

Gliding motility of the diatom *Craspedostauros australis* coincides with the intracellular movement of raphid-specific myosins

Metin Gabriel Davutoglu¹, Veikko F. Geyer¹, Lukas Niese¹, Johannes Müller², Marcelo Leomil Zoccoler², Valeria Sabatino¹, Robert Haase², Nils Kröger^{1,2,3}, Stefan Diez^{1,2, 4*}, Nicole Poulsen^{1*}

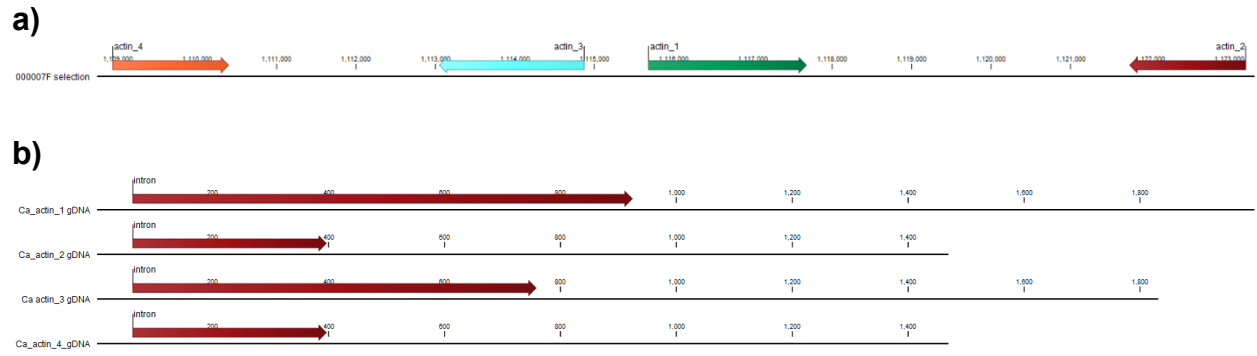
¹ B CUBE - Center for Molecular Bioengineering, TUD Dresden University of Technology, 01307 Dresden, Germany

² Cluster of Excellence Physics of Life, TUD Dresden University of Technology, 01062 Dresden, Germany

³ Faculty of Chemistry and Food Chemistry, TUD Dresden University of Technology, 01062 Dresden, Germany

⁴ Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany

*correspondence to: stefan.diez@tu-dresden.de, nicole.poulsen@tu-dresden.de



Supplementary Figure 1: (a) Genomic arrangement of the four *C. australis* actin genes on the Pacbio Genome assembly contig 7F¹ **(b)** Gene structure of the four *C. australis* actin genes showing the position of the intron.

actin_1	MSDDEDIAALVIDNGSGMCKAGFAGDDAPRSVFP SLIGRARQPGIMVGMEQRDAYVVGDEA	60
actin_2	MSDDEDIAALVIDNGSGMCKAGFAGDDAPRSVFP SLIGRARQPGIMVGMEQRDAYVVGDEA	60
actin_3	MSDDEDIAALVIDNGSGMCKAGFAGDDAPRSVFP SLIGRARQPGIMVGMEQRDAYVVGDEA	60
actin_4	MSDDEDIAALVIDNGSGMCKAGFAGDDAPRSVFP SLIGRARQPGIMVGMEQRDAYVVGDEA	60

actin_1	QAKRGVLTLYKPIEHGIVTNWDDMEKIWHHTFY NELRVAPPAHPVLLTEAPQNPKANRER	120
actin_2	QAKRGVLTLYKPIEHGIVTNWDDMEKIWHHTFY NELRVAPPAHPVLLTEAPQNPKANRER	120
actin_3	QAKRGVLTLYKPIEHGIVTNWDDMEKIWHHTFY NELRVAPPAHPVLLTEAPQNPKANRER	120
actin_4	QAKRGVLTLYKPIEHGIVTNWDDMEKIWHHTFY NELRVAPPAHPVLLTEAPQNPKANRER	120

actin_1	MTQIMFETFNVPAMYVNIQAVLSLYASGRTTGC VLDSDGVDGVTHTVPIYEGYALPHAVIRL	180
actin_2	MTQIMFETFNVPAMYVNIQAVLSLYASGRTTGC VLDSDGVDGVTHTVPIYEGYALPHAVIRL	180
actin_3	MTQIMFETFNVPAMYVNIQAVLSLYASGRTTGC VLDSDGVDGVTHTVPIYEGYALPHAVIRL	180
actin_4	MTQIMFETFNVPAMYVNIQAVLSLYASGRTTGC VLDSDGVDGVTHTVPIYEGYALPHAVIRL	180

actin_1	DLAGRDLTEYLMKILTERGYSLTTAEREIVRDI KESLCFVAVDFEEMKKAESSALEK	240
actin_2	DLAGRDLTEYLMKILTERGYSLTTAEREIVRDI KESLCFVAVDFEEMKKAESSALEK	240
actin_3	DLAGRDLTEYLMKILTERGYSLTTAEREIVRDI KESLCFVAVDFEEMKKAESSALEK	240
actin_4	DLAGRDLTEYLMKILTERGYSLTTAEREIVRDI KESLCFVAVDFEEMKKAESSALEK	240

actin_1	SFELPDGNIIVIGNERFRCPEVLFQPNLTGLEMD GIGDSTFQTIMKCDVDIRKDLIANIV	300
actin_2	SFELPDGNIIVIGNERFRCPEVLFQPNLTGLEMD GIGDSTFQTIMKCDVDIRKDLIANIV	300
actin_3	SFELPDGNIIVIGNERFRCPEVLFQPNLTGLEMD GIGDSTFQTIMKCDVDIRKDLIANIV	300
actin_4	SFELPDGNIIVIGNERFRCPEVLFQPNLTGLEMD GIGDSTFQTIMKCDVDIRKDLIANIV	300

actin_1	LSGGTTFMFPGISERMSKEITALAPASIKVKIV APPERKYSVWIGGSILASLSTFQSMWIS	360
actin_2	LSGGTTFMFPGISERMSKEITALAPASIKVKIV APPERKYSVWIGGSILASLSTFQSMWIS	360
actin_3	LSGGTTFMFPGISERMSKEITALAPASIKVKIV APPERKYSVWIGGSILASLSTFQSMWIS	360
actin_4	LSGGTTFMFPGISERMSKEITALAPASIKVKIV APPERKYSVWIGGSILASLSTFQSMWIS	360

actin_1	KEEYDESGPSIVHRKCF*	377
actin_2	KEEYDESGPSIVHRKCF*	377
actin_3	KEEYDESGPSIVHRKCF*	377
actin_4	KEEYDESGPSIVHRKCF*	377

Supplementary Figure 2: Clustal Omega alignment of the four *C. australis* actin protein sequences showing a single amino acid exchange. Identical amino acid residues are highlighted in grey and indicated with an asterisk (*), a : (colon) indicates conservation between groups of strongly similar properties.

actin_4_cDNA	ATGTCTGACGATGAAGATATCGCCGCCTTGTTCATCGACAACGGTTCGGAATGTGCAAA	60
actin_3_cDNA	ATGTCTGACGACGAAGATATCGCCGCCTTGTTCATCGACAACGGTTCGGAATGTGCAAG	60
actin_1_cDNA	ATGTCTGACGACGAAGATATCGCCGCCTTGTTCATCGACAACGGTTCGGAATGTGCAAA	60
actin_2_cDNA	ATGTCTGACGATGAAGATATCGCCGCCTTGTTCATCGACAACGGTTCGGAATGTGCAAA *****	60
actin_4_cDNA	GCCGGATTTCGCTGGAGATGACGCCCCACGCTCCGTCTTCCCATCCCTCATCGGACGTGCC	120
actin_3_cDNA	GCCGGATTTCGCTGGAGATGACGCCCCACGCTCCGTCTTCCCATCCCTCATCGGACGTGCC	120
actin_1_cDNA	GCCGGATTTCGCTGGAGATGACGCCCCACGCTCCGTCTTCCCATCCCTCATCGGACGTGCC	120
actin_2_cDNA	GCCGGATTTCGCTGGAGATGACGCCCCACGCTCCGTCTTCCCATCCCTCATCGGACGTGCC *****	120
actin_4_cDNA	CGCCAGCCAGGTATCATGGTTGGTATGGAGCAGCGTGACGCATACGTCGGTGACGAGGCA	180
actin_3_cDNA	CGCCAGCCAGGTATCATGGTTGGTATGGAGCAGCGTGACGCATACGTCGGTGATGAGGCA	180
actin_1_cDNA	CGCCAGCCAGGTATCATGGTTCGGTATGGAGCAGCGTGACGCATACGTCGGTGACGAGGCA	180
actin_2_cDNA	CGCCAGCCAGGTATCATGGTTCGGTATGGAGCAGCGTGACGCATACGTCGGTGACGAGGCA *****	180
actin_4_cDNA	CAGGCCAAGCGTGGTGTCTCACACTGAAGTACCCAATCGAGCACGGTATCGTCACCAAC	240
actin_3_cDNA	CAGGCCAAGCGTGGTGTCTCACACTGAAGTACCCAATCGAGCACGGTATCGTCACCAAC	240
actin_1_cDNA	CAGGCCAAGCGTGGTGTCTCACACTGAAGTACCCAATCGAGCACGGTATCGTCACCAAC	240
actin_2_cDNA	CAGGCCAAGCGTGGTGTCTCACACTGAAGTACCCAATCGAGCACGGTATCGTCACCAAC *****	240
actin_4_cDNA	TGGGACGATATGGAGAAGATCTGGCATCACACCTTCTACAACGAGTCCGTGTGCGCCCA	300
actin_3_cDNA	TGGGACGATATGGAGAAGATCTGGCATCACACCTTCTACAACGAGTCCGTGTGCGCCCA	300
actin_1_cDNA	TGGGACGATATGGAGAAGATCTGGCATCACACCTTCTACAACGAGTCCGTGTGCGCCCA	300
actin_2_cDNA	TGGGACGATATGGAGAAGATCTGGCATCACACCTTCTACAACGAGTCCGTGTGCGCCCA *****	300
actin_4_cDNA	GAGGCCACCCAGTCTGTCTCACCGAGGCCCCACAGAACCCAAGGCCAACCCGGAACGT	360
actin_3_cDNA	GAGGCCACCCAGTCTGTCTCACCGAGGCCCCACAGAACCCAAGGCCAACCCGGAACGT	360
actin_1_cDNA	GAGGCCACCCAGTCTGTCTCACCGAGGCCCCACAGAACCCAAGGCCAACCCGGAACGT	360
actin_2_cDNA	GAGGCCACCCAGTCTGTCTCACCGAGGCCCCACAGAACCCAAGGCCAACCCGGAACGT *****	360
actin_4_cDNA	ATGACCCAGATCATGTTTCGAGACCTTCAACGTCCCGCCATGTACGTCAACATCCAGGCC	420
actin_3_cDNA	ATGACCCAGATCATGTTTCGAGACCTTCAACGTCCCGCCATGTACGTCAACATCCAGGCC	420
actin_1_cDNA	ATGACCCAGATCATGTTTCGAGACCTTCAACGTCCCGCCATGTACGTCAACATCCAGGCC	420
actin_2_cDNA	ATGACCCAGATCATGTTTCGAGACCTTCAACGTCCCGCCATGTACGTCAACATCCAGGCC *****	420
actin_4_cDNA	GTCCTCTCCCTGTACGCCTCTGGACGTACCACCGGATGCGTGTGGACTCCGGTGATGGT	480
actin_3_cDNA	GTCCTCTCCCTGTACGCCTCTGGACGTACCACCGGATGCGTGTGGACTCCGGTGATGGT	480
actin_1_cDNA	GTCCTCTCCCTGTACGCCTCTGGACGTACCACCGGATGCGTGTGGACTCCGGTGATGGT	480
actin_2_cDNA	GTCCTCTCCCTGTACGCCTCTGGACGTACCACCGGATGCGTGTGGACTCCGGTGATGGT *****	480
actin_4_cDNA	GTGACACACACTGTCCCAATCTACGAGGGATACGCCCTTCCACACGCCGTATCCGTCTC	540
actin_3_cDNA	GTGACACACACCGTTCCAATCTACGAGGGATACGCCCTTCCACACGCCGTATCCGTCTC	540
actin_1_cDNA	GTGACACACACCGTTCCAATCTATGAGGGATACGCCCTTCCACACGCCGTATCCGTCTC	540
actin_2_cDNA	GTGACACACACCGTTCCAATCTATGAGGGATACGCCCTTCCACACGCCGTATCCGTCTC *****	540
actin_4_cDNA	GATTTGGCCGGACGTGACCTGACCGAGTACCTGATGAAGATCCTCACAGAGCGTGGTTAC	600
actin_3_cDNA	GATTTGGCCGGACGTGACCTGACCGAGTACCTGATGAAGATCCTCACAGAGCGTGGTTAC	600
actin_1_cDNA	GATTTGGCCGGACGTGACCTGACCGAGTACCTGATGAAGATCCTCACAGAGCGTGGTTAC	600
actin_2_cDNA	GATTTGGCCGGACGTGACCTGACCGAGTACCTGATGAAGATCCTCACAGAGCGTGGTTAC *****	600
actin_4_cDNA	TCGCTGACAACCACCGCCGAGCGCAAATCGTCCGTGACATCAAGGAGAGTCTCTGCTTC	660
actin_3_cDNA	TCGCTGACAACCACCGCCGAGCGCAAATCGTCCGTGACATCAAGGAGAGTCTCTGCTTC	660
actin_1_cDNA	TCGCTGACAACCACCGCCGAGCGCAAATCGTCCGTGACATCAAGGAGAGTCTCTGCTTC	660
actin_2_cDNA	TCGCTGACAACCACCGCCGAGCGCAAATCGTCCGTGACATCAAGGAGAGTCTCTGCTTC *****	660
actin_4_cDNA	GTCGCCGTGACTTCGAGGAGGAGATGAAGAAGGCCGCTGAGTCTTCTGCTCTTGAGAAG	720
actin_3_cDNA	GTCGCCGTGACTTCGAGGAGGAGATGAAGAAGGCCGCTGAGTCTTCTGCTCTTGAGAAG	720
actin_1_cDNA	GTCGCCGTGACTTCGAGGAGGAGATGAAGAAGGCCGCTGAGTCTTCTGCTCTTGAGAAG	720
actin_2_cDNA	GTCGCCGTGACTTCGAGGAGGAGATGAAGAAGGCCGCTGAGTCTTCTGCTCTTGAGAAG *****	720

actin_4_cDNA	TCTTTTCGAGCTTCCAGATGGTAACATCATCGTGATTGGAAACGAGCGCTTCCGTTGCCCA	780
actin_3_cDNA	TCTTTTCGAGCTTCCAGATGGTAACATCATCGTGATTGGAAACGAGCGCTTCCGTTGCCCA	780
actin_1_cDNA	TCTTTTCGAGCTTCCAGATGGTAACATCATCGTGATTGGAAACGAGCGCTTCCGTTGCCCA	780
actin_2_cDNA	TCTTTTCGAGCTTCCAGATGGTAACATCATCGTGATTGGAAACGAGCGCTTCCGTTGCCCA	780

actin_4_cDNA	GAGGTCTGTTCAGCCTAACTTGACTGGTCTCGAAATGGACGGTGTGGCGACTCCACC	840
actin_3_cDNA	GAGGTCTGTTCAGCCTAACTTGACTGGTCTCGAAATGGACGGTGTGGCGACTCCACC	840
actin_1_cDNA	GAGGTCTGTTCAGCCTAACTTGACTGGTCTCGAAATGGACGGTGTGGCGACTCCACC	840
actin_2_cDNA	GAGGTCTGTTCAGCCTAACTTGACTGGTCTCGAAATGGACGGTGTGGCGACTCCACC	840

actin_4_cDNA	TTCCAGACCATCATGAAGTGCATGTGCGATATCCGCAAGGATCTGTACGCCAACATCGTC	900
actin_3_cDNA	TTCCAGACCATCATGAAGTGCATGTGCGATATCCGCAAGGATCTGTACGCCAACATCGTC	900
actin_1_cDNA	TTCCAGACCATCATGAAGTGCATGTGCGATATCCGCAAGGATCTGTACGCCAACATCGTC	900
actin_2_cDNA	TTCCAGACCATCATGAAGTGCATGTGCGATATCCGCAAGGATCTGTACGCCAACATCGTC	900

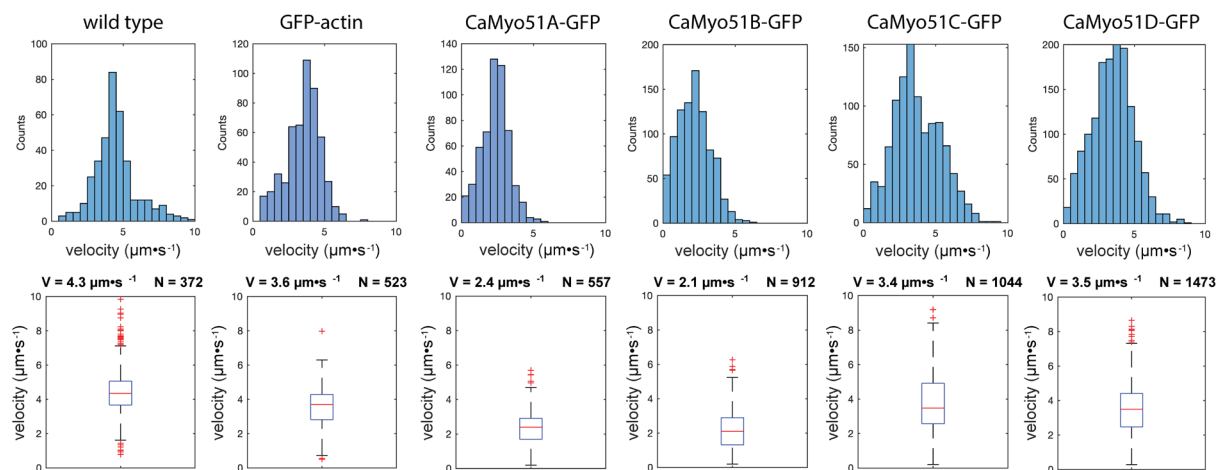
actin_4_cDNA	CTTTCTGGAGGAACCACCATGTTCCAGGAATCAGCGAGCGCATGTGGAAGGAGATCACC	960
actin_3_cDNA	CTTTCTGGAGGAACCACCATGTTCCAGGAATCAGCGAGCGCATGTGGAAGGAGATCACC	960
actin_1_cDNA	CTTTCTGGAGGAACCACCATGTTCCAGGAATCAGCGAGCGCATGTGGAAGGAGATCACC	960
actin_2_cDNA	CTTTCTGGAGGAACCACCATGTTCCAGGAATCAGCGAGCGCATGTGGAAGGAGATCACC	960

actin_4_cDNA	GCCCTCGCCCCAGCCTCCATCAAGGTCAAGATCGTGGCCCCACCGGAGCGAAAGTACTCC	1020
actin_3_cDNA	GCCCTCGCCCCAGCCTCCATCAAGGTCAAGATCGTGGCCCCACCGGAGCGAAAGTACTCC	1020
actin_1_cDNA	GCCCTCGCCCCAGCCTCCATCAAGGTCAAGATCGTGGCCCCACCGGAGCGAAAGTACTCC	1020
actin_2_cDNA	GCCCTCGCCCCAGCCTCCATCAAGGTCAAGATCGTGGCCCCACCGGAGCGAAAGTACTCC	1020

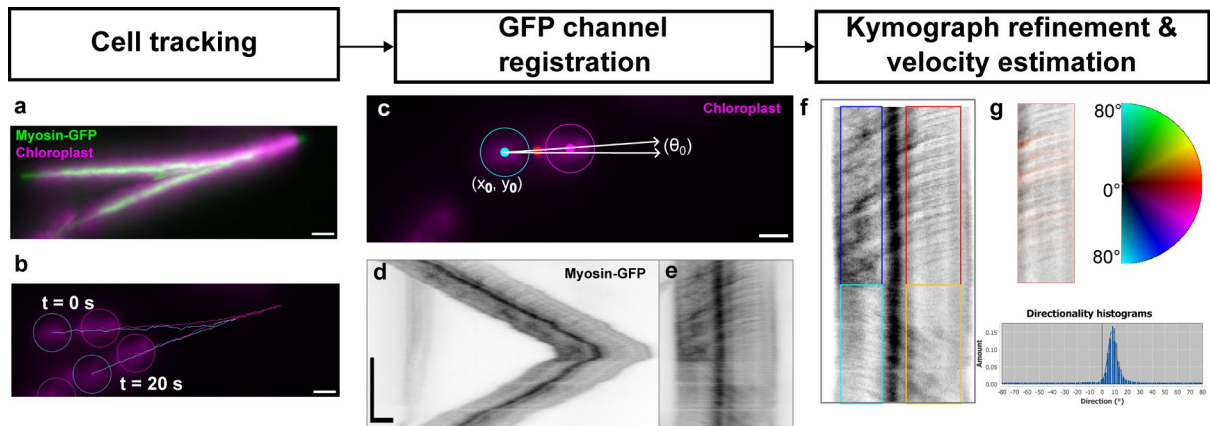
actin_4_cDNA	GTGTGGATCGGAGGATCCATCCTCGCTTCCCTGTCTACATTCCAGAGCATGTGGATCAGC	1080
actin_3_cDNA	GTGTGGATCGGAGGATCCATCCTCGCTTCCCTGTCTACATTCCAGAGCATGTGGATCAGC	1080
actin_1_cDNA	GTGTGGATCGGAGGATCCATCCTCGCTTCCCTGTCTACATTCCAGAGCATGTGGATCAGC	1080
actin_2_cDNA	GTGTGGATCGGAGGATCCATCCTCGCTTCCCTGTCTACATTCCAGAGCATGTGGATCAGC	1080

actin_4_cDNA	AAGGAGGAGTACGACGAGTCCGGTCCATCCATCGTCCACCGCAAGTGCTTCTAA	1134
actin_3_cDNA	AAGGAGGAGTACGACGAGTCCGGTCCATCCATCGTCCACCGCAAGTGCTTCTAA	1134
actin_1_cDNA	AAGGAGGAGTACGACGAGTCCGGTCCATCCATCGTCCACCGCAAGTGCTTCTAA	1134
actin_2_cDNA	AAGGAGGAGTACGACGAGTCCGGTCCATCCATCGTCCATCGCAAGTGCTTCTAA	1134

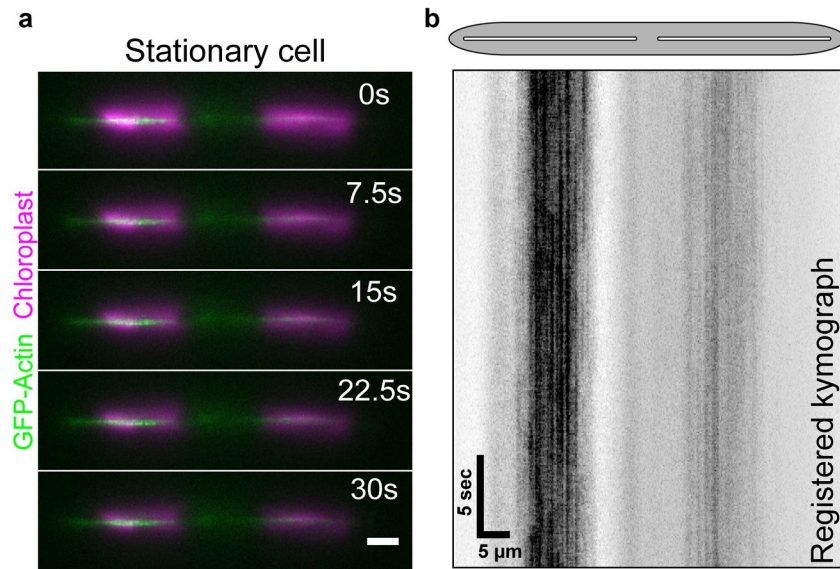
Supplementary Figure 3: Clustal Omega alignment of the four *C. australis* mRNA sequences. Identical nucleotides are highlighted in grey and indicated with an asterisk (*).



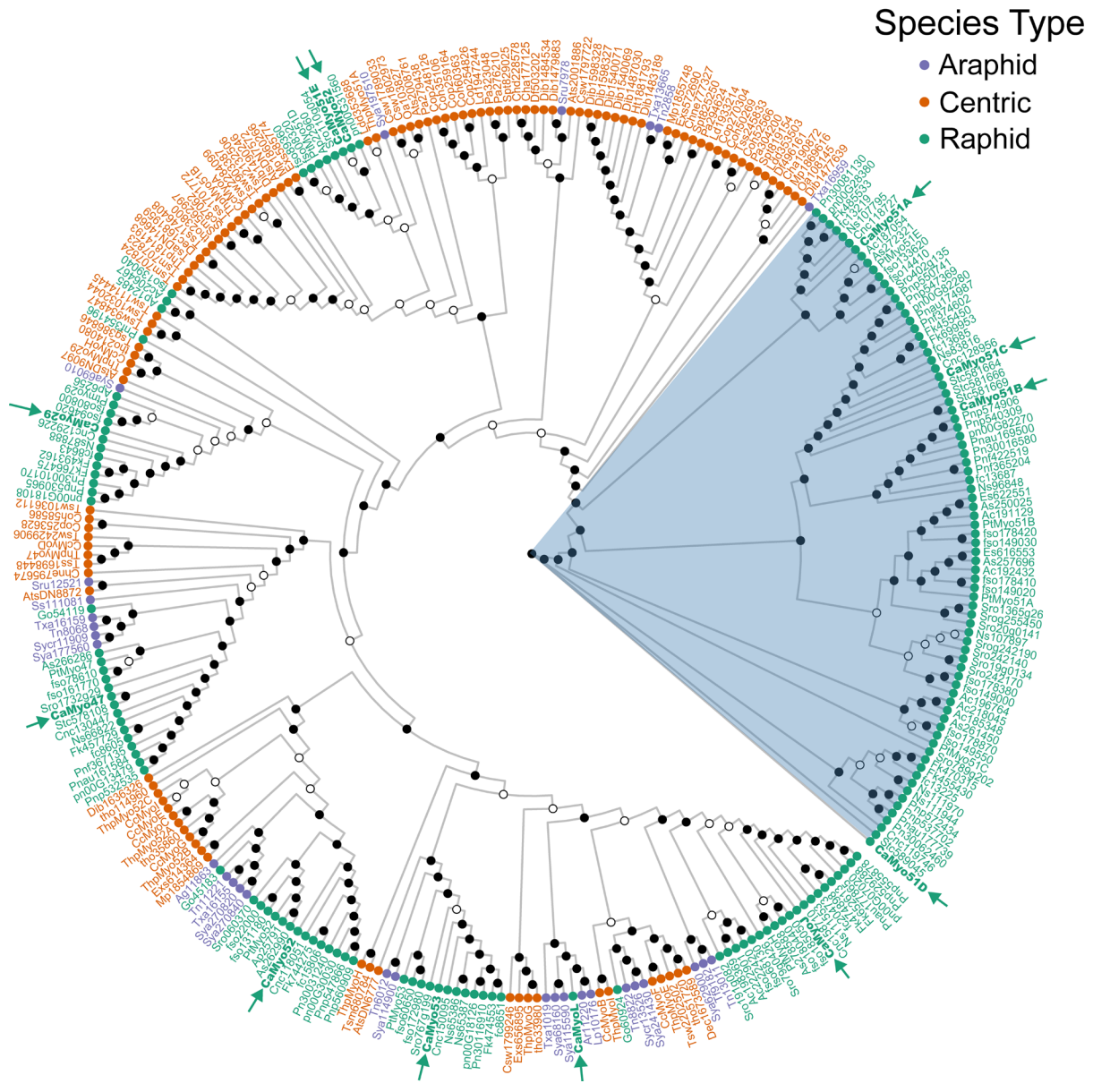
Supplementary Figure 4: Cell motility analysis of *C. australis* cell lines. Histograms (upper panels) and corresponding box plots (lower panels) showing the distribution of cell velocities in populations of wild type *C. australis* (far left) and *C. australis* cell lines expressing actin or myosins as GFP fusion proteins used for TIRFM analyses of actin and myosin dynamics. In the box plots, the red center lines denote the median (50th percentile) values, with the blue boxes having lower and upper bounds of the 25th and 75th, respectively. The black whiskers indicate the 5th and 95th percentiles, and outlier values outside of this range are indicated by red plus signs. N represents the number of investigated cells.



Supplementary Figure 5: Image analysis pipeline. **a** Maximum projection of a 20 s dual-channel movie of a gliding cell. **b** Overlay of tracks generated using TrackMate using the chloroplast autofluorescence signal in the far red optical spectrum. Chloroplast positions (circles) are shown only for the first and last frames. **c** Chloroplast positions for each frame were used to determine the position of the midpoint between the two chloroplasts (red dot) and the angle of the cell relative to a horizontal origin. **d, e** Kymographs of myosin-GFP signals (in black) generated by reslicing and projecting each row of pixels of the image before (**d**) and after (**e**) registration using displacement and angle values. **f** Refined kymograph after performing a Fast-Fourier-Transformation, masking horizontal traces, and inverting the Fourier transformation. Indigo, red, cyan and yellow boxes indicate regions-of-interest (ROIs) used for directionality measurements. **g** Orientation map showing the preferred orientation of the myosin-GFP signals in the red ROI shown in **f** (top left), color wheel showing gradient of colors used to indicate angle (top right) and histogram of estimated angles generated by the directionality plugin in Fiji. Scale bars – horizontal 5 μm , vertical 5 s.

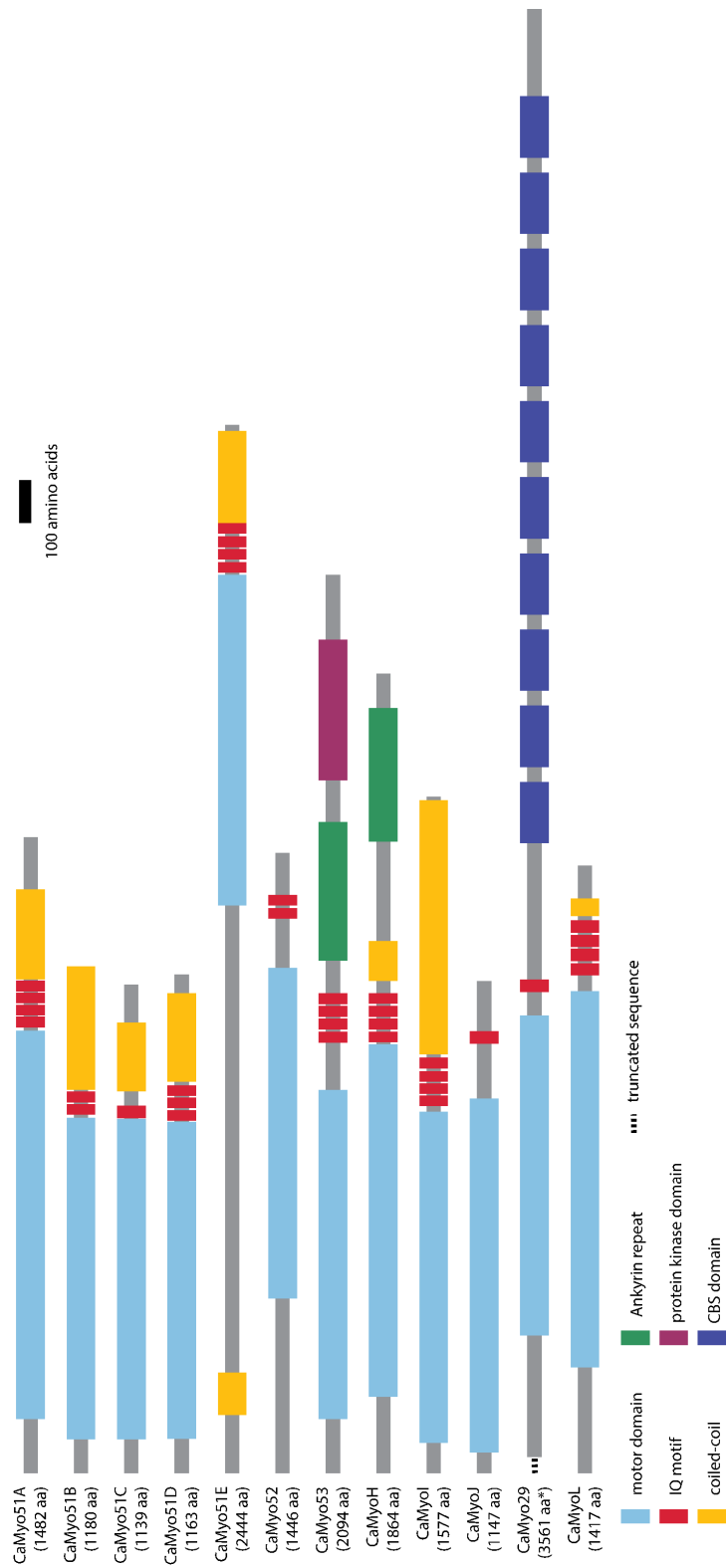


Supplementary Figure 6: TIRFM based kymograph analysis of GFP-actin in stationary cells. **a** Montage generated for dual-channel TIRFM data showing the positions of a stationary cell at 7.5 s intervals (GFP-actin in green and chloroplast autofluorescence in magenta, generated from Supplementary Movie S2). **b** Kymograph analysis of GFP-actin movement. GFP-channel data from the **Supplementary Movie S2** shown in **a** was registered using cell tracking data (GFP-actin in black). Gray ellipse and white slits above kymograph approximate the positions of the cell body and raphe openings, respectively. (Scale bars: horizontal = 5 μm, vertical = 5 s)



Supplementary Figure 7: Phylogenomic analysis of diatom myosins. Cladogram of 309 diatom myosins. The position of *C. australis* myosins (CaMyoX, 12 sequences) are in bold and labeled with a green arrow. A large clade containing only raphid diatom myosins, including the four *C. australis* myosins investigated in this study (CaMyo51A-D) is highlighted in blue. Node tips are color coded by species type. Internal nodes indicate Ufboot bootstrap scores of <95 (hollow nodes) or ≥ 95 (solid nodes). Species name are abbreviated as follows as the full sequence names can be found in Supplementary Table 1. *Amphora coffeaeformis* (Ac); *Asterionellopsis glacialis* (Ag); *Amphiprora paludosa* (Ap) *Astrosyne radiata* (Ar); *Amphiprora sp* (As); *Attheya septentrionalis* (Ats); *Craspedostauros australis* (Ca); *Cyclotella cryptica* (Cc); *Chaetoceros affinis* (Cha); *Chaetoceros debilis* (Chd); *Chaetoceros neogracile* (Chne); *Cylindrotheca closterium* (Cnc); *Corethron hystrix* (Coh); *Corethron pennatum* (Cop); *Coscinodiscus wailesii* (Csw); *Detonula confervacea* (Dec); *Dactyliosolen fragilissimus* (Df); *Ditylum brightwellii* (Dib); *Entomoneis sp* (Es); *Extubocellulus spinifer* (Exs); *Fragilariopsis*

cylindrus (Fc); *Fragilariopsis kerguelensis* (Fk); *Fistulifera solaris* (Fso); *Grammatophora oceanica* (Go); *Helicotheca tamesis* (Ht); *Leptocylindrus danicus* (Ld); *Leptocylindrus danicus* var *danicus* (Ldd); *Licmophora paradoxa* (Lp); *Minutocellus polymorphus* (Mp); *Nitzschia* sp (Ns); *Odontella aurita* (Oa); *Proboscia alata* (Pa); *Pseudo-nitzschia multiseriata* (Pn); *Pseudo-nitzschia australis* (Pnau); *Pseudo-nitzschia fraudulenta* (Pnf); *Pseudo-nitzschia pungens* (PnP); *Phaeodactylum tricornutum* (Pt); *Skeletonema costatum* (Sc); *Skeletonema menzelii* (Sm); *Stephanopyxis turris* (Spt); *Seminavis robusta* (Sro); *Striatella unipunctata* (Sru); *Staurosira* sp. (Ss); *Stauroneis constricta* (Stc); *Synedro acus* (Sya); *Synedropsis cf recta* (Sycr); *Thalassionema frauenfeldii* (Tf); *Thalassiosira pseudonana* (Thp); *Thalassionema nitzschioides* (Tn); *Thalassiosira antarctica* (Tsa); *Thalassiosira gravida* (Tsg); *Thalassiosira miniscula* (Tsm); *Thalassiosira* sp. (Tss); *Thalassiosira oceanica* (Tho); *Thalassiosira weissflogii* (Tsw); *Thalassiothrix antarctica* (Txa)

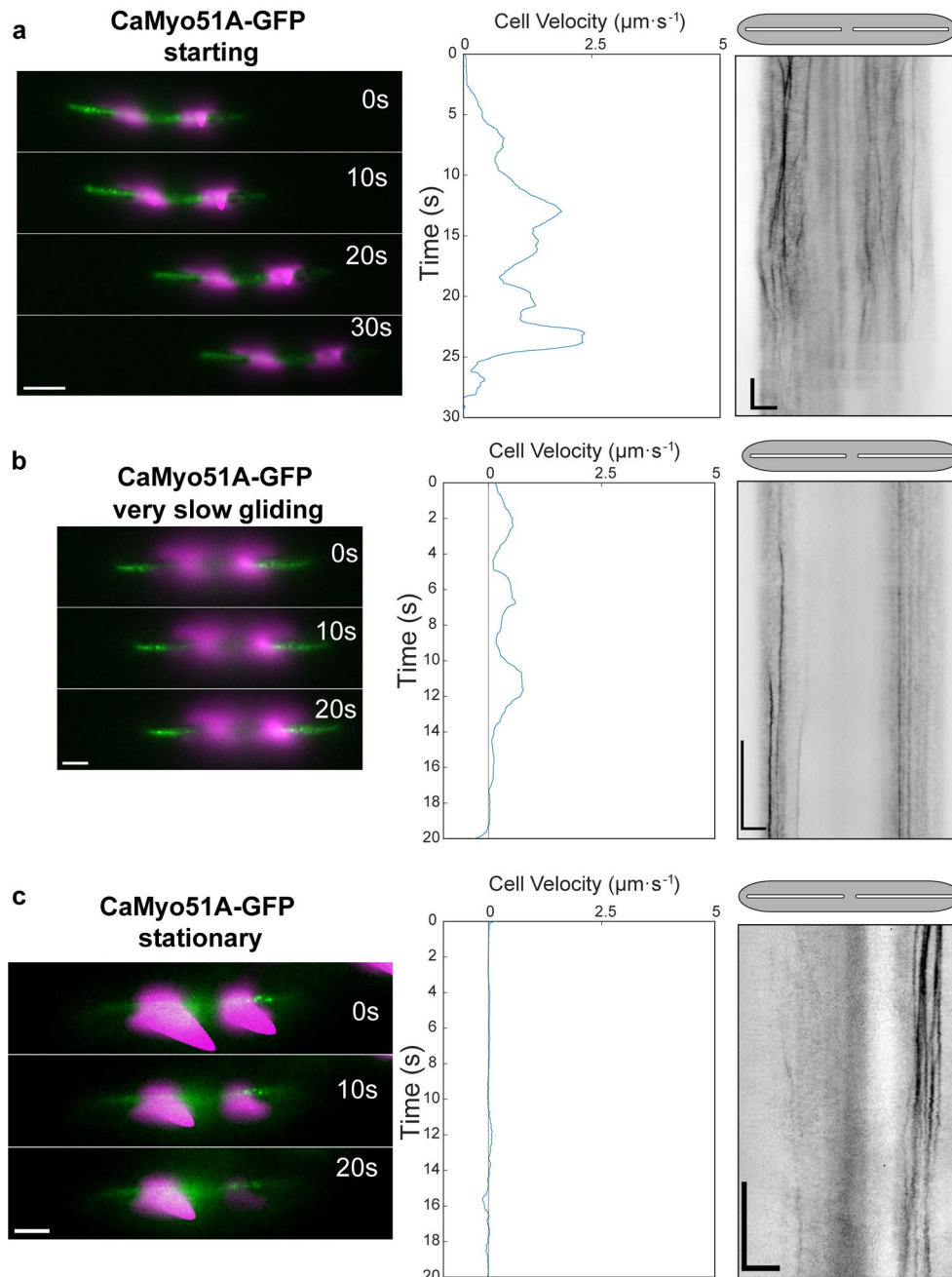


Supplementary Figure 8. Schematic of twelve predicted *C. australis* myosins

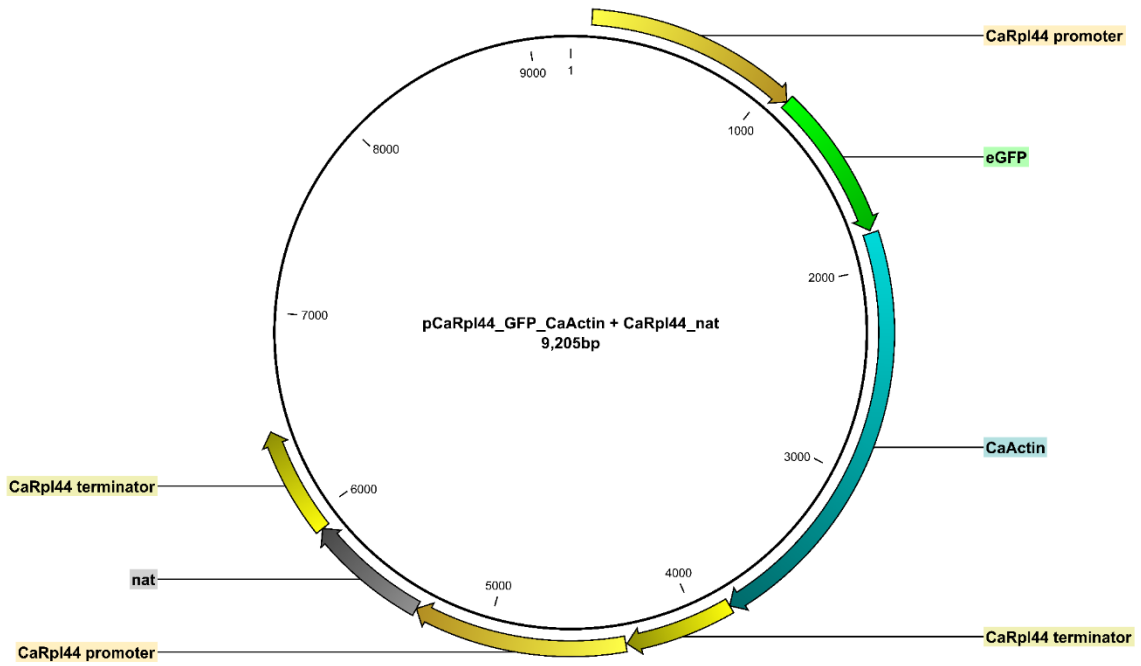
CaMyo51A	LHEAAILYNLKHARHV-QALPYTRTGDIVIAVNPYQWFHVLVYSHDNQNTYSEALIWKNRSN	59
CaMyo51D	LHEAAILYNLKERHC-SAKPYTRVSDIVIAVNPYQWLTDLYTEKVRLLYYAQKLVWEQSE-	58
CaMyo51B	LHEAAILYNLKERHL-NNHPYTRTGDIVIAVNPYQWFKEIYTEKVRARFANKLVWEDHE-	58
CaMyo51C	LHEPAILFNLKERHG-SGKPYTRTGDIIIAVNPYQWFKEIYTEKVRNRYARILVWENVN-	58
CeMyo5	LHEPAVLHNLQVRFVKGSSIIYTYCGIVLVAINPYADCSHIYGEEIIQVYRGA-----	52
DdMyo1E	ITENAFIENLTMRHK-SDNIYTYIGDVVISTNPFKNL-NIYKESDIKAYNGR-----	50
HsMhc2	LHEPAVLYNLKERYA-AWMIYTYSGLFCVTVNPKWL-PVYKPEVVTAYRGG-----	50
OcMhc1	LHEPAVLYNLKEXYA-AWMIYTYSGLFCVTVNPKWL-PVYNAEVVTAYRGG-----	50
	: * * .: ** . ** . . : : **: : * . :	
CaMyo51A	ITKRKSTSKPNSNDNSNSNNSNNQTKSQTPQSQGEGASPTADISDPRNGLVPHVYEAS	119
CaMyo51D	-----QDPRQGLPPHIYEVS	73
CaMyo51B	-----GDARVLVEPHVYEAS	73
CaMyo51C	-----GDQRSQVEPHVYETS	73
CeMyo5	-----GKSAREMDPHIFAVA	67
DdMyo1E	-----YKYEMPPHIYALA	63
HsMhc2	-----KRQEAPPHIFISIS	63
OcMhc1	-----KRQEAPPHIFISIS	63
	***: :	
	P-loop	
CaMyo51A	ALAYQGLAARGDDQSILVSGESGAGKTETVKICMSHLASIQNLTPKQRQKQVRLNAAVSS	179
CaMyo51D	SLCYRGLAYEGVNSILVSGESGAGKTETVKICMNHIASVQKDP TAV-----	120
CaMyo51B	SLAYKGLAFEGSDQSILVSGESGAGKTETVKICMNHIASVQRGPTSD-----	120
CaMyo51C	ALAYNGLAFGGNDQSILVSGESGAGKTETVKICMNHIASCQQGPTAK-----	120
CeMyo5	EEAHFDMGAFGKSQSIIVSGESGAGKTVSAKFMRYLASVAASKTRNGG-----	116
DdMyo1E	NDAYRSMRQSQENQCVIISGESGAGKTEASKKIMQFLT FVSSNQSPNG-----	111
HsMhc2	DNAYQFMLTDRENQSILITGESGAGKTVNTKRVIQYFATIAVTGKKE-----	112
OcMhc1	DNAYQFMLTDRENQSILITGESGAGKTVNTKRVIQYFATIAVTGDKKE-----	112
	.: : .*.:::***** * : .: :	
	Switch-1	
CaMyo51A	ASSVASSALSTTGSASAPANSEIVQRVLDNSNPLLEAFGNAKTLRNDNSSRFGKYIQLQFD	239
CaMyo51D	-----SEID---ADAVNRI LDSNPLLEAFGNAKTRRNDNSSRFGKYIVLQFD	165
CaMyo51B	-----ADGGAKFSDPIVQRIMESNPLLEAFGNAKTRRNDNSSRFGKYTQLQFD	168
CaMyo51C	-----GGGN---NVVVNRVLESNPLLEAFGNAKTRRNDNSSRFGKYCQLQFD	164
CeMyo5	-----T-----TSIEARVLASNPIMESIGNAKTIRNDNSSRFGKFIQINFC	157
DdMyo1E	-----ERISKMLLDSNPLLEAFGNAKTLRNDNSSRFGKYMENQFN	151
HsMhc2	-----EITSGKIQGTLEDQIISANPLLEAFGNAKTVRNDNSSRFGKFIRIHFG	160
OcMhc1	-----EATSGKMQGTLEDQIISANPLLEAFANAKTVRNDNSSRFGKFIRIHFG	160
	: : : **:* : : .*** *****: : *	
CaMyo51A	NPLSHQLHNASSIDDMAVAATGMSSRQQVLAQSTGTASRCILAGSKCEVYLLEKSRVISH	299
CaMyo51D	RKEPKN-----VGDNPRAIMSGSKCEVYLLEKSRVVYH	198
CaMyo51B	MGDKSS-----KAFATDKASRKNLAGSKCEVYLLEKNRVTTH	206
CaMyo51C	MGTKAE-----QEFRG-RDKATCVLAGSKCDVYLLEKNRVVMH	201
CeMyo5	ERGR-----RIVGAEMKTYLLEKSRLVFQ	181
DdMyo1E	AVGS-----PIGKITNYLLEKSRVVG	174
HsMhc2	TTGK-----LASADIETYLLEKSRVVFQ	183
OcMhc1	TTGK-----LASADIETYLLEKSRVTFQ	183
	... *****.*: :	
CaMyo51A	AEGERTFHIFFYQLLAAPPEYKATVWSGLEDTDNESFCYVGWTDNTIEGRSDEDRFQLTL	359
CaMyo51D	DETERTYHIFFYQLLNAPESEKVKIWKGLKGTKNESFAYVVGKSPEVRIDGKEDSEHFPETK	258
CaMyo51B	DTTERTYHSFYQLLAAPENEKVEIWSSELKGTNYESFAYIGTPTDVIIEGKKDGDHWHHTL	266
CaMyo51C	DPEERTYHIFFYQIIASK--HKTKYWSHLAGT TNESFRFVKGAPDTIEGMDGDHFNETH	259
CeMyo5	APGERNYHIFFYQLCAARNHQVLKD--LHLGPC-ESYSYLTQGGDSRIPGVDDKADFEALL	238
DdMyo1E	TQGERSFHIFFYQMLKGLSQSKLNE--LGLTPNAPAYEYLLKSGCFDVSTIDDSGEFKIIV	232
HsMhc2	LKAERSYHIFFYQITSNKKPELIEM--LLITNPNYDYPFV-SQGEISVASIDDQEELMATD	240
OcMhc1	LKAERSYHIFFYQIMSNNKPPDLIEM--LLITNPNYDYAFV-SQGEITVPSIDDQEELMATD	240
	.:* *: : : : : *	
CaMyo51A	NSLALVGVIDDKLLKALLRAICIVLQLGNLEFAP--LDD---DKCVISTDEELMELAELE	413
CaMyo51D	KALEIVGIKDDKYITLFAQIAAVLQLGNLSFAPKPGND---EESVVTSKKEFKELAELE	314
CaMyo51B	EILDITINVKGNDRRDLMRAMCIVLQLGNVSEFMKDPND---DRSVVKDRKEFDKLAALM	322
CaMyo51C	KVLEKVGVDNDNMLTLMQAI IIVLQIGNLTFGPDPKDD---ERSVVTSTKDFALAELE	315
CeMyo5	KALQLLGFDEKQMSDVFRLLAGLLLLGNVHFENGEGSSAVSASSC--QEISR--LCREFW	294
DdMyo1E	KAMETLGLKESDQNSIWRILAAIHLHIGNITFAEAAEQRTGTMTT-VKVSDTKSLAAASCL	291
HsMhc2	SAIDILGFTNEEKVSIYKLTGAVMHYGNLKFQKQREEQAE PDGTEVAD-----KAAYLQ	295
OcMhc1	SAIDILGFTSDERVSIYKLTGAVMHYGNMKEKQKQREEQAE PDGTEVAD-----KAAYLQ	295
	.: : . . : : : : **:*	

CaMyo51A	DVPPDELAVALLVRTVTAKGE----	TFQVPLKAETAKDSTDAFAKEIYAKTFLWLVRTIN	469
CaMyo51D	GCEEDLLGRAFTERTIKTRGE----	EYKVPLKSDVSKESADAFKAIYAKIFLWLVRIN	370
CaMyo51B	GVTNQDLVDCLTERTMKTRNE----	SYKVP LNADVAKESADAFKAIYSKLF LWLVRALN	378
CaMyo51C	CVPEEAMECSFTERTMKTRNE----	TYKVP LNATVAKESAFAKAIYGR LFLWLVRIN	371
CeMyo5	KISESDLRIWLTRREIRAVNE----	IVTKPLTKNEAVRSRDALTKMLYSHLFGWLVDKIN	350
DdMyo1E	KTDQQSLSIALCYRSISTGVGKRCSSVISPMDCNQAAYS RDALAKALYERLFNWLVS KIN		351
HsMhc2	SLNSADLLKALCYPRVKVGN EYVT----	KGQTVQVSNVAVGALAKAVYEKMFLWVVARIN	351
OcMhc1	SLNSXDLLKALCYPRVKVGN EYVT----	KGQTVQQVYNAV GALAKAVYEKMFLWVTRIN	351
	:	:	:
		Switch-2	
CaMyo51A	DATCAENNYRNASFDHQRFVIGL	LDIFGFETFEVNRFEQLCIN YANEK LQQKFTQDIFR	529
CaMyo51D	DATCAENNYPGADAN-TEFGV	VGLLDIFGFESFPINGFEQLCIN YCNEK LQAKFTHDIFQ	429
CaMyo51B	DATCAELNYEGG-MD-KHFGV	IGLLDIFGFESFVRNRFEQLCIN YANEK LQAKFTEDIFR	436
CaMyo51C	DATSAELNYKSG-LM-EHFGV	IGLLDIFGFESFVRNRFEQLCIN YANEK LQAKFTEDIFR	429
CeMyo5	EALNEKDKLDGNTQKKRPDRF	IGLVDIYGFETFDVNSFEQFSIN YANEK LQQQFNQHVFK	410
DdMyo1E	TIINCT-----TEKGPVIGL	LDIYGFVFNNSFEQLNINFCNEK LQQLFIELTLK	402
HsMhc2	QQLDTK-----QPRQYF	IGLVDIAGFEIFDFNSLEQLCIN FTNEK LQQFFNHHMFV	402
OcMhc1	QQLDTK-----QPRQYF	IGLVDIAGFEIFDFNSLEQLCIN FTNEK LQQFFNHHMFV	402
		: : * : * : * * : *	
CaMyo51A	SVQTEYEFEGIELGEITYDDNTDV	-LDLVEGGMGLLALLNEECV RPKGSDKTFVFKAQAM	588
CaMyo51D	TVQEEYKAEGLNLDIQYDDNSDV	-LELIEGKFGI IKQLNEECV RPKGNDEAFVSKALQS	488
CaMyo51B	SVQEEYEFEGIPLEIKYDDNTDV	-LDLIEGKAGLLAMLNEECV RPKGGDDAFVAKALAA	495
CaMyo51C	SVQDEYEKEGIPLEIKYDDNTDV	-LDLIEGKGLLAMLNEECV RPKGTDQAFVTKALAA	488
CeMyo5	LEQEEYIREEIEWVVRVDFHDN	QPA-IDLIEGPVGMINLLDEQCKRLNGSDADWLSQLQNS	469
DdMyo1E	SEQEEYVREGIEWKNI EYFNKPI	CELIEKKPIGLISLLDEACLI AKSTDTQTFDLSICKQ	462
HsMhc2	LEQEEYKKEGIEWTFIDFGMDL	AACIELIEKPMGIFSILEEEECMFPKATDTSFKNKLYDQ	462
OcMhc1	LEQEEYKKEGIEWEFIDFGMDL	AACIELIEKPMGIFSILEEEECMFPKATDTSFKNKLYEQ	462
		* * * : : : : : * : * * : . * : .	
CaMyo51A	NK-----DNAVWIREPHYSATSE	GIRHYAGKVTYDSNGFVTKNMDTLP TDLQDCAAKS	641
CaMyo51D	NK-----TVPCLIAKNTFKRSE	FGINHYAGAVIYTAQDFVIRNTD TLPDLQVCAKAC	541
CaMyo51B	NK-----KSPCLIAKNTFKRSE	FGINHYAGAVIYTAQDFVIRNTD TLPDLQVCAKAC	541
CaMyo51C	NK-----QSPCLFPAK-VNRMG	FGIHHYAGKVMYDAEEFVIRNQD TLPDLSDLASLC	540
CeMyo5	TELRKNPQLAFPK-----VRSN	DFIVRHFADVTYSTDGFVEKNRDAI GEQLLDVVVAS	523
DdMyo1E	FE-KN-PHLQSYVSKDRSIGDTC	FRLKHYAGDVTYDVRGFLDKNKD TLFGLISSMQSS	520
HsMhc2	HL-GKSANFQKPK--VVKGAEA	HFALIH YAGVDYDYNITGWLEKNK DPLNETVVGLYQKS	519
OcMhc1	HL-GKSNNFQKPK--PAKRKVEA	HFSLVHYAGTVDYDYNITGWLDKNK DPLNETVVGLYQKS	519
		* : * . * * : : * * : :	
CaMyo51A	SNAILAKELTNEAMNPFAPVQS	STKAPETA AAAAAAPK KKKRKP KGARGRPGAPTASSA	701
CaMyo51D	KNEILAKHLNDNDKCTNQESK	-----AP-----AP-----	565
CaMyo51B	NNEIICKHLNNDKSTNLTKA	-----P-----T-----	569
CaMyo51C	SNEIIAKHMTNDSCNTNFESK	-----T-----KS-----	563
CeMyo5	KFPFIRTIVIGSTAPTSVSS	-----	542
DdMyo1E	SDPLVQGLFPPTREPDS	-----	537
HsMhc2	AMKTLAQLFSGAQTAELEG	-----	538
OcMhc1	AMKTLAFLFTGTAAAEAE	-----	537
	:	:	:
		actin binding	
CaMyo51A	SAPTSRPGHHMPRRSVKSNLVAE	TVWTKFKTQLSRLMTQLNDTRTRYIRCIKPN SKKRPR	761
CaMyo51D	---AA-KSGRGP RRRAKSNLVAE	TVMTKFKTQLSSLMKGLSTTKSRYIRCVKPN VLK KKL	621
CaMyo51B	---P-GAAKAP RRRAKSNLVAE	TVWTKYKQLMSLMNMLAQTNSRYIRCIKPN TYKKPS	624
CaMyo51C	---AQ-VAKKAP RRRAKSNLVAE	TVWAKYKQLMSLMNMLKTTE SKYIRCIKPNAPKKPL	619
CeMyo5	SSSSSTP-----GKRTIKKT	VASQFRDSLKELMSVLCSTRPHYVRCIKPNDSKISF	593
DdMyo1E	-----KK-----RPETAGS	QFRNAMNALITLLACSPHYVRCIKSNDNKQAG	579
HsMhc2	AGGGAKKGG-----KKGSSFQ	TVSALFRENLNKLMTNLRSTHSPHFVRCIIPNETKTPG	592
OcMhc1	--GGGKGG-----KKGSSFQ	TVSALFRENLNKLMTNLRSTHSPHFVRSIIPNETKTPG	589
		* . : : : * . * : : * : *	

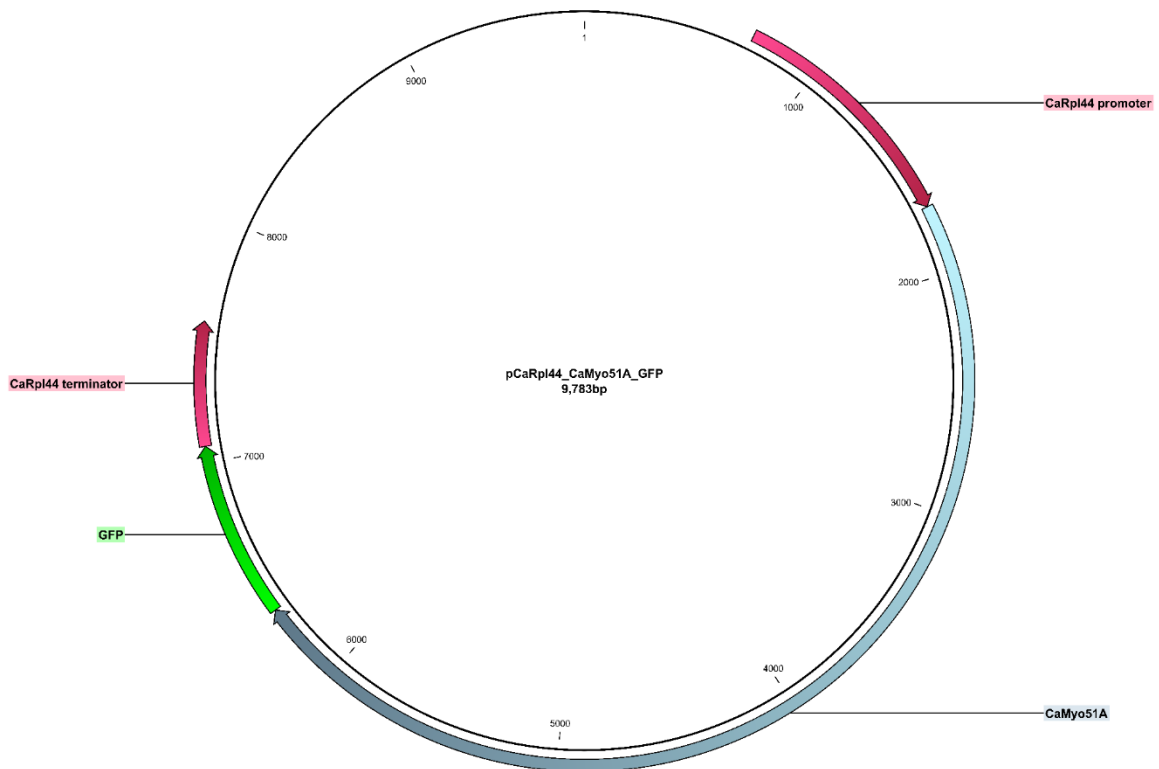
Supplementary Figure 9: Sequence alignment of the motor domain of the *C. australis* myosins CaMyo51A-D and myosins from *Oryctolagus cuniculus* (Genbank ID AAA74199.1); *Homo sapiens* (Gene bank ID: AAD29950.1), *Caenorhabditis elegans* (Genbank ID: AAA97926.1), *Dictyostelium discoideum* (Genbank ID: EAL63071.1). Identical amino acids are indicated by (*) and highlighted in grey, a : (colon) indicates conservation between groups of strongly similar properties; a . (period) indicates conservation between groups of weakly similar properties



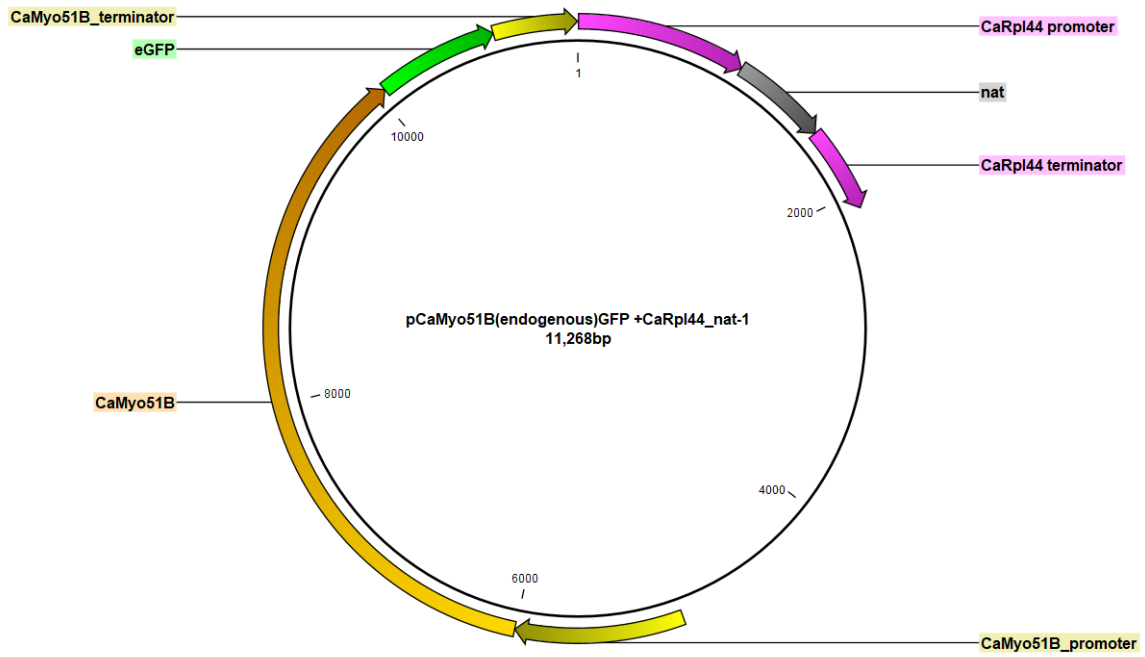
Supplementary Figure 10: Kymograph analysis of CaMyo51A-GFP in starting, slowly-moving and stationary cells. Analysis of 25-30 s time-lapse segments of CaMyo51A-GFP expressing cells upon **a** starting (Supplementary Movie S7), **b** slowly-moving (Supplementary Movie S8) and **c** being stationary (Supplementary Movie S9). (Left panels) Montages showing the position of cells at 10 s intervals (GFP in green, chloroplast autofluorescence in magenta, scale bars: 5 μm). (Middle panels) Cell velocity as function of time, generated from chloroplast tracking data. (Right panels) Registered kymographs generated from GFP-channel data (black) showing movement of myosins relative to the cell. Gray ellipses and white slits above kymographs approximate the positions of the cell body and raphe openings, respectively. (Scale bars: horizontal = 5 μm , vertical = 5 s)



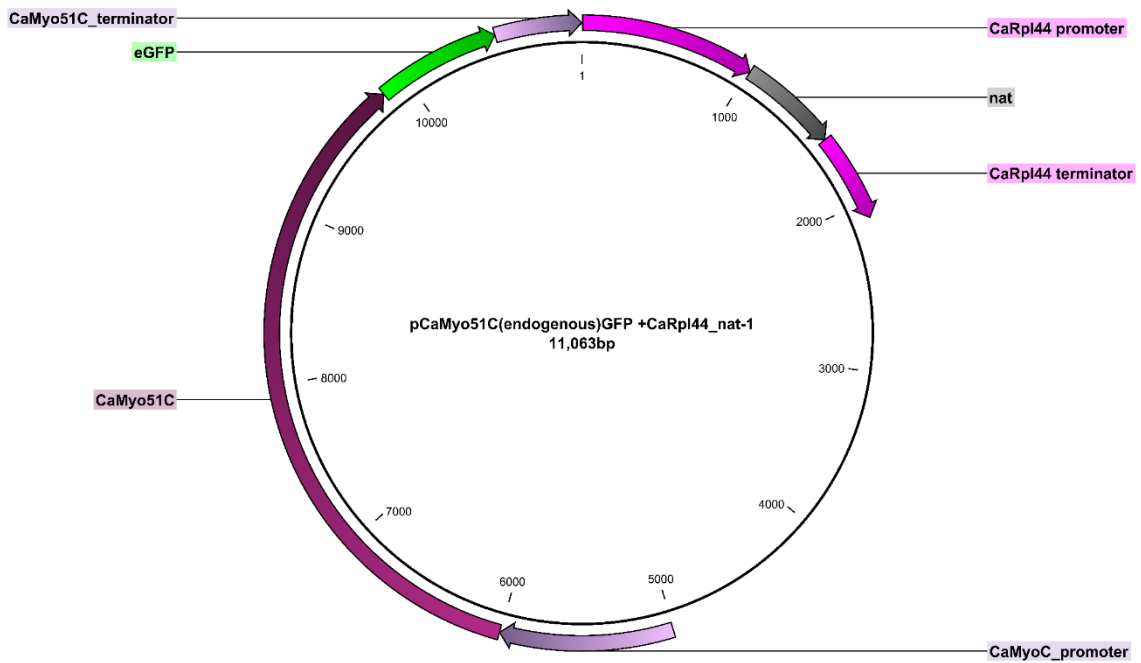
Supplementary Figure 11: Plasmid map of the DNA construct for N-terminal eGFP tagging of the *C. australis* Actin. nat = nourseothricin (ClonNat) resistance gene; eGFP = enhanced green fluorescent protein.



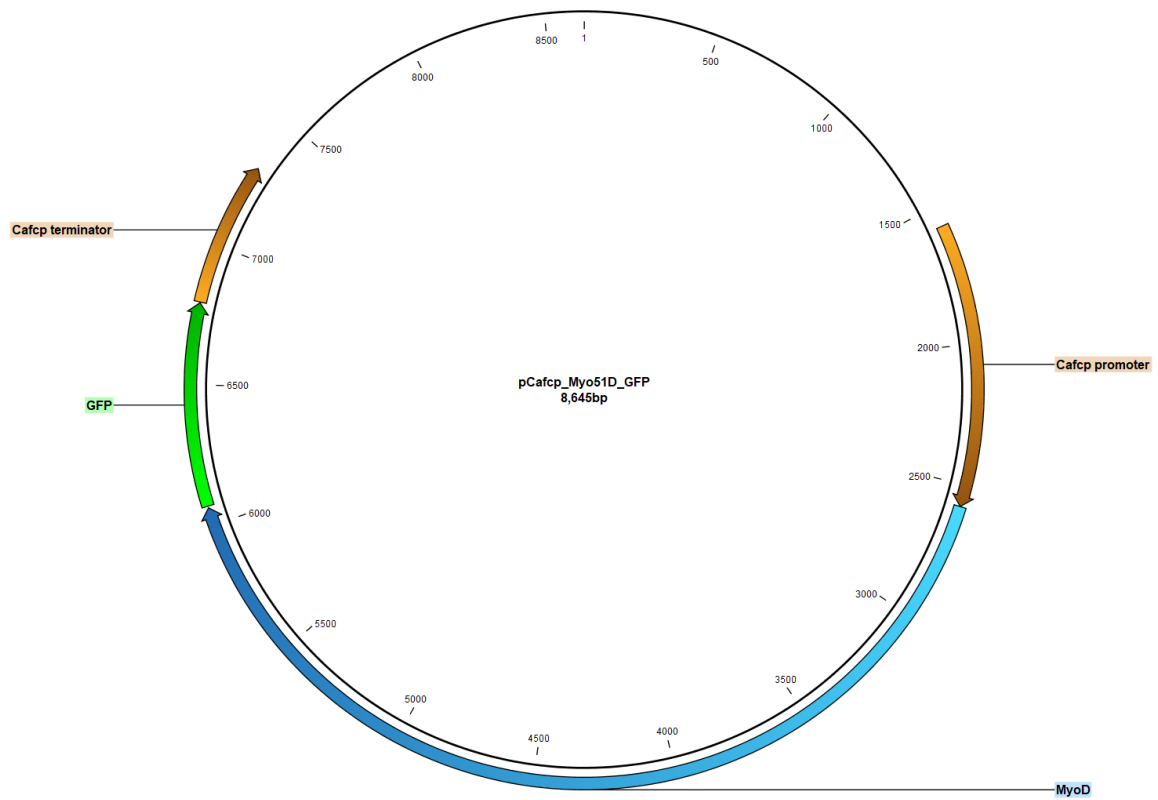
Supplementary Figure 12: Plasmid map of the DNA construct for C-terminal eGFP tagging of the *C. australis* CaMyo51A under the control of the CaRpl44 regulatory sequences. eGFP = enhanced green fluorescent protein.



Supplementary Figure 13: Plasmid map of the DNA construct for C-terminal eGFP tagging of the *C. australis* CaMyo51B under the control of its endogenous CaMyo51B regulatory sequences. nat = nourseothricin (ClonNat) resistance gene; eGFP = enhanced green fluorescent protein.



Supplementary Figure 14: Plasmid map of the DNA construct for C-terminal eGFP tagging of the *C. australis* CaMyo51C under the control of its endogenous CaMyo51B regulatory sequences. nat = nourseothricin (ClonNat) resistance gene; eGFP = enhanced green fluorescent protein.



Supplementary Figure 15: Plasmid map of the DNA construct for C-terminal eGFP tagging of the *C. australis* CaMyo51D under the control of the Cafcp regulatory sequences. eGFP = enhanced green fluorescent protein.

Supplementary Tables

Supplementary Table 1: Primers used to determine full-length gene models

Gene	Sequence
CaMyo51A	
3'RACE	1 ST PCR: 5'-AGGGTCCTCGATTCAAACCCGCTCCTCGAAGCA-3'; 2 nd PCR: 5'-GTGCGTCCGAAGGGCTCCGACAA-3'
5'RACE	1 ST PCR: 5'-CAGGGTACCCTATTCGAAACCGAAGATGTTCGAG-3'; 2 nd PCR: 5'-GCCAGC GCGGATGCTTCGTAGA-3'
Full-length gene model	sense: 5'-CCACAGACAGCGCTGCGACT-3' antisense 5'-CAGGGTACCCTATTCGAAACCGAAGATGTTCGAG-3'
Full-length gene model	sense: 5'-AGGGAATTCGATTCAAACCCGCTCCTCGAAGCA-3' antisense: 5'-GGC GATGTAGGTTGCCTT-3'
CaMyo51B	
3'RACE	1st PCR: 5'-GCAGCCAGATCATCGCTTACC-3' 2nd PCR: 5'-CGCAAGGACTTCGACATCATG-3'
5' RACE	1 st PCR: 5'-CCCTCGTGGTCTCCAGACG-3' 2 nd PCR: 5'-GGGCAAGTTGACCATGTCCGG -3'
Full-length gene model	sense: 5'-AAACAACAATTACAACAACCTACATGGGCAAGAAAAAGGCC-3' antisense: 5'- CCTCGCCCTTGCTCACCATTCTAGAAGCTACGGATTTCGGCAGC-3'
CaMyo51C	
3'RACE	1st PCR: 5'-CCTCATCCAGGCCAACGAGTC-3' 2nd PCR: 5'-CGTTCGAGATGTTGAACAACC-3'
5' RACE	1st PCR: 5'-GCTCCACCATGTCTGGGAAGG-3' 2nd PCR: 5'-ACTGTGGCACGCGGACTTGGG-3'
Full-length gene model	sense: 5'-AAACAACAATTACAACAACCTACGTATGGGCGAGAAAAAGTCCCAATACG-3'; antisense: 5'-CCTCGCCCTTGCTCACCATTCTAGACAGGCCTGGGGCGGATGC-3'
CaMyo51D	
3'RACE	1ST PCR: 5'-TGCGAGAGCATGAAGCGTGAC-3' 2nd PCR: 5'-CGCGAAGGGATTGAACGCAAC-3'
5' RACE	1 st PCR: 5'-GATACGGTTGACGACCGCATC-3' 2 nd PCR: 5'- GTCGCCGACGTTCTTGGGCTC-3'
Full-length gene model	sense: 5'-AAACAACAATTACAACAACCTACATGCCAAAGGAAAAGGAC-3' antisense: 5'- CCTCGCCCTTGCTCACCATTCTAGAATCGGAATCAGAGTCCGAG-3'
CaMyo51E	
3'-end	manual annotation
5' RACE	1 st PCR: 5'-GATACCTGTTGCTTCCTTTTG-3'

	2 nd PCR: 5'-TCGCAGCAGCGCTCTTGTTGC-3'
CaMyo52	
3'RACE	1 st PCR: 5'-CAGGAATGCAAACGACTT-3' 2 nd PCR: 5'-CGCGTTGCACCAATATCA-3'
5' RACE	1 st PCR: 5'-TTGTCTACCTGCTCTTCCG-3' 2 nd PCR: 5'-GCACGACATTGCCATCCTC-3'
CaMyo53	
3'RACE	1 st PCR: 5'-CAGCAACAACAGCAGCAA-3' 2 nd PCR: 5'-GAGAACGGAGTGCCAAAG-3'
5' RACE	1 st PCR: 5'-ATTTCTGGTTCTCCAAGC-3' 2 nd PCR: 5'-GCTTGAAGGGGTTGATGG-3'
CaMyo47	
3'RACE	1 st PCR: 5'-GGCATTTTGGGACGACAT-3' 2 nd PCR: 5'-CGATATTTTGCCGTCAC-3'
5' RACE	1 st PCR: 5'-GAAACCAGGATCGATTGG-3' 2 nd PCR: 5'-GCATCGTAGCTGAGGGAGA-3'
CaMyoI	
3'RACE	1 st PCR: 5'-CGTCGATGAGGATGAGAT-3' 2 nd PCR: 5'-CGAAGATGTGAAGGCCAA-3'
5' RACE	1 st PCR: 5'-GGATGGATTGAGATTTGC-3' 2 nd PCR: 5'-CGATGGATGAAAGATGCC-3'
CaMyoJ	
3'RACE	1 st PCR: 5'-CAAGGAAGGCAAGGAGCA-3' 2 nd PCR: 5'-GTCATGAATGGGGCAACC-3'
5' RACE	1 st PCR: 5'-AGAATGGCCTGGTTGCAG-3' 2 nd PCR: 5'-CTTGTCCCAGTAGCGCTT-3'
CaMyo29	
3'RACE	1 st PCR: 5'-ATGAGCCGCGTGTCCATGAAC-3' 2 nd PCR: 5'-GTCTCAAGAAGGCCTTCTGCG-3'
5'-end	truncated, manual annotation
CaMyoL	
3'RACE	1 st PCR: 5'-CAACGATTTGCGAGAGGG-3' 2 nd PCR: GAAACAGTCAAGGCCGAAG-3'
5' RACE	1 st PCR: 5'-GATTGGTTGACTTGCACG-3' 2 nd PCR: 5'-TGGAGATGGGTCAAGGTG-3'

References

1. Zackova Suchanova, J. *et al.* Diatom adhesive trail proteins acquired by horizontal gene transfer from bacteria serve as primers for marine biofilm formation. *New Phytol.* **240**, 770–783 (2023).
2. Priyam, A. *et al.*, Sequenceserver: A Modern Graphical User Interface for Custom BLAST Databases, *Molecular Biology and Evolution*, **36**, 2922–2924(2019).
3. Keeling PJ, Burki F, Wilcox HM, Allam B, Allen EE, et al. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing. *PLOS Biology* **12**, e1001889. (2014)
4. Osuna-Cruz, C.M., Bilcke, G., Vancaester, E. *et al.* The *Seminavis robusta* genome provides insights into the evolutionary adaptations of benthic diatoms. *Nat Commun* **11**, 3320 (2020).