Cation and anion channelrhodopsins: Sequence motifs and taxonomic distribution

Elena G. Govorunova[1], Oleg A. Sineshchekov[1], Hai Li[1], Yumei Wang[1], Leonid S. Brown[2], Alyssa Palmateer[2], Michael Melkonian[3], Shifeng Cheng[4], Eric Carpenter[5], Jordan Patterson[5], Gane K.-S. Wong[5,6], and John L. Spudich[1#]

[1]Center for Membrane Biology, Department of Biochemistry & Molecular Biology, The University of Texas Health Science Center at Houston McGovern Medical School, Houston, Texas, USA

[2]Department of Physics and Biophysics Interdepartmental Group, University of Guelph, Guelph, Ontario, Canada

[3]Max Planck Institute for Plant Breeding Research, Integrative Bioinformatics, Cologne, Germany

[4]Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China

[5]Departments of Biological Sciences and of Medicine, University of Alberta, Edmonton, Alberta, Canada

[6]Beijing Genomics Institute-Shenzhen, Shenzhen, China

Running title: Cation and anion channelrhodopsins

[#]Address correspondence to John L. Spudich, john.l.spudich@uth.tmc.edu.

Word count Abstract: 227

Word count Text: 4,360

ABSTRACT

Cation and anion channelrhodopsins (CCRs and ACRs, respectively) primarily from two algal species, *Chlamydomonas reinhardtii* and *Guillardia theta*, have become widely used as optogenetic tools to control cell membrane potential with light. We mined algal and other protist polynucleotide sequencing projects and metagenomic samples to identify 75 channelrhodopsin homologs from three channelrhodopsin families, including one revealed in dinoflagellates in this study. We carried out electrophysiological analysis of 33 natural channelrhodopsin variants from different phylogenetic lineages and 10 metagenomic homologs in search of sequence determinants of ion selectivity, photocurrent desensitization, and spectral tuning in channelrhodopsins. Our results show that association of a reduced number of glutamates near the conductance path with anion selectivity depends on a wider protein context, because prasinophyte homologs with the identical glutamate pattern as in cryptophyte ACRs are cation-selective. Desensitization is also broadly context-dependent, as in one branch of stramenopile ACRs and their metagenomic homologs its extent roughly correlates with phylogenetic relationship of their sequences. Regarding spectral tuning, two prasinophyte CCRs exhibit red-shifted spectra to 585 nm, although their retinal-binding pockets do not match those of previously known similarly red-shifted channelrhodopsins. In cryptophyte ACRs we identified three specific residue positions in the retinal-binding pocket that define the wavelength of their spectral maxima. Lastly, we found that dinoflagellate rhodopsins with a TCP motif in the third transmembrane helix and a metagenomic homolog exhibit channel activity.

IMPORTANCE

Channelrhodopsins are widely used in neuroscience and cardiology as research tools and are considered as prospective therapeutics, but their natural diversity and mechanisms remain poorly characterized. Genomic and metagenomic sequencing projects are producing an ever-increasing wealth of data, whereas biophysical characterization of the encoded proteins lags behind. In this study we used manual and automated patch clamp recording of representative members of four channelrhodopsin families including a family that we report in this study in dinoflagellates. Our results contribute to a better understanding of molecular determinants of ionic selectivity, photocurrent desensitization, and spectral tuning in channelrhodopsins.

INTRODUCTION

Channelrhodopsins (ChRs) are light-gated ion channels initially discovered in chlorophyte algae, in which they serve as photoreceptors guiding phototactic orientation (1-3). Subsequently, ChRs have also been found in the genomes/transcriptomes of cryptophyte and haptophyte algae (4, 5), the heterotrophic protists known as Labyrinthulea (5) and giant viruses that infect marine microorganisms (6, 7). Ongoing polynucleotide sequencing projects provide a rich hunting ground for further exploration of ChR diversity and taxonomic distribution.

Functionally, ChRs are divided into cation and anion channelrhodopsins (CCRs and ACRs, respectively) (8). Both ChR classes serve for photocontrol of excitable cells, such as neurons and cardiomyocytes, via a biotechnique known as optogenetics (9, 10). However, structural determinants for cation and anion selectivity in ChRs remain poorly understood. X-ray crystal structures (11-15) indicate that the ion conductance path in algal ChRs is formed by transmembrane helices (TM) 1, 2, 3 and 7. All so far known ACRs contain a non-carboxylate residue in the position of the protonated Schiff base counterion in bacteriorhodopsin (Asp85), whereas in nearly all CCRs this carboxylate is conserved. However, this sequence feature cannot be regarded as a sole indicator of anion selectivity, because some chlorophyte CCRs also show a non-carboxylate residue in the counterion position (e.g. *Ds*ChR1 from *Dunaliella salina* (16)).

Most chlorophyte CCRs contain five Glu residues in TM2 and the TM2-TM3 loop (Glu82, Glu83, Glu90, Glu97, and Glu101 in ChR2 from *Chlamydomonas reinhardtii, Cr*ChR2), whereas in all so far known ACRs most or even all of the corresponding positions are occupied with non-carboxylate residues. Therefore, it has been proposed that negative electrostatic potential of the channel pore defines cation selectivity (17, 18). Indeed, mutagenetic remodeling of the pore to reduce electronegativity yielded permeability for anions in chlorophyte CCRs (17-21). Yet, some of the TM2 glutamates are conserved in ACRs and apparently do not interfere with their anion conductance.

Other biophysical properties of ChRs relevant for optogenetic applications are their desensitization under continuous or pulsed illumination (also called “inactivation” in the literature) and spectral sensitivity. In an earlier study, a group of ACRs discovered in the TARA marine transcriptomes demonstrated particularly rapid and strong desensitization (22). As their source organisms were not known, these proteins were named MerMAIDs (Metagenomically discovered, Marine, Anion-conducting and Intensely Desensitizing channelrhodopsins). However, strong desensitization cannot serve as a characteristic of a single ChR family, because it was also observed in some “bacteriorhodopsin-like” CCRs (BCCRs) from cryptophytes that show very little sequence homology with MerMAIDs (23).

To gain more insight into the taxonomic distribution and structure-function relationships of ChRs, we identified 75 ChR homologs from several phylogenetic lineages and metagenomic samples, and tested 27 of them along with 16 previously reported sequences by heterologous expression in cultured mammalian cells followed by patch clamp recording. We show that the same pattern of conserved Glu residues may accompany cation or anion conductance in ChRs from different taxa, and that the degree of desensitization in MerMAID homologs is the greater, the closer their sequences are to those of the first reported MerMAIDs. We report two prasinophyte CCRs with red-shifted spectra and confirm that three specific residues in the retinal-binding pocket are responsible for wavelength regulation in cryptophyte ACRs. Finally, we demonstrate that some dinoflagellate rhodopsins possess channel activity.

RESULTS

Prasinophyte CCRs

Only a few of >150 so far identified chlorophyte ChRs (Fig. 1, and Data Sets S1 and S2) have been tested by heterologous expression. Both *C. reinhardtii* ChRs conduct cations (2, 3), so other chlorophyte ChRs were also assumed to be CCRs. However, a recent study has demonstrated that two ChRs from the prasinophyte genus *Pyramimonas* in fact conduct anions (6), which called for a more detailed functional analysis of chlorophyte ChRs.

Three ChR homologs derived from the prasinophytes *Crustomastix stigmatica*, *Mantoniella squamata* and *Pyramimonas melkonianii* (6) exhibit a residue pattern typical of cryptophyte ACRs, i.e. they display conserved Glu82 and Glu90 with non-carboxylate residues in the positions of Glu83, Glu97, Glu101 and Glu123 of *Cr*ChR2 (Fig. 2A). In a *Cymbomonas tetramitiformis* sequence (6), Glu90 and Glu97 are conserved, whereas Glu82 is replaced with Gln (Fig. 2A). We synthesized mammalian codon-adapted versions of these rhodopsin domains, fused them with C-terminal enhanced yellow fluorescent protein (EYFP), expressed in HEK293 (human embryonic kidney) cells and analyzed by manual whole-cell patch clamp.

Three of these ChRs generated photocurrents (Fig. 2B and C) in our standard buffer system (for solution compositions see Table S1), whereas the homolog from *P. melkonianii* that we named *Pyme*ChR was non-electrogenic. The action spectra of photocurrents generated by the *M. squamata* and *C. stigmatica* homologs were red-shifted (the rhodopsin maxima at ~580 and 585 nm, respectively; Fig. 2D). Their retinal-binding pockets are nearly identical, but differ from those of previously known red-shifted ChRs (Fig. S1A). Both spectra exhibited a second band at ~520 nm that reflected a Förster resonance energy transfer (FRET) from EYFP to rhodopsin, as was earlier shown in RubyACRs from Labyrinthulea (5). The efficiency of FRET was even greater in the homolog from *C. stigmatica*, the rhodopsin peak of which was observed in the green spectral region and could not be accurately resolved because of the FRET contribution (Fig. 2D, olive line).

To test the relative permeability of the prasinophyte homologs for H^+, Na^+ and Cl^- we varied the concentration of each of these ions in the bath (for solution compositions see Table S1), measured the current-voltage relationships and determined the reversal potentials (E_{rev}). *Cr*ChR2 was included in this experiment for comparison. Fig. 2E shows that under all tested conditions E_{rev} for all three homologs was close to the equilibrium potential of H^+, indicating that they are H^+-selective channels with negligible permeability for Na^+ and Cl^-. We named them *Ct*CCR, *Ms*CCR, and *Cs*CCR. A less positive E_{rev} of *Ms*CCR photocurrents probed under the H^+ gradient does not result from the permeability for Na^+ as it does in *Cr*ChR2, and most likely reflects a contribution of intramolecular charge transfers, as previously found in other CCRs (24).

Four sequences from Chlorophyceae have only the Glu82 homolog, as do prasynophyte ACRs (Fig. S1B), but show no close sequence homology to them. Four sequences from

Chlorodendrophyceae contain no glutamate residues in any of the six analyzed positions (Fig. S1D) and form a separate branch on the phylogenetic tree (Fig. 1). Very unusually, in the latter sequence group the Asp residue corresponding to Asp212 of bacteriorhodopsin is located not four residues upstream, as in most known microbial rhodopsins, but three residues upstream of the Schiff base lysine (Fig. S1E). None of these eight proteins, nor a ChR homolog from the streptophyte *Coleochaete* generated channel currents (see Supplemental Text and Fig. S1C for details).

Stramenopile ACRs and their metagenomic homologs

The first MerMAIDs reported were seven homologous ACRs identified in metagenomic samples (22). Recently, close homologs were found in unclassified stramenopile species (5, 25), which suggests that the original MerMAIDs also originate from stramenopiles. We have identified 20 additional MerMAID homologs, nine haptophyte ACR homologs and two Labyrinthulea ACR homologs (Data Set S1) in metagenomic databases (Data Set S4). We tested EYFP fusions of five metagenomic MerMAID homologs (abbreviated by the lower case "mg" in the protein name), two closely related sequences from the unclassified stramenopile strain TOSAG23-3 (abbreviated by "*sT*" (5)), and three sequences from the bicosoecid stramenopile *Cafeteria roenbergensis* (abbreviated by "*Car*" to distinguish them from *C. reinhardtii* ChRs). In most sequences of this group both Glu82 and Glu90 (*Cr*ChR2 numbering) are conserved, as in the previously known cryptophyte ACRs and MerMAIDs (Fig. 3A). The five tested MerMAID homologs and those from TOSAG23-3 clustered together with the first reported MerMAIDs (Fig. S2), whereas Cafeteria homologs formed a separate branch related to haptophyte ACRs (Fig. 1). Each of these homologs generated photocurrents in HEK293 cells (Fig. 3B). As shown below, these ChRs conduct anions, so we designated them ACRs.

Out of all tested homologs, *sT*ACR2 is the most closely related to the first reported MerMAIDs, which exhibit nearly complete desensitization (Fig. S2). Similarly, *sT*ACR2 photocurrents showed nearly complete desensitization (Fig. 3C, cyan), whereas photocurrents from *sT*ACR1, the most distant homolog (Fig. S2), exhibited only ~40% desensitization (Fig. 3C, dark yellow). The values for other homologs were intermediate (Fig. 3C). All ChRs of this group demonstrated exclusively anion permeability (Fig. 3D). The action spectra of their photocurrents are shown in Fig. 3E and F. The shape of some spectra (e.g. mgACR2 and mgACR5) indicated a contribution of FRET from EYFP.

Cryptophyte ACRs

Cryptophytes are the taxon in which the first natural ACRs were discovered (4). To explore the diversity of cryptophyte ACRs further, we have analyzed transcriptomes of 20 additional cryptophyte strains (Table S2) and identified 15 transcripts homologous to previously known cryptophyte ACRs (Data Set S1). As no species names have been assigned to their source organisms, we used the numbers 3-8 in the abbreviated protein names to designate different *Rhodomonas* strains (the numbers 1 and 2 have already been assigned to the previously

analyzed strains). The Glu82 homologs is conserved in all, and the Glu90 homolog, in most of these proteins, whereas all other analyzed positions are occupied with non-carboxylate residues (Fig. 4A). Thirteen homologs generated photocurrents upon expression in HEK293 cells (Fig. 4B and C).

To verify permeability for anions in the three cryptophyte ACR homologs that were well-expressed and generated photocurrents in the nA range by manual patch clamp (Fig. 4B), we used a high-throughput automated patch clamp (APC) instrument, SyncroPatch 384i, with solutions provided by the manufacturer (for their full compositions see Table S1). The internal solution was predominantly CsF to promote formation of gigaseals, and the external solution was predominantly NaCl. Representative series of photocurrent traces recorded from *R3*ACR1 under incremental voltage using the SyncroPatch 384i and AxoPatch 200B with the same solutions are shown in Figs. 4D and E. *Gt*ACR1 and *Cr*ChR2, well-characterized by manual patch clamp, were included in the SyncroPatch experiment as ACR and CCR controls, respectively. With the SyncroPatch solutions, the E_{rev} of *Gt*ACR1 photocurrents was negative, whereas that of *Cr*ChR2 was positive (Fig. 4D). In all tested homologs the E_{rev} was close to that of *Gt*ACR1 (Fig. 4D), which confirmed their anion selectivity.

Previously, we and others demonstrated that Cys133, Ser156 and Tyr207 in *Gt*ACR1 (absorption maximum 515 nm) corresponding to Arg129, Gly152 and Phe203 in *Gt*ACR2 (absorption maximum 470 nm) define the spectral difference between these two proteins (15, 26)(E.G. Govorunova, O.A. Sineshchekov and J.L. Spudich, manuscript in preparation). According to *Gt*ACR1 crystal structures, the side chains of Cys133 and Ser156 are located near the β-ionone ring of the chromophore (Fig. 5A), whereas the hydroxyl group of Tyr207 forms a hydrogen bond with Asp234 in the photoactive center. Comparative analysis of these positions (Fig. 5C) and action spectra of photocurrents (Fig. 5B and D) in the 13 functional ACR homologs has revealed that only those proteins in which the residues match those of *Gt*ACR2 exhibit blue-shifted absorption maxima. When Cys or Met are found at position 133, or Ser or Ala at position 156, the spectrum is shifted to longer wavelengths.

Dinoflagellate ChRs

Dinoflagellates exhibit genuine phototactic orientation (27-29), and their genomes encode multiple type I rhodopsins (30-32). However, to the best of our knowledge, none of these rhodopsins has so far been reported to exhibit channel function. Some rhodopsin sequences from dinoflagellates of the genera *Ansanella, Pelagodinum,* and *Symbiodinium* (6, 33-36) contain the TCP motif in the middle of TM3 that is conserved in most so far known channelrhodopsins (Fig. S3A). This motif is also conserved in 17 proteins encoded by the deep-ocean TARA marine transcriptomes that cluster together with these dinoflagellate rhodopsins and form a distinct branch of the phylogenetic tree (Fig. 1). A very unusual feature of this entire sequence cluster is that Asp212 of bacteriorhodopsin, highly conserved in all so far known channelrhodopsins, is replaced with Asn or, in one homolog, Leu (Fig. S3B).

Only one of the five tested metagenomic rhodopsin domains of this group was electrogenically photoactive upon expression in HEK293 cells, producing photocurrents barely resolved from the noise level (Fig. 6A, black bar). The fusion protein formed disk-shaped fluorescent aggregates within the cells (Fig. S3C, top). The addition of the trafficking signal (TS) between rhodopsin and EYFP, and the endoplasmic reticulum export motif (ER) at the C terminus of the fusion protein (37) reduced formation of the aggregates (Fig. S3C, bottom) and significantly increased the photocurrents, although they still reached only ~20-pA level at best (Fig. 6A, blue bar). In our standard buffer system with nearly symmetrical ionic concentrations in the bath and pipette the sign of the photocurrents reversed at positive voltages indicating passive ion transport (Fig. 6B, top). We named this protein mgdChR1 (for metagenomic dinoflagellate homolog ChannelRhodopsin 1).

A homologous rhodopsin domain from the coral endosymbiont *Symbiodinium microadriaticum* has a ~300-residue N-terminal extension (Fig. S4), which is much longer than that found in other known ChRs, including mgdChR1. An expression construct encoding residues 1-600 produced no tag fluorescence. However, when the N-terminal extension was deleted, and TS and ER export motifs added, fluorescence was observed, and passive photocurrents of a similarly small amplitude as from mgdChR1 were recorded (Fig. 6B, bottom). We named this protein Δ*Sm*ChR1 to emphasize truncation of the N-terminal extension. A homologous protein Δ*Sm*ChR2 from the same organism also generated photocurrents, but their amplitudes were even smaller. Channel currents from Δ*Sm*ChR1, but not from mgdChR1, were preceded with a fast negative current, the sign of which did not reverse at positive voltages (Fig. 6B, bottom). Such currents have previously been recorded from several other ChRs and interpreted as intramolecular charge displacement associated with isomerization of the retinal chromophore (24). The photocurrent action spectra of mgdChR1 and Δ*Sm*ChR1 peaked in the green spectral region (Fig. 6C). Small amplitudes of mgdChR1 and Δ*Sm*ChR1 photocurrents make accurate measurements of the reversal potentials problematic, so we were not able to determine their ionic selectivity.

DISCUSSION

We report functional testing of 43 ChR homologs from prasinophytes, stramenopiles, cryptophytes, dinoflagellates and metagenomic samples. An unexpected result is that ACRs appear to be more widely spread among protist taxa than CCRs. Another unexpected result is that the same residue pattern comprising conserved Glu82 and Glu90 with non-carboxylate residues in the positions of Glu83, Glu97, Glu101, and Glu123 (*Cr*ChR2 numbering) found in most ACRs from stramenopiles, cryptophytes, haptophytes, and metagenomic samples is also found in the CCRs, *Cs*CCR and *Ms*CCR, from prasinophytes. Out of five Glu residues in TM2 and TM2-TM3 loop, Glu82 is most conserved across the entire ChR family. According to our empirical calculations using PROPKA3 (38), the pK_a of the Glu82 homolog is acidic in all so far published X-ray crystal structures of ChRs, including that of *Gt*ACR1 in which it apparently does

not prevent anion conductance. In *Cr*ChR2, replacement of Glu82 with Ala strongly inhibited expression in mammalian cells judged by the tag fluorescence and correspondingly reduced photocurrents (39), which suggests that this residue is needed for correct protein folding and/or membrane targeting.

Glu90 appears to be essential for cation conductance in *Cr*ChR2, as mutation of this residue to Lys or Arg confers permeability for anions (19). Yet, this Glu is conserved in most ACRs except those from prasinophytes and Labyrinthulea. In the unphotolyzed state of both *Cr*ChR2 (40) and *Gt*ACR1 (41) this residue is neutral at neutral pH. Glu90 deprotonates during the photocycle of *Cr*ChR2 (42, 43). Photoinduced protonation changes of the Glu90 homolog in *Gt*ACR1 (Glu68) have not been studied by time-resolved molecular spectroscopy, but electrophysiological and UV-vis flash-photolysis data indicate that it also deprotonates upon photoexcitation (44). Further research is needed to clarify the role of this residue in anion conductance.

Photocurrent desensitization in different ChR families correlates with accumulation of different intermediates of the photocycle. In MerMAID1 desensitization is correlated with the M intermediate (22), but in *Rhodomonas* BCCRs it is with a novel extremely blue-shifted intermediate (23). Finally, desensitization in *Cr*ChR2 is correlated with accumulation of blue-absorbing P480 that is considered either as a late intermediate in a single branched photocycle (45) or the initial state of a parallel photocycle (43). Desensitization was reduced in the E44Q and C84T mutants of MerMAID1 (22). However, the mutated residues (corresponding, respectively, to Glu90 and Cys128 of *Cr*ChR2) are not the sole cause of strong desensitization in MerMAIDs, because they are conserved in many ChRs that do not show strong desensitization, including the closely related *sT*ACR1 characterized here.

According to quantum mechanical/molecular mechanical calculations using the *Gt*ACR1 crystal structure, replacement of Ser156 with Gly or Ala stabilizes S_0, predicting a 11-12-nm blue shift of the absorption maximum (46). All tested cryptophyte ACRs that contain Gly in this position exhibited blue-shifted spectra. The spectra of two ACRs that contain Ala in this position (R7ACR1 and R8ACR3) were ~25 nm blue-shifted from that of *Gt*ACR1, whereas the spectra of the other two (C2ACR and R8ACR2) were very similar to that of *Gt*ACR1, suggesting that the expected phenotypic effect of Ser to Ala substitution in these proteins was compensated for by other changed residues.

Our results and those of other groups suggest that most biophysical properties of ChRs relevant for their optogenetic applications cannot be assigned to a few individual residues, but rather reflect interactions between many of them. A cumulative larger set of electrophysiological data to which our study contributes might be used in the future to train machine learning algorithms to identify sequence motifs that define ionic selectivity, desensitization and absorption spectra. Implementation of such algorithms has already helped to improve plasma membrane targeting and light sensitivity of ChRs (47, 48).

Protein sequences of dinoflagellate ChRs and their metagenomic homologs are distantly related to ChRs from giant viruses (Fig. 1), two of which have been shown recently to passively conduct cations upon heterologous expression (7). However, Asp212 of bacteriorhodopsin is conserved in these viral CCRs, as in most other known microbial rhodopsins, whereas it is replaced with Asn in dinoflagellate ChRs. Analysis of *Symbiodinium* transcriptomes reveals potential latent infection by large dsDNA viruses (49), so viral origin of dinoflagellate ChRs cannot be excluded.

So far, the function of ChRs as photoreceptors guiding phototaxis has been verified directly only in the chlorophyte *C. reinhardtii*, the model organism for which methods of gene silencing and knockdown have been developed (1, 50). Several other chlorophyte and one cryptophyte species have been shown to generate photoreceptor currents, very similar to those in *C. reinhardtii* and likely resulting from ChR photoexcitation (51-54). The direction of photoreceptor currents recorded in both freshwater and marine flagellates is depolarizing, which reflects cation influx or anion efflux. Both *C. reinhardtii* phototaxis receptors are CCRs (2, 3), but ACRs might also contribute to depolarizing photoreceptor currents even in marine flagellates, if their membrane potential is sufficiently low. To the best of our knowledge, the membrane potential has not been estimated in any ACR-containing organism, but it is -170 mV in the giant marine unicellular alga *Acetabularia mediterranea* (55).

Based on the action spectra of dinoflagellate phototaxis, rhodopsins have been suggested as photoreceptors that guide this behavior (56). Our demonstration of channel activity in dinoflagellate rhodopsins with the TCP motif in TM3 strongly supports this hypothesis. The spectral sensitivity of dinoflagellate ChRs matches that of phototactic accumulation observed in *Symbiodinium* and unclassified coral symbiotic dinoflagellates (57, 58). The latter studies suggest that coral larvae use GFP fluorescence to attract dinoflagellate symbionts that are necessary for their survival.

Manual patch clamp is time-consuming and requires considerable skill. We sought to test whether APC can be used for characterization of hundreds of ChR variants that evolved in various protists. The planar-array principle implemented in the SyncroPatch 384i allows seal formation on micron-size orifices in the glass bottom of microwell plates (chips) into which cell suspension is pipetted, thus bypassing pipette fabrication and offering the option for recording multiple cells in parallel (59). APC is mostly used for drug screening, especially cardiac safety testing, in stably transfected cell lines. However, generation of such lines for ChR screening is not practical. We found that even upon chemical transfection that yielded only 30-70% visibly fluorescent cells depending on the construct, using the SyncroPatch 384i considerably sped up data collection, as compared to manual patch clamp.

MATERIALS AND METHODS

Bioinformatics

To identify metagenomic homologs of MerMAIDs, haptophyte ACRs and Labyrinthulea ACRs, we first searched selected datasets of the Integrated Microbial Genomes and Microbiomes at the Department of Energy's Joint Genome Institute (JGI) (Data Set S4) using the keyword "rhodopsin", and then performed blastp (protein-protein BLAST) search using RubyACR sequences as a query. A similar procedure was used to identify rhodopsin genes in the dinoflagellate genomes from various sources listed in Data Set S4. *Cafeteria roenbergensis* ChRs were identified in the National Center for Biotechnology Information (NCBI) non-redundant protein database using blastp and *Gt*ACR1 sequence as a query.

To explore the diversity of cryptophyte ACRs, we analyzed transcriptomes of 20 cryptophyte strains each sequenced on the Illumina HiSeq 2000 platform and assembled with the Bridger algorithm (60). Using a hidden Markov model (HMM) (61) based on known cryptophyte ACRs, we identified 15 novel transcripts for experimental characterization. We also analyzed 136 deep-ocean metatranscriptomic libraries from the TARA Oceans Expedition (62) assembled with the Plass protein-level algorithm (63). Four distinct HMMs were built using previously known sequences of cryptophyte ACRs, cryptophyte BCCRs, chlorophyte CCRs and MerMAIDs. While many transcripts could be uniquely assigned to one of these four HMMs, some aligned weakly but equally well to two or more HMMs, and could not be assigned unambiguously. Remarkably, 17 of these "ambiguous" sequences turned out to be close homologs of dinoflagellate ChRs that were not included among our HMMs.

Rhodopsin sequences from Data Set S1 were aligned using MUSCLE with default parameters as implemented in MegAlign Pro software v. 17.1.1 (DNASTAR Lasergene, Madison, WI) and truncated after the end of TM7. Phylogeny was analyzed with IQ-TREE v. 2.1.2 (64) using automatic model selection and ultrafast bootstrap approximation (1000 replicates) (65). The best tree was visualized and annotated with iTOL v. 5.7 (66).

Molecular biology and HEK293 transfection

DNA polynucleotides encoding the opsin domains optimized for human codon usage were synthesized and cloned by GenScript (Piscataway, NJ) into the mammalian expression vector pcDNA3.1 (Life Technologies, Grand Island, NY) in frame with an EYFP tag for expression in HEK293 cells. The cells were transfected using the ScreenFectA transfection reagent (Waco Chemicals USA, Richmond, VA). All-*trans*-retinal (Sigma) was added at the final concentration of 3 µM immediately after transfection.

Manual patch clamp recording

Photocurrents were recorded 48-96 h after transfection in whole-cell voltage clamp mode with an AxoPatch 200B amplifier and digitized with a Digidata 1440A using pClamp 10 software (all from Molecular Devices, Union City, CA). Patch pipettes with resistances of 2-4 MΩ were fabricated from borosilicate glass. The ionic compositions of the bath and pipette solutions are shown in Table S1. For determination of E_{rev}, K^+ in the pipette solution was replaced with Na^+ to minimize the number of ionic species in the system, and the holding voltages were corrected

for liquid junction potentials calculated using the Clampex built-in calculator. Continuous light pulses were provided by a Polychrome V (T.I.L.L. Photonics GMBH, Grafelfing, Germany) in combination with a mechanical shutter (Uniblitz Model LS6, Vincent Associates, Rochester, NY; half-opening time 0.5 ms). The maximal photon density at the focal plane of the 40× objective was 5.2-8.5 mW mm^{-2} depending on the wavelength. The action spectra were constructed by calculation of the initial slope of photocurrent and corrected for the photon density measured at each wavelength (5). Further analysis was performed using Origin Pro software (OriginLab Corporation, Northampton, MA). The images were taken with a CoolSNAP HQ2 monochrome camera (Photometrics, Tucson, AZ).

Automated patch clamp recording

Automated patch clamp recording was conducted with a SyncroPatch 384i (Nanion Technologies) using planar borosilicate glass medium-resistance chips in a 384 microtiter plate format with one or four holes per well and Nanion Standard solutions (for their composition see Table S1). Transfected cells were dissociated using TrypLE™ Express, diluted with CHO-S-SFM-II medium (both from ThermoFisher) and resuspended in External Physiological solution (Nanion Technologies) at 10^5-4×10^5 cells ml^{-1}. Each well was filled with 30 µl Chip Fill solution, to which 20 µl of the cell suspension was added. Seal formation was enhanced by the addition of 40 µl of NMDG 60 solution with 10 mM $CaCl_2$ (final concentration). After capturing the cells, 50 µl of the external solution was replaced with 40 µl of NMDG 60 solution, and 40 µl of the mixture was removed. For data acquisition and analysis, respectively, PatchControl384 and DataControl384 software v. 1.9.0 was used (both Nanion Technologies). Illumination was provided with LUXEON Z blue LEDs LXZ1-PB01 (470 ± 20 nm) controlled by custom-built software.

Statistics

Descriptive statistics was used as implemented in Origin software. The data are presented as mean ± sem values; the data from individual replicates are also shown when appropriate. The sample size was estimated from previous experience and published work on similar subjects, as recommended by the NIH guidelines. No normal distribution of the data was assumed; when a specific statistics hypothesis was tested, non-parametric tests were used.

Data availability

The polynucleotide sequences of ChR homologs reported in this study have been deposited to GenBank (accession numbers MW557552-MW557594).

Acknowledgements

We thank Dr. Tim Strassmaier, Leo Angelo Morada, Stephan Holzhauser and Rasmus Gönner (Nanion Technologies) for their expert help with the SyncroPatch 384i.

Funding

This work was supported by National Institutes of Health Grants R01GM027750 and U01NS118288, and Endowed Chair AU-0009 from the Robert A. Welch Foundation to J.L.S, and by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant RGPIN-2018-04397 to L.S.B. A.P. was supported by a NSERC USRA scholarship, S.C. was supported by the Agricultural Science and Technology Innovation Program (ASTIP). The access to the SyncroPatch 384i was provided by a research grant from Nanion Technologies. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest

The authors declare no conflict of interest.

REFERENCES

1. Sineshchekov OA, Jung K-H, Spudich JL. 2002. Two rhodopsins mediate phototaxis to low- and high-intensity light in *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA 99:8689-8694.
2. Nagel G, Ollig D, Fuhrmann M, Kateriya S, Musti AM, Bamberg E, Hegemann P. 2002. Channelrhodopsin-1: a light-gated proton channel in green algae. Science 296:2395-8.
3. Nagel G, Szellas T, Huhn W, Kateriya S, Adeishvili N, Berthold P, Ollig D, Hegemann P, Bamberg E. 2003. Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. Proc Natl Acad Sci USA 100:13940-5.
4. Govorunova EG, Sineshchekov OA, Liu X, Janz R, Spudich JL. 2015. Natural light-gated anion channels: A family of microbial rhodopsins for advanced optogenetics. Science 349:647-650.
5. Govorunova EG, Sineshchekov OA, Li H, Wang Y, Brown LS, Spudich JL. 2020. RubyACRs, non-algal anion channelrhodopsins with highly red-shifted absorption. Proc Natl Acad Sci USA 117:22833-22840.
6. Rozenberg A, Oppermann J, Wietek J, Fernandez Lahore RG, Sandaa RA, Bratbak G, Hegemann P, Béjà O. 2020. Lateral gene transfer of anion-conducting channelrhodopsins between green algae and giant viruses. Curr Biol 30:4910-4920.
7. Zabelskii D, Alekseev A, Kovalev K, Rankovic V, Balandin T, Soloviov D, Bratanov D, Savelyeva E, Podolyak E, Volkov D, Vaganova S, Astashkin R, Chizhov I, Yutin N, Rulev M, Popov A, Eria-Oliveira AS, Rokitskaya T, Mager T, Antonenko Y, Rosselli R, Armeev G, Shaitan K, Vivaudou M, Buldt G, Rogachev A, Rodriguez-Valera F, Kirpichnikov M, Moser T, Offenhausser A, Willbold D, Koonin E, Bamberg E, Gordeliy V. 2020. Viral rhodopsins 1 are an unique family of light-gated cation channels. Nat Commun 11:5707.
8. Govorunova EG, Sineshchekov OA, Li H, Spudich JL. 2017. Microbial rhodopsins: Diversity, mechanisms, and optogenetic applications. Annu Rev Biochem 86:845-872.
9. Deisseroth K. 2015. Optogenetics: 10 years of microbial opsins in neuroscience. Nat Neurosci 18:1213-1225.

10. Sasse P, Funken M, Beiert T, Bruegmann T. 2019. Optogenetic termination of cardiac arrhythmia: Mechanistic enlightenment and therapeutic application? Front Physiol 10:675.
11. Kato HE, Zhang F, Yizhar O, Ramakrishnan C, Nishizawa T, Hirata K, Ito J, Aita Y, Tsukazaki T, Hayashi S, Hegemann P, Maturana AD, Ishitani R, Deisseroth K, Nureki O. 2012. Crystal structure of the channelrhodopsin light-gated cation channel. Nature 482:369-374.
12. Volkov O, Kovalev K, Polovinkin V, Borshchevskiy V, Bamann C, Astashkin R, Marin E, Popov A, Balandin T, Willbold D, Buldt G, Bamberg E, Gordeliy V. 2017. Structural insights into ion conduction by channelrhodopsin 2. Science 358:eaan8862.
13. Oda K, Vierock J, Oishi S, Rodriguez-Rozada S, Taniguchi R, Yamashita K, Wiegert JS, Nishizawa T, Hegemann P, Nureki O. 2018. Crystal structure of the red light-activated channelrhodopsin Chrimson. Nat Commun 9:3949.
14. Kim YS, Kato HE, Yamashita K, Ito S, Inoue K, Ramakrishnan C, Fenno LE, Evans KE, Paggi JM, Dror RO, Kandori H, Kobilka BK, Deisseroth K. 2018. Crystal structure of the natural anion-conducting channelrhodopsin *Gt*ACR1. Nature 561:343-348.
15. Li H, Huang CY, Govorunova EG, Schafer CT, Sineshchekov OA, Wang M, Zheng L, Spudich JL. 2019. Crystal structure of a natural light-gated anion channelrhodopsin. Elife 8:e41741.
16. Zhang F, Vierock J, Yizhar O, Fenno LE, Tsunoda S, Kianianmomeni A, Prigge M, Berndt A, Cushman J, Polle J, Magnuson J, Hegemann P, Deisseroth K. 2011. The microbial opsin family of optogenetic tools. Cell 147:1446-57.
17. Berndt A, Lee SY, Ramakrishnan C, Deisseroth K. 2014. Structure-guided transformation of channelrhodopsin into a light-activated chloride channel. Science 344:420-424.
18. Wietek J, Beltramo R, Scanziani M, Hegemann P, Oertner TG, Wiegert JS. 2015. An improved chloride-conducting channelrhodopsin for light-induced inhibition of neuronal activity in vivo. Sci Rep 5:14807.
19. Wietek J, Wiegert JS, Adeishvili N, Schneider F, Watanabe H, Tsunoda SP, Vogt A, Elstner M, Oertner TG, Hegemann P. 2014. Conversion of channelrhodopsin into a light-gated chloride channel. Science 344:409-412.
20. Berndt A, Lee SY, Wietek J, Ramakrishnan C, Steinberg EE, Rashid AJ, Kim H, Park S, Santoro A, Frankland PW, Iyer SM, Pak S, Ahrlund-Richter S, Delp SL, Malenka RC, Josselyn SA, Carlen M, Hegemann P, Deisseroth K. 2016. Structural foundations of optogenetics: Determinants of channelrhodopsin ion selectivity. Proc Natl Acad Sci USA 113:822-829.
21. Wietek J, Rodriguez-Rozada S, Tutas J, Tenedini F, Grimm C, Oertner TG, Soba P, Hegemann P, Wiegert JS. 2017. Anion-conducting channelrhodopsins with tuned spectra and modified kinetics engineered for optogenetic manipulation of behavior. Sci Rep 7:14957.
22. Oppermann J, Fischer P, Silapetere A, Liepe B, Rodriguez-Rozada S, Flores-Uribe J, Peter E, Keidel A, Vierock J, Kaufmann J, Broser M, Luck M, Bartl F, Hildebrandt P, Simon Wiegert J, Beja O, Hegemann P, Wietek J. 2019. MerMAIDs: a family of metagenomically discovered marine anion-conducting and intensely desensitizing channelrhodopsins. Nat Commun 10:3315.

23. Sineshchekov OA, Govorunova EG, Li H, Wang Y, Melkonian M, Wong GK-S, Brown LS, Spudich JL. 2020. Conductance mechanisms of rapidly desensitizing cation channelrhodopsins from cryptophyte algae. mBio 11:e00657-20.
24. Sineshchekov OA, Govorunova EG, Wang J, Li H, Spudich JL. 2013. Intramolecular proton transfer in channelrhodopsins. Biophys J 104:807-817.
25. Labarre A, Lopez-Escardo D, Latorre F, Leonard G, Bucchini F, Obiol A, Cruaud C, Sieracki ME, Jaillon O, Wincker P, Vandepoele K, Logares R, Massana R. 2021. Comparative genomics reveals new functional insights in uncultured MAST species. ISME J doi:10.1038/s41396-020-00885-8.
26. Kojima K, Miyoshi N, Shibukawa A, Chowdhury S, Tsujimura M, Noji T, Ishikita H, Yamanaka A, Sudo Y. 2020. Green-sensitive, long-lived, step-functional anion channelrhodopsin-2 variant as a high-potential neural silencing tool. J Phys Chem Lett 11:6214-6218.
27. Forward RB. 1974. Phototaxis by the dinoflagellate *Gymnodinium splendens* Lebour. J Protozool 21:312-315.
28. Liu S-M, Häder D-P, Ullrich W. 1990. Photoorientation in the freshwater dinoflagellate, *Peridinium gatunense* Nygaard. FEMS Microbiol Lett 73:91-102.
29. Horiguchi T, Kawai H, Kubota M, Takahashi T, Watanabe M. 1999. Phototactic responses of four marine dinoflagellates with different types of eyespot and chloroplast. Phycol Res 47:101-107.
30. Slamovits CH, Okamoto N, Burri L, James ER, Keeling PJ. 2011. A bacterial proteorhodopsin proton pump in marine eukaryotes. Nat Commun 2:183.
31. Marchetti A, Schruth DM, Durkin CA, Parker MS, Kodner RB, Berthiaume CT, Morales R, Allen AE, Armbrust EV. 2012. Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. Proc Natl Acad Sci USA 109:E317-25.
32. Coesel SN, Durham BP, Groussman RD, Hu SK, Caron DA, Morales RL, Ribalet F, Armbrust EV. 2021. Diel transcriptional oscillations of light-sensitive regulatory elements in open-ocean eukaryotic plankton communities. Proc Natl Acad Sci USA 118.
33. Shoguchi E, Shinzato C, Kawashima T, Gyoja F, Mungpakdee S, Koyanagi R, Takeuchi T, Hisata K, Tanaka M, Fujiwara M, Hamada M, Seidi A, Fujie M, Usami T, Goto H, Yamasaki S, Arakaki N, Suzuki Y, Sugano S, Toyoda A, Kuroki Y, Fujiyama A, Medina M, Coffroth MA, Bhattacharya D, Satoh N. 2013. Draft assembly of the *Symbiodinium minutum* nuclear genome reveals dinoflagellate gene structure. Curr Biol 23:1399-408.
34. Aranda M, Li Y, Liew YJ, Baumgarten S, Simakov O, Wilson MC, Piel J, Ashoor H, Bougouffa S, Bajic VB, Ryu T, Ravasi T, Bayer T, Micklem G, Kim H, Bhak J, LaJeunesse TC, Voolstra CR. 2016. Genomes of coral dinoflagellate symbionts highlight evolutionary adaptations conducive to a symbiotic lifestyle. Sci Rep 6:39734.
35. Shoguchi E, Beedessee G, Tada I, Hisata K, Kawashima T, Takeuchi T, Arakaki N, Fujie M, Koyanagi R, Roy MC, Kawachi M, Hidaka M, Satoh N, Shinzato C. 2018. Two divergent *Symbiodinium* genomes reveal conservation of a gene cluster for sunscreen biosynthesis and recently lost genes. BMC Genomics 19:458.

36. Grigoriev IV, Hayes RD, Calhoun S, Kamel B, Wang A, Ahrendt S, Dusheyko S, Nikitin R, Mondo SJ, Salamov A, Shabalov I, Kuo A. 2021. PhycoCosm, a comparative algal genomics resource. Nucleic Acids Res 49:D1004-D1011.
37. Gradinaru V, Zhang F, Ramakrishnan C, Mattis J, Prakash R, Diester I, Goshen I, Thompson KR, Deisseroth K. 2010. Molecular and cellular approaches for diversifying and extending optogenetics. Cell 141:154-65.
38. Olsson MHM, Sondergaard CR, Rostkowski M, Jensen JH. 2011. PROPKA3: Consistent treatment of internal and surface residues in empirical pK(a) predictions. J Chem Theory Comput 7:525–537.
39. Sugiyama Y, Wang H, Hikima T, Sato M, Kuroda J, Takahashi T, Ishizuka T, Yawo H. 2009. Photocurrent attenuation by a single polar-to-nonpolar point mutation of channelrhodopsin-2. Photochem Photobiol Sci 8:328-36.
40. Ritter E, Stehfest K, Berndt A, Hegemann P, Bartl FJ. 2008. Monitoring light-induced structural changes of Channelrhodopsin-2 by UV-visible and Fourier transform infrared spectroscopy. J Biol Chem 283:35033-35041.
41. Yi A, Mamaeva NV, Li H, Spudich JL, Rothschild KJ. 2016. Resonance Raman study of an anion channelrhodopsin: Effects of mutations near the retinylidene Schiff base. Biochemistry 55:2371-2380.
42. Lorenz-Fonfria VA, Resler T, Krause N, Nack M, Gossing M, Fischer von Mollard G, Bamann C, Bamberg E, Schlesinger R, Heberle J. 2013. Transient protonation changes in channelrhodopsin-2 and their relevance to channel gating. Proc Natl Acad Sci USA 110:E1273-81.
43. Kuhne J, Vierock J, Tennigkeit SA, Dreier MA, Wietek J, Petersen D, Gavriljuk K, El-Mashtoly SF, Hegemann P, Gerwert K. 2019. Unifying photocycle model for light adaptation and temporal evolution of cation conductance in channelrhodopsin-2. Proc Natl Acad Sci USA 116:9380-9389.
44. Sineshchekov OA, Li H, Govorunova EG, Spudich JL. 2016. Photochemical reaction cycle transitions during anion channelrhodopsin gating. Proc Natl Acad Sci USA 113:E1993-2000.
45. Lorenz-Fonfria VA, Heberle J. 2014. Channelrhodopsin unchained: structure and mechanism of a light-gated cation channel. Biochim Biophys Acta 1837:626-42.
46. Tsujimura M, Noji T, Saito K, Kojima K, Sudo Y, Ishikita H. 2020. Mechanism of absorption wavelength shifts in anion channelrhodopsin-1 mutants. Biochim Biophys Acta Bioenerg 1862:148349.
47. Bedbrook CN, Yang KK, Rice AJ, Gradinaru V, Arnold FH. 2017. Machine learning to design integral membrane channelrhodopsins for efficient eukaryotic expression and plasma membrane localization. PLoS Comput Biol 13:e1005786.
48. Bedbrook CN, Yang KK, Robinson JE, Mackey ED, Gradinaru V, Arnold FH. 2019. Machine learning-guided channelrhodopsin engineering enables minimally invasive optogenetics. Nat Methods 16:1176-1184.
49. Lawrence SA, Floge SA, Davy JE, Davy SK, Wilson WH. 2017. Exploratory analysis of *Symbiodinium* transcriptomes reveals potential latent infection by large dsDNA viruses. Environ Microbiol 19:3909-3919.

50. Greiner A, Kelterborn S, Evers H, Kreimer G, Sizova I, Hegemann P. 2017. Targeting of photoreceptor genes in *Chlamydomonas reinhardtii* via zinc-finger nucleases and CRISPR/Cas9. Plant Cell 29:2498-2518.
51. Litvin FF, Sineshchekov OA, Sineshchekov VA. 1978. Photoreceptor electric potential in the phototaxis of the alga *Haematococcus pluvialis*. Nature 271:476-478.
52. Sineshchekov OA, Govorunova EG, Jung K-H, Zauner S, Maier U-G, Spudich JL. 2005. Rhodopsin-mediated photoreception in cryptophyte flagellates. Biophys J 89:4310-4319.
53. Govorunova EG, Spudich EN, Lane CE, Sineshchekov OA, Spudich JL. 2011. New channelrhodopsin with a red-shifted spectrum and rapid kinetics from *Mesostigma viride*. mBio 2:e00115-11.
54. Govorunova EG, Sineshchekov OA, Li H, Janz R, Spudich JL. 2013. Characterization of a highly efficient blue-shifted channelrhodopsin from the marine alga *Platymonas subcordiformis*. J Biol Chem 288:29911-29922.
55. Saddler HDW. 1970. The membrane potential of *Acetabularia mediterranea*. J Gen Physiol 55:802-821.
56. Foster K-W, Smyth RD. 1980. Light antennas in phototactic algae. Microbiol Rev 44:572-630.
57. Hollingsworth LL, Kinzie RA, Lewis TD, Krupp DA, Leong JAC. 2015. Phototaxis of motile zooxanthellae to green light may facilitate symbiont capture by coral larvae. Coral Reefs 24:523.
58. Aihara Y, Maruyama S, Baird AH, Iguchi A, Takahashi S, Minagawa J. 2019. Green fluorescence from cnidarian hosts attracts symbiotic algae. Proc Natl Acad Sci USA 116:2118-2123.
59. Obergrussberger A, Friis S, Bruggemann A, Fertig N. 2021. Automated patch clamp in drug discovery: major breakthroughs and innovation in the last decade. Expert Opin Drug Discov 16:1-5.
60. Chang Z, Li G, Liu J, Zhang Y, Ashby C, Liu D, Cramer CL, Huang X. 2015. Bridger: A new framework for de novo transcriptome assembly using RNA-seq data. Genome Biol 16:30.
61. Eddy SR. 2011. Accelerated profile HMM searches. PLoS Comput Biol 7:e1002195.
62. Carradec Q, Pelletier E, Da Silva C, Alberti A, Seeleuthner Y, Blanc-Mathieu R, Lima-Mendez G, Rocha F, Tirichine L, Labadie K, Kirilovsky A, Bertrand A, Engelen S, Madoui MA, Meheust R, Poulain J, Romac S, Richter DJ, Yoshikawa G, Dimier C, Kandels-Lewis S, Picheral M, Searson S, Tara Oceans C, Jaillon O, Aury JM, Karsenti E, Sullivan MB, Sunagawa S, Bork P, Not F, Hingamp P, Raes J, Guidi L, Ogata H, de Vargas C, Iudicone D, Bowler C, Wincker P. 2018. A global ocean atlas of eukaryotic genes. Nat Commun 9:373.
63. Steinegger M, Mirdita M, Söding J. 2019. Protein-level assembly increases protein sequence recovery from metagenomic samples manyfold. Nat Methods 16:603-606.
64. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol 37:1530-1534.
65. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: Improving the ultrafast bootstrap approximation. Mol Biol Evol 35:518-522.

66. Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 47:W256-W259.
67. Sineshchekov OA, Govorunova EG, Li H, Spudich JL. 2017. Bacteriorhodopsin-like channelrhodopsins: Alternative mechanism for control of cation conductance. Proc Natl Acad Sci USA 114:E9512-E9519.
68. Klapoetke NC, Murata Y, Kim SS, Pulver SR, Birdsey-Benson A, Cho YK, Morimoto TK, Chuong AS, Carpenter EJ, Tian Z, Wang J, Xie Y, Yan Z, Zhang Y, Chow BY, Surek B, Melkonian M, Jayaraman V, Constantine-Paton M, Wong GK, Boyden ES. 2014. Independent optical excitation of distinct neural populations. Nat Methods 11:338-346.
69. Tashiro R, Sushmita K, Hososhima S, Sharma S, Kateriya S, Kandori H, Tsunoda SP. 2021. Specific residues in the cytoplasmic domain modulate photocurrent kinetics of channelrhodopsin from the alga *Klebsormidium nitens*. Commun Biol 4:235.

FIGURE LEGENDS

Figure 1. An unrooted phylogenetic tree of ChRs. The nodes are color-coded as following: red, confirmed anion selectivity; blue, confirmed cation selectivity; gray, non-functional; black, ion selectivity not determined. Thicker nodes show ChRs characterized in this study. Gray circles show ultrafast bootstrap support values above 95%. A tree file in the Newick format is available as Data Set S2, and the corresponding protein alignment, as Data Set S3.

Figure 2. Prasinophyte CCRs. (A) Amino acid residues corresponding to the indicated positions in *Cr*ChR2. ChRs characterized in this study are in bold (black, functional; gray, non-functional). Conserved glutamates are highlighted red. (B) Peak photocurrent amplitudes generated at -60 mV in response to 1-s light pulses at the wavelength of the spectral maximum. (C) Desensitization of photocurrents after 1-s illumination. (D) Action spectra of photocurrents. The data points show mean ± sem (n = 4-8 scans). (E) Reversal potentials of photocurrents. The bars in B, C and E show mean ± sem; diamonds, data from individual cells.

Figure 3. Stramenopile ACRs and their metagenomic homologs. (A) Amino acid residues in the ion conductance pathway, corresponding to the indicated positions in *Cr*ChR2. ChRs characterized in this study are in bold face. Conserved glutamates are highlighted red. (B) Peak photocurrent amplitudes generated at -60 mV in response to 1-s light pulses at the wavelength of the spectral maximum. (C) Desensitization of photocurrents after 1-s illumination. (D) Reversal potentials of photocurrents. In B-D the bars show mean ± sem; empty diamonds, data from individual cells. (E and F) Action spectra of photocurrents. The data points are mean ± sem (n = 4-6 scans).

Figure 4. Cryptophyte ACRs. (A) Amino acid residues corresponding to the indicated positions in *Cr*ChR2. ChRs characterized in this study are in bold, non-functional homologs are in grey. Conserved glutamates are highlighted red. (B) Peak photocurrent amplitudes generated at -60 mV in response to 1-s light pulses at the wavelength of the spectral maximum. (C) Desensitization of photocurrents after 1-s illumination. (D and E) Series of photocurrent traces

recorded from *R3*ACR1 upon incremental voltage increase with the SyncroPatch 384i (D) and AxoPatch 200B (E) at 470-nm excitation. Note the smaller amplitude and slower kinetics of the SyncroPatch traces, as expected from the lower stimulus intensity. (F) Reversal potentials of photocurrents. In B, C and F the bars show mean ± sem; empty diamonds, data from individual cells.

Figure 5. Color tuning in cryptophyte ACRs. (A) A crystal structure of *Gt*ACR1 (6EDQ) showing the three side chains that contribute to the spectral difference between *Gt*ACR1 and *Gt*ACR2. (B and D) Action spectra of photocurrents generated by the indicated cryptophyte ACRs. The data points show mean ± sem (n = 4-8 scans). (C) Amino acid residues involved in color tuning in the functional cryptophyte homologs. The numbering is according to the *Gt*ACR1 sequence.

Figure 6. Δ*Sm*ChR1 and its metagenomic homolog mgdChR1. (A) Peak photocurrent amplitudes generated at -60 mV in response to 1-s light pulses at the wavelength of the spectral maximum. The bars show mean ± sem; empty diamonds, data from individual cells. The asterisk indicates $p < 0.01$ by the Mann-Whitney test. "mgdChR1no" denotes the construct without TS and ER motifs. (B) Photocurrent traces recorded at -60 and 60 mV from mgdChR1 (top) and Δ*Sm*ChR1 (bottom). The Δ*Sm*ChR1 at 60 mV was shifted 50 ms to the right relative to the trace at -60 mV to show the fast negative peak. (C) Action spectra of photocurrents. The data points show mean ± sem (n = 10-12 scans).

SUPPLEMENTAL TEXT

Supplemental Results

"Core" chlorophyte and streptophyte ChR homologs

Only one Glu residue (corresponding to Glu82 of *Cr*ChR2) is conserved in TM2 of the functionally characterized *Pyramimonas* ACRs (6). Four sequences from Chlorophyceae (two from *Chlamydomonas noctigama,* one from *Chlamydomonas* sp. and one from *Chloromonas subdivisa*) also exhibit this residue pattern (Fig. S1B). Upon expression of three of these polynucleotides, small hyperpolarizing photocurrents that did not reverse at positive voltages were recorded (Fig. S1C). They likely reflect intramolecular transfer of the Schiff base proton to an outwardly located acceptor, as previously found in other ChRs (24, 67). The *C. subdivisa* sequence was very poorly expressed and generated no currents.

Out of the four Chlorodendrophyceae sequences with the misplaced Asp212 homolog, one (named here *Tch*ChR) has already been tested earlier and found to be non-electrogenic (68). We have synthesized and tested two other sequences of this group, *Ps*ChR4 from *P. subcordiformis* and *Ta*ChR from *T. astigmatica*. Neither generated photocurrents upon expression in HEK293 cells, although normal tag fluorescence was observed.

Functional CCRs have previously been reported in the streptophyte classes Mesostigmatophyceae (53) and Klebsormidiophyceae (69). A ChR homolog has also been

identified in *Coleochaete irregularis* (6) from the class Coleochaetophyceae, which is more closely related to land plants. However, we could not detect any photocurrents upon its expression in HEK293 cells.

Supplemental Figure Legends

Figure S1. (A) The residues of the retinal-binding pockets. Variants tested in this study are in bold. The residues in *Ms*CCR and *Cs*CCR that differ from those in Chrimson and RubyACRs, earlier known red-shifted ChRs, are highlighted red. The numbers are according to bacteriorhodopsin sequence. (B and D) The residues in the positions of the conserved glutamates (highlighted red) in the ion conductance pathway in the indicated homologs from Chlorophyceae (B) and Chlorodendrophyceae (D). Variants tested in this study are in bold, non-functional, in gray. The numbers are according to *Cr*ChR2 sequence. (C) Photocurrent traces recorded from *C1*ChR in response to 1-s illumination at -60 and 60 mV. (E) Alignment of the part of TM7 of the indicated Chlorodendrophyceae homologs. The Schiff base Lys is highlighted blue; the upstream Glu, red.

Figure S2. A section of the phylogenetic tree from Fig. 1 redrawn in a rectangular format. Red nodes show variants with proven anion selectivity, thicker nodes show variants tested in this study. Red numbers are the values of photocurrent desensitization from Fig. 3C in the main text (mean ± sem, n = 5-6 cells).

Figure S3. (A and B) Alignments of TM3 (A) and TM7 (B) of dinoflagellate ChRs and their metagenomic homologs. The TCP motif and the Asn residue corresponding to Asp212 of bacteriorhodopsin are highlighted red. (C) EYFP tag fluorescence in cells expressing mgdChR1 constructs schematically shown on top of the images. TS, trafficking signal, ER, endoplasmic reticulum export motif.

Figure S4. Alignment of *S. microadriaticum* ChR1 with and without the N-terminal extension.

Supplemental Table Legends

Table S1. Compositions of the solutions used in patch clamp recording.

Table S2. List of cryptophyte strains analyzed.

Supplemental Data Set Legends

Data Set S1. Genbank accession numbers, abbreviated protein names, source organisms, habitats, transcript names and amino acid sequences used to construct the phylogenetic tree in Fig. 1. Note that only sequences that cover the entire N-terminal and rhodopsin domains are included. Literature references to identification and electrophysiological characterization of the sequences are also provided. The sequences identified and characterized in this study are shown in bold.

Data Set S2. A Newick file of the tree shown in Fig. 1 in the main text.

715 **Data Set S3.** An alignment of the C-truncated sequences used to construct the tree in Fig. 1 in
716 the main text.

717 **Data Set S4.** A list of sequence databases searched.

718

719 **Table S1.** Solution compositions in AxoPatch and SyncroPatch recordings.

720 Abbreviations: Asp, aspartate; EGTA, ethylene glycol tetraacetic acid; HEPES, 4-(2-
721 hydroxyethyl)-1-piperazineethanesulfonic acid; LJP, liquid junction potential; NMDG, N-Methyl-
722 D-glucamine. All concentrations are in mM.

	NaCl	KCl	CsCl	CsF	$CaCl_2$	$MgCl_2$	EGTA	HEPES	NMDG*	Glucose	NaAsp	pH	LJP pipette
Pipette standard	—	126			0.5	2	5	25	—	—	—	7.4	—
Pipette NaCl	126	—			0.5	2	5	25	—	—	—	7.4	—
Bath standard	150	—			1.8	1	—	10	—	5	—	7.4	1.1
Bath NaAsp	—	—			1.8	1	—	10	—	5	150	7.4	-13
Bath pH 6.4	150	—			1.8	1	—	10	—	5	—	6.4	1.1
Bath NMDG	1.4	—			1.8	1	—	10	148.6	5	—	7.4	7
Nanion Internal CsF	10	—	10	110	—	—	10	10	—	—	—	7.2	—
Nanion External Chip Fill	140	4	—	—	—	—	—	10	—	5	—	7.4	—
Nanion External Physiological	140	4	—	—	2	1	—	10	—	5	—	7.4	—
Nanion External NMDG 60	80	4	—	—	2	1	—	10	60	5	—	7.4	—

723 *NMDG stock (pH 9) was prepared from 1M NMDG and ~700 mM HCl.

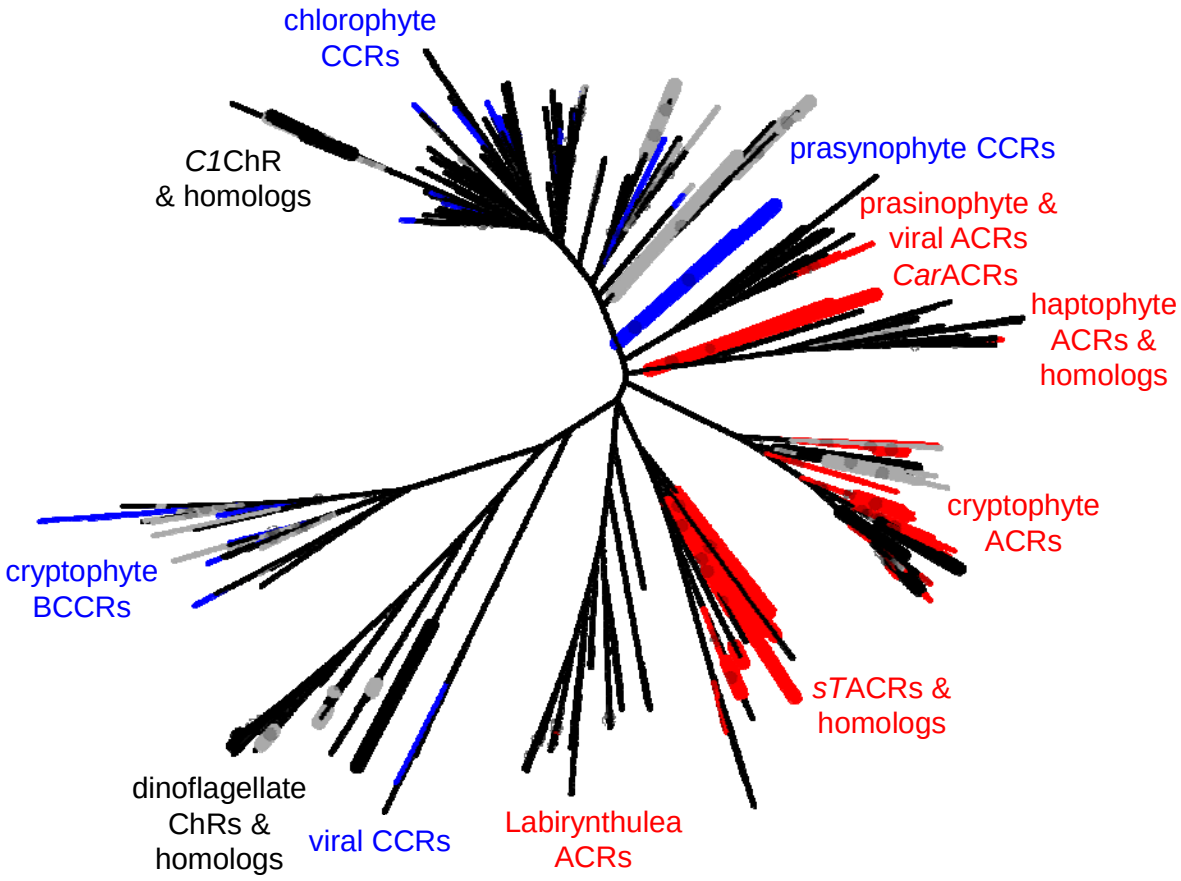
Table S2. List of cryptophyte strains analyzed

Abbreviations: BEA, Banco Español de Algas, of the Universidad of Las Palmas de Gran Canaria, Spain; CCAC, Culture Collection of Algae at the University of Cologne, Germany; SCCAP, Scandinavian Culture Collection of Algae and Protozoa at the University of Copenhagen, Denmark; CCMP, Culture Collection of Marine Phytoplankton at the Provasoli-Guillard National Center for Marine Algae and Microbiota at Woods Hole Oceanographic Institution, USA; NIES, National Institute for Environmental Studies, Tsukuba, Japan.

	Catalog number	Genus and species names	Habitat (growth medium)	ACRs found?
1.	CCMP 0268	*Chroomonas sp.*	Marine (ASP-12)*	No[#]
2.	CCAC 4037 B	*Chroomonas sp.*	Marine (ASP-12)*	No[#]
3.	NIES 1370	*Chroomonas sp.*	Marine (ASP-12)*	No[#]
4.	CCAC 0173 B	*Chroomonas sp.*	Marine (ASP-H)*	No[#]
5.	CCAC 1627 B	*Chroomonas sp.*	Marine (ASP-H)*	Yes
6.	CCAC 3782 B	*Chroomonas sp.*	Freshwater	No[#]
7.	CCAC 3670 B	*Chroomonas sp.*	Marine (ASP-H)*	No[#]
8.	BEA 0199B	*Cryptomonas cf. pyrenoidifera*	Freshwater	No[#]
9.	CCAC 0108	*Cryptomonas gyropyrenoidosa*	Freshwater	No[#]
10.	CCAC 0031	*Cryptomonas obovoidea*	Freshwater	No[#]
11.	CCAC 0064	*Cryptomonas ovata*	Freshwater	No[#]
12.	SCCAP K0416	*Geminigera sp.*	Marine (ASP-12)*	No[#]
13.	CCAC 1074 B	*Komma caudata*	Freshwater	No[#]
14.	CCMP 0760	*Rhodomonas sp.*	Marine (ASP-12)*	Yes
15.	CCAP 979/6	*Rhodomonas sp.*	Marine (ASP-12)*	Yes
16.	CCAC 1480 B	*Rhodomonas sp.*	Freshwater	Yes
17.	CCAC 3787 B	*Rhodomonas sp.*	Marine (ASP-12)*	Yes
18.	CCAC 3407 B	*Rhodomonas sp.*	Marine (ASP-12)*	Yes
19.	CCAC 3799 B	*Rhodomonas sp.*	Marine (ASP-12)*	Yes
20.	BEA 0603B	*Urgorri complanatus*	Marine (ASP-12)*	No[#]

*For media recipes see the CCAC website (https://www.uni-due.de/biology/ccac/)

[#]Our finding of no ACR transcripts in this species does not necessarily mean that they are not encoded by its genome and expressed; the transcript abundance may simply have been below the detection limit of our method.



736 **Figure 1.** An unrooted phylogenetic tree of ChRs. The nodes are color-coded as following: red,
737 confirmed anion selectivity; blue, confirmed cation selectivity; gray, non-functional; black, ion
738 selectivity not determined. Thicker nodes show ChRs characterized in this study. Gray circles
739 show ultrafast bootstrap support values above 95%. A tree file in the Newick format is available
740 as Data Set S2, and the corresponding protein alignment, as Data Set S3.

741

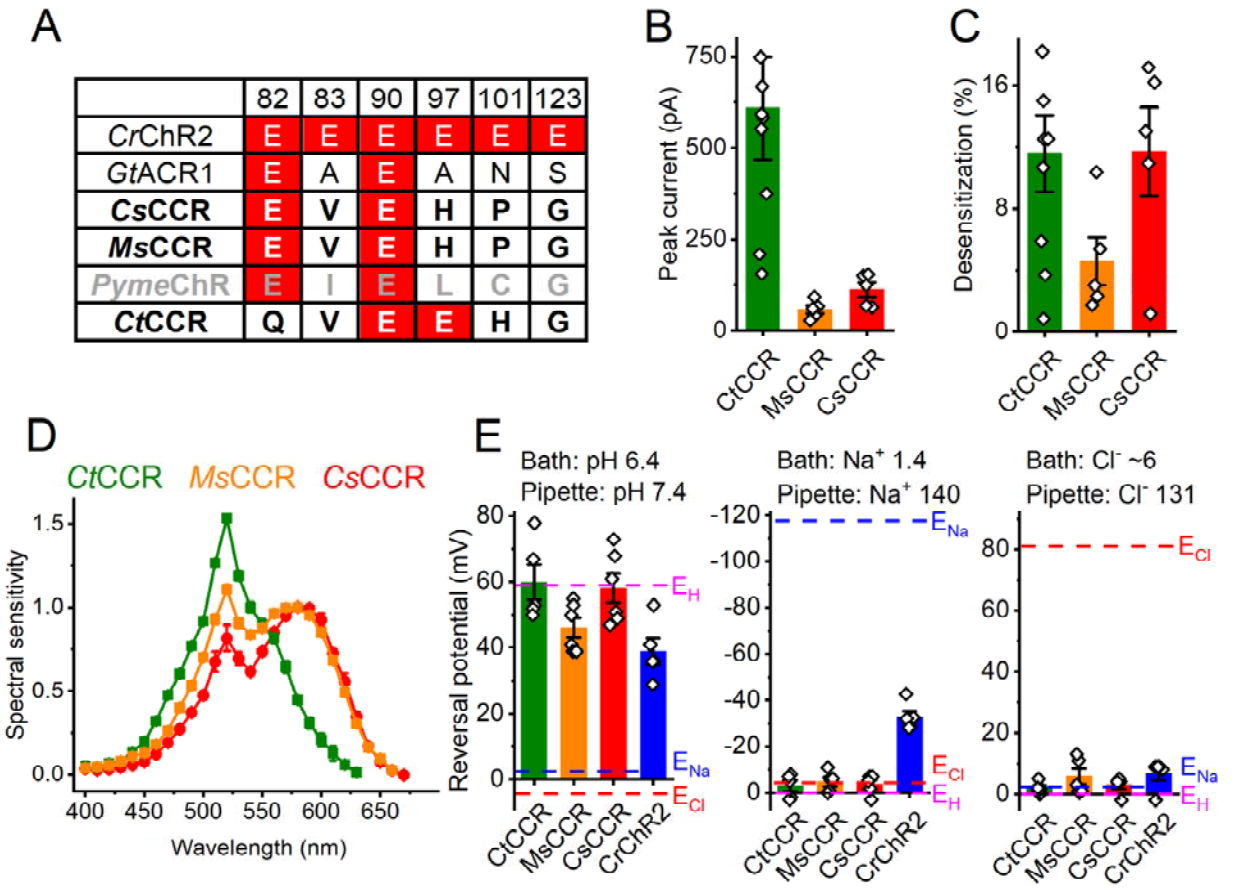
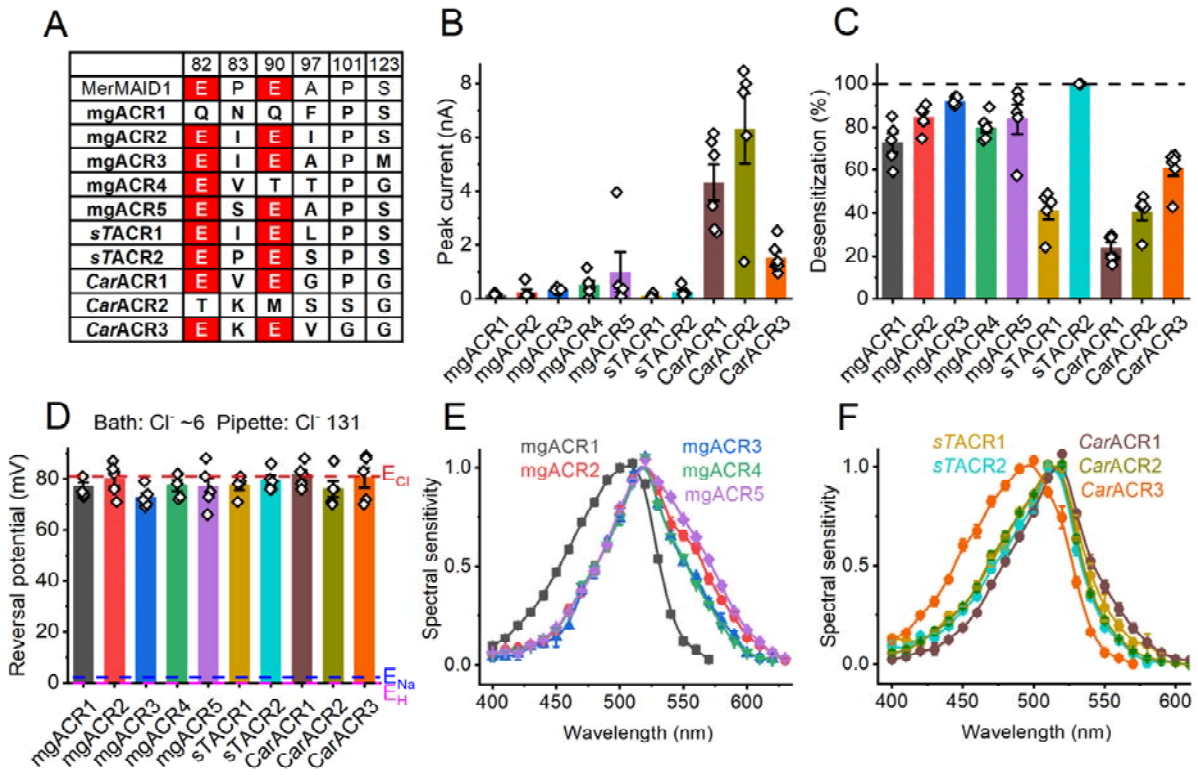
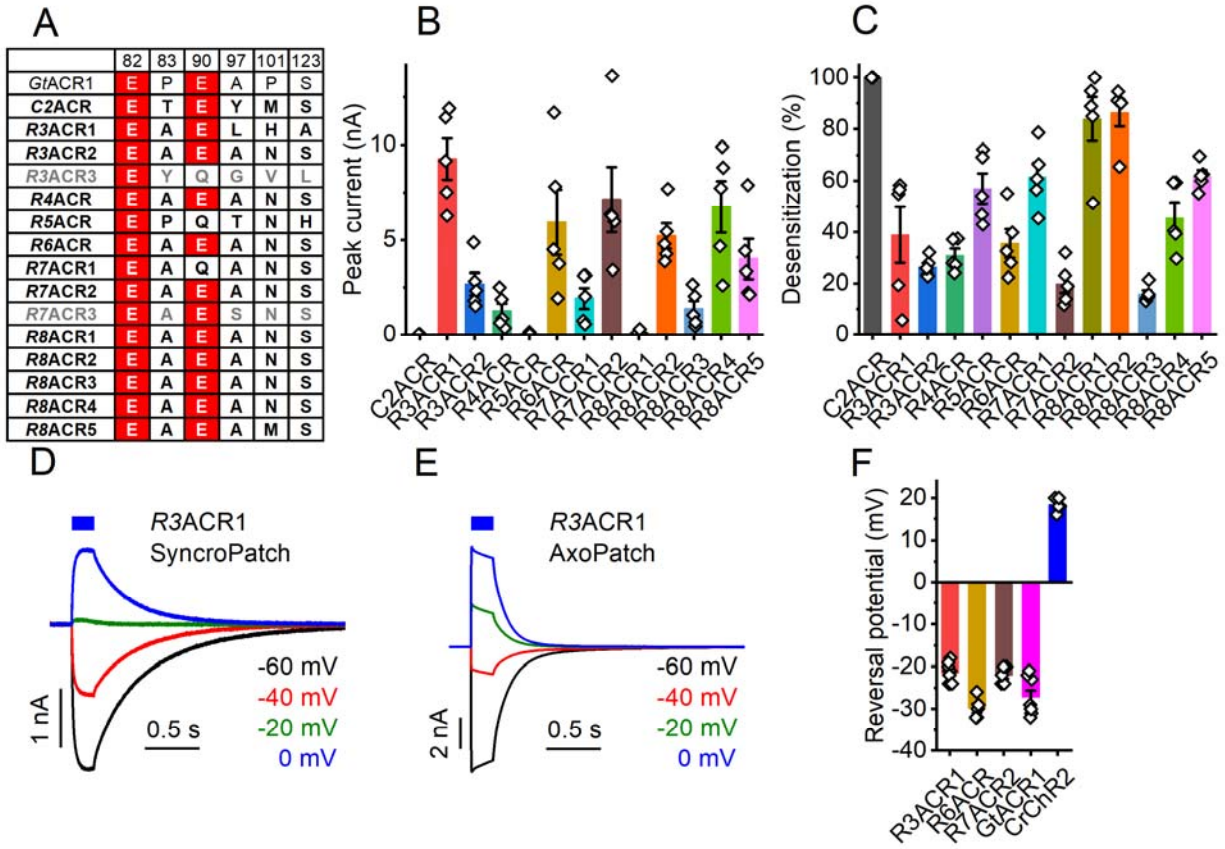


Figure 2. Prasinophyte CCRs. (A) Amino acid residues corresponding to the indicated positions in *Cr*ChR2. ChRs characterized in this study are in bold (black, functional; gray, non-functional). Conserved glutamates are highlighted red. (B) Peak photocurrent amplitudes generated at -60 mV in response to 1-s light pulses at the wavelength of the spectral maximum. (C) Desensitization of photocurrents after 1-s illumination. (D) Action spectra of photocurrents. The data points show mean ± sem (n = 4-8 scans). (E) Reversal potentials of photocurrents. The bars in B, C and E show mean ± sem; diamonds, data from individual cells.



	82	83	90	97	101	123
MerMAID1	E	P	E	A	P	S
mgACR1	Q	N	Q	F	P	S
mgACR2	E	I	E	I	P	S
mgACR3	E	I	E	A	P	M
mgACR4	E	V	T	T	P	G
mgACR5	E	S	E	A	P	S
***s*TACR1**	E	I	E	L	P	S
***s*TACR2**	E	P	E	S	P	S
***Car*ACR1**	E	V	E	G	P	G
***Car*ACR2**	T	K	M	S	S	G
***Car*ACR3**	E	K	E	V	G	G

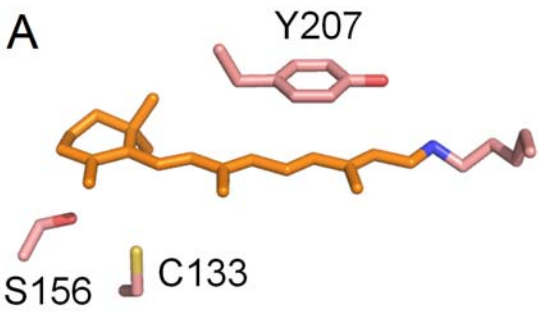
Figure 3. Stramenopile ACRs and their metagenomic homologs. (A) Amino acid residues in the ion conductance pathway, corresponding to the indicated positions in *Cr*ChR2. ChRs characterized in this study are in bold face. Conserved glutamates are highlighted red. (B) Peak photocurrent amplitudes generated at -60 mV in response to 1-s light pulses at the wavelength of the spectral maximum. (C) Desensitization of photocurrents after 1-s illumination. (D) Reversal potentials of photocurrents. In B-D the bars show mean ± sem; empty diamonds, data from individual cells. (E and F) Action spectra of photocurrents. The data points are mean ± sem (n = 4-6 scans).

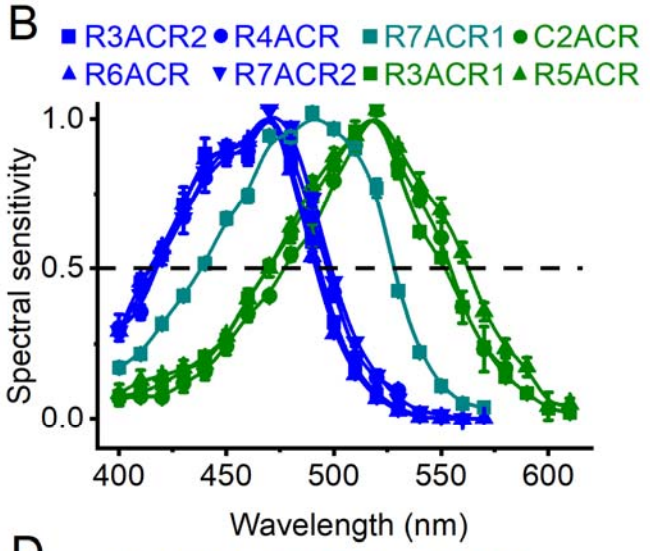


	82	83	90	97	101	123
*Gt*ACR1	E	P	E	A	P	S
C2ACR	E	T	E	Y	M	S
***R3*ACR1**	E	A	E	L	H	A
***R3*ACR2**	E	A	E	A	N	S
*R3*ACR3	E	Y	Q	G	V	L
***R4*ACR**	E	A	E	A	N	S
***R5*ACR**	E	P	Q	T	N	H
***R6*ACR**	E	A	E	A	N	S
***R7*ACR1**	E	A	Q	A	N	S
***R7*ACR2**	E	A	E	A	N	S
*R7*ACR3	E	A	E	S	N	S
***R8*ACR1**	E	A	E	A	N	S
***R8*ACR2**	E	A	E	A	N	S
***R8*ACR3**	E	A	E	A	N	S
***R8*ACR4**	E	A	E	A	N	S
***R8*ACR5**	E	A	E	A	M	S

759 **Figure 4.** Cryptophyte ACRs. (A) Amino acid residues corresponding to the indicated positions in
760 *Cr*ChR2. ChRs characterized in this study are in bold, non-functional homologs are in grey.
761 Conserved glutamates are highlighted red. (B) Peak photocurrent amplitudes generated at -60
762 mV in response to 1-s light pulses at the wavelength of the spectral maximum. (C)
763 Desensitization of photocurrents after 1-s illumination. (D and E) Series of photocurrent traces
764 recorded from *R3*ACR1 upon incremental voltage increase with the SyncroPatch 384i (D) and
765 AxoPatch 200B (E) at 470-nm excitation. Note the smaller amplitude and slower kinetics of the
766 SyncroPatch traces, as expected from the lower stimulus intensity. (F) Reversal potentials of
767 photocurrents. In B, C and F the bars show mean ± sem; empty diamonds, data from individual
768 cells.

769





C

	133	156	207
*Gt*ACR1	C	S	Y
*Gt*ACR2	R	G	F
C2ACR	**M**	**A**	**Y**
R3ACR1	**C**	**G**	**F**
R3ACR2	**R**	**G**	**F**
R4ACR	**R**	**G**	**F**
R5ACR	**R**	**S**	**F**
R6ACR	**R**	**G**	**F**
R7ACR1	**R**	**A**	**F**
R7ACR2	**R**	**G**	**F**
R8ACR1	**R**	**S**	**F**
R8ACR2	**R**	**A**	**F**
R8ACR3	**R**	**A**	**F**
R8ACR4	**R**	**G**	**F**
R8ACR5	**R**	**G**	**F**

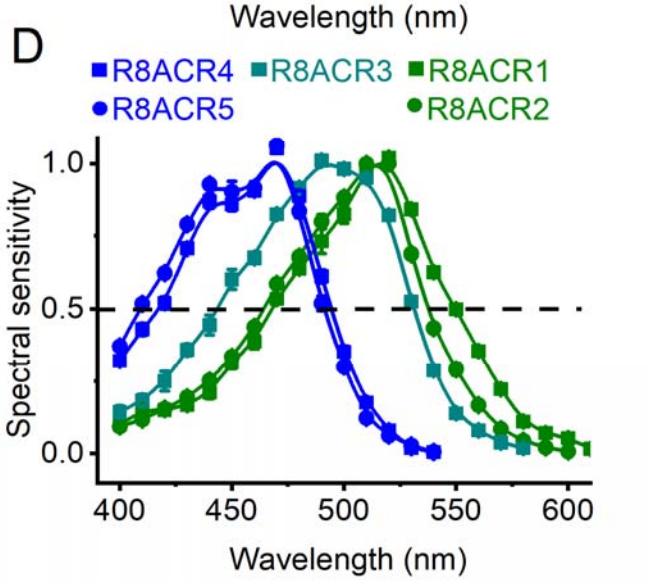
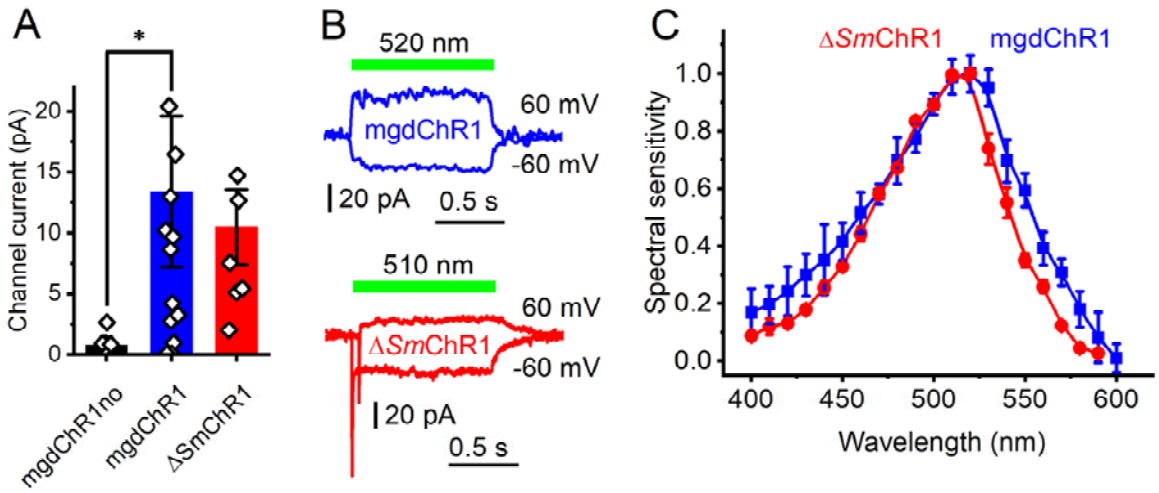


Figure 5. Color tuning in cryptophyte ACRs. (A) A crystal structure of *Gt*ACR1 (6EDQ) showing the three side chains that contribute to the spectral difference between *Gt*ACR1 and *Gt*ACR2. (B and D) Action spectra of photocurrents generated by the indicated cryptophyte ACRs. The data points show mean ± sem (n = 4-8 scans). (C) Amino acid residues involved in color tuning in the functional cryptophyte homologs. The numbering is according to *Gt*ACR1 sequence.



776 **Figure 6.** Δ*Sm*ChR1 and its metagenomic homolog mgdChR1. (A) Peak photocurrent amplitudes
777 generated at -60 mV in response to 1-s light pulses at the wavelength of the spectral maximum.
778 The bars show mean ± sem; empty diamonds, data from individual cells. The asterisk indicates p
779 < 0.01 by the Mann-Whitney test. "mgdChR1no" denotes the construct without TS and ER
780 motifs. (B) Photocurrent traces recorded at -60 and 60 mV from mgdChR1 (top) and Δ*Sm*ChR1
781 (bottom). The Δ*Sm*ChR1 at 60 mV was shifted 50 ms to the right relative to the trace at -60 mV
782 to show the fast negative peak. (C) Action spectra of photocurrents. The data points show mean
783 ± sem (n = 10-12 scans).

784

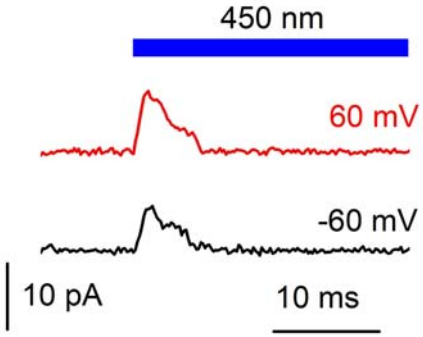
A

	49	53	57	85	89	90	93	115	138	141	142	145
***Ms*CCR**	**I**	**E**	**Y**	**G**	**T**	**T**	**L**	**N**	**F**	**G**	**C**	**M**
***Cs*CCR**	**I**	**E**	**Y**	**G**	**T**	**T**	**I**	**N**	**F**	**G**	**C**	**M**
Chrimson	V	E	V	E	S	C	I	C	Y	S	C	G
RubyACRs	V	T	Y	F	T	C	I	D	N	S	Y	C

B

	82	83	90	97	101	123
*Cr*ChR2	E	E	E	E	E	E
*Pyme*ACR	E	V	G	K	A	G
***Cn*ChR3**	**E**	**V**	**V**	**Q**	**P**	**Q**
***Cn*ChR4**	**E**	**V**	**V**	**Q**	**P**	**Q**
***C1*ChR**	**E**	**V**	**V**	**V**	**P**	**Q**
*Cs*ChR2	E	V	V	Q	P	Q

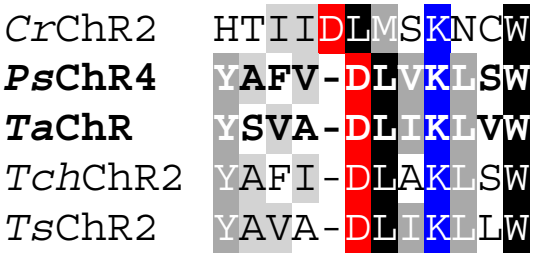
C



D

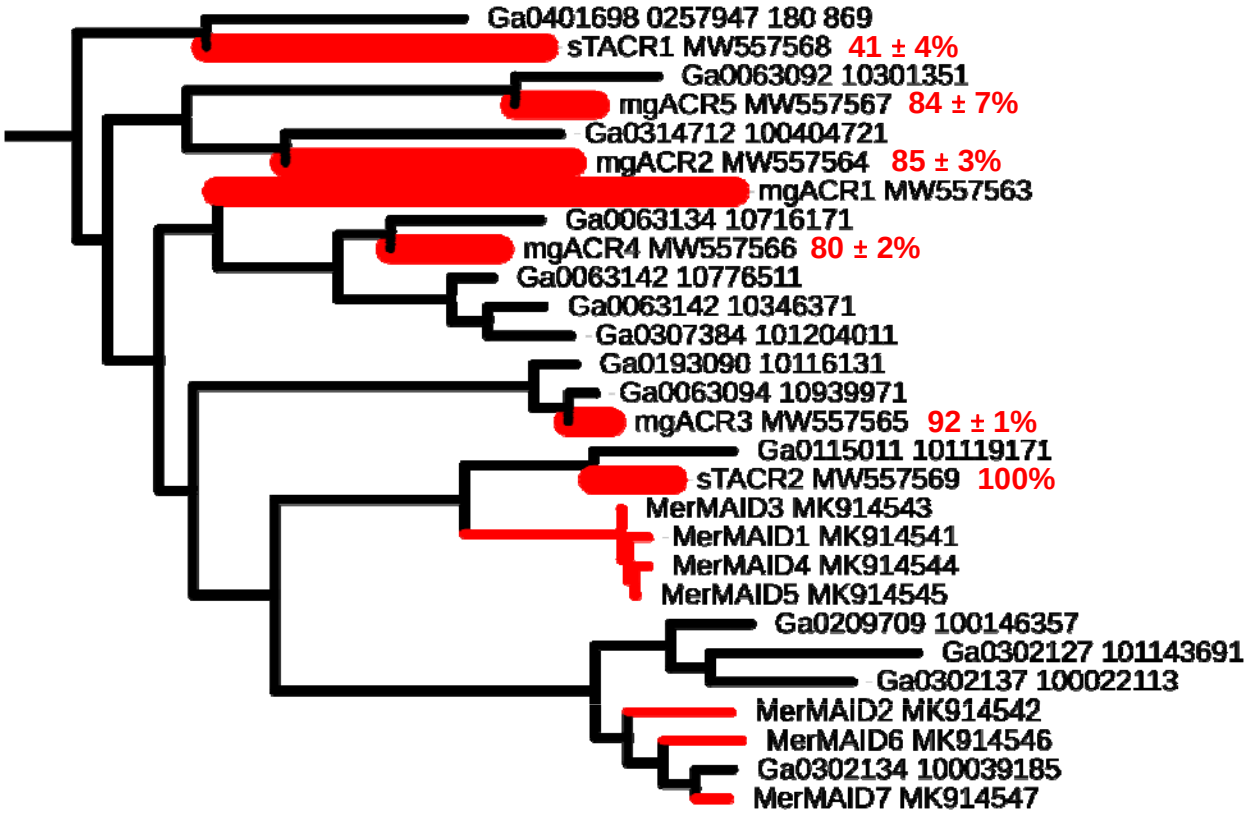
	82	83	90	97	101	123
*Cr*ChR2	E	E	E	E	E	E
***Ps*ChR4**	**T**	**L**	**V**	**S**	**T**	**A**
***Ta*ChR**	**V**	**I**	**V**	**A**	**P**	**S**
*Tch*ChR	T	L	V	S	T	A
*Ts*ChR2	T	L	V	S	N	A

E



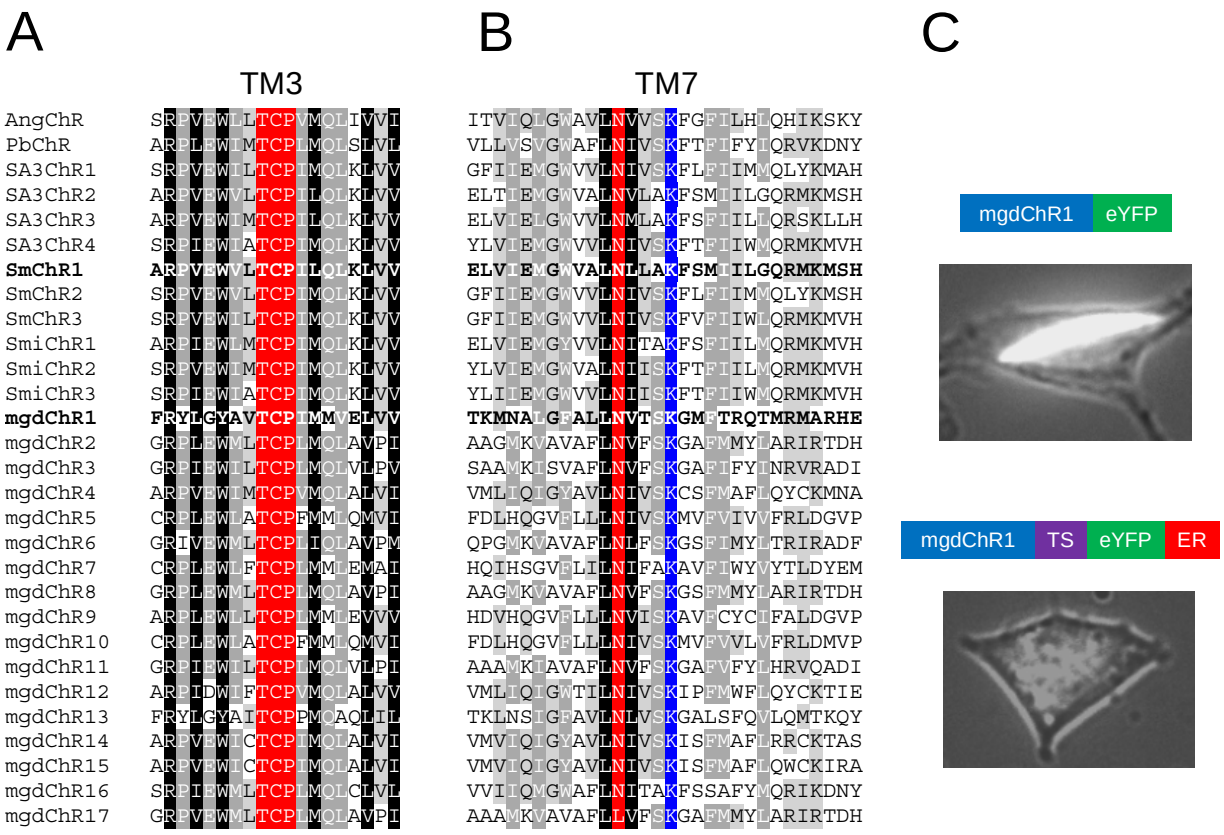
785 **Figure S1.** (A) The residues of the retinal-binding pockets. Variants tested in this study are in
786 bold. The residues in *Ms*CCR and *Cs*CCR that differ from those in Chrimson and RubyACRs,
787 earlier known red-shifted ChRs, are highlighted red. The numbers are according to
788 bacteriorhodopsin sequence. (B and D) The residues in the positions of the conserved
789 glutamates (highlighted red) in the ion conductance pathway in the indicated homologs from
790 Chlorophyceae (B) and Chlorodendrophyceae (D). Variants tested in this study are in bold, non-
791 functional, in gray. The numbers are according to *Cr*ChR2 sequence. (C) Photocurrent traces
792 recorded from *C1*ChR in response to 1-s illumination at -60 and 60 mV. (E) Alignment of the part
793 of TM7 of the indicated Chlorodendrophyceae homologs. The Schiff base Lys is highlighted
794 blue; the upstream Glu, red.

795



796 **Figure S2**. A section of the phylogenetic tree from Fig. 1 redrawn in a rectangular format. Red
797 nodes show variants with proven anion selectivity, thicker nodes show variants tested in this
798 study. Red numbers are the values of photocurrent desensitization from Fig. 3C in the main text
799 (mean ± sem, n = 5-6 cells).

800



801 **Figure S3**. (A and B) Alignments of TM3 (A) and TM7 (B) of dinoflagellate ChRs and their
802 metagenomic homologs. The TCP motif and the Asn residue corresponding to Asp212 of
803 bacteriorhodopsin are highlighted red. (C) EYFP tag fluorescence in cells expressing mgdChR1
804 constructs schematically shown on top of the images. TS, trafficking signal, ER, endoplasmic
805 reticulum export motif.

806

807 **Figure S4**. Alignment of *S. microadriaticum* ChR1 with and without the N-terminal extension.