               ||||| ||||||||  |  | ||| | || ||| |  | ||||| ||||| || |||||
MAT3    2461   AAGGTGCTGAAGCTCACGTCGTTCCGCCTGGCGCTGCTGTGCGAGAACTTCGATTTCAGC  2520

RBR1    2290   CCCTTGCAACGTCTAGATGTCAATACAAAGGTGTATCAGACCATCGAGCATGCTCTGTTC  2349
               ||  |  |  | |  || ||||| |  ||||||||  || |||| |||||||| |||| |
MAT3    2521   CCGCTAGAGGGGCCCGAGGTCAACAGCAAGGTGTACGAGGCCATTGAGCATGCGCTGTAC  2580

RBR1    2350   TTCCAGACACACTTGTTCTATAATCGGCACATCGACCAGATAATCCTCTCGGCACTTTAT  2409
                  ||||| ||| ||||||| || || |||||||||||||| || || ||| |||| ||
MAT3    2581   AAGCAGACGCACCTGTTCTACAACCGCCACATCGACCAGATCATGCTGTCGACACTGTAC  2640

RBR1    2410   GGTTTTTGCAAAGTTCACAGGCTGGCCCAGGTTTCATTTCGTGAAATAATCGCACATTAC  2469
               || |  ||||| || |||| |||| | ||||| || || || || || ||||| || |||
MAT3    2641   GGCTACTGCAAGGTGCACAAGCTGTCGCAGGTCTCCTTCCGCGAGATCATCGCGCAGTAC  2700

RBR1    2470   CGCAAGCAGCCTCAAGCACAACCGTCCATATTCCGTTCGGTCGTGGTTGAGCAGTCAAAT  2529
               ||||||||||| || || || | ||| || ||||| ||||||||  | || |||
MAT3    2701   CGCAAGCAGCCGCAGGCGCAGCAGTCGATCTTCCGCTCGGTCGTCATCGACCAGGTGCTG  2760

RBR1    2530   CCAAGCTTGCAGGTGGGGAGGCGTGCCGATATTATTGCGTTCTACAA  2576
               ||||   ||||| |   |  ||| ||||| || || |  ||||||||
MAT3    2761   CCAACGCTGCAGATCCAGTCGCGCGCCGACATCATCGGCTTCTACAA  2807
```

```
Identities: 39 of 48 nucleotides (81%), gaps: 0 of 48 nucleotides (0%).

RBR1    3004   GAGCACCAAATACCAGATGGCTTAGCGGCACTGCTGCAGGCACTGGAC  3051
               |||||  | || ||||| ||||| || ||||||||||||||  |||||
MAT3    3319   GAGCATAAGATTCCAGAGGGCTTGGCCGCACTGCTGCAGGCGTTGGAC  3366
```

Supplemental Figure 4. Amino Acid Sequence Alignment of *Volvox* RBR1 and *Chlamydomonas* MAT3.

BLASTp result (basic local alignment search tool, protein) using the National Center for Biotechnology Information (NCBI) databases. Numbers on both sides of the sequence alignment correspond to amino acid residue numbering of *Volvox* RBR1 and *Chlamydomonas* MAT3. Identical and similar (“+”) amino acid residues are given between the sequences.

```
Identities: 381 of 639 residues (59%), similarities: 471 of 639 residues (73%),
gaps: 25 of 639 residues (3%).

RBR1  17    LLKEVNFGDAIS-QEGFLQLYRELRASYECQENASPASDYAGEQDEEVKHLCTQASLLFV  75
            L+K +  G+  + Q   L+LY EL+ +YE QE A+P  +YAG+QDE+VK  C QASLLFV
MAT3  14    LIKGLTIGNVPALQPDVLELYAELKGAYEAQERAAP-REYAGDQDEDVKQACKQASLLFV  72

RBR1  76    ARVLFAAQAQ---GGRT-VSNFGCPISRIVAAAGINLLDFFREVNVVVSKLSSYFEARGG  131
            ARVL AAQ +   G  T  SN G  ISRIVA AGINLLDFFREV+VVVS+LS+YFE+RG
MAT3  73    ARVLSAAQKKEKDGSETQPSNLGLSISRIVAVAGINLLDFFREVHVVVSELSAYFESRGS  132

RBR1  132   SSRLFAQQAQLKENSETVVVMGLLAKKYKDNFGMLLHQLDFFKQIVLKLGWNAFLVLRIK  191
            SS+ F+QQA LKENSETVVVMGLLAKKYKDNF   LHQLDF+KQ+VL+LGW+AFLVLR+K
MAT3  133   SSKQFSQQALLKENSETVVVMGLLAKKYKDNFNTFLHQLDFYKQVVLRLGWSAFLVLRVK  192

RBR1  192   LLSSFPDVVSCVELLPCVFAILVSHAPQLPDCLSHVSGAEYRSALLKSLSEICKADYSRV  251
            LLS+FPDVVSCVELLPC+FAIL SHAP+LPDCL+H++  +    LLKS+++ CKADY RV
MAT3  193   LLSAFPDVVSCVELLPCIFAILASHAPRLPDCLTHITREDRGKFLLKSMADTCKADYGRV  252

RBR1  252   QARIPSVEALLSHVLTSAVPEWREAASHAQLESNLAKSATAAEKCTFPGVVDLVSSRLLE  311
            QAR+PSVEALL+ V TSAVPEWR A + A+     AK+ + A      G VDLV+S +LE
MAT3  253   QARMPSVEALLTQVFTSAVPEWRTAVAEAK-----AKALSDASGPAAVGGVDLVASPVLE  307

RBR1  312   GLVTDGNRMQRALSALETEYAQHYTRGGSELDEREFLFTDFTKFASPRFSPRHMQSAVMK  371
            GLVTD +RM R ++ALE EY QHY+RG +ELDEREFL TDFTKFASPRFSP HM S + K
MAT3  308   GLVTDTDRMNRVVAALEREYEQHYSRGATELDEREFLSTDFTKFASPRFSPGHMHSTMAK  367

RBR1  372   LRAGPLPIRQGSLLGPGAHTTASYAISPSSMNPKTSLLA--RMHSPLPKLNLGVETGPPV  429
            LR+G +P+R G  LGPG HT      +P+   P+   L    +HSPLP ++L    G P
MAT3  368   LRSGSMPLRSGGPLGPGLHTA-----TPAHPGPQQLHLVPLGLHSPLPMMHLNAGPGVPG  422

RBR1  430   TPVSEIMSTSAWLRGITASLAAEPSPSVMRYLAEVAVPPDGEQGACVTVTLSAAEQLNQR  489
            TP+SE M  SAWLRG+T+++ AEPSP++ R+LA  A+P +        V     +QL +R
MAT3  423   TPISEAMGASAWLRGVTSNMVAEPSPTLQRFLA--ALPINSNSSGPTPV-----QQLMKR  475

RBR1  490   VQELVSCIMPEEKIPSLLGSFPLLQPSLVSERRIEVTKVYYLSLERILQAEEKNNGVAGV  549
            V++LV+ ++P+E  PSLLG FPLL  SL  ERRIE TK+Y  SL+ ILQ E++ NG+
MAT3  476   VRDLVASVIPDEAAPSLLGPFPLLNTSLGVERRIEATKLYCHSLDTILQNEQRTNGLPAA  535

RBR1  550   TSLLSAGKFHRALVACSAEVVTACYRMVSCTFPKVLDALHIKAFDLAKMIQCFVRSITTL  609
             +LL + KF R L+AC  EVV ACYRMVSC FP VLDAL IKAFD+A +I  FV+SI TL
MAT3  536   LALLGSAKFQRGLIACCIEVVAACYRMVSCAFPNVLDALRIKAFDMANIIGTFVKSIATL  595

RBR1  610   PRDLKRHLFLVEEKILESLAWEADSPLYSLICSTVAQDE  648
            PR+LKRHLFL+EEKILE LAWE  S LY LI +   ++E
MAT3  596   PRELKRHLFLIEEKILECLAWEPGSSLYHLIVNVHTENE  634
```

Identities: 116 of 194 residues (59%), similarities: 142 of 194 residues (73%), gaps: 3 of 194 residues (1%).

```
RBR1   695  FSPVKKAR-MEGSQHAVHRYLRKLPPCIGAKFVESAHTDGCAAALNHFCCKVLKLAAFRL  753
             SP KKAR  +G  H+   Y+  LP  +G     +    G A  L  F  KVLKL +FRL
MAT3   772  MSPAKKARGTDGGAHSTQSYIESLPNAVGFP-PAAGAAQGQAGLLFEFLRKVLKLTSFRL  830

RBR1   754  AVLCDNFDFSPLQRLDVNTKVYQTIEHALFFQTHLFYNRHIDQIILSALYGFCKVHRLAQ  813
            A+LC+NFDFSPL+  +VN+KVY+ IEHAL+ QTHLFYNRHIDQI+LS LYG+CKVH+L+Q
MAT3   831  ALLCENFDFSPLEGPEVNSKVYEAIEHALYKQTHLFYNRHIDQIMLSTLYGYCKVHKLSQ  890

RBR1   814  VSFREIIAHYRKQPQAQPSIFRSVVVEQSNPSLQVGRRADIIAFYNQVFVPCMKSFLLRS  873
            VSFREIIA YRKQPQAQ SIFRSVV++Q  P+LQ+  RADII FYN VFVP M++FLL+S
MAT3   891  VSFREIIAQYRKQPQAQQSIFRSVVIDQVLPTLQIQSRADIIGFYNAVFVPAMRNFLLKS  950

RBR1   874  GKGEYPGTIGPTIG  887
             +    G  GP +G
MAT3   951  -ESNGSGASGPGLG  963
```

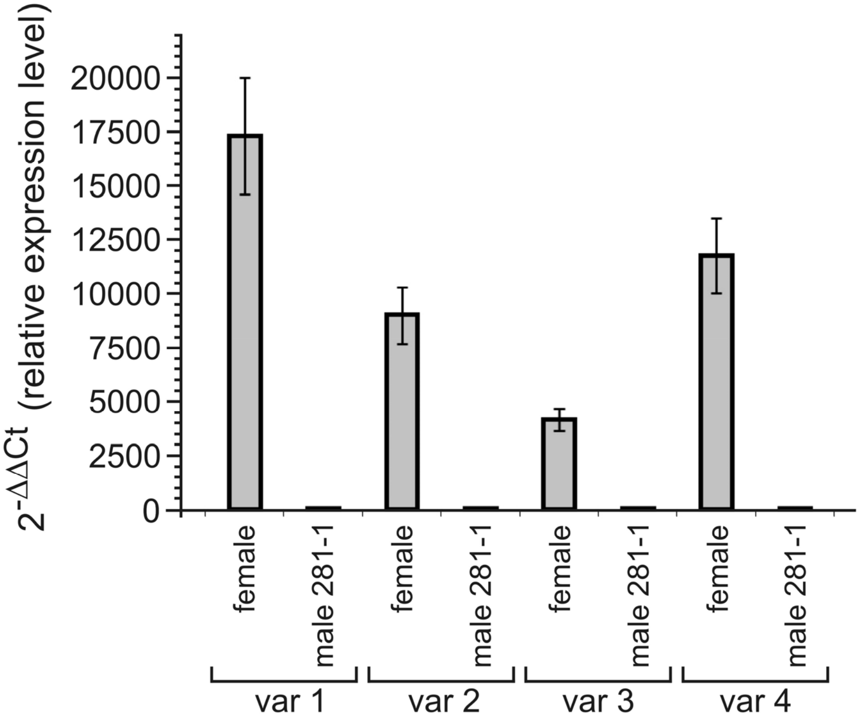
Identities: 41 of 111 residues (36%), similarities: 56 of 111 residues (50%), gaps: 13 of 111 residues (11%).

```
RBR1   1002  EHQIPDGLAALLQALDSQQGGADNCGTESQQYCEVGSRLNAGPGWVLRPVGLDGGPKSCT  1061
             EH+IP+GLAALLQALDSQ+   +N   E ++  EV +        V R           T
MAT3   1107  EHKIPEGLAALLQALDSQKAAEENTAEEEEEEEEVDADEEQPQARVTR--------SGRT  1158

RBR1   1062  GLDQSFNGQAACAGLDVFDDAD--TYDCSQFQKTC---RRQRTPNRKHGSD  1107
               D+S   +      +  DD D  + D    Q+T    RRQRTPNR++G+D
MAT3   1159  ARDRSTGRERRGRHTEARDDMDEASMDAEFSQQTVVTGRRQRTPNRRYGAD  1209
```

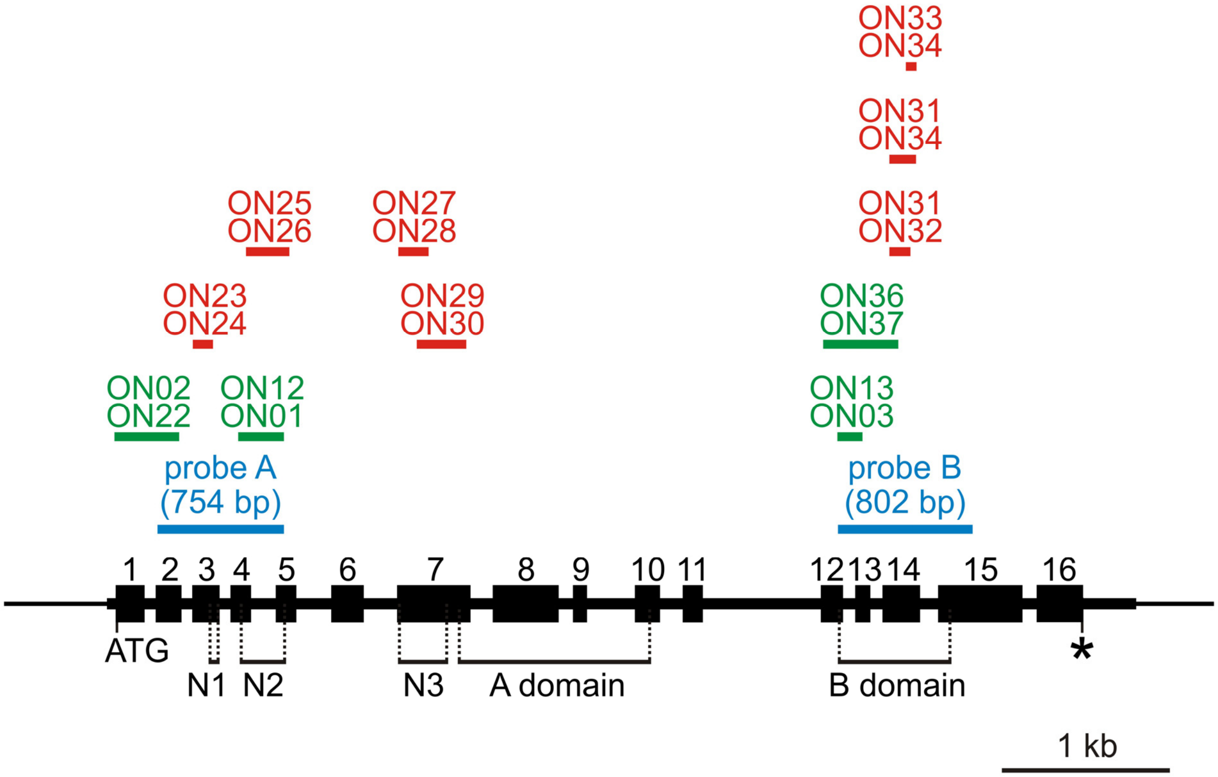
Supplemental Figure 5. Expression Level of *RBR1* Splice Variants in Female *Volvox* Relative to the Background Noise Obtained in Real-Time PCR Experiments with Male *Volvox*.

The expression level (prior to embryogenesis) of the four *RBR1* splice variants (var1 to var4) in female *Volvox* was compared to the background noise obtained in real-time PCR experiments with male 281-1. Error bars refer to the standard deviation (S.D.) of the y-value. All real-time RT-PCR experiments were carried out in triplicate from technical repeats.



Supplemental Figure 6. Localization of PCR Primers and DNA Gel Blot Probes within the *RBR1* Gene Used to Confirm the Absence of the *RBR1* Gene in Males.

The localization of sequences coding for N1-N3 domains as well as A and B domains are indicated within the physical map of exon and intron segments in the *Volvox RBR1* gene (splice variant 1). Standard PCR primers are given in green and degenerate PCR primers are given in red. Horizontal bars indicate the relative position of the primers given above the bars. Probes for DNA gel blot analyses are indicated in blue. The length of each probe is indicated in bp. ATG, translation start site; asterisk (*), translation stop site; thick horizontal bars, exons; numbers above exons, numbering of exons; thinner horizontal bars, introns and 5'- and 3'-UTRs; thinnest horizontal bars, upstream and downstream sequences.



Supplemental Figure 7. Presence of *RBR1* in Male Transformants: PCR Results with *RBR1*-Specific Primers.

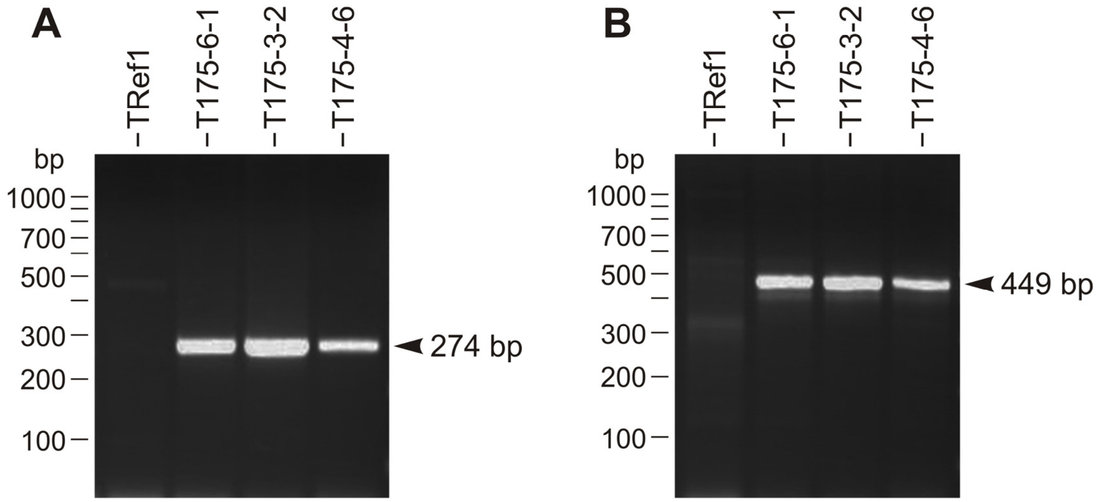
Genomic DNA of the males TRef1, T175-6-1, T175-3-2 and T175-4-6 was analyzed by PCR using *RBR1*-specific primers (sequences presented in Supplemental Table 2 online). TRef1 is a transgenic male descendant of 281-1 that has been transformed only with the *nit*A selectable marker gene, and transformants T175-6-1, T175-3-2 and T175-4-6 are descendants of 281-1 that have been transformed with both the *nit*A gene and the female *RBR1* gene.

The expected sizes of *RBR1* fragments were as follows:

(A) 274 bp (primers ON12 and ON01).

(B) 449 bp (primers ON36 and ON37).

The localization of PCR primers is indicated on the map of the *RBR1* gene in Supplemental Figure 6 online.



Supplemental Table 1. Sequence Comparison of Putative Orthologs Identified Both at the Mating Type Locus of *C. reinhardtii* and on a Small Genomic Fragment of the Female *V. carteri* Genome (Scaffold 43) that Contains the *RBR1* Gene.

Number (As given in the text)	Gene product / Protein	Protein ID *V. carteri*	Protein ID *C. reinhardtii*	Percent identity (Number of identical amino acid residues / total number of residues)	Percent similarity (Number of similar amino acid residues / total number of residues)
1	Spermine synthase	76291	206416	76.36% (239/313)	87.54% (274/313)
2	Protein of unknown function	127490	182394	76.33% (129/169)	83.43% (141/169)
3	Predicted Na^+-dependent cotransporter	82799	112377	78.88% (239/303)	88.45% (268/303)
4	AMP-activated protein kinase	82801	112765	83.40% (216/259)	91.12% (236/259)
5	N-myristoyl transferase	106391	97790	78.06% (306/392)	85.97% (337/392)
6	Splicing factor 3a, subunit 2	106397	196072	73.87% (212/287)	79.09% (227/287)
7	Vacuolar ATP synthase, subunit C	106399	187322	73.99% (276/373)	87.94% (328/373)
8	Cytoplasmic dynein 1b heavy chain	64869	24009	86.61% (3765/4347)	92.96% (4041/4347)
9	Hydroxyproline-rich glycoprotein	118743	195937	34.63% (160/462)	44.59% (206/462)
10	Retinoblastoma-related protein RBR1/MAT3	127148	187248	45.99% (566/1231)	61.1% (752/1231)
11	Predicted sugar kinase	82819	126745	88.37% (304/344)	93.02% (320/344)

Protein IDs are as given in the sequence databases maintained by the DOE Joint Genome Institute. See JGI *C. reinhardtii* genome portal 3.0 (http://genome.jgi-psf.org/Chlre3/Chlre3.home.html) and JGI *V. carteri* genome portal 1.0 (http://genome.jgi-psf.org/Volca1/Volca1.home.html).

Supplemental Table 2. Oligonucleotide Primers.

Primer name	Primer sequence	Comment
ON01	5'-CGTGACTAACGAGGATGGC	*RBR1*, 19-mer
ON02	5'-GTATGTGTTTCGCTCGACGG	*RBR1*, 20-mer
ON03	5'-ATCGCATAAGACCGCAAGCC	*RBR1*, 20-mer
ON04	5'-CGCAGCAGGCTCAGCTGAAG	*RBR1*, 20-mer
ON05	5'-ATGAGCTGCTGACAGTTCGC	*RBR1*, 20-mer
ON06	5'-TGACTGCATGCTATCGCATG	*RBR1*, 20-mer
ON07	5'-TTGGTTACCGTCCTCGGTTC	*RBR1*, 20-mer
ON08	5'-AGCCCTGACAATGTTACTGG	*RBR1*, 20-mer
ON09	5'-CGTCCTTCATATGCTCGAAG	*RBR1*, 20-mer
ON10	5'-GACGAAGCATTGTATCATCTTG	*RBR1*, 22-mer
ON11	5'-TAGAGATTAACCGACCTAACG	*RBR1*, 21-mer
ON12	5'-GCAGATAGTGCTTAAGCTTGG	*RBR1*, 21-mer
ON13	5'-CGGCAGCATTAAACCATTTCTG	*RBR1*, 22-mer
ON14	5'-CGATGGTCTGATACACCTTTGT	*RBR1*, 22-mer
ON15	5'-AACCTGGGCCAGCCTTTGT	*RBR1*, 19-mer
ON16	5'-CTGTGGCGAAGCAGTCAAATC	*RBR1*, 21-mer
ON17	5'-CTGCGCAAACAATCGCGAAC	*RBR1*, 20-mer
ON18	5'-CTGCTGCGGGCATTAATCTTC	*RBR1*, 21-mer
ON19	5'-GATTGACCCGAAAGATAGCAGG	*RBR1*, 22-mer
ON20	5'-TGAGAAGACGTACGAGCTGC	β-actin, 20-mer
ON21	5'-CCTCCATGCCGATTAGGCTA	β-actin, 20-mer
ON22	5'-CAATCCGTGATATTGGACAGC	*RBR1*, 21-mer
ON23	5'-CTKGAYTTYWTYCGCGAGGTG	*RBR1*, 21-mer, 32-fold degeneracy
ON24	5'-TCTCCTKSAGCWGMGCCTGCT	*RBR1*, 21-mer, 16-fold degeneracy

ON25	5'-GCTYYTGAGYKCSTTTCCCGA	*RBR1*, 21-mer, 32-fold degeneracy
ON26	5'-TGSSYRAGGCAGTCYGGAAG	*RBR1*, 20-mer, 32-fold degeneracy
ON27	5'-TGCTGGAGGGCCTKGTSAC	*RBR1*, 19-mer, 4-fold degeneracy
ON28	5'-GASTGCATGTGCCSYGGYG	*RBR1*, 19-mer, 16-fold degeneracy
ON29	5'-GAGTTCYTGTYYACRGACTTCAC	*RBR1*, 23-mer, 16-fold degeneracy
ON30	5'-AGCCAYGCRGACGYGCYCATT	*RBR1*, 21-mer, 16-fold degeneracy
ON31	5'-TTCTAYAAYCGSCACATCGACCA	*RBR1*, 23-mer, 8-fold degeneracy
ON32	5'-GCYTGMGGCTGCTTGCGGTA	*RBR1*, 20-mer, 4-fold degeneracy
ON33	5'-TACCGCAAGCAGCCKCARGC	*RBR1*, 20-mer, 4-fold degeneracy
ON34	5'-TCRAYSACGACCGARCGGAA	*RBR1*, 20-mer, 16-fold degeneracy
ON35	5'-GAGCAGGATGAAGAGGTCAAAC	*RBR1*, 22-mer
ON36	5'- AGCCAGCACGCTGTTCATC	*RBR1*, 19-mer
ON37	5'- CAAAAACCATAAAGTGCCGAGAGG	*RBR1*, 24-mer

Supplemental Table 3. Comparison of a Subset of Sex-Related and Sex-Unrelated Proteins of *C. reinhardtii* with the Corresponding Putative Orthologs of *V. carteri*.

C. reinhardtii protein	Putative orthologous protein in *V. carteri*	Percent identity (Number of identical amino acid residues / total number of residues)	Percent similarity (Number of similar amino acid residues / total number of residues)	Sex-related?
Retinoblastoma-related protein MAT3 (Acc. No. AF375824)	Retinoblastoma-related protein RBR1 (Acc. No. EF123087)	45.99% (566/1231)	61.10% (752/1231)	Yes
Plus mating structure glycoprotein FUS1 (Acc. No. U49864)	--- (no significant putative ortholog)	---	---	Yes
Minus dominance protein MID (Acc. No. U92071)	Putative transcription factor with RWP-RK motif (*V. carteri* protein ID 127314)	34.21% (26/76)	76.31% (58/76)	Yes
Minus agglutinin SAD1 (Acc. No. AY450929)	Putative forkhead transcription factor (*V. carteri* protein ID 94393)	17.45% (493/2826)	25.30% (715/2826)	Yes
Small G-protein YPTC1/RABD1 (Acc. No. U13168)	Small G-protein YptV1 (Acc. No. M93438)	95.10% (193/203)	99.00% (201/203)	No
β1-Tubulin (Acc. No. M10064)	β1-Tubulin (Acc. No. X12855)	99.70% (442/443)	99.70% (442/443)	No
Actin (Acc. No. D50838)	Actin (Acc. No. M33963)	98.10% (370/377)	99.50% (375/377)	No
Nitrate reductase NIT1 (Acc. No. XM_001696645)	Nitrate reductase NitA (Acc. No. X64136)	80.70% (697/864)	95.90% (829/864)	No
Oxygen-evolving enhancer protein 1, PSBO (Acc. No. XM_001694647)	Oxygen-evolving enhancer protein 1, PsbO (Acc. No. AF110780)	88.00% (257/292)	96.90% (283/292)	No
Light harvesting protein LHCA9 (Acc. No. XM_001692496)	Light harvesting protein LhcA (Acc. No. AF110787)	87.80% (187/213)	97.60% (208/213)	No
Bi-ubiquitin UBQ1a (Acc. No. XM_001701327)	Polyubiquitin UbiA (Acc. No. X74214)	100.00% (152/152)	100.00% (152/152)	No