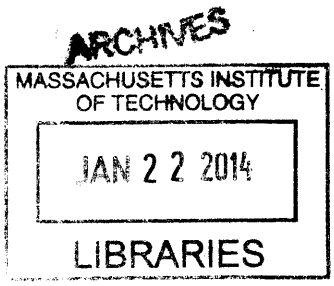
Independent Two-Color Optogenetic Excitation of Neural Populations

by

Nathan Cao Klapoetke

B.S. Electrical Engineering
University of Wisconsin - Madison, 2007



Submitted to the Department of Biological Engineering,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Biological Engineering
at the Massachusetts Institute of Technology

July 2013

Author: _____

Nathan C. Klapoetke
Department of Biological Engineering, MIT

Certified by: _____

Edward S. Boyden, III, Ph.D.
Associate Professor
Program in Media Arts and Sciences, MIT
Thesis Supervisor

Accepted by: _____

Forest M. White, Ph.D.
Associate Professor
Department of Biological Engineering, MIT
Graduate Program Committee Chair

This thesis was examined and approved by:

Edward F. Delong, Ph.D.
Professor
Department of Civil and Environmental Engineering, MIT
Department of Biological Engineering, MIT
Thesis Chair

Edward S. Boyden, III, Ph.D.
Associate Professor
Program in Media Arts and Sciences, MIT
Thesis Supervisor

Alice Y. Ting, Ph.D.
Associate Professor
Department of Chemistry, MIT
Thesis Reader

Abstract

The optical modulation of neurons with channelrhodopsins, a class of genetically encoded light-gated ion channels, has enabled the spatiotemporally precise interrogation of the roles individual cell types play in neural circuit dynamics. A topic of great interest to the neuroscience community is the independent optical excitation of two distinct neuron populations with different wavelengths, which would enable the interrogation of emergent phenomena such as circuit dynamics, plasticity, and neuromodulation. Previous implementations have focused on maximizing spectral separation by driving one channelrhodopsin in the violet (405 nm) and the other in the yellow (590 nm), yet it has not been possible to achieve independent violet excitation without eliciting spikes from both populations, due to the intrinsic UV-blue light sensitivity of the retinal chromophore.

This thesis designs and implements an improved two-color excitation scheme where effective light sensitivity is utilized to achieve independent optical excitation in blue (470 nm) and red (625 nm) channels. Zero post-synaptic crosstalk is demonstrated in acute murine slice, using two novel channelrhodopsins identified from a systematic screen of 80 naturally occurring, previously uncharacterized opsins in primary neuron culture. Gene88 is the first known yellow-peaked channelrhodopsin, with a peak 45 nm more red-shifted than any previous channelrhodopsin, while Gene90 has the fastest channel turn on, turn off, and recovery kinetics of any known channelrhodopsin. These opsins' novel properties enable the first known demonstration of post-synaptic crosstalk-free two-color excitation with temporally precise modulation of spatially inseparable neuron populations.

Acknowledgements

There are innumerable people who have helped me on my journey. I would like to thank:

My advisor, Ed Boyden, for inspiring me to take interest in neuroscience, showing me how to think broadly about science, and providing an intellectual environment.

Martha Constantine-Paton, for serving as an unofficial "advisor" and sharing her wisdom on neural plasticity and physiology.

Brian Chow for his mentorship in my early years and teaching me that most experiments are not intrinsically hard if you put enough effort into it.

The Boyden lab in general, but especially: My UROP, Tania Morimoto, for her indispensible help early on with algae and HEK cell patching when I was still figuring out how to do research. Amy Chuong for help with molecular biology and her company as a fellow lab rat. Yongku Cho for providing invaluable feedback and providing independent validations. Daniel Schmidt for always providing engaging intellectual conversations and challenging me.

The Constantine-Paton lab: Yasunobu Murata for being the best collaborator and generously doing all of the in utero electroporations. Amanda Birdsey-Benson for the initial 2-color slice patch validations. Andrew Bolton for sharing his slice knowledge.

1KP consortium: Gane Wong for providing access to 1KP and his support for optogenetics. Melkonian lab for generously prepping majority of the algae samples. Most of this work would not have been possible without them.

Numerous collaborators for providing immensely useful feedback on tools in a neuroscience context, and teaching me new things. In particular, Daniel Hochbaum, for many late evening discussions on optical tools and sharing amusing stories about science.

My thesis committee: Ed Delong and Alice Ting for making themselves available and providing outside perspectives.

All my friends and fellow classmates for providing distraction from research. Especially Raj for being a great roommate and engaging me in basketball.

My mom and grandparents for supporting me and teaching me the value of education.

Last but not least, my girlfriend for her unwavering support in finishing my PhD and keeping me sane.

Table of Contents

1. Background

Microbial rhodopsins are a family of light-sensitive integral membrane proteins found in many kingdoms of life including archaebacteria, eubacteria, fungi, and protista[1-4]. These proteins perform diverse roles ranging from ion pumping for photophosphorylation to sensory photoreception for phototaxis[5, 6]. Microbial rhodopsins consist of a retinal chromophore bound in the pore region of a seven-transmembrane alpha helical apoprotein called opsin. Retinal is covalently attached via Schiff base linkage to a lysine residue in the seventh helix. The rhodopsin reaction cycle, known as the photocycle, begins with the absorption of light by retinal, which causes the retinal to photoisomerize (typically from *all-trans* to 13-*cis*). The energy stored in the retinal is then coupled to the opsin to thermally drive a series of conformational changes in order to translocate ions across the cell membrane and/or change the signaling state of a cytoplasmic transducer.

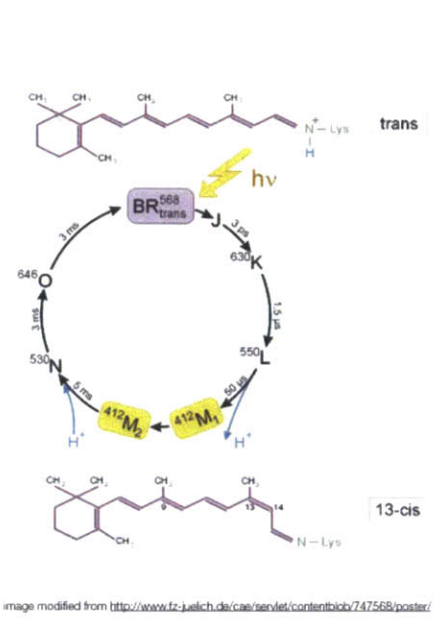


image modified from http://www.fz-juelich.de/cae/servlet/contentblob/747568/poster/

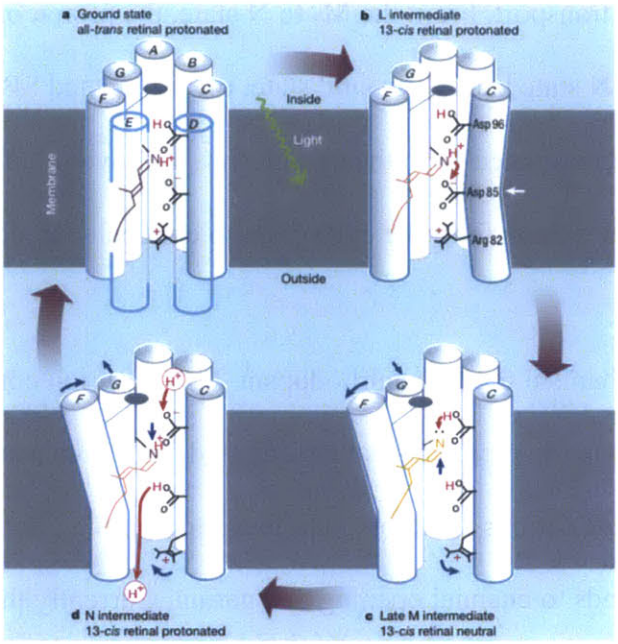


image modified from Kühlbrandt 2000 Nature

The archetypical photocycle model along with cartoon structure shown above is based on the archaeal proton pump bacteriorhodopsin (bR) from the species Halobacterium sp. SG1. Each photocycle intermediate is denoted with a letter (from J through O). In bR, the retinal is located approximately in the middle of the pore, effectively partitioning the opsin into cytoplasmic and extracellular halves. Initially in the ground state, both Schiff base (SB) and proton donor D96 (on the cytoplasmic side) are protonated[7] but upon photon absorption, the retinal isomerizes from *all-trans* to 13-*cis* within a few picoseconds to form the K intermediate[8]. Due to the retinal movement, the local environment around SB becomes energetically unfavorable and the proton on SB transfers to the proton acceptor 85 (on the extracellular side) to form the L state. The M intermediate is divided into two phases, early and late denoted as M_1 and M_2 respectively. During the transition from L to M_1, the proton on D85 is released to the extracellular surface[7]. Then between M_1 and M_2, the Schiff base nitrogen switches orientation from facing the extracellular side to the cytoplasmic side, changing the accessibility of Schiff base and enforcing

vectorial transport. From the M_2 to N state, the proton on D96 is donated to the Schiff base, then from the N state back to ground state, the protonated SB isomerizes back to all-trans form and D96 is reprotonated from the cytoplasm[7]. The overall effect of one photocycle is the movement of a single proton from the cytoplasm to extracellular side.

The mechanism for channelrhodopsin-2 (ChR2)[9] ion conduction is currently debated, but spectroscopy experiments have suggested a P520 intermediate corresponds to the conduction state where cations are allowed to leak through the channel[10]. It is not known which intermediate corresponds to channel opening and gating. Currently the only channelrhodopsin crystal structure that has been solved at high resolution (2.3A) is the ground state form of a chimaera of ChR1 and ChR2 known as C1C2[11]. Comparison with the bacteriorhodopsin (bR) crystal structure shows C1C2's helices 1, 2, and 7 are more outward tilted on the extracellular side than bR. In its crystal form, C1C2 is also a dimer with disulfide bonds between N-terminal domain and helix 3, 4. Since the crystal structure is only for the ground state, there is no direct information on the channel gating mechanism or the open-pore size. Potential hypotheses for cation conduction pathway are based on the alignment of charged residues lining the pore surface, with the majority of negative charged residues in helix 2[11].

Despite the lack of mechanistic understanding of channelrhodopsin function, numerous mutants and chimaeras have been made in attempt to shift the peak wavelength sensitivity, improve channel closing kinetics, and reduce channel inactivation. The general strategy for spectrally shifting channelrhodopsins involves helical swaps between channelrhodopsins possessing different spectral peaks, such as ChR1, ChR2, and VChR1. While single mutations can spectrally

tune the peak by up to 30 nm, large spectral shifts from a blue to green peak is achieved by helical swaps, suggesting color may be a distributed property[12, 13]. In addition there have not been any spectral shifts exceeding 10 nm beyond the natural range between the ChR2 and VChR1 peaks (470 nm to 550 nm): all mutant and chimaera peak wavelengths which have been characterized to date have been bounded by the parent channelrhodopsins[13].

There has been additionally been a great interest in expanding the range of channel closing kinetics, both in accelerating as well as slowing. Slow closing kinetic mutants stabilize the channel open-state and are used for bistable modulation experiments[13, 14], while fast closing kinetic mutants are thought to destabilize the channel open-state and are used to drive temporally precise high frequency neuronal spiking[12]. The main limitation, however, is that modulating the closing-kinetics also correspondingly changes the effective light sensitivity or channel open time such that fast kinetic mutants are less light sensitive, and slow kinetic mutants are more light sensitive[15]. This suggests an intrinsic tradeoff between speed and light sensitivity, which may be indicative of limitations in ChR2 and VChR1 function.

Finally, channel inactivation has severely limited the reliability of channelrhodopsin optical responses. Most efforts to reduce channel inactivation have focused on screening for an improved steady state photocurrent, which generally has resulted in slower closing kinetics which require additional mutations to ameliorate[16, 17].

While previous work in engineering channelrhodopsin variants has led to a better understanding of tuning individual properties, the complexity of channelrhodopsin function has meant there is

yet to be a clear biophysically improved channelrhodopsin without a trade-off in either conductance or kinetics. To date, the most widely used channelrhodopsin is still the original blue light sensitive ChR2 with the H134R mutation[18], suggesting that the current scaffolds used for engineering may be functionally limited, and it may perhaps be useful to search for other natural channelrhodopsins with distinct spectral and conductance properties.

2. Sequencing

Channelrhodopsin-1 from the green algae *Chlamydomonas reinhardtii* was the first reported microbial rhodopsin light-gated ion channel in 2002[19]. Since then less than ten channelrhodopsin sequences have been published, yet only *Chlamydomonas reinhardtii* channelrhodopsin-2 (ChR2) and *Volvox carteri* channelrhodopsin-1 (VChR1) have been shown as capable of optically driving neuronal spiking, primarily due to the fact that the majority of other channelrhodopsins expressed poorly under heterologous conditions[20, 21]. Further, the peak spectral diversity of the known channelrhodopsins is limited between blue (450 nm) and green (550 nm) wavelengths[22]. There is therefore an urgent need to rapidly expand the known number of channelrhodopsin sequences in order to explore and leverage their natural functional diversity (e.g., color sensitivity, trafficking, kinetics) particularly in the context of optically exciting cells.

Current high-throughput sequencing technologies have a finite sequencing capacity (typically 1-10 gigabases per lane), and there is an inherent tradeoff between depth/coverage of sequencing and the number of unique organisms which can be sequenced. Here we present a general two stage sequencing strategy termed TRACE (Transcriptomics and Rapid Amplification of cDNA Ends) (Fig. 2.1). In the first step, standard transcriptomic (or genomic) sequencing is performed at low or high coverage to sample a large number of organisms, allowing us to identify which organism or sample contain relevant genes of interest which can be used to generate unique primers. In the second step, RNA ligase mediated RACE is performed on both the 5' and 3' end of mRNA using the gene specific primers designed in step one.

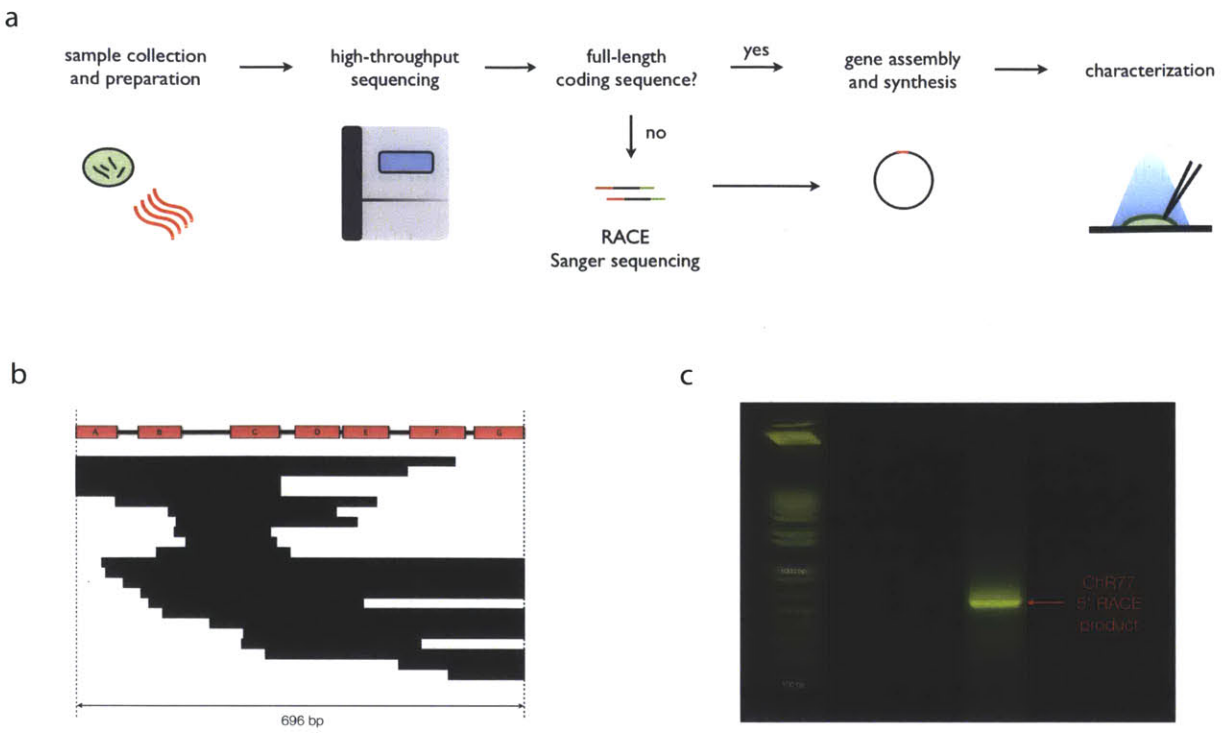


Figure 2.1: Transcriptomics and Rapid Amplification of cDNA Ends (TRACE). a) TRACE pipeline. b) Alignment of truncated transcripts bearing homology to ChR2 sequence. Each black row represents a different sequence, with white space indicating missing sequence. c) Agarose gel showing sequence specific 5' RACE product for gene77.

The OneKP consortium, in partnership with the Beijing Genomics Institute (BGI), took advantage of the decreasing cost of sequencing to sequence one thousand plant transcriptomes on the Illumina Genome Analyzer II. The paired-end sequencing yielded 2 gigabases per lane, which was then de novo assembled with BGI's SOAPdenovo and the Broad Institute's Trinity assembler. Despite more than 20-fold coverage of each transcriptome, many opsin-like genes were missing 5' or 3' sequencing data for the putative opsin transmembrane domains (Fig. 2.1). While truncated transcripts may be suitable for bioinformatics analysis, it is not acceptable for

characterizing protein function; since we were only interested in a small subset of genes relative to the whole transcriptome, it was also not reasonable to re-sequence an entire transcriptome in the hopes of better coverage.

We performed RNA ligase mediated RACE on algae total-RNA to enrich and sequence the missing regions from transcriptome sequencing of six opsin-like genes (see appendix for protocol). Unlike conventional RACE, we selectively ligated an RNA primer to the 5' mRNA end to act as an universal primer for subsequent 5' cDNA amplification. We thus selectively enriched 5' end sequences and subsequently blunt-end cloned the fragments for Sanger sequencing (Fig. 2.1). The entire protocol from algae RNA extraction to sequencing requires four to five days and is therefore not practical for all one thousand transcriptomes, but was manageable for only subsets of genes that contain unusual motifs or are homologous to existing high performing channelrhodopsins.

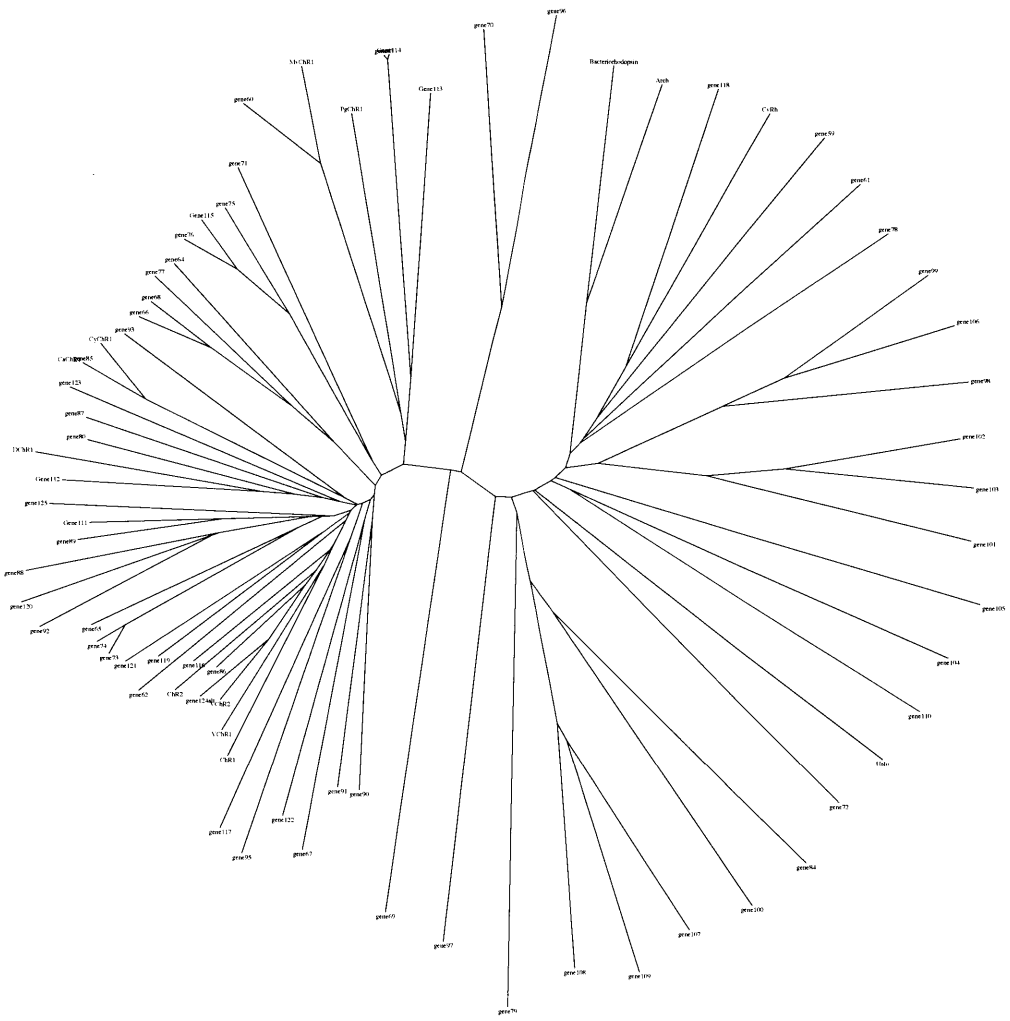


Figure 2.2: Unrooted phylogenetic tree for novel opsin sequences.

We synthesized all 57 full-length opsin-like genes obtained from transcriptome sequencing, along with the 6 found by RACE. We compiled full-length genes from more than 200 assembled algal transcriptomes based on the criteria that these proteins (1) contain Schiff base lysine for binding retinal, and (2) possessed seven transmembrane domains when aligned to the ChR2 sequence. These new opsin-like genes have 42-83% protein sequence similarity to ChR2, and 38-61% protein sequence similarity to bacteriorhodopsin (Fig. 2.2). Interestingly E123 and H134 position found in ChR2 are not strictly conserved but the highly charged glutamic acid clusters in helix two appear to be generally conserved (Table 3.1). While there are numerous other sequence differences among these opsins and many have low transmembrane sequence homology to known opsins, the functional importance of sequence divergence can only be validated through direct characterization. Our TRACE approach represents a broad search and reliable refinement methodology for obtaining full-length coding sequences for screening.

3. Screening

Through the high-throughput sequencing of algal transcriptomes, we identified over 50 novel channelrhodopsin homologs but it is impossible to predict their electrogenicity based on sequence homology alone. It is therefore necessary to experimentally determine whether these genes encode electrogenic proteins, and if any possess improved or novel biophysical properties. This chapter describes the rationale for the screening setup in neurons and the precautions we have taken to avoid user selection bias in patch-clamp characterization. We also compared the top hits from the screen to existing channelrhodopsins to evaluate differences in biophysical properties and optical spiking performance in neurons.

3.1 Preliminary Screening Insights

The initial screening goals were to determine which opsin-like genes encoded electrogenic proteins and merited more in-depth characterization. We decided to use whole-cell patch clamp for characterization due to its picoampere sensitivity and sub-millisecond temporal resolution, as no chemical or genetically-encoded voltage sensors currently exist that can rival patch clamp on either front. Many electrical characterization of microbial rhodopsin genes in intact cells is carried out in Xenopus oocytes, because their large surface area (~1 mm diameter) supports high total levels of opsin expression which allow for photocurrent amplitudes in the nanoampere to microampere range, substantially above the noise floor of most electrophysiology setups. However, a major downside to oocyte use is that current densities (current normalized by cell membrane capacitance) often do not translate to mammalian cells due to differences in trafficking, making opsin comparisons difficult to extrapolate to other systems[16]. Since the

intended optogenetic application is in mammalian cells, we decided to conduct our initial characterization in HEK293FT cells.

We synthesized opsin-like genes with mammalian codon optimization and subcloned them into a modified version of the pEGFP-N3 vector behind the CMV promoter with a C-terminal GFP fusion. In the first round of HEK293 cell screening, we examined 9 new opsin-like sequences (gene62-69, 80) along with two known electrogenic channelrhodopsins ChR2 and MvChR1 as positive controls (see table). We measured >200 pA photocurrent for ChR2, but to our surprise we did not detect any photocurrent in 4 out of 9 new opsin-like genes, despite high sequence homology in transmembrane regions to ChR2. Furthermore, for the positive control MvChR1 which was previously reported to have >100 pA photocurrent in HEK293 cells, we did not detect any current (Fig. 3.1). Upon consultation with the Spudich lab (who discovered MvChR1), we realized all-trans retinal supplementation might be necessary. We patched MvChR1 again with 1 uM retinal and measured photocurrents comparable to those previously reported[22]. We proceeded to rescreen the opsin-like genes with retinal supplementation.

Retinal supplementation is generally unnecessary for mammalian cells and organisms because endogenous retinal production is sufficient for functional opsin expression[20]. However, it is possible that opsin-like genes requiring retinal supplementation in HEK293 cells may have lower retinal binding affinities. We compared the effects of retinal supplementation in primary hippocampal neuron culture for ChR2 and Gene66 and did not observe differences between supplement-free and supplemented media (Fig. 3.1). However, it should be noted that the neurons were cultured in media containing B27 supplements, which has the retinal precursor

vitamin A. It is therefore possible that the discrepancy between retinal supplementation in HEK293 and cultured neurons for Gene66 is the result of higher retinal levels in the neuron culture media.

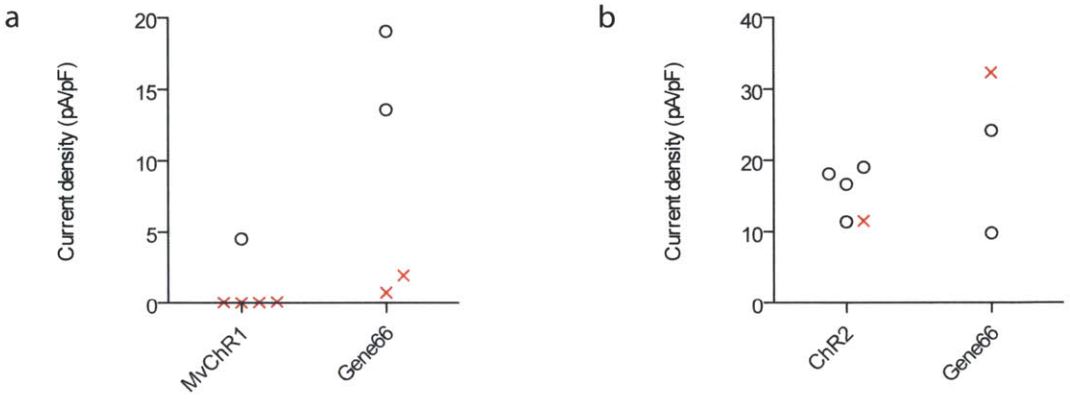


Figure 3.1: Effect of all-trans retinal (ATR) supplementation on current density. a) HEK293 cell comparison of MvChR1 and Gene66 (MvChR1 543 or 575 nm; Gene66 434 nm). b) Cultured neurons comparison of ChR2 and Gene66 (470 nm for both constructs). Peak current density was measured in response to a 1 second light pulse at 10 mW/mm^2 with cells voltage clamp to -65 mV. Each cell is denoted as either a black circle for 1 to 5 uM ATR supplementation or a red X for no retinal supplementation.

We next characterized opsin photocurrent decay kinetics (tau off) at -60 mV in HEK293 cells and cultured neurons (Fig. 3.2). ChR2 was used as a positive control, and we did a single exponential fit for photocurrent decay after a short 5-10 ms pulse illumination. We observed no difference in the ChR2 tau off between HEK293 and neurons and the measured tau off values were within the 8-20ms range reported in the literature[16]. However, we observed the tau off in neurons as compared to HEK293 cells for Gene65 was more than two times greater (*$P<0.05$),

despite blocking neurons with tetrodotoxin (TTX) and 2-Amino-5-phosphonopentanoic acid (AP5). Given that the tau off kinetics is an opsin biophysical property and that the membrane properties of HEK293 and neurons are similar, the discrepancy is unlikely to be the result of opsin structure or membrane biomechanics. Previous literature have observed prolonged tau off kinetics for ChR2 in neurons due to a combination of space clamp error and prolonged opsin depolarization leading to recruitment of secondary calcium currents[23]. It is highly likely that the slower Gene65 tau off kinetics observed in neurons is also due to recruitment of endogenous receptor/channels in neurons, meaning that the functional tau off is not necessarily the same as biophysical tau off and that it may be better to measure kinetics in the intended cell type of application (in this case neurons).

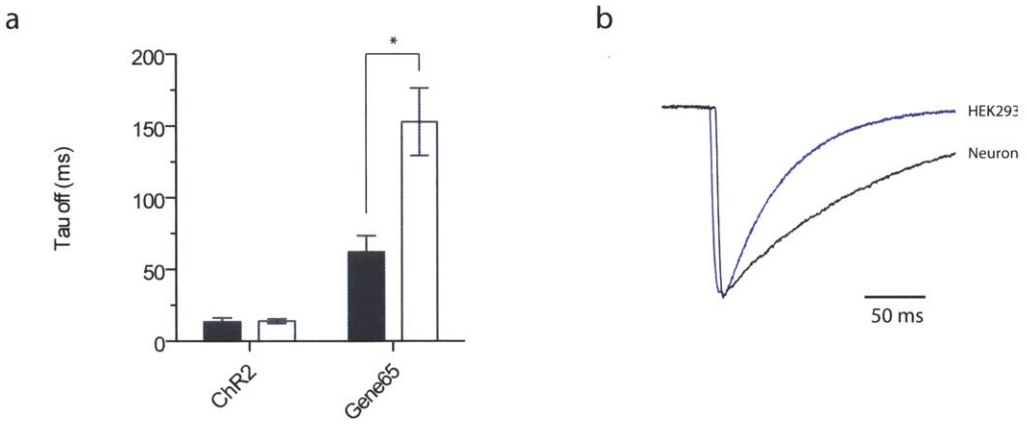


Figure 3.2: Channelrhodopsin tau off kinetics comparison between HEK293 cells and cultured neurons. a) ChR2 has the same kinetics in both HEK293 cells (filled in) and cultured neurons (hollow), whereas Gene65 has significantly slower kinetics in neurons. b) Normalized traces of Gene65 tau off in response to 10 ms and 5 ms pulse of blue (470 nm) light in HEK293 and neurons respectively. Traces are aligned to the time when the light is turned off.

Another key opsin property is membrane trafficking, which is often a dominant factor in neuroscience applications. While HEK293 cells can provide information about whether an opsin molecule can traffic to the plasma membrane, the degree of membrane trafficking (plasma membrane bound versus cytosolic) is not generally representative of neuronal membrane trafficking. In addition, HEK293 cells cannot provide any information about the axonal and dendritic membrane trafficking which is frequently desired in many neuroscience contexts. We therefore chose to assess membrane trafficking solely in neurons.

Previously, multiple labs have shown that trafficking motifs and even fluorophore choice can impact functional opsin membrane expression[24-26]. We tested one of the most widely used trafficking motifs adapted from the Kir2.1 channel[24, 25, 27, 28] and found ChR2 has similar photocurrent with or without Kir trafficking sequences (Fig. 3.3), but bright protein aggregates were observed for the Kir-trafficked version. Gene66, on the other hand, has significant intracellular aggregation without trafficking motifs, and the addition of Kir2.1 motifs removed puncta. However, the Kir trafficking version of Gene66 with mCherry fusion did not remove puncta, suggesting interference between fluorophore and trafficking motifs. The highly context dependent effect of Kir trafficking motifs on opsin trafficking renders comparisons between different trafficking variants difficult to interpret.

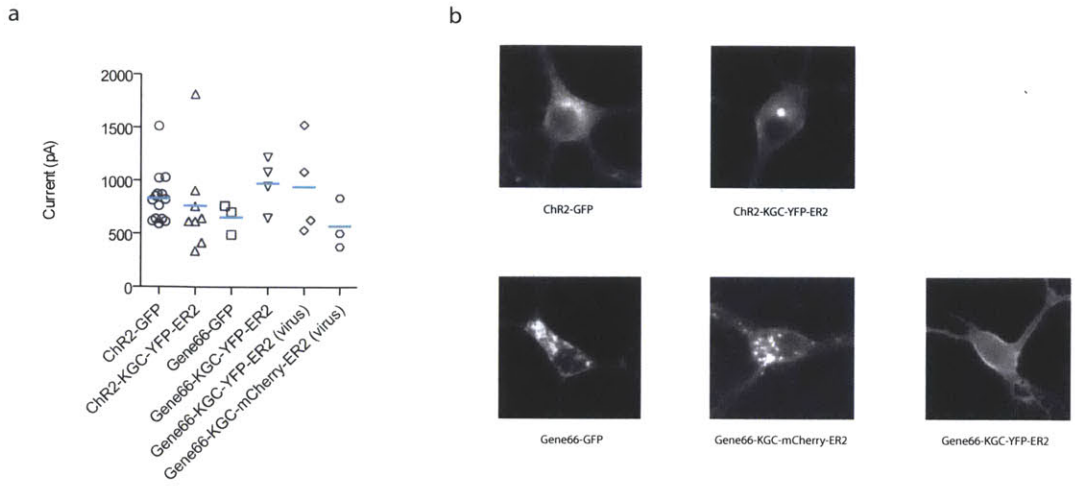


Figure 3.3: Complex interaction of channelrhodopsin, Kir2.1 trafficking sequence, and fluorophore on membrane trafficking and function. a) Photocurrent comparison in cultured neurons of ChR2 and Gene66 trafficking variants shows no significant difference in function. KGC and ER2 are Kir2.1 trafficking motifs for golgi and endoplasmic reticulum export respectively. ChR2 and gene66 trafficking variants were illuminated with 470 nm or 406 nm light respectively at 10 mW/mm^2. b) Representative wide-field images to illustrate subcellular expression pattern. ChR2 with Kir2.1 trafficking has bright perinuclear puncta. Gene66 membrane trafficking can be improved with Kir2.1 trafficking but is dependent on the choice of fluorophore. Neurons were transfected with calcium phosphate unless otherwise noted.

Another issue that arose from screening Gene66 was the difficulty in finding opsin-expressing neurons. Gene66 generally expressed at significantly lower levels than ChR2 and the corresponding fused YFP fluorescence was not visible by eye even in fully lentivirally-transduced coverslips. Because the highest expressing cells were the only ones which could be

visualized and thus patched, the dataset is greatly biased (see gene66 photocurrent data in Fig. 3.3 and Fig. 3.5 for comparison of biased versus unbiased neuron selection).

These initial screening results yielded several important points: 1) the retinal concentrations in culture media may be insufficient for saturation of the opsin binding site, 2) functional kinetics may differ depending on cell type due to recruitment of secondary channels, 3) the effect of membrane trafficking motifs is highly context dependent, and 4) an alternative method for visualizing opsin-expressing cells is needed for unbiased characterization.

gene #	Phylum	Genus	Species	Functional in	Added retina	E123 analog	H134 analog	Comment
59		Phytolacca	americana	Yes				
60	Streptophyta	Mesostigma	viride	No	Yes	E	A	
61		Dioscorea	villosa	No				
62		Neochlorosa	sp.	Yes		E	H	
63	Chlorophyta	Monomastix	opisthostigma	No		E	A	
64	Chlorophyta	Scherffelia	dubia	Yes		E	H	
65	Chlorophyta	Brachiomona	submarina	Yes		E	H	
66	Chlorophyta	Tetraselmis	striata	Yes		E	H	
67	Chlorophyta	Tetraselmis	chui	No		E	Q	
68	Chlorophyta	Tetraselmis	chui	Yes		E	H	
69	Chlorophyta	Tetraselmis	chui	No		A	A	
70	Chlorophyta	Spermatozop	exsultans	No		Q	A	
71	Chlorophyta	Pedinomona	minor	No	Yes	E	H	
72	Glaucophyta	Cyanophora	paradoxa	Yes		D	D	Inward pump?
73	Chlorophyta	Stephanosph	pluvialis	Yes		E	H	
74	Chlorophyta	Haematococ	droebakensis	Yes		E	H	
75	Chlorophyta	Tetraselmis	striata	No	Yes	E	H	
76	Chlorophyta	Tetraselmis	chui	No	Yes	E	H	
77	Chlorophyta	Tetraselmis	cordiformis	Yes		E	H	
78	Haptophyta	Pavlova	lutheri	No	Yes	D	D	
79	Chlorophyta	Scherffelia	dubia	No	Yes	T	A	
80	Chlorophyta	Brachiomona	submarina	Yes		E	H	
84	Cryptophyta	Rhodomonas	sp.			D	D	
85	Chlorophyta	Chloromona	reticulata-A	Yes		E	H	
86	Chlorophyta	Chloromona	oogama	Yes		E	H	
87	Chlorophyta	Chloromona	subdivisa	Yes		E	H	
88	Chlorophyta	Chlamydomc	noctigama	Yes		E	K	
89	Chlorophyta	Chlamydomc	noctigama	Yes		E	H	
90	Chlorophyta	Stigeocloniu	helveticum	Yes		M	H	
91	Chlorophyta	Microthamni	kuetzigianum-A	Yes		E	H	
92	Chlorophyta	Chlamydomc	bilatus-A	No		E	H	
93	Chlorophyta	Heterochlam	inaequalis	Yes		E	H	
95	Chlorophyta	Chloromona	reticulata-A	Yes		E	H	
96	Chlorophyta	Chlamydomc	noctigama			Q	L	
97	Cryptophyta	Proteomona	sulcata	Yes		A	Q	reversal potential near 0mV
98	Cryptophyta	Cryptomona	curvata			E	V	
99	Cryptophyta	Chroomonas	sp.			E	L	
100	Cryptophyta	Chroomonas	sp.			D	D	
101	Cryptophyta	Chroomonas	sp.			D	L	
102	Cryptophyta	Chroomonas	sp.			D	V	
103	Cryptophyta	Proteomona	sulcata			D	Q	
104	Cryptophyta	Chroomonas	sp.			D	T	
105	Glaucophyta	Gloeochaete	wittrockiana			A	A	
106	Cryptophyta	Hemiselmis	virescens			D	D	
107	Cryptophyta	Proteomona	sulcata			D	D	
108	Cryptophyta	Proteomona	sulcata	Yes		D	D	reversal potential near 0mV
109	Cryptophyta	Proteomona	sulcata			D	D	
110	Cryptophyta	Rhodomonas	sp.			D	T	
111	Chlorophyta	Chlamydomc	bilatus-A	Yes		E	H	
112	Chlorophyta	Asteromona	gracilis-B	Yes		S	H	
113	Chlorophyta	Pyramimona	parkeae	No		E	A	
114	Chlorophyta	Monomastix	opisthostigma	No		E	A	
115	Chlorophyta	Tetraselmis	striata	No		E	H	
116	Chlorophyta	Lobomonas	rostrata			E	P	
117	Chlorophyta	Lobomonas	rostrata			E	H	
118	Chlorophyta	Stichococcus	bacillaris			D	D	
119	Chlorophyta	Hafniomona	reticulata			E	H	
120	Chlorophyta	Chlamydomc	bilatus-A	No		E	H	
121	Chlorophyta	Hafniomona	reticulata			E	H	
122	Chlorophyta	Carteria	crucifera			E	H	
123	Chlorophyta	Carteria	crucifera			E	H	
124	Chlorophyta	Volvox	aureus			E	H	
125	Chlorophyta	Phacotus	lenticularis			E	H	

Table 3.1: Summary of opsin screening in HEK293 cells and definition of gene codenames.

3.2 Systematic Screen in Neuron Culture

Based on the results of our initial screen, it is clear that in order to gather relevant trafficking information, we needed to conduct further screens in neurons using a standardized backbone and we also needed an unbiased selection method to quantify different molecules' efficacies. The simplest method for removing selection bias is to use a secondary fluorophore uncorrelated in intensity to the opsin-GFP intensity. Several possible expression methods for co-expressing a separate fluorophore with the opsin exist, including the 2A peptide, IRES, and co-transfection with separate fluorophores[25, 29-31]. We decided not to use 2A peptides or IRES because the fluorophore expression is likely correlated with the opsin expression[32], while co-transfecting with a fluorophore under a different promoter and backbone should generate an expression profile uncorrelated with that of the opsin-GFP backbone. For the systematic screen in neurons, we transfected the opsin-GFP under CaMKII promoter and tdTomato under the ubiquitin promoter at a 10:1 ratio and patch-clamped 7-10 days later (Fig. 3.4), selecting neurons based on tdTomato fluorescence and cell morphology. Quantitative imaging of both GFP and tdTomato were performed in order to compare fluorescence across all opsins. We illuminated with 470nm, 530nm, and 625nm light using 5 ms pulses at equal photon flux (2.5×10^{20} photons/s/mm^2) to determine primary color sensitivity, and additionally used 1 s pulse of 660 nm light at 10 mW/mm2 to screen for red-light sensitivity. All protocols included a 30 second dark period to allow the opsin to recover to ground state.

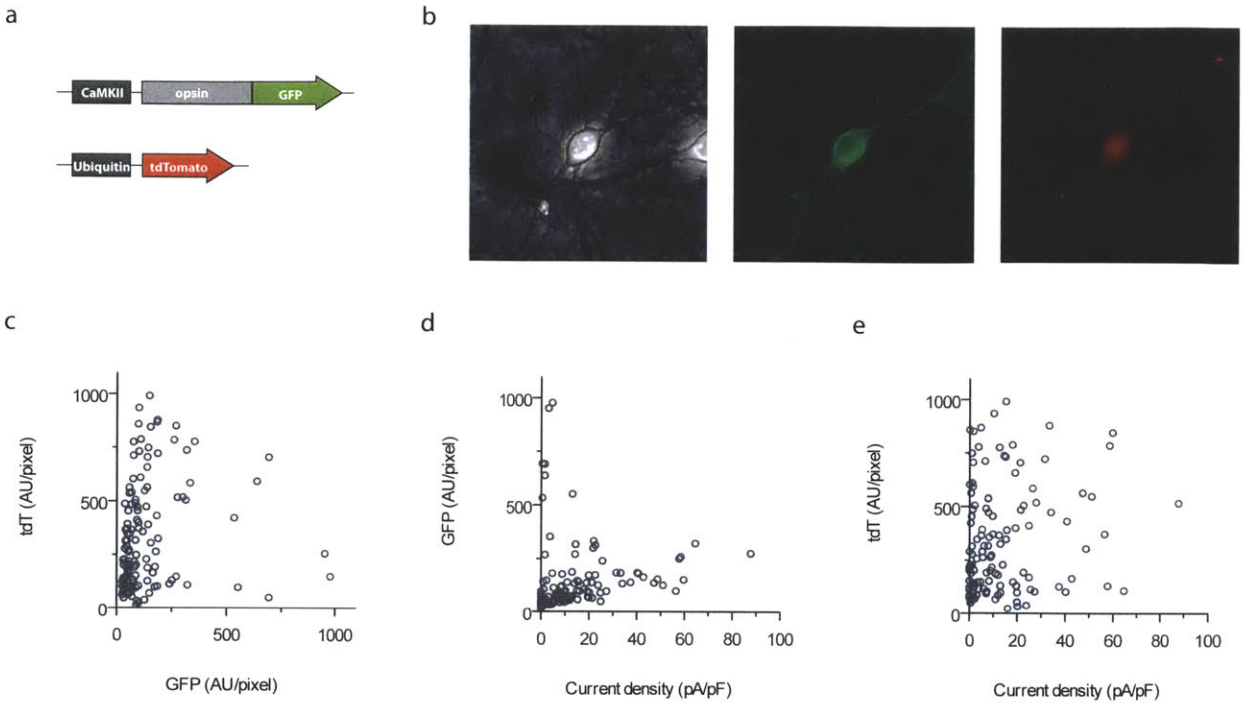


Figure 3.4: Unbiased screening in cultured neurons. a) Double plasmid transfection scheme to enable independent visualization of transfected cells and opsin-expressing cells. b) Representative widefield images of the same neuron taken in phase contrast (left), GFP (middle), and tdT (right). Yellow dashed lines in the phase contrast image is the mask boundary used for quantifying soma fluorescence. c-e) Comparison of aggregate fluorescence and current density for all opsins screened.

The aggregate opsin-GFP and tdTomato imaging data measured over soma is shown in figure 3.4c. We did not observe a correlation between GFP and tdTomato intensity when data from all constructs was pooled. Both fluorophores spanned 2 logs of intensity, suggesting a high expression variance. Globally, there was a positive correlation between GFP intensity and current density, although this was not necessarily true for individual constructs (Fig. 3.6). A subset of low and high GFP fluorescing cells had no photocurrents, indicating poor trafficking

and/or conductance. However, there were no occurrences of low GFP intensity with high current density, suggesting no opsin with an obvious increase in conductance. As expected, tdTomato was uncorrelated with current density as a population or across individual constructs, confirming that the experimenter cannot bias toward cells with high current densities by selecting for tdTomato expression.

Individual opsin fluorescence and peak wavelength photocurrent are shown in figure 3.6. As previously mentioned, GFP and current are not always correlated, most likely due to the fact that fluorescence measurements were gathered by widefield rather than confocal microscopy, meaning measurements were carried out on total cell fluorescence rather than membrane fluorescence alone. The predominant bias in fluorescence appears to be the presence of bright puncta and general perinuclear aggregates (e.g. gene89). Additionally, the highest photocurrent cells did not necessarily have the best membrane expression (e.g. $C1V1_{TT}$, gene64), but mainly had an increase in total fluorescence. Only ChR2, C1V1TT, gene87, gene88, and gene90 have a significant GFP and photocurrent correlation ($P<0.05$, Spearman correlation).

It is interesting to note that the GFP fluorescence range is different for each opsin, based on an unbiased sampling. For example, gene108, consistently has more than four times brighter GFP density than other opsin genes with many cells actually saturating the camera. Since all experiments were conducted with the same promoter/vector and from the same batches of neuron culture, the respective neuron protein production capacities are likely to be comparable, suggesting the differences in GFP fluorescence are due to protein folding efficiency and/or how well the opsin molecule is degraded and tolerated in neurons.

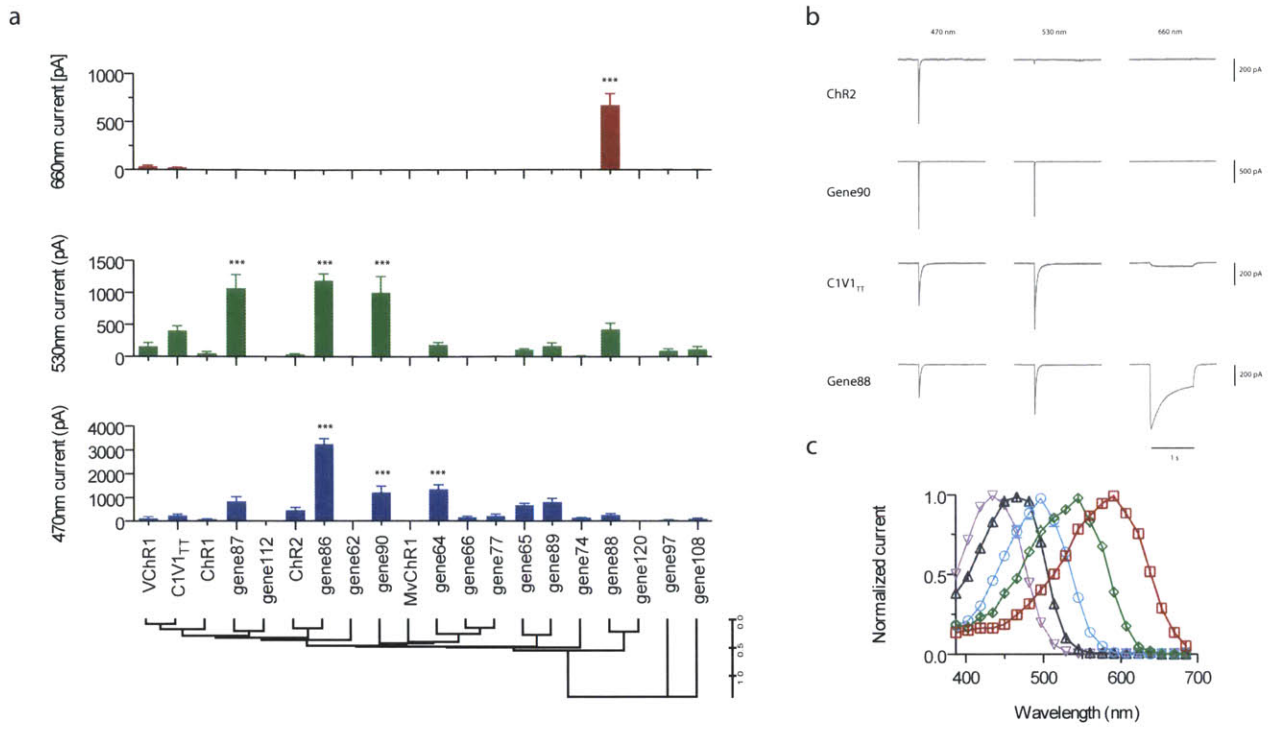


Figure 3.5: Novel spectral classes of algal channelrhodopsins discovered through functional screening. a) Channelrhodopsin photocurrents in cultured neurons in response to 1 s red (top), 5 ms green (middle) and 5 ms blue light (bottom). Light powers are 10, 3.66, and 4.23 mW/mm^2 for red, green and blue respectively (matched photon flux for blue and green). Channelrhodopsins are clustered based on sequence homology. b) Representative voltage-clamped traces of channelrhodopsin currents from screening data. c) Action spectra measured in HEK293 cells at equal photon flux from 387 nm to 695 nm. From left to right: gene66, ChR2, gene90, $C1V1_{TT}$, and gene88.

Photocurrent measurements show that all opsins have functional blue and green light sensitivity, with gene64, gene86, and gene90 having significantly higher current than ChR2 in the blue

(P<0.0001) and gene86, gene87, and gene90 having significantly higher current than $C1V1_{TT}$ in the green (P<0.0001). To our surprise, we found gene88 also had red light sensitivity and did not have increased current in the green or blue relative to $C1V1_{TT}$, indicating the red light sensitivity was not due to increased overall amplitude but rather spectral shift. We measured action spectra in HEK cells and observed gene88 was 45 nm more red-shifted than any previously natural or engineered channelrhodopsin. We also tested gene120, the closest homolog to gene88 in sequence, but did not observe any photocurrents most likely due to protein trafficking or folding issues (Fig. 3.6).

We also measured current decay kinetics (tau-off) using the 5 ms photocurrent data, and to our surprise we observed that the kinetics of naturally occurring opsins span 3 logs in range. Gene89 has the slowest tau-off of 1609 ± 635 ms, but does not possess the analogous C128 or D156 mutations stabilizing the open channel state in ChR2 mutants[13, 14], suggesting the mechanism for open state stabilization may be different than ChR2 and further mutations can be possibly introduced to make gene89 bistable. On the other hand, gene90 has the fastest tau-off at 3.59 ± 0.21 ms, 1 ms lower than the previous lowest reported tau-off[12]. Interestingly, gene90 has a methionine instead of glutamic acid in the E123 position, although previous literature has shown that mutating E123 to alanine, threonine, or serine decreases the ChR2 tau-off[12]. Out of all opsins with greater than 1nA current in response to blue or green light pulses, gene86 is the only opsin with tau-off greater than 20 ms. However, these slow kinetics are probably not biophysical differences, but rather poor space clamp in neuronal processes: when we measured gene86 tau-off at lower irradiances, we observed faster tau-off kinetics and lower current (Fig. 3.7). We additionally used CsCl2 pipette solutions to reduce space clamp error and observed fast tau-off

kinetics comparable to ChR2, suggesting gene86 has a slow apparent tau-off kinetics due to secondary channel recruitment in neurons.

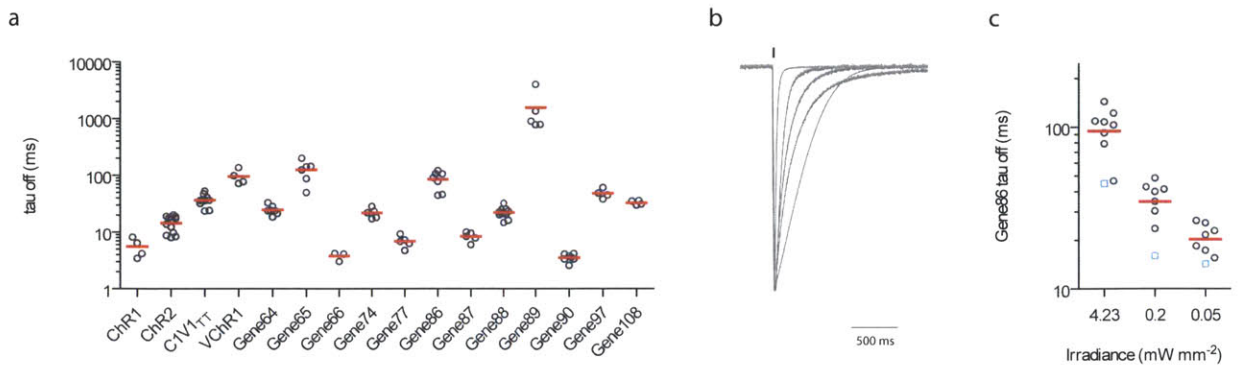


Figure 3.7: Channelrhodopsin tau-off kinetics in cultured neurons. a) Single exponential fit of tau-off kinetics from different opsins in response to 5 ms light at peak sensitivity. b) Representative traces from left to right: gene90, ChR2, gene88, $C1V1_{TT}$, gene86. c) Gene86 tau-off measured as function of blue (470 nm) irradiance. Black circle and blue square cells were recorded with potassium and cesium chloride intracellular solutions respectively.

For opsins with more than 1nA current in the screen, we additionally measured inactivation and recovery kinetics in order to characterize opsin photocurrent fidelity. Previous channelrhodopsin literature characterizes recovery via paired pulse illumination, with a fixed duration between sets of paired pulse illuminations to allow dark-state recovery, but with a varying time interval between the first and second pulse to study the photocurrent recovery rate[15, 33]. The recovery rate is then measured by dividing the second pulse amplitude by the first pulse amplitude for a given set. However, the first pulse in each paired-pulse illumination from these protocols redefines the ground-state amplitude within the context of that set, guaranteeing all opsins will appear to fully recover, even if this is only an artifact resulting from the first pulse inactivating to the same

degree as the second pulse over time and multiple sets. In order to characterize the true recovery rate, we do not redefine the ground state each time. Rather, we simply use one initial pulse as baseline, and then measure photocurrent response after 1 and 30 seconds in the dark. The raw traces and recovery ratios are shown in figure 3.9. Note that no genes recovered to 100% after 30 seconds in the dark including ChR2 and this is not due to errors in patch clamp (i.e. access resistance, membrane resistance and capacitance were the same throughout protocol).

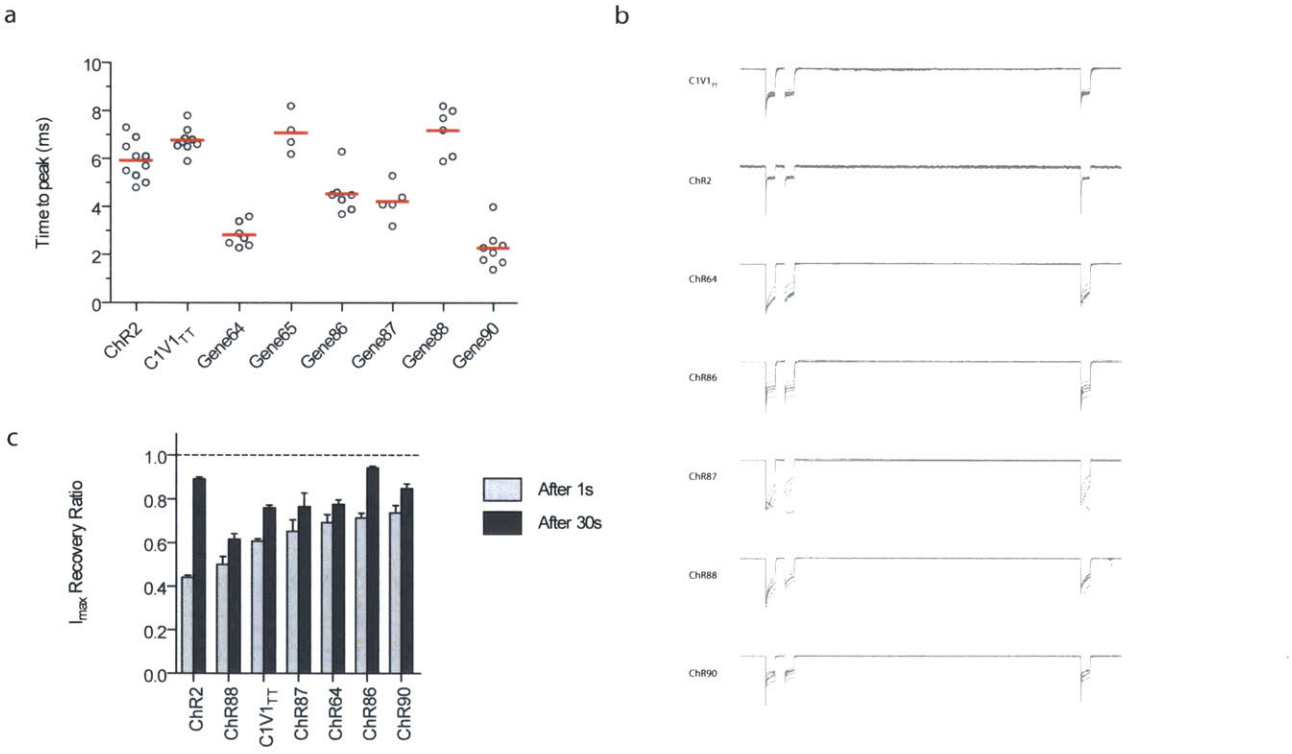


Figure 3.9: Channelrhodopsin turn-on, inactivation, and recovery kinetics. a) Time-to-peak is measured as the time from light onset to 90% of peak value. b) Normalized traces of recovery kinetics, with each patched cell overlaid. All three pulses are 1 second long, with 1 second wait

in the dark between the first and the second pulse, and 30 seconds wait in the dark between the second and the third pulse. c) Recovery ratio of peak current using the first pulse as baseline.

The first pulse in the recovery traces show each opsin's inactivation kinetics. Of these, ChR2 has the most consistent inactivation and recovery kinetics, with peak-normalized traces from different ChR2 expressing cells perfectly aligning. All other examined opsins had large cell-to-cell variances, with some gene87-expressing cells even showing increasing current in the first pulse rather than the decreasing current that would be expected from inactivation. The cell-to-cell variation in different traces suggests each cell may have started with a different light:dark adaptation ratio and that not all cells were initially in the ground state. No genes other than ChR2 reached a true steady state value within the first pulse, so there will also be some bias in the recovery rate since not all cells were driven to the same state at the end of the first pulse. However, the illumination and dark period lengths are similar to those used in optogenetic experiments, suggesting these results are reflective of actual performance in a neuroscientific context.

ChR2 has the worst inactivation of all measured-opsins with tau desensitization of 36 ms (Fig. 3.7). ChR2 additionally has the worst short-term recovery, with the transient current recovering to only 44% after waiting only one second in the dark, but ChR2 fully recovers to 89% after 30s in the dark. The new red sensitive gene88 has slower inactivation with a tau desensitization of 343 ms , but it has a very poor recovery with only 62% recovery ratio after 30s in the dark. However, gene88 eventually recovers to >90% after 60 to 120 seconds wait in the dark (data not

shown): practically speaking, this means gene88 will be more history-dependent than ChR2 without sufficient time for recovery in the dark.

We finally measured the light-onset-to-90%-of-peak time to quantify channel opening kinetics. Fast turn on kinetics are usually desired because most optically driven spikes are done with 5ms or shorter pulses. We found gene88 to have the slowest turn-on kinetics of 7.2 ms, comparable to the previously published $C1V1_{TT}$, but this data was gathered at 625 nm, more than 30 nm off-peak. Based on action spectrum data, the gene88 turn-on kinetics at 590nm are estimated to be 5-6 ms. Gene64 and gene90 had the fastest turn on kinetics at 2.8 and 2.3 ms, more than twice as fast as ChR2 ($P<0.0001$).

From our systematic opsin screening in neurons, we have identified the first known yellow-peaked channelrhodopsin gene88 and many improved green- and blue-light sensitive channelrhodopsins. Gene90, in particular, was found to have the fastest channel turn on, turn off, and recovery kinetics of all known opsins, suggesting gene90 may have the best spike frequency and fidelity. Further experiments in non-TTX treated neurons were necessary to validate the optical spiking performance of these newly discovered genes.

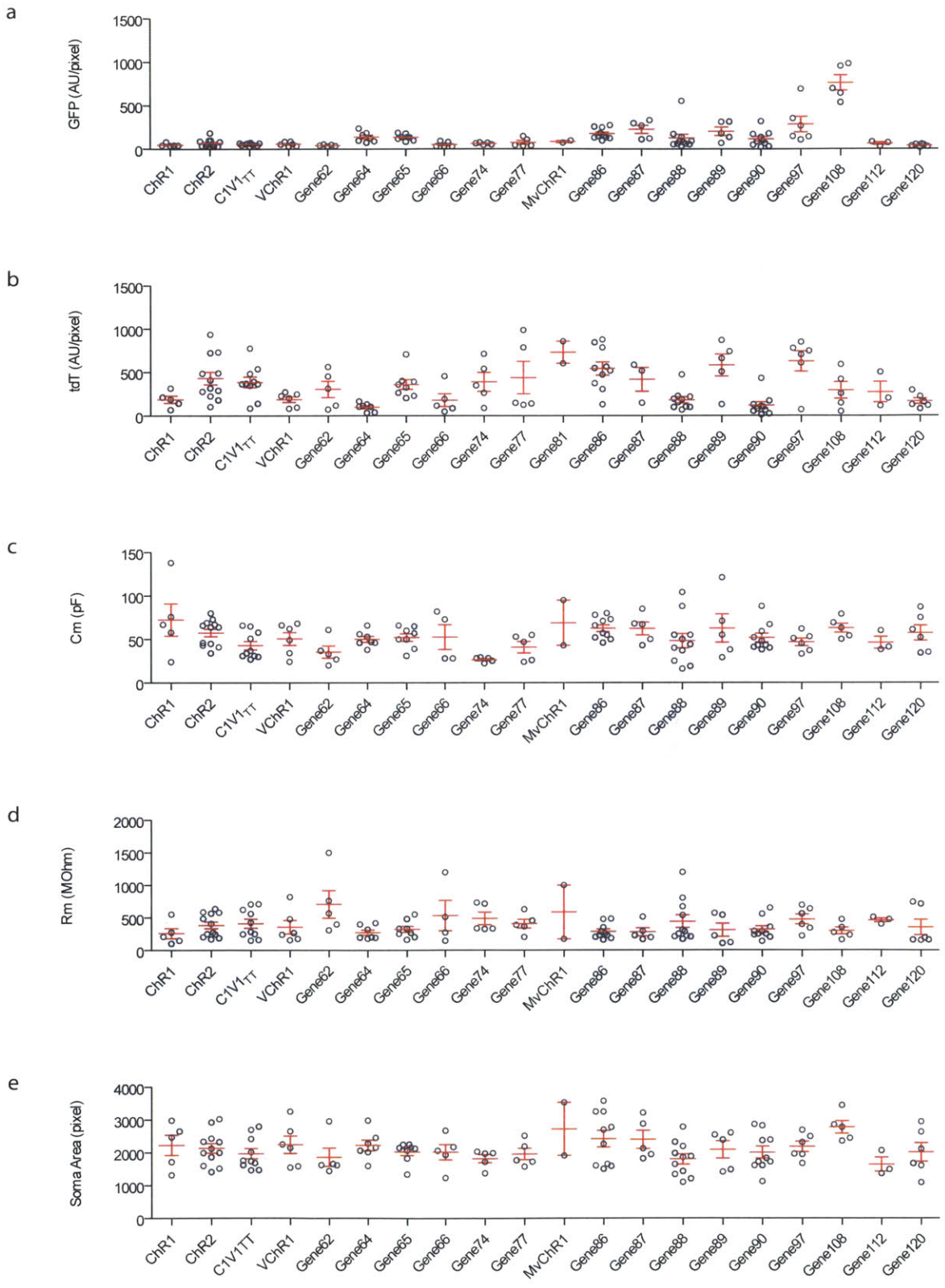


Figure 3.8: Neuron and patch-clamp parameters for opsin screening data.

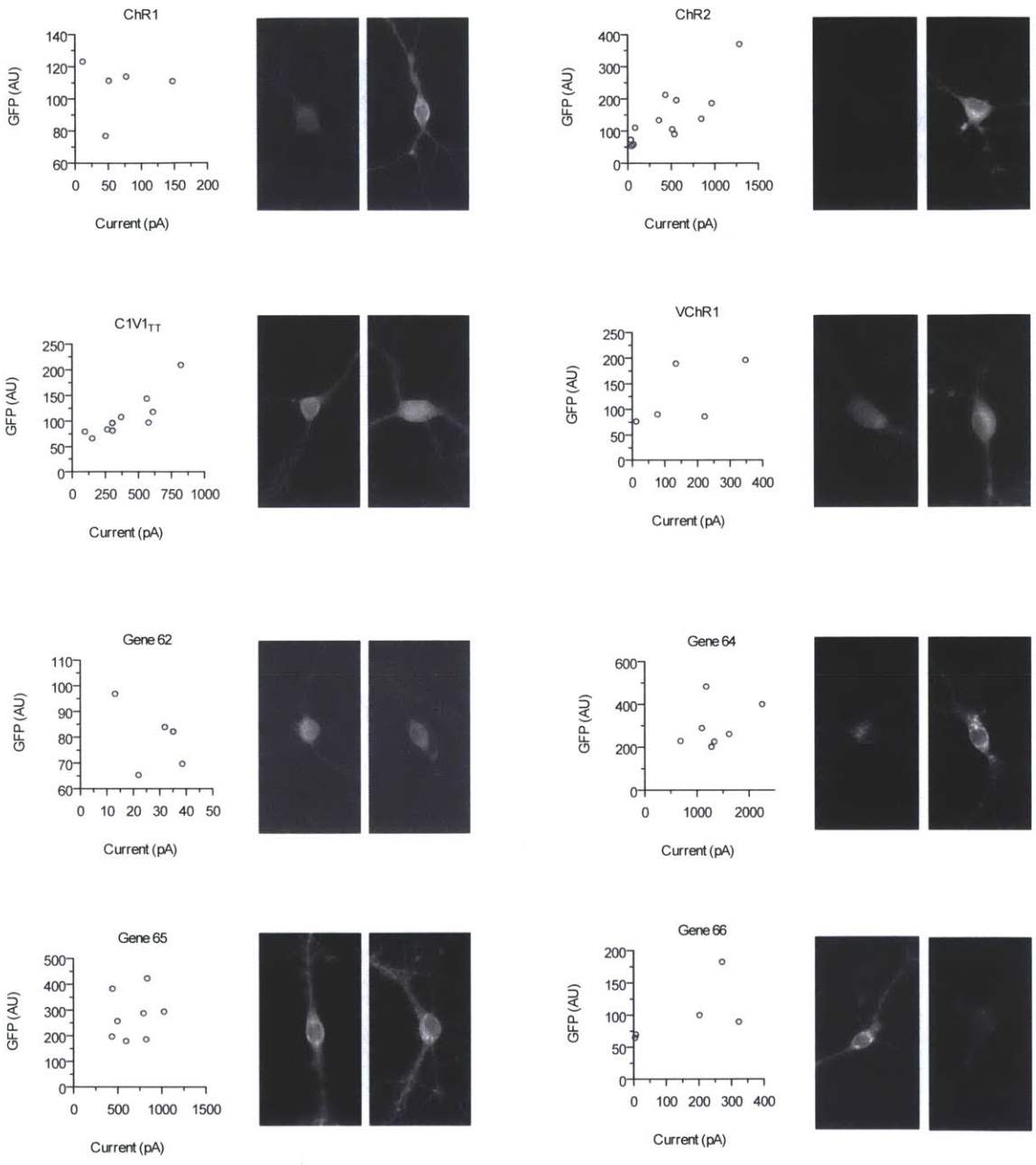


Figure 3.6 continued

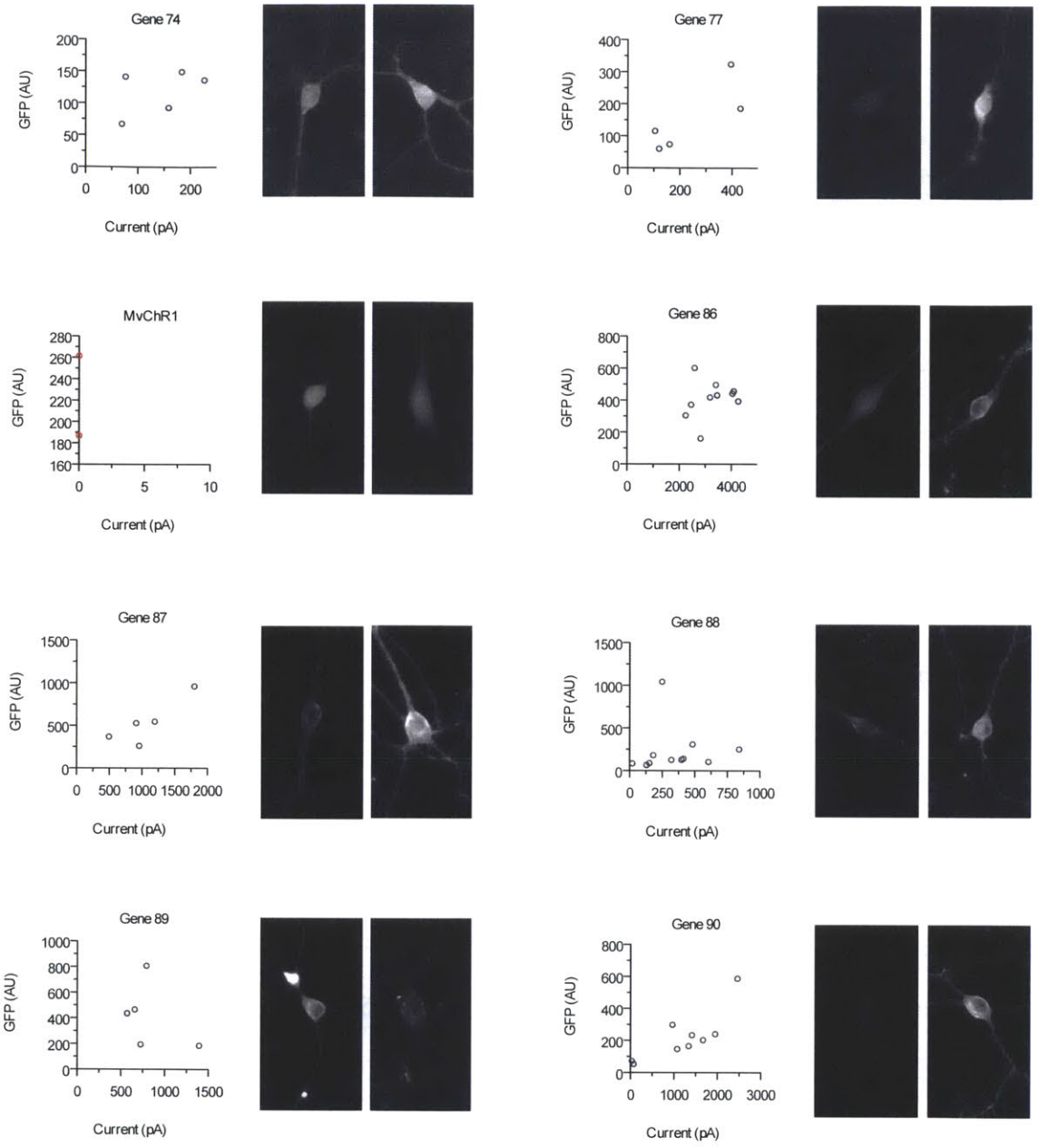


Figure 3.6 continued

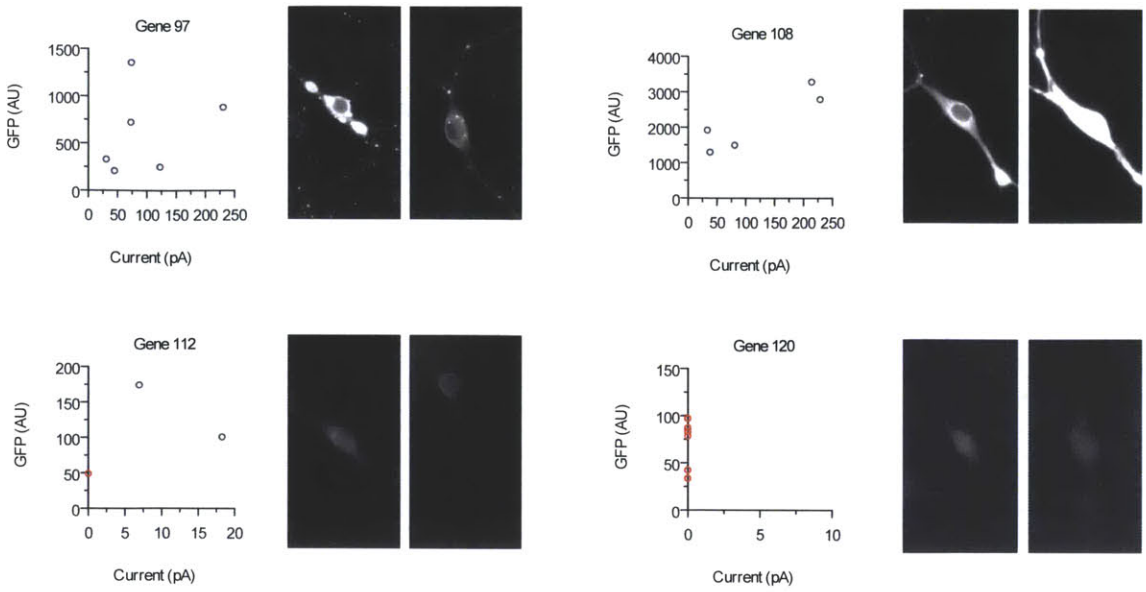


Figure 3.6: Channelrhodopsin trafficking and photocurrent comparison from cultured neuron screening. Quantitative soma GFP fluorescence versus peak sensitivity current (left), GFP images for the median current (middle) and the maximum current (right) cells are shown for each opsin (left). Note that in the GFP versus current plot, the GFP intensity values are absolute and can be compared between different opsins. However due to two logs variance in absolute GFP intensity, the brightness and contrast settings are varied for different opsins images shown, so the GFP images should not be used to compare brightness between different constructs. However the brightness and contrast settings for each opsin's median and maximum current cells are matched, to illustrate whether higher expression correlates with increased photocurrent.

3.4 High-Frequency Green Light Driven Spiking

Results from our screen in neurons found only gene86, 87, and 90 to have over 1 nA current in response to 5ms pulses of green light at 3.66 mW/mm^2. Previous literature on channelrhodopsin-mediated spiking concluded channel closing kinetics and recovery rate to be the crucial parameters in determining spike fidelity and frequency[12, 33]. Due to gene86 apparent slow tau-off

kinetics, we decided to only test spiking performance of gene87 and 90 along with $C1V1_{TT}$, the fastest published green light sensitive channelrhodopsin. Both Gene87 and 90 are significantly faster than $C1V1_{TT}$ in tau-off kinetics and recovery rate, so we expected to see significant improvements in green light driven spiking performance.

Given that optical spiking is inherently dependent on opsin expression levels, we again decided to use a double plasmid transfection method similar to that used in figure 3.4 to avoid selection bias. We selected neurons based on cell morphology and presence of tdTomato fluorescence. Neurons were only included for analysis if they were able to drive spikes with 100% probability with 10 pulses of $5mW/mm^2$ green light at 1Hz; cells that failed to drive spikes at 1Hz are shown in figure 3.11a, all of which have significantly lower GFP expression level than cells that successfully drove 100% spiking at 1Hz. The tdTomato expression levels for both groups were comparable.

We evaluated spike frequency performance using trains of 40 pulses at 5 to 60Hz at three different irradiances (Fig. 3.11c). At the lowest irradiance ($0.5\ mW/mm^2$), all but one gene90 cell had 100% spiking fidelity at or below 40Hz, while no gene87 cells were capable of driving spikes at 100%, and no $C1V1_{TT}$ cells were capable of driving spikes. In general, when stimulation frequencies are within the opsin's biophysical capacity (less than 10 Hz for most opsins), the spike probability performance increases at higher light power, suggesting spike fidelity is highly photocurrent-dependent. At higher frequencies exceeding that capacity, the spiking performance is dependent both on photocurrent and tau-off kinetics such that higher irradiances (or correspondingly higher photocurrents) paradoxically do not necessarily mean

better spike fidelity (gene87 in Fig. 3.11c), due to the fact that increased photocurrents at higher light powers can lead to prolonged membrane depolarization, which results in a depolarization block and sodium channel inactivation as previously reported[15, 17]. Overall, gene90 had significantly higher spike probabilities across all irradiances and frequencies than both gene87 and $C1V1_{TT}$. There were no significant differences between the neurons' resting potential and spike threshold, indicating that the differences in spike probability is not due to neuronal variability but rather opsin expression and kinetic properties.

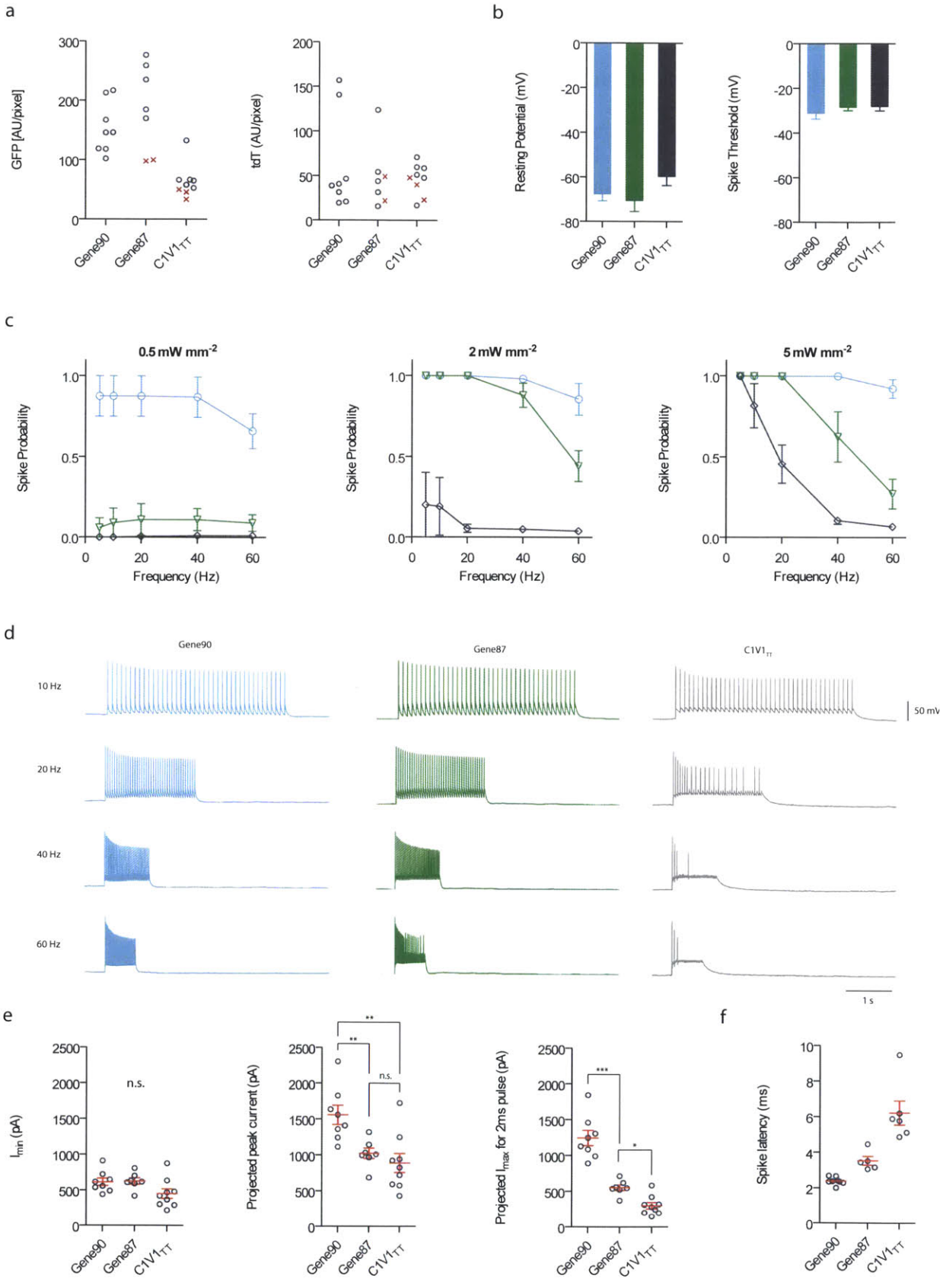

a
GFP [AU/pixel]
tdT (AU/pixel)
Gene90
Gene87
C1V1TT
b
Resting Potential (mV)
Spike Threshold (mV)
c
0.5 mW mm-2
2 mW mm-2
5 mW mm-2
Spike Probability
Frequency (Hz)
d
Gene90
Gene87
C1V1TT
10 Hz
20 Hz
40 Hz
60 Hz
50 mV
1 s
e
Imin (pA)
n.s.
Projected peak current (pA)
Projected Imax for 2ms pulse (pA)
f
Spike latency (ms)

Figure 3.11: Green light driven spike frequency and fidelity analysis. a) tdT was co-expressed with opsin-GFP for unbiased selection of neurons to patch. Any neuron that could not drive a train of 1 Hz 10 pulses of green (530nm) light at 5 mW/mm^2 with 100% spike probability are excluded from further spike frequency analysis. Generally neurons that could not drive spikes at 1Hz (red X) have lower GFP intensity than the neurons that could (black circle). b) Neuron resting and spike threshold controls. c) Spike frequency population average as a function of irradiance. d) Representative spike frequency traces at 5 mW/mm^2. e) Measured minimum current (Imin) in response to 1 second of green light (left). The projected peak current (middle) and maximum current for 2 ms light pulse (right) are calculated based on the measured Imin and previously measured photocurrent ratios done under the same conditions but with TTX. f) Spike latency at 5 mW/mm^2 5 Hz. All spiking protocols used 2 ms light pulses with 30 s wait in the dark.

Previous papers have shown C1V1$_{TT}$ can drive 100% spiking at 20 Hz or less, but we were unable to observe 100% spiking even at 10 Hz. Part of this discrepancy may be due to lower expression levels, as we did not append KGC trafficking sequences to boost membrane expression as those papers did[13]. To investigate whether photocurrent is a primary factor in determining spike fidelity at high irradiances and low frequency, we also measured the minimum photocurrent (I_{min}) at the end of a 1 second pulse of 5mW/mm^2 light (Fig 3.11e). Due to escaped sodium spikes during illumination, we were unable to directly measure the peak photocurrent, and had to predict the peak based on experiments done under similar conditions but with added TTX.

People typically use peak and I_{min} (steady state) photocurrents to estimate the best and worst case scenarios for opsin-driven spiking, as they respectively correspond to the full activated and fully desensitized states. In this case, I_{min} is not statistically different between the opsins, indicating that the opsins are not being driven purely in the steady state. Curiously, the projected peak photocurrent was only higher for gene90, contradicting our spike frequency data showing gene87 is better than $C1V1_{TT}$ at 20Hz and lower. Upon closer inspection, we realized the projected peak photocurrent is not reflective of the actual opsin-generated driving force because none of the tested opsins were able to reach their peak during the measured short 2 ms pulses (Fig 3.9a). When opsin turn-on time is also taken into account, the differences in spike fidelity at or below 10Hz can be explained by the predicted 2 ms photocurrent, suggesting $C1V1_{TT}$ membrane expression (and consequently photocurrent) would need to be boosted by more than four-fold in order to match gene90 spike fidelity.

The unreliability of using peak and steady state photocurrent to predict spike fidelity has several implications. First, channelrhodopsins are not driven in a purely active or desensitized state. The use of peak and steady state photocurrent often respectively over- and underestimate real opsin performances, so they should only be used as demarcating possible boundaries for performance. Second, tau-on kinetics significantly impact spike performance, since opsins are not like electrodes which can generate driving forces in less than 1 ms: this explains why molecules reported to have greater than 1nA peak photocurrent can have poor spike fidelity, even at low frequencies, due to an inability to reach their peak current within the 1-5 ms pulse typically used for neuroscience experimental protocols. It is therefore desirable to have both fast turn-on and turn-on kinetics for high fidelity spiking.

Faster turn-on kinetics are also reflected in spike latency (defined here as light onset to spike peak). Gene90 has the fastest turn-on time and the shortest spike latency at 2.39±0.08 ms: the short jitter suggests gene90 is overshooting the spike threshold and may be at a fundamental biophysical limit, given that spike initiation and width are both 1 ms long. $C1V1_{TT}$ has the longest turn-on time and correspondingly the highest jitter. The high latency jitter indicates $C1V1_{TT}$ does not overshoot the spike threshold during illumination and relies on post-illumination charge integration to elicit spikes, as can be seen in the fact that the spike peak for $C1V1_{TT}$ occurs at 4 to 10 ms, well after the end of illumination at 2 ms. Spike timing and fidelity is therefore generally a function of significant depolarization above the spike threshold can be reliably achieved.

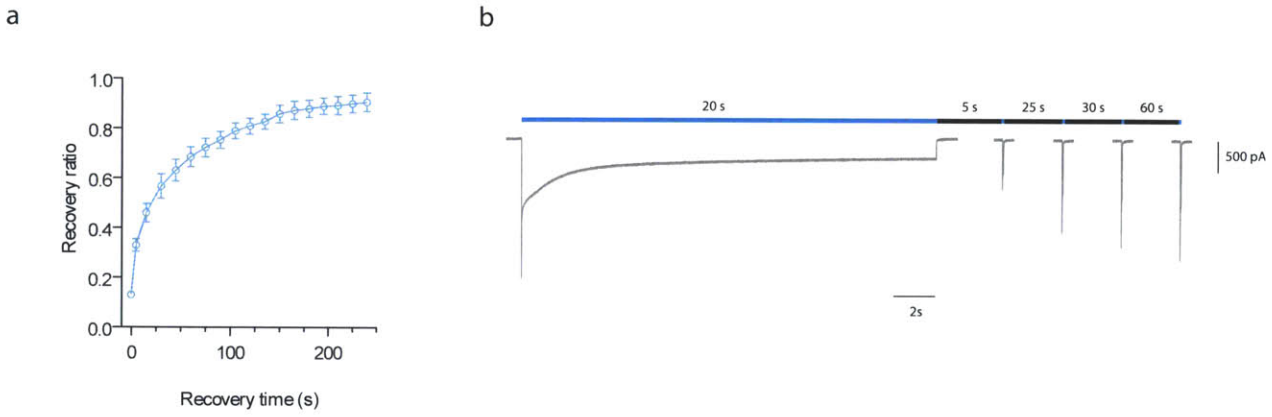


Figure 3.12: Gene90 full inactivation and recovery kinetics in cultured neurons. a) Peak current recovery ratio versus waiting time in darkness. The zero time corresponds to the end of the 20 s illumination. b) Representative traces show gene90 inactivation has a fast and a slow component under continuous illumination. Recovery progress is monitored with 50 ms pulses. All illumination are done with blue (470 nm) light at 5 mW/mm^2.

The spiking protocols used here consist of only 40 pulses, whereas scientific experiments sometimes require stimulation at high frequencies on a seconds-to-minutes timescale. The presented data therefore does not reflect performance over long stimulations, and it is impractical to gather spiking data where each protocol is minutes long. Instead, we conducted separate experiments for gene90 using long 20 s pulses to evaluate the complete inactivation state and recovery time (Fig. 3.12). We found that it takes roughly 10 seconds to inactivate gene90 and the steady state current is only 13% of the peak current. The long continuous illumination provides a lower bound on high frequency stimulation (e.g. 200Hz), demonstrating that one should not expect sustained high frequency spiking performance beyond 5 to 10 seconds if the steady state current is not high enough to drive spikes. The peak current recovery time is fast and 50% recovery can be achieved within 25 seconds, but >80% recovery will take more than 100 seconds suggesting gene90 may not be able to sustain continuous high frequency stimulation for long periods, and will require 1-5 seconds wait in the dark for partial recovery.

The spike frequency data presented here shows gene90 to be the fastest green light drivable channelrhodopsin reported to date. We determined the critical factors for gene90's spike fidelity are its fast turn-on and turn-off kinetics, mimicking the temporal precision of electrical spiking. However, the primary downside to gene90 is its low steady state current, although this will only be problematic for prolonged high frequency stimulation, since gene90 also has the fastest recovery rates among channelrhodopsins (Fig. 3.9). Further engineering to improve gene90's steady state photocurrents may be possible, as most engineering efforts with ChR2 have generally tried to do the same[16, 18, 33]. These results indicate gene90 may be the most optimal

green light drivable channelrhodpsin both for neuroscientific applications and as a scaffold for further improvements.

3.5 Red Light Driven Spiking

Gene88 is the first yellow-peaked channelrhodopsin discovered, with a peak wavelength sensitivity more than 45 nm red-shifted than previous channelrhodopsins[22, 34]. Because opsins possess broad action spectra, even green peaked channelrhodopsins such as C1V1 have been demonstrated to depolarize the membrane and elicit spikes in the red (630 nm) under continuous illumination[13]. In theory, it may be possible to elicit spikes in the far-red to near infrared if a green-peaked channelrhodopsin had a 1-2 order of magnitude improvement in photocurrent, so the distinction between colors is not readily apparent, as spiking is inherently a function of the degree of membrane depolarization. In practice, no green-peaked channelrhodopsins can actually drive precisely timed red-light elicited spikes with pulse lengths shorter than 5 ms pulses, due to the slow turn-on kinetics, since action spectrum reflects different turn-on time in response to different colors. Gene88 has a turn-on time around 7.2 ms in the red, comparable to 5.9 ms for ChR2 in the blue, so as long as expression levels are sufficiently high, gene88 can drive precisely timed spikes at 625nm (Fig. 3.14). In higher expressing cells, gene88 is even able to drive timed spikes at 660nm with 5 ms pulses (Fig. 3.14d).

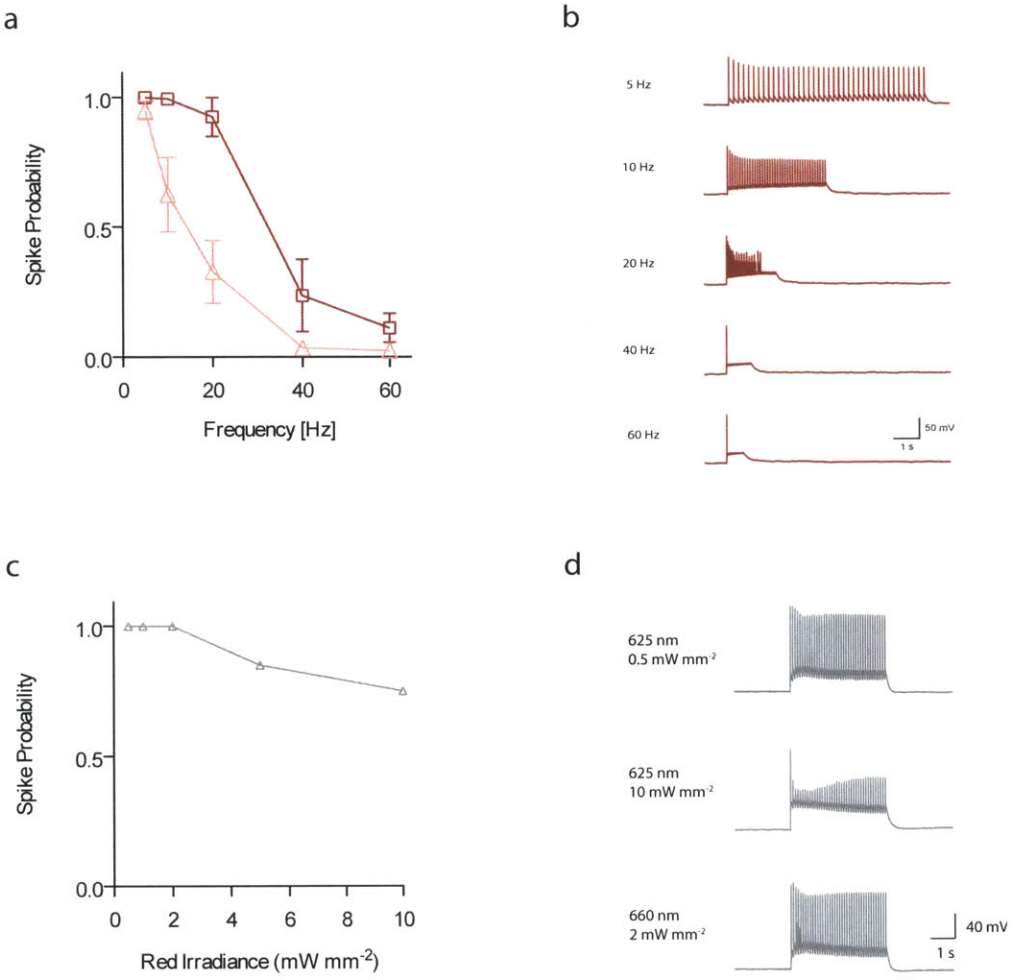


Figure 3.14: Gene88 red light spiking. a) Red (625 nm) light spike frequency data for gene88 neurons selected based on tdT (pink triangle) or gene88-GFP expression (red square). The neurons selected based on gene88-GFP expression did not have tdT co-transfected. b) Representative traces showing gene88 can only reliably drive spikes at lower than 20 Hz. c-d) Irradiance plot for a gene88 neuron optically stimulated with 5 ms pulses of 625 nm light at 5 Hz. Representative traces to illustrate depolarization block by gene88 at 10 mW/mm^2. Gene88 is also able to drive spikes at 660 nm with 5 ms pulse train at 5 Hz.

However, Gene88 cannot sustain higher frequency spiking above 10 Hz even in high expressing cells because (1) gene88 slow turn-off kinetics can cause depolarization block which is exacerbated with higher photocurrents (Fig. 3.14d) and (2) gene88 has poor inactivation recovery rate and higher frequency stimulation generally more greatly inactivates channelrhodopsins. The practical spike frequency limitation for gene88 is roughly 5 to 10Hz .

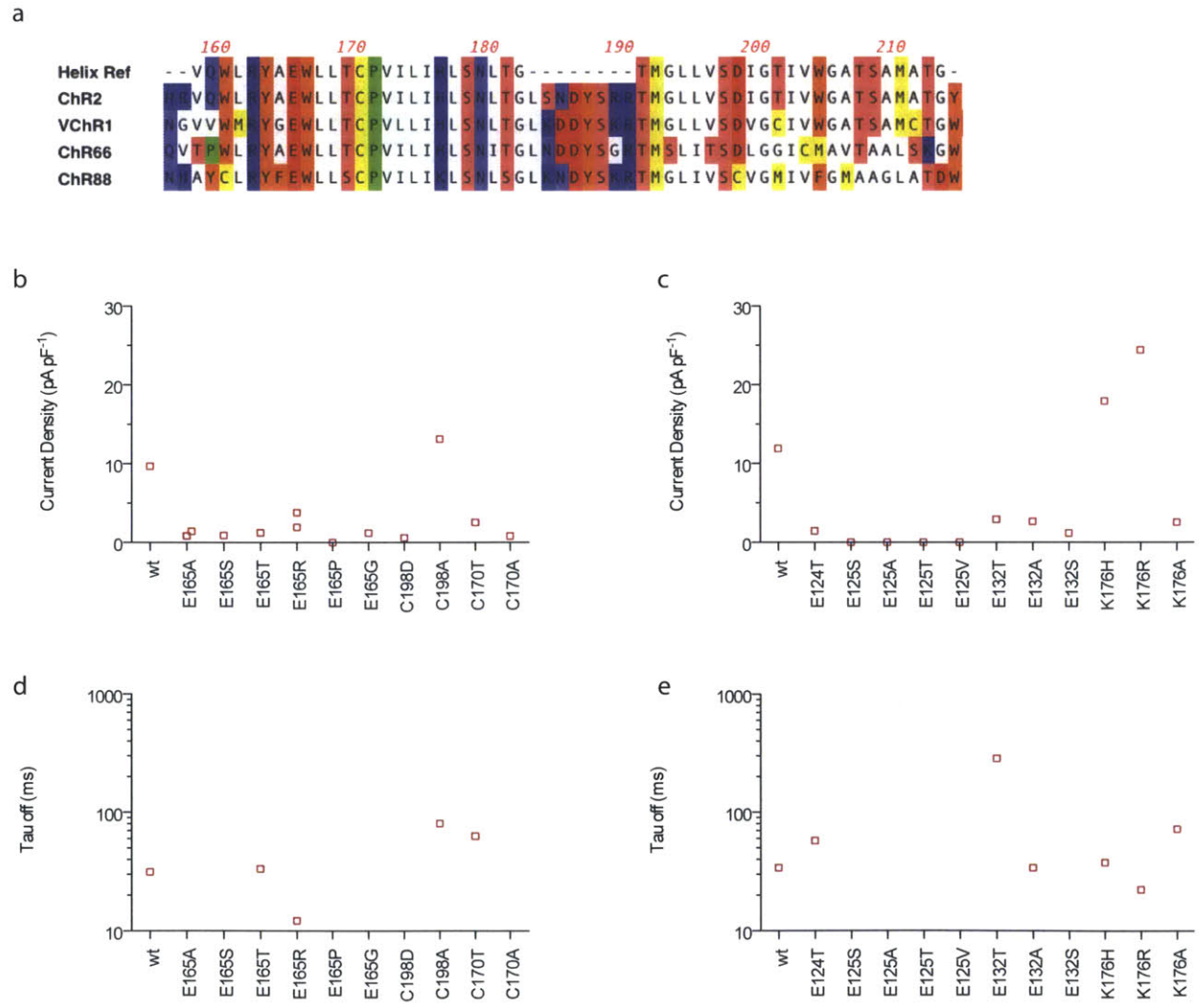


Figure 3.15: Gene88 mutagenesis screen for faster tau-off kinetics in HEK293 cells. a) Channelrhodopsin sequence alignment for the third and fourth helix. Residues were selected for

mutagenesis based on existing ChR2 literature. b,d) Current density and tau-off kinetics from the first round of screening. Gene88 E165 corresponds to E123 in ChR2. c,e) Current density and tau-off kinetics from the second round of screening. Gene88 K176 corresponds to H134 in ChR2.

Both turn-off and recovery kinetics need to be addressed to improve gene88's spiking reliability at higher frequencies. We screened seven residues in gene88 using HEK293 cells based on analogous residues which modulated kinetics in ChR2 or VChR1 derivatives, and measured current and tau-off kinetics based on 1s red-light pulses (Fig. 3.15). We only characterized one or two of the best expressing cells for each mutant, so while the data may skew towards higher values, negative results should indicate actual disruptions of function. We did not report kinetic values for mutants with current densities of less than 2 pA/pF. We found only two residues which improved tau-off kinetics when mutated: E165, which is analogous to the E123 "ChETA" residue in ChR2, and K176 which is analogous to H134 in ChR2. The "ChETA" mutation improves ChR2 kinetics by more than 2 fold, but lowers the "ChETA" mutant's current density due to faster channel closing[15]. We observed a similar trend with E165: E165 mutants have more than five times lower current density than the wildtype. The K176R mutant, on the other hand, improved both tau-off kinetics and current density. The increase in density is probably due to high expression as oppose to conductance (data not shown). The analogous residue H134R in ChR2 improves steady state current, but actually slows turn-off kinetics[18]. The opposite effect of mutating K176 and H134 into arginine reflect the differences in structure between ChR2 and gene88, as would be expected based on their highly divergent spectral sensitivities.

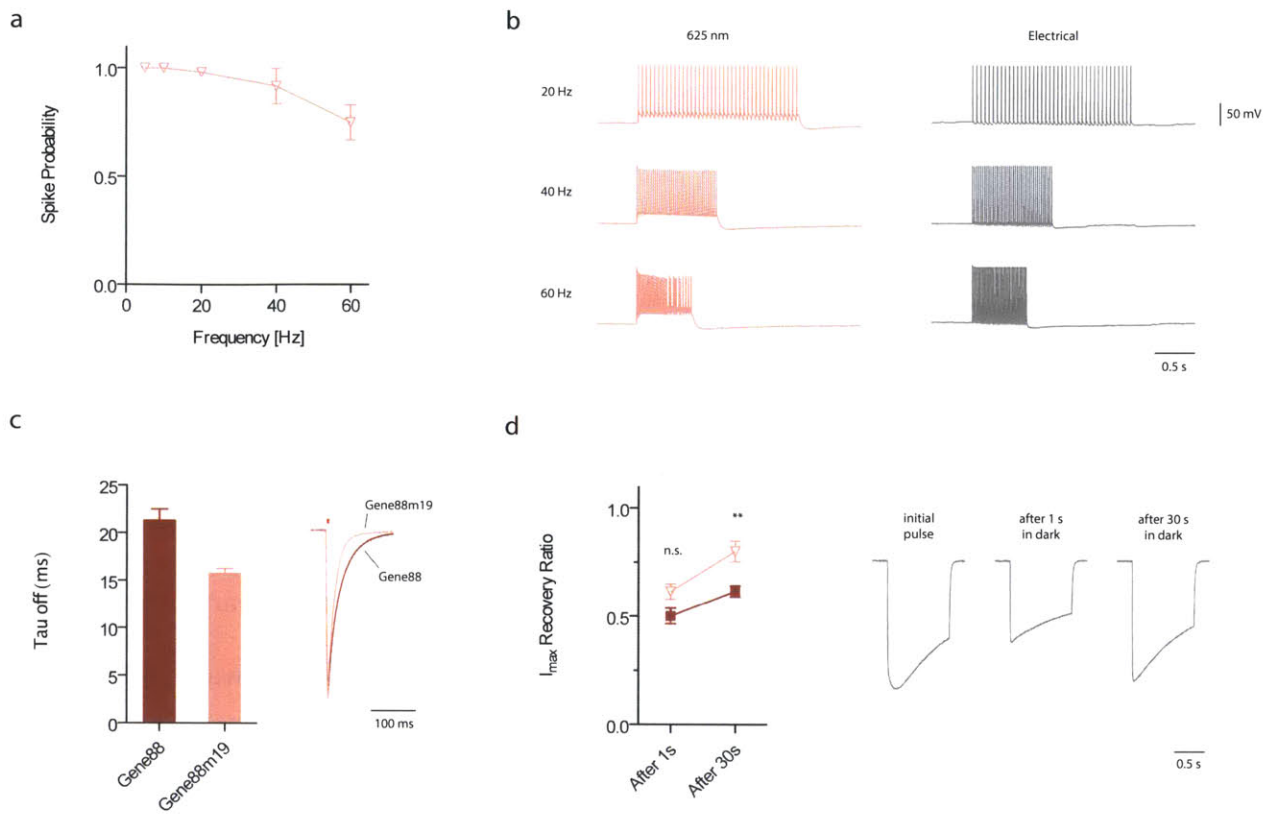


Figure 3.16: Gene88m19 characterization in neurons. a) Gene88m19 spike frequency response with red (625 nm) light 5 ms pulse trains at 5 mW/mm^2. Gene88m19 neurons were selected based on GFP expression. b) Representative traces show gene88m19 cannot sustain spiking at 60 Hz in the second half of the pulse train, whereas electrical spiking shows consistent response throughout entire train. c) Gene88m19 has significantly faster tau-off kinetics than gene88 wildtype in neurons. d) Recovery kinetics comparison shows gene88m19 has significantly higher recovery after 30 s than the wildtype (left). Gene88m19 recovery trace normalized to the peak of the initial pulse is shown on the right.

We named the K176R mutant gene88m19 and tested its spiking performance in neurons. Gene88m19 had 15.77±0.45 ms tau-off kinetics in neurons, similar to the corresponding values measured in HEK293 cells. Unexpectedly the inactivation recovery rate improved compared to

the wildtype, with 80 % recovery after 30 seconds (Fig. 3.16d). Gene88m19's improved biophysical kinetics is additionally reflected in its red spiking performance at high frequency, where it can sustain even 60Hz spiking with 100% fidelity in some cells. However, the spike fidelity is generally still limited by the immediate recovery rate (after 1s I_{max} recovery ratio). As shown by the 60Hz spike trace, failures only occur in the latter half of the train, indicative of channel inactivation within the train. Further engineering should be done to improve gene88m19 recovery kinetics. In its present state, gene88m19 is the only channelrhodopsin capable of driving precisely timed high frequency spikes with red light and should enable neuroscience experiments in which red light is necessary to avoid a visual response.

4. Two-color Excitation

The ability to independently optogenetically excite two distinct populations of neurons using different wavelengths of light would enable researchers to assess how multiple populations of neurons work together to result in emergent phenomena such as circuit dynamics, plasticity, and neuromodulation. Previous implementations of two-color excitation are based on the principle of spectral separation between the blue channelrhodopsin ChR2 and the green channelrhodopsin VChR1 (or its chimeric variant $C1V1_{TT}$). However, the action spectra of channelrhodopsins all have long tails of excitation towards the blue end of the spectrum, a fundamental property of the retinal chromophore. Due to this spectral overlap, previous implementations could not reliably drive spikes with ChR2 without also eliciting spikes with VChR1/$C1V1_{TT}$ in the blue[13, 34, 35].

This chapter describes strategies for minimizing blue light crosstalk, channelrhodopsin selection criteria for two-color excitation, failure modes for two-color excitation, and our implementation of two-color with gene88 and gene90 in acute murine cortical slice to demonstrate zero post-synaptic crosstalk.

4.1 Crosstalk

Neuronal crosstalk can be operationally defined on two levels: subthreshold membrane perturbation and suprathreshold spiking. Spike threshold varies between cells and will certainly depend on the state of the neuron: any voltage deflection will likely drive spiking in neurons that are active or always near a spike threshold. Spike threshold is the one component that the user cannot directly control without altering the cell's natural state and the user will need to be mindful about the amount of voltage deflection which drives spiking.

In the context of optical excitation of neurons, this means a "blue" channelrhodopsin expressing neuron should not have any voltage deflection in response to red light and a "red" channelrhodopsin expressing neuron should not have any voltage deflection in response to blue light. However, it is impossible to have zero crosstalk because the retinal chromophore of all channelrhodopsins naturally absorbs in ultraviolet (UV) to blue wavelengths. This inherent blue light sensitivity can be observed in the action spectrum (Fig. 3.5c), such that all opsins irregardless of their peak have a non-zero current component even at the UV 385nm. Due to this biophysical limitation, subthreshold crosstalk will always occur: the goal is to eliminate suprathreshold crosstalk in order to achieve zero post-synaptic crosstalk. While this may not satisfy neuroscientists interested in subthreshold neural computation, the independent manipulation of separate pathways and observation of interaction between pathways is an important step forward.

4.2 Gene88 Blue Crosstalk

We have previously shown gene88 to be the only channelrhodopsin capable of eliciting precisely timed spikes in the red, and its action spectrum shows it to have less blue light sensitivity than green channelrhodopsins such as $C1V1_{TT}$ and gene87, suggesting it as an optimal candidate for the "red" channelrhodopsin role. We characterized gene88 blue light crosstalk by illuminating with a 5 ms pulse train at 5 Hz with 470nm light at various light powers (Fig. 4.1). At blue irradiance less than 0.6 mW/mm^2, the crosstalk voltage is approximately linear with respect to light power, so doubling irradiance results in a doubled voltage crosstalk. At a given irradiance, the spread in cell-to-cell peak voltage crosstalk is due to differences in opsin expression level, so doubling functional expression levels should also double the corresponding crosstalk. While it is

possible to purposely reduce the overall expression level of gene88 in order to reduce blue crosstalk, this is in general not desirable since spike fidelity requires higher expression levels. In addition, higher overall expression is generally needed for good neuronal processes and pathway projection labeling, as there is no reliable method for differentially boosting expression in subcellular regions[36-38].

Modulating pulse duration has a much more dramatic and nonlinear effect, such that a neuron will actually depolarize enough to spike at a given light power using a 1 s pulse instead of a train of 5 ms pulses. For longer pulse duration, the blue crosstalk voltage is integrated over time until the opsin steady state current is reached. In practice, long pulse durations should not be used since there is no control of spike timing or guarantee that depolarization block will not occur. However the continuous pulse illumination scenario corresponds to the high frequency limit (i.e. 200 Hz) and approximates the worse case crosstalk scenario, demonstrating that crosstalk is also a function of stimulation frequency.

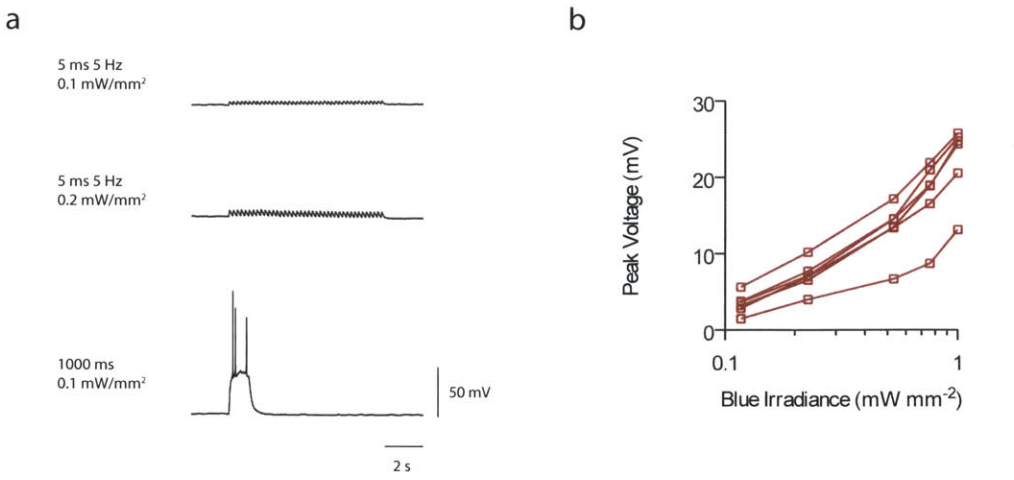


Figure 4.1: Gene88 blue (470 nm) light crosstalk characterization in cultured neurons. a) Pulsed versus constant illumination traces at low blue light irradiance from the same neuron. b) Peak voltage response at various blue irradiance (5 ms 5 Hz). Each curve represents a different neuron.

Gene88's crosstalk relationship places some severe constraints on usable blue light stimulation powers and frequencies. Operationally for an acceptable subthreshold crosstalk of less than 15 mV, one would have to use less than 0.5 mW/mm^2. In addition, high frequency stimulation in the blue would most likely not be possible due to charge integration. One potential method to reduce frequency induced charge integration and crosstalk voltage is to engineer a red channelrhodopsin with a faster tau off. If the molecule can turn off immediately post-illumination, the neuron membrane should repolarize to baseline (limited by the membrane time constant) such that subsequent light pulses will not accumulate in voltage. The ideal red channelrhodopsin for two-color excitation would therefore have low blue or overall light sensitivity, excellent membrane trafficking, and fast tau off kinetics.

4.3 Blue Channelrhodopsin Selection

Any two-color excitation scheme with channelrhodopsins requires a separation in effective blue light sensitivity. Given that gene88 can elicit crosstalk in excess of 15 mV with a 5 ms blue pulse at 0.5mW/mm^2, the complementary blue channelrhodopsin must be able to reliably drive spikes at less than 0.5mW/mm^2. Previous ChR2 engineering efforts on improving effective light sensitivity have resulted in opsins with significantly slower channel turn off kinetics[14, 16, 17]. In fact when all major ChR2 variants were compared side-by-side, the effective light sensitivity was correlated with turn-off kinetics[15]. A possible explanation for the observed slower turn-off kinetics in light sensitive mutants is that the channel open state is being stabilized. So in the extreme case where the channel becomes bistable, the tau-off kinetics is more than 25 minutes[13]. However the main issue with improving effective light sensitivity by simply increasing turn-off kinetics is the loss of temporal precision. In the case of a bistable channelrhodopsin, you need to either use long continuous pulses (relying on the photon integration property) or high intensity short pulses[13]. Neither approach is acceptable for use with gene88 in two-color excitation. To date the wildtype ChR2 is still one of the most light-sensitive channelrhodopsins and none of the mutants can actually drive spikes with short pulses at significantly lower light power than ChR2[15, 17].

Our screening results found gene64, 86, 87, and 90 to have higher blue photocurrent than ChR2 at ~3 mW/mm^2, but it was unknown whether they are more light-sensitive than ChR2 and if they would have sufficient photocurrent to drive spikes at <0.5 mW/mm^2. We selected gene90 as the most promising blue candidate due to it having both the fastest turn-on and turn-off kinetics, which would suggests its higher photocurrent than ChR2 may due to increased light-sensitivity

instead of merely stabilizing the channel open state like previous ChR2 mutants. We compared gene90 and ChR2 in cultured neurons by varying blue irradiance over three orders of magnitude and measuring spike probability over trains of 10 pulses, using the same conditions used to characterize blue crosstalk in gene88. In order to account for variations in expression, we purposely sampled the same range of GFP expression for both molecules. Figure 4.2 shows the spike probability of individual neurons at blue irradiances between 0.01 to 10 mW/mm^2, with the minimum irradiance threshold for 100% spike probability (MIT_{100}) plotted as a function of GFP density. Representative current-clamped traces of individual neurons at various irradiances show that gene90 neither inactivates nor accumulates charge at high light powers (Fig. 4.2b). Within each group, both gene90 and chr2 have MIT_{100} values spanning more than one log, predominantly due to variance in expression levels. Across similar expression levels, gene90 consistently drives spikes at lower light powers than ChR2, indicating that gene90 is functionally more light sensitive. At the highest expression level achieved in cultured neurons, ChR2 can drive spikes at 0.36 mW/mm^2 while gene90 can drive spikes at 0.06 mW/mm^2. At the lowest expression level observed in cultured neurons, ChR2 cannot even drive spikes at 20 mW/mm^2 while gene90 required 0.76 mW/mm^2 to drive spikes.

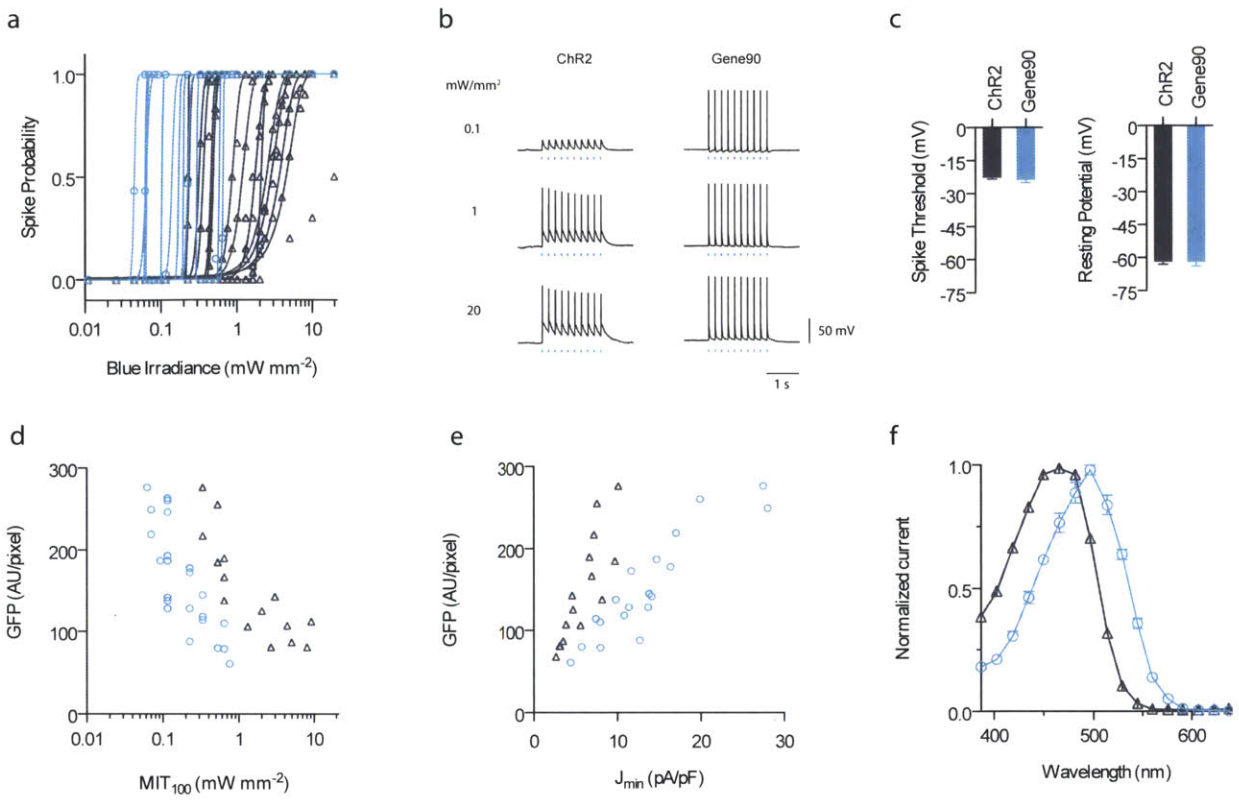


Figure 4.2: Comparison of ChR2 and gene90 blue (470 nm) light sensitivity in cultured neurons. a-b) Spike probability of ChR2 (black) and gene90 (cyan) based on 10 pulses of 5 ms 5 Hz blue light at various irradiance. c) Spike threshold and resting potential controls show no significant difference between ChR2 and gene90. d) Minimum irradiance spike threshold for 100% spike probability (MIT100) is shown against soma GFP expression for each neuron. e) Minimum current density in response to 1 s pulse of 5 mW/mm^2 blue light. f) Action spectrum.

Gene90's marked improvement in effective light sensitivity is not due to differences in neuron resting potential or spike threshold (Fig. 4.2c). We measured the minimum current density at 5mw/mm2 for a 1 blue light pulse and plotted it against GFP intensity (Fig. 4.2e). Both gene90 and ChR2 have positive correlations between GFP and current density (J_{min}), indicating the within molecule variance in J_{min} is predominantly due to expression. At any given GFP intensity,

gene90 has higher current than ChR2, as is consistent with the MIT_{100} trend. However, this does not directly explain the improved spiking performance at low light powers since the current is measured at high light power, and GFP intensity does not necessarily directly correspond to the number of functional opsin molecules in the membrane.

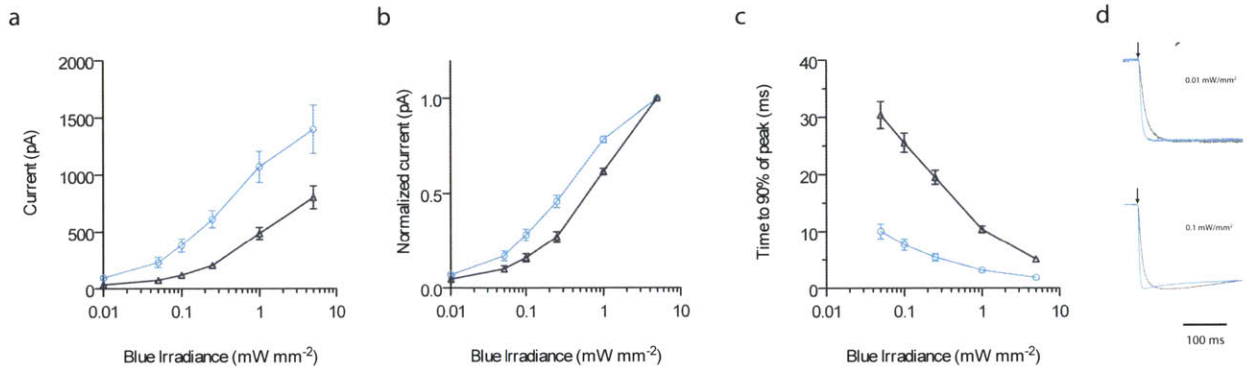


Figure 4.3: Comparison of ChR2 and gene90 blue irradiance photocurrent and turn-on kinetics. a-b) Photocurrent for 5 ms pulse of blue (470nm) light illuminated from low to high irradiance. c) Time-to-peak based on 1 s pulse of blue light defined from the onset of light to 90% of peak or maximum current. d) Averaged traces from all recorded ChR2 and gene90 neurons at indicated irradiance. In all panels, cyan circles denote gene90, grey triangles denote ChR2.

In order to determine whether gene90's improved blue light sensitivity is due to biophysical differences or merely improved membrane trafficking, we measured photocurrent irradiance curves in separate sets of gene90 and ChR2 expressing neurons. In figure 4.3a,b we measured average and normalized current response to 5 ms light pulse under similar conditions as figure 4.2. The average current shows gene90 has 2.5 times higher current than ChR2 at less than 1 mW/mm^2, although this could be mainly due to expression. Gene90 appears to be more sensitive

in the normalized current response curve than ChR2, but we did not sample saturating current conditions for either opsin: the normalized current response therefore does not indicate the true optical dose response curve. We also measured time to peak for 1s pulse in figure 4.3c, which shows gene90 has 3 to 4 times faster turn on time than chr2 at less than 0.5 mW/mm^2, meaning that even if ChR2 and gene90 have identical expression levels, ChR2 current will always be 3-4 times lower than gene90 in response to a 5ms light pulse simply because ChR2 does not reach its peak current within the pulse. The observed improvement in gene90's light sensitivity compared to ChR2 is therefore due to biophysical turn-on time, rather than expression levels.

It is important to note that while the time-to-peak parameter is not strongly influenced by expression level, it is a not a single molecule property (i.e. not photocycle state transition time). Rather, time-to-peak is a reflection of the photocycle dynamics averaged over a population of molecules. Previous study has shown that ChR2 has the fastest turn on time among all channelrhodopsin mutants[15], but the time-to-peak time is correlated to the channel inactivation time, suggesting that ChR2 peak is reached faster primarily due to faster inactivation. In the case of gene90 the inactivation is tri-phasic with a fast 7.8 ms inactivation occurring first (Fig. 3.12). However at low light levels (<0.5 mW/mm^2), there is no clear channel inactivation so the faster time-to-peak cannot be due solely to gene90 saturating faster than ChR2 (Fig. 4.3d). In addition gene90's current onset time consistently occurs earlier than ChR2 under the same conditions at low light levels. This would suggest gene90's improved time-to-peak may be due to improved absorption cross section, quantum yield in converting photons to ionic movement, and/or single channel conductance.

Together, the spike irradiance data and the kinetic data demonstrate gene90 is functionally more blue light sensitive than ChR2, and that gene90 is the more optimal blue channelrhodopsin choice for two-color excitation. Interestingly, gene90 is actually spectrally 30 nm red-shifted relative to ChR2. If the criteria for 2 color were based on pure spectral separation, it would not be obvious that selecting a red-shifted channelrhodopsin to partner with gene88 would result in less blue light crosstalk. However, the downside to using a red-shifted channelrhodopsin as the blue partner is that there will be minor red light induced crosstalk (Fig. 4.4). Gene90 has up to 4mV crosstalk in response to 20 mW/mm^2 red light powers at 5Hz, unlike ChR2, which has no current in response to red light (data not shown). In the worse case scenario, gene90 may have up to 16 mV crosstalk in response to continuous illumination (Fig. 4.4b). Under normal gene88 red light excitation conditions (less than 5 mW/mm2, less than 10 Hz) the red crosstalk should not cause any suprathreshold perturbations to neurons, but it may not be acceptable for two-color applications where one might need zero crosstalk in the red channel.

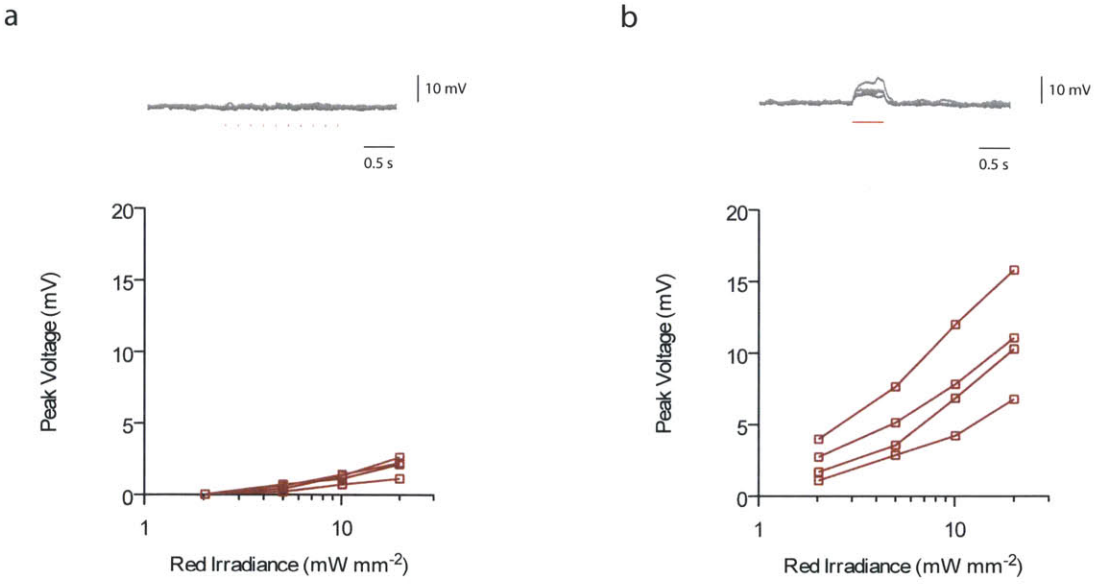


Figure 4.4: Gene90 red (625 nm) light crosstalk characterization. a) Peak voltage crosstalk for 5 ms pulses at 5 Hz b) and for continuous 500 ms pulse. Overlaid traces for neurons excited with 5 mW/mm^2 of red light are shown for both pulsed and continuous illumination.

Finally, spike irradiance data has shown that effective light sensitivity will always be partly dependent on expression level. Theoretically it may be possible that ChR2 is a better blue chop choice for 2-color excitation if one can guarantee the highest expression level in tissue (e.g. transgenic[26, 39]). In practice, however, no one has demonstrated ChR2 spiking with short pulses at less than 0.5 mW/mm^2 in slice or in vivo. A similar critique can be made for gene90: while the expression dynamic range window for spiking at less than 0.5 mW/mm^2 may be large, it may not be possible to achieve sufficiently high expression levels in animals. Therefore irradiance characterization must also be performed in tissue to validate gene90 as the optimal choice for blue channelrhodopsin in two-color excitation applications.

4.4 Individual Channelrhodopsin Validation in Slice

It has been long observed that the main difficulty in implementing two-color excitation is the UV-blue sensitivity common to all channelrhodopsins[34]. Our neuron culture evaluation found the most important parameter for zero post-synaptic blue crosstalk to be the effective blue light sensitivity, and we found gene88 and gene90 to be the most optimal pair for minimizing crosstalk to less than 15 mV. In principle, the potential blue irradiance dynamic range that can guarantee less than 15 mV crosstalk is from 0.05 to 0.5 mW/mm^2, with the lower bound corresponding to the lowest light power that can be used to drive spike with gene90 and the upper bound corresponding to the highest light power that can be used to depolarize gene88 by

less than 15 mV. However, the stated dynamic range is highly dependent on expression levels of both molecules. In order to determine the practical dynamic range in tissue, we slice patched gene88 and gene90 separately from *in utero* electroporated layer 2/3 excitatory neurons, and we patched either opsin-expressing neurons or non-opsin expressing cells post-synaptic to opsin-expressing neurons. Yasunobu Murata generously performed electroporation for all of the slice experiments.

The aim of the first set of slice experiments was to replicate the spike irradiance and the crosstalk characterizations previously conducted in neuron culture. Slice spike irradiance data from gene88 and gene90 cells were gathered from four mice per construct and the results are overlaid in figure 4.5. Gene88 robustly drives spikes in the red starting at 1 mW/mm^2 and as expected, no spike is observed in any gene90 neurons. However, we observed double spikes at higher light power in the red for gene88 due to its slow tau-off kinetics (data not shown). The subthreshold red crosstalk in gene90 cells is less than 2 mV at 6 mW/mm^2. In the blue gene88 has the same crosstalk voltage as measured in cultured neurons and the lowest blue light power that can be used to drive 100% spiking in gene90 neuron is 0.05 mW/mm^2. However the minimum blue light power needed to drive 100% spiking in all gene90 expressing neurons is 0.2 mW/mm^2. This means the practical blue irradiance range for less than 15mV crosstalk is from 0.2 to 0.5 mW/mm^2 (highlighted in Fig. 4.5c). While this range is significantly smaller than the theoretical 0.05 to 0.5 mW/mm^2 range, it is nonetheless sufficiently large enough to account for variation in expression and light scattering.

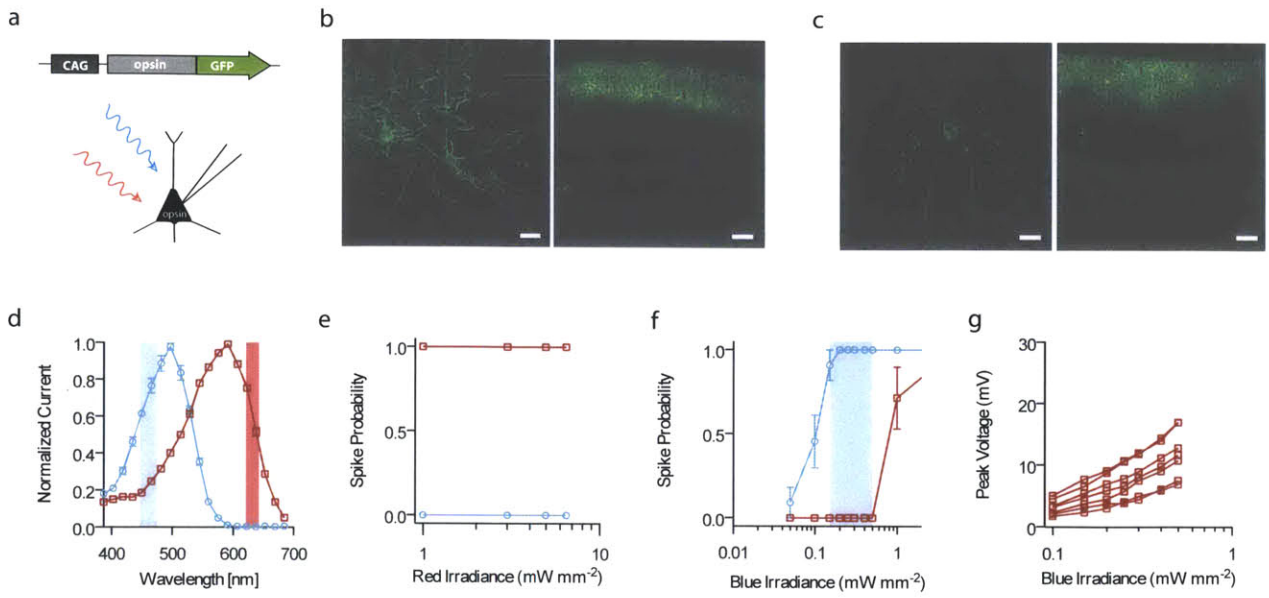


Figure 4.5: Two-color excitation characterization of gene88 and gene90 expressed in separate acute slices. a) Diagram of expression vector used for in utero electroporation. Opsin expressing neurons were patched based on presence of GFP fluorescence. b) Histology of gene88 expression in layer 2/3 cortex. c) Histology of gene90 expression in layer 2/3 cortex. d) Illumination wavelengths used (vertical bars) overlaid on top of the action spectra of gene88 (red) and gene90 (cyan). e-f) Blue and red light driven spike probability (5 pulses at 5 ms 5 Hz) at various irradiance from separate gene88 and gene90 expressing slices. Blue vertical bar represents the range where gene90 drove spikes at >90% and no crosstalk spike was observed for gene88 neurons. g) Peak voltage blue crosstalk measured in gene88 neurons. Histology scale bars, 20 um (left) and 123 um (right) image. Amanda Birdsey-Benson patched all neurons shown in this figure.

We next slice patched non-expressing neurons that were post-synaptic to opsin-expressing cells. Again gene88 and gene90 were expressed in 3-4 separate mice per construct and both the pre- and the post-synaptic neurons are in layer 2/3. Unlike the previous experiments where we could

visualize and select opsin-expressing cells based on GFP fluorescence, we could only use opsin-GFP fluorescence to guess where synaptically connected neurons may be. We randomly patched non-fluorescent neurons that were surrounded by densely labeled fluorescent processes. In addition we also used full field illumination with 20x objective, corresponding to 1.6 mm diameter spot size respectively, in order to increase the probability that we would optically excite all synaptically connected processes. Doing full field illumination allows us to assess the worse case scenario for blue crosstalk, however it also has couple downsides (Fig. 4.6). First the wide-field illumination will drive multiple synapses, often resulting in overlapping responses that have variable time delay or amplitude variation that are not possible to parse into individual events. However it is worth noting that even subcellular illumination (5um spot size) in slice with ChR2 can recruit multiple synapses[40, 41]. Second, there can also be strong inhibitory network feedback after the first post-synaptic response, which will appear as excitatory synaptic responses below the chloride reversal potential. Due to these limitations, it is not possible to track and analyze individual post-synaptic events with confidence, so we treated all post-synaptic events as aggregate and examined only the first 20 millisecond post light onset for the initiation of direct synaptic events.

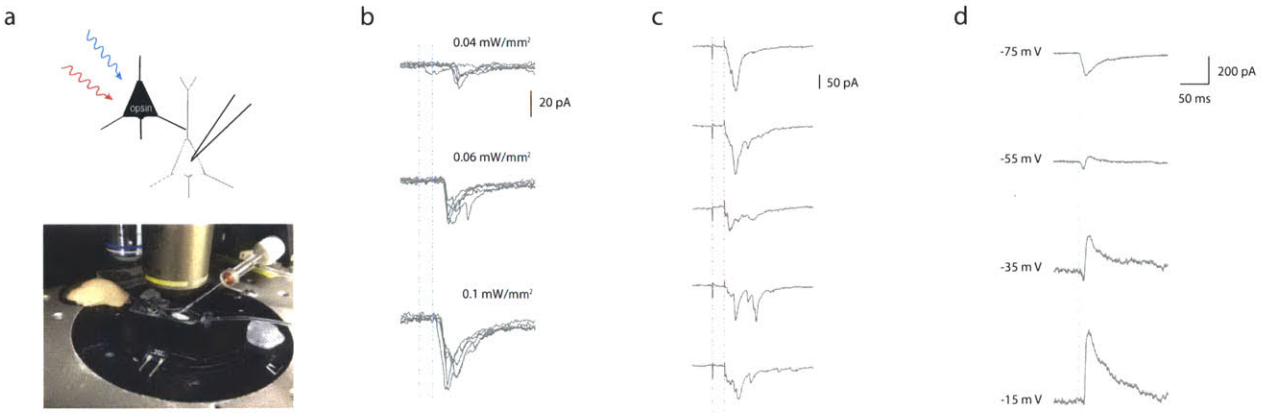


Figure 4.6: Optically evoked post-synaptic current (PSC) measured in non-opsin expressing neurons. a) Photo of slice patching setup. Widefield illumination at 20x (1.6 mm diameter spot size) was used to optically drive spikes in the opsin expressing pre-synaptic neurons. Both pre- and post-synaptic neurons are in layer 2/3. b) Gene90 PSC from the same neuron shows faster onset and multiple PSC peaks at higher blue irradiance. Trials are overlaid for each irradiance. c) Trial-to-trial variability for gene88 PSC at 4 mW/mm^2 of red light. d) Gene90 PSC at various holding voltage shows immediate excitatory PSC followed by strong inhibitory PSC in response to 1 mW/mm^2 of blue light. Vertical dashed lines denote the start and end of 5 ms light pulse for all traces. All stimulations were done at 0.2 Hz.

Based on the data from direct patching, we chose 0.3 mW/mm^2 in the blue for driving gene90 and 1-3 mW/mm^2 in the red for driving gene88. As expected, we were able to drive robust post-synaptic responses and did not observe any crosstalk in the blue or red. Representative traces for optically driven responses are shown for a single cell in figure 4.7. All traces from every cell patched are overlaid for the optical crosstalk channel, clearly showing there is zero post-synaptic crosstalk in the blue and red. We also did illumination at >2 mW/mm^2 in the blue to purposely elicit crosstalk from gene88 slices. We observed greater than 0.6 mW/mm2 of blue light power was required to elicit any post-synaptic crosstalk in gene88 slices and the crosstalk was generally not consistent at less than 1 mW/mm^2, occurring in less than 50% of trials.

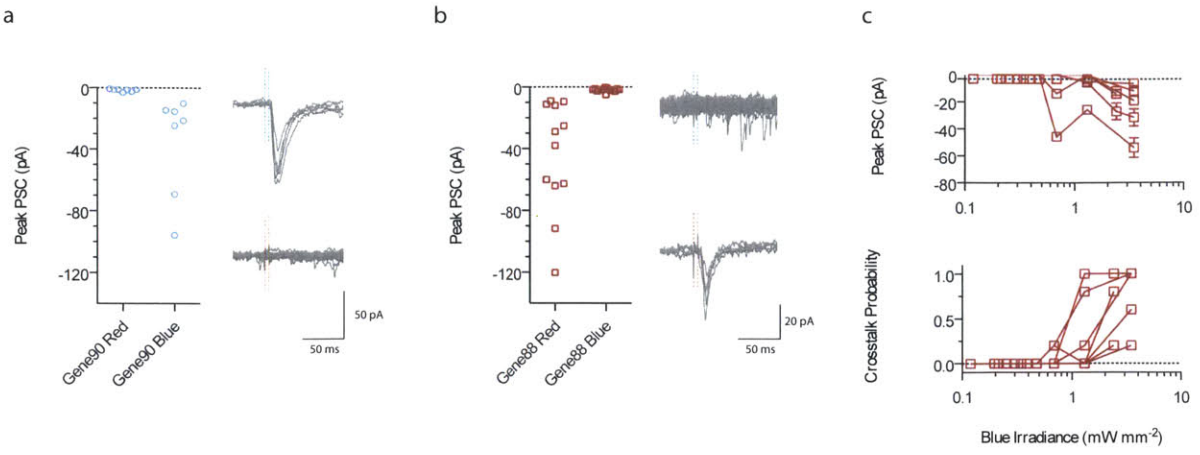


Figure 4.7: Demonstration of zero post-synaptic crosstalk with two-color excitation of gene88 and gene90 expressed in separate slices. a-b) Averaged peak post-synaptic current (PSC) in response to 0.3 mW/mm^2 of blue light or 1-4 mW/mm^2 of red light. Gene90 traces show blue light evoked PSC from a representative neuron and no red light evoked PSC was observed from all neurons recorded overlaid. Similarly gene88 traces show red light evoked PSC from a representative neuron and no blue light evoked PSC was observed from all neurons recorded overlaid. c) Blue irradiance PSC was recorded for a subset of gene88 post-synaptic neurons from panel b). No blue post-synaptic crosstalk was observed at less than 0.6 mW/mm^2. In all panels, optical stimulation was done with 5 ms pulses at 0.2 Hz using widefield illumination (1.6 mm diameter spot size) to drive opsin expressing pre-synaptic neurons.

The characterizations thus far have only discussed 2-color excitation robustness in terms of amplitudes on the pre- and post-synaptic sides, but spike and post-synaptic response timings are another metric for quantifying spike fidelity. We compiled the latency to spike peak and the post-synaptic current onset time from all slice experiments (Fig. 4.9). In panel a, the population average for latency to spike peak is plotted as a function of irradiance. For both gene88 and

gene90, the latency decreases as irradiance is increased: this is intuitive since at higher light power, increased photocurrents will depolarize neurons to spike threshold faster. Gene90 does not saturate spike latency until 5 mW/mm^2, after which it is biophysically impossible for spikes to have shorter latency due to spike rise time and spike width. The population spike latency standard deviation increases as the irradiance is decreased, reflecting the 5 ms pulse duration does not over drive gene90 beyond spike threshold as gene90 turn-on time becomes greater than the pulse width. So at lower irradiances, gene90 takes longer to depolarize the cell membrane and the spike timing will strongly depend on variability in spike threshold and gene90 expression.

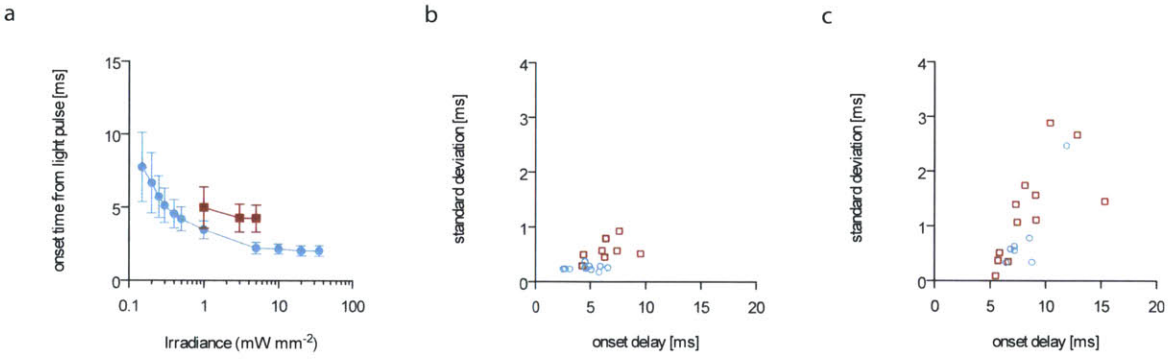


Figure 4.9: Comparison of timings for optically evoked spiking and post-synaptic currents. a) Time-to-spike-peak as a function of irradiance for optically driving gene88 neurons in the red and gene90 neurons in the blue. Values shown are population average ± standard deviation. b) Time-to-spike-peak for individual gene90 (0.3 mW/mm^2 blue light) and gene88 (5 mW/mm^2 red light) neurons. Values shown are trial average. c) Post-synaptic current onset time for individual neurons post-synaptic to gene90 (0.3 mW/mm^2 blue light) or post-synaptic to gene88 (1-4 mW/mm^2 red light). Values shown are trial averages. In all panels, red squares denote gene88, blue circles denote gene90.

For the two-color excitation irradiance conditions, it is not possible to drive gene90 at saturation hence the gene90 spike jitter is larger than what it could be, but still comparable to all other channelrhodopsins driven at saturation conditions. The trial mean and standard deviations for individual pre- or post-synaptic cells using light powers chosen for two-color excitation are shown in figure 4.9b,c. In the opsin-expressing neurons, the trial-to-trial jitter and the mean latency are significantly smaller for gene90 than for gene88 ($P<0.0001$). In the post-synaptic neurons, both latency and jitter are bigger than for the opsin-expressing cells, partly due to variances in synaptic transmission, although the longer latency may also be in part due to poly-synaptic transmission. The relative timing between synaptic events is the most important aspect for spike timing experiments such as spike timing plasticity[42]: it is therefore acceptable to have longer spike and post-synaptic response latency so long as the jitter is small and predictable.

These slice experiments have validated gene88 and gene90's abilities to drive robust spiking and post-synaptic response without eliciting any post-synaptic optical crosstalk under identical experimental conditions, and these results indicate we will be able to independently excite distinct neural populations in the same slice.

4.5 Double Post-Synaptic Validation in Slice

This final set of slice experiments aims to demonstrate that distinct neural populations expressing gene88 or gene90 can be independently driven by different wavelengths. We used Cre-dependent plasmids to guarantee mutually exclusive expression of gene88 and gene90 via *in utero* electroporation. Due to the difficulty of using *in utero* electroporation to target both excitatory

and inhibitory neurons that are directly pre-synaptic to a common target region, we decided to target layer2/3 excitatory neurons and to artificially segregate them into distinct populations with cre-dependent vectors. The expression scheme is shown in figure 4.10a. We did triple plasmid electroporation with gene88 in "Cre-OFF" form, gene90 in "Cre-ON" form, and separate Cre plasmids to control the ratio of the number of cells expressing gene88 to gene90. Gene88 and gene90 are C-terminally fused with mOrange2 and GFP respectively to allow for identification. We did not observe any neuron expressing both fluorophores in the slices recorded (data not shown).

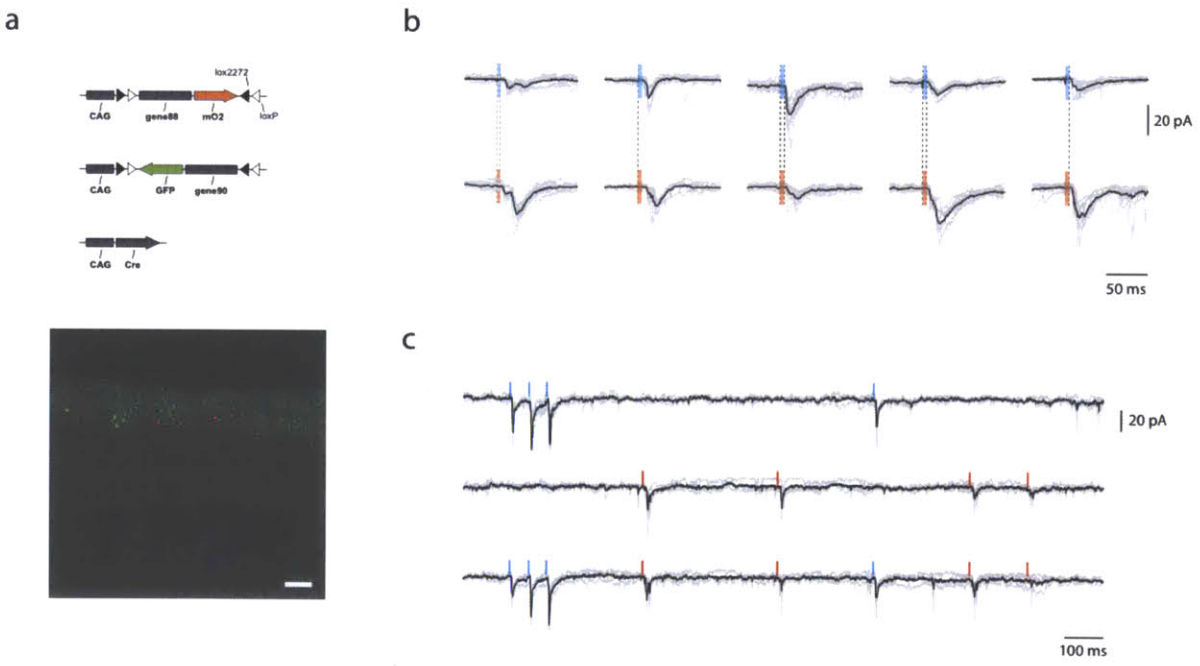


Figure 4.10: Independent optical excitation of distinct neural populations expressing gene88 and gene90. a) Triple plasmid electroporation scheme to guarantee mutually exclusive expression of gene88 and gene90 (top). Histology of intermingled gene88 and gene90 expressing neurons in layer 2/3 (bottom). b) Red and blue light driven post-synaptic responses from five different neurons. c) Poisson optical stimulation demonstrate blue and red post-synaptic currents are

independent. In all panels, optical stimulation was done with 5 ms pulses of blue (0.3 mW/mm^2) or red (1-4 mW/mm^2) light. Scale bar is 123 um.

We patched post-synaptic non-expressing neurons in layer 2/3 and again used wide-field illumination at 20x to stimulate gene88 synapses with 1-4 mW/mm^2 of red light and gene90 synapses with 0.3 mW/mm^2 of blue light. Raw traces from six neurons in response to single red or blue pulse of light are shown in figure 4.10b. Given that both synaptic responses are excitatory, we next explored using paired-pulse stimulation to distinguish red-drivable from blue-drivable synapses. Paired-pulse stimulation protocols are a metric for examining short-term synaptic plasticity where closely timed pulses (typically less than 50 ms apart) would cause the first pulse to enhance or depress the second pulse due to pre-synaptic terminal adaption (i.e., accumulation in calcium or other factors). Red-red and blue-blue paired-pulses should exhibit facilitation, depression or no change, while red-blue or blue-red pulses should be a linear summation of the single red and single blue pulse responses; if red and blue driven synapses are truly independent, then no history dependent effects should occur at the pre-synaptic terminals. We performed paired-pulse stimulation with a 50 ms gap at 0.2 Hz in six neurons and some representative traces are shown in figure 4.8. Unfortunately protocols were not temporally shuffled, so the absolute amplitudes in all but one neuron drifted due to patch quality and could not be quantitatively evaluated. In general we observed red-red pulses tended to be depressed or the second red-pulse failed, probably due to gene88's inability to drive high frequency pulses without high expression levels (Fig. 4.8c,d). Thus, the observed paired-pulse depression may not not reflect synaptic-level changes, but rather gene88 inactivation dynamics. Gene90, on the other hand, showed blue-blue pulse facilitation in two neurons and no change in four neurons. Across

all neurons, the red-blue pulses were linear summations of the individual color response, but the blue-red pulses sometimes resulted in lower red responses (Fig. 4.8c,d).

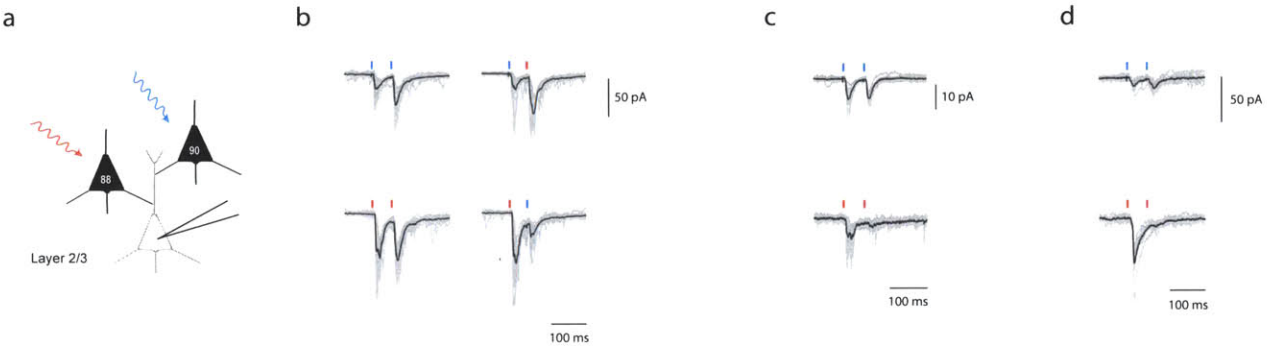


Figure 4.8: Paired-pulse optical stimulations showed history dependent post-synaptic response. a) Gene88 and gene90 expressed in separate neurons in the same slice using triple plasmid in utero electroporation. Patched neurons are post-synaptic to both gene88 and gene90 based on optical response to blue and red light. b) Paired-pulse responses for an exemplar neuron. Blue-blue stimulation exhibited facilitation, while other paired pulses were linear summation of individual pulse response. c-d) Blue-blue and red-red responses from the same neuron showed no differences in the blue paired PSC, while the second red pulse response often failed. In the paired-pulse protocol, the second pulse is delayed by 50ms. Black trace is the averaged response, grey traces are individual trials.

We investigated the blue-red pulse interference in single-opsin expressing slices and repeated a temporally shuffled version of the protocol used in double post-synaptic experiments. Gene90 did not show any interference in red-blue and blue-red responses; however, some gene88 cells exhibited the interference in blue-red pulses as observed in double post-synaptic experiments. Figure 4.11 shows the clearest example with gene88 expressed pre-synaptically in an inhibitory

neuron. These traces demonstrate that although blue light does not drive spiking responses in gene88 neuron, the blue light does activate gene88 molecules and can sometimes inactivate gene88 such that the subsequent red pulse will not drive spikes reliably. Thus the observed interference observed is not the result of the synapse, but the poor temporal precision of gene88.

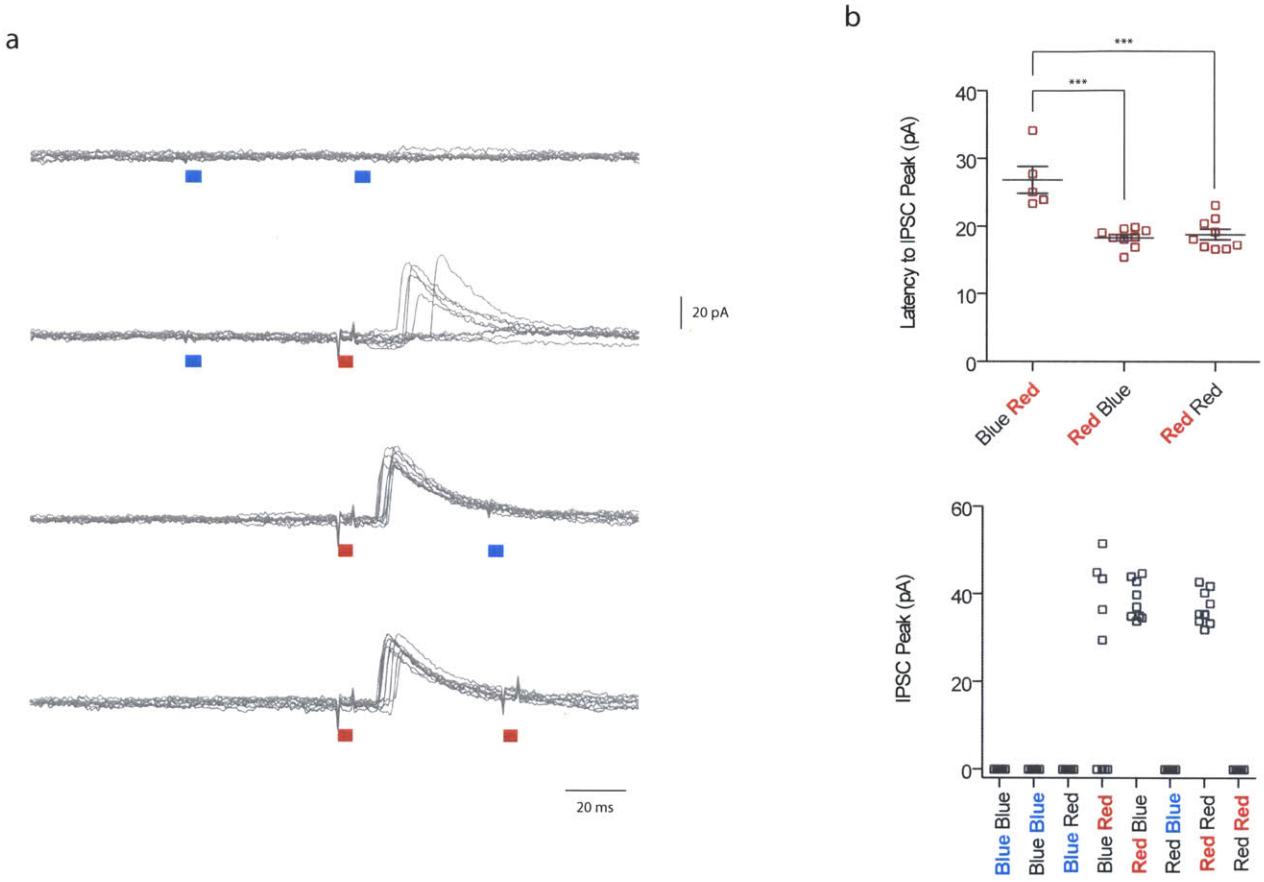


Figure 4.11: Optical interference in gene88 post-synaptic response. a) Paired-pulse illumination (50 ms gap) of an inhibitory pre-synaptic neuron expressing gene88. b) Inhibitory post-synaptic current latency and peak amplitude. Bolded and colored labels refer to the illumination color used in the paired-pulse stimulation.

Given the limitation of not being able to reliably target constructs to separate excitatory and inhibitory neurons, we could not directly prove color independence in double post-synaptic neurons. However the interference observed in paired-pulse responses clearly reflects the kinetic properties of gene88 and gene90. These double-post synaptic demonstrations, taken along with previous single-opsin expressing slice experiments strongly suggest that different synapses were independently excited.

4.7 Two-color Excitation Limitations

The two-color excitation slice experiments have demonstrated zero post-synaptic crosstalk with gene88 and gene90. However the current implementation is limited to low frequency (less than 10 Hz) modulation due to the slow kinetics of gene88 and the frequency dependence of blue crosstalk. We have already engineered the gene88m19 mutant that can robustly drive great than 40 Hz spikes in cultured neurons, but further validation is necessary in slice. On the other hand, reducing frequency dependence of blue crosstalk requires finding or engineering more light sensitive blue channelrhodopsin or less light sensitive yellow channelrhodopsin. Currently there are no obvious methods to improve on blue light sensitivity without sacrificing temporal precision, but it may be possible to engineer gene88 to be less light sensitive.

Although we have minimized blue crosstalk by exploiting biophysical differences in channelrhodopsin turn-on kinetics, channelrhodopsin expression level is still the most important factor in determining the effective irradiance dynamic range for two-color excitation. Since the minimum blue irradiance threshold for 100% spiking is determined by gene90, it would be highly desirable to have gene90 stably expressed in order to reduce variances in expression.

Finally the irradiance dynamic range in the blue (0.2 to 0.5 mW/mm^2) is sufficient for two-color excitation in slice, but may not be acceptable for in vivo applications where light scattering from an optical fiber light source is an issue. One potential way to precisely control irradiance in vivo is to reduce illumination to small volumes using an optical waveguide, which can effectively reduce illumination to 2D addressable surfaces[43]. While this may not be a viable strategy for every experimental paradigm, it is likely to be sufficient as the main reason for using two-color excitation is to investigate spatially inseparable neural populations, where the addressed tissue volume will likely be small.

5. Conclusion

This thesis has characterized over 40 novel opsins from transcriptomic sequencing of more than 200 algae by the 1000 plants transcriptome sequencing consortium (OneKP), and discovered many channelrhodopsins with improved kinetics, current, recovery, and spectral range. Two novel channelrhodopsins in particular, gene88 and gene90, possess unusual biophysical properties. Gene88 is the first known yellow-peaked channelrhodopsin, with a peak more than 45 nm red-shifted than any previous channelrhodopsin while Gene90 has the fastest channel turn on, turn off, and recovery kinetics of known channelrhodopsins.

We also describe and implement an alternative two-color excitation scheme where effective light sensitivity, as oppose to maximizing spectral separation, is used to achieve independent optical excitation in the blue (470 nm) and red (625 nm) using gene88 and gene90. We implemented this scheme in acute murine slice and found the irradiance dynamic range for zero post-synaptic crosstalk to be 0.2-0.6 mW/mm^2 in the blue and 1-6 mW/mm^2 in the red. However, due to the widefield illumination setup, we were not able to reliably resolve individual post-synaptic events due to excitation of multiple synapses. Alternative experiment setup using a scanning beam or digital micromirror device should eliminate these issues and enable analysis of spike timing plasticity.

Finally while this thesis has mainly focused on gene88 and gene90 due to their optimized properties for two-color excitation, many of the other opsin genes may be of general interest to other neuroscience applications. Gene66 is the most blue-shifted channelrhodopsin reported with peak at 435nm, however further optimization is necessary to improve its poor membrane trafficking in mammalian cells. Gene86 has the highest photocurrent response in cultured neurons and it may be more light sensitive than gene90 in applications where submillisecond

temporal resolution is not needed. Gene87 is one of the best membrane trafficking channelrhodopsin and has faster functional expression in vivo than ChR2 and gene90 (data not shown), so it may be appropriate for time restricted studies such as looking at neural circuit dynamics during development. There are numerous other opsin genes that did not function well in neurons, but may possess novel biophysical properties. This thesis provides the basic framework for future studies to understand the diversity of sequence, structure, and function relationships of these novel channelrhodopsins.

Appendix

Sequences

The following are opsin protein sequences truncated to the 7 transmembrane region based on alignment with ChR2. We also tested longer sequences (i.e. with more of the cytoplasmic region) for a subset of opsins and we did not observe any changes in function.

>gene59
MGNDALNVNGFTNTVTTDNHITVRGSDWYYAVTAVMAVSTFAFMGLSFTKPREQRIFHYITAAITLVASI
AYFSMGSNLGWTGIQVEFERANPKVGGNIRQIFYVRYIDWVITTPLLLLDLLLTCGLPMPTIIYTILIDE
VMIVTGLVGALVKSSYKWGYYVFGCVAFLFVAWTVVFEGRAHARALGQDINKVYTICGVWTIFLWFLYPI
AWGVSEGGNVIPSDSEAVFYGILDLLAKPVFGALLIWGHRNIDLARLGLHINDPAKTGTPGSSNNEKRGI
NTNNNTA

>gene60
MADVVPPATPVSDICFAPCQEDCITIRYFIENDFEGCLPGHFDQYSSFGVLPDVSKAALYICMVVSVIQI
AFYGFQWWRKTCGWEVWFVAVIETAIYIINITSEADSPFTLYLTNGRLSPQVRYMEWLMTCPVILIALSN
ITGMAEEYNKRTMTLLTSDVCCIVLGMMAAVSKPKLKGILYAVGWMFGLLQFWTAFMVYHDAHKAVPKPL
AWYVRAMGYVFFASWSTFPAFFLLGPEGLEVFSGTVSTIAHACSDMVSKNLWGFMDWHLRVLVARYHRRL
FKAEEEHAAKKGQTLDAGLPRSTSFVRGLGDDVEIDP

>gene61
MIVHPDMIKSAVATTSVAPIPTVIPTPPIFVNAGDTGQRTLWVVVVLMALSSIVFYAMAFRVPVQKRLFH
ILTAFITTFAFLSYFAMATGDGINYNGYTVTESHKKVPDTHQDYYRQIYWARYVDWSVTTPLLLLDLALL
AGLSGANILVAIVADVIMILTGLFAAYGKTDGAKWGWYAWACIAYLVVVYQLASAGRAAVATKDSKTKAF
YGSIAGFTLILWTIYPIIWALSEGAHIIGVDAEILSYAVLDVLAKPVFGFWLLFTHDSMSSTSPSLEGFW
SNGFGSQGALRVGDDD

>gene62
MADFVWQGAGNGGPSAMVSHYPNGSVLLESSGSCYCEDWYTSRGNHVEHSLSNACDWFAFAISVIFLVYY
AWAAFNSSVGWEEIYVCTVELIKVSIDQFLSSNSPCTLYLSTGNRVLWIRYGEWLLTCPVILIHLSNVTG
LKDNYSKRTMALLVSDIGTIVFGVTSAMCTGYPKVIFFILGCCYGANTFFNAAKVYLEAHHTLPKGSCRT
LIRLMAYTYYASWGMFPILFVLGPESFGHMNMYQSNIAHTVIDLMSKNIWGMLGHFLRHKIREHILIHGD
LRTTTTVNVAGEEMQVETMVAAEDADETTV

>gene63
MRRRALLLLLTFICLVALGLASTPAPVVSTSPPPATAASGDAHGAPAAGESGGNHAAPTEHGGEGHSSEN
GGEHGGGHSSEHHSSSDGELYPPTFSEEEAARALQWVCFMISLMLALFYFWQYYIGTSKWEVLYVCFFES
LAYIFDIWYLHKNPISLVFKQGHHAEWLRYMEWLLTCPILLIALSRVGVADGAYSKRTMRLLTSDQGTIL
MGVTAAFAQGYVRLGFYLVGCAFGFNTYYTAAAVYLEAWRNVPDAEKDVVKWMATIYYFSWVMFPVLFAI
GSEGLGYITLSGSVIGHCIADLLSKNVWGLLEWYLEFCVHSRYHAELAALEEEEEEEEGEEGAYHIPGVV
VFAVGSAKVLDSLKEQFEGFADEVVAG

>gene64
MGGAPAPDAHSAPPGNDSAGGSEYHAPAGYQVNPPYHPVHGYEEQCSSIYIYYGALWEQETARGFQWFAV
FLSALFLAFYGWHAYKASVGWEEVYVCSVELIKVILEIYFEFTSPAMLFLYGGNITPWLRYAEWLLTCPV
ILIHLSNITGLSEEYNKRTMALLVSDLGTICMGVTAALATGWVKWLFYCIGLVYGTQTFYNAGIIYVESY
YIMPAGGCKKLVLAMTAVYYSSWLMFPGLFIFGPEGMHTLSVAGSTIGHTIADLLSKNIWGLLGHFLRIK
IHEHIIMYGDIRRPVSSQFLGRKVDVLAFVTEEDKV

>gene65
MEAYAYPELLGSAGRSLFAATVPENISESTWVDAGYQHFWTQRQNETVVCEHYTHASWLISHGTKAEKTA
MIACQWFAFGSAVLILLLYAWHTWKATSGWEEVYVCCVELVKVLFEIYHEIHHPCTLYLVTGNFILWLRY
GEWLLTCPVILIHLSNITGLKNDYNKRTMQLLVSDIGCVVWGVTAALCYDYKKWIFFCLGLVYGCNTYFH
AAKVYIEGYHTVPKGECRIIVKVMAGVFYCSWTLFPLLFLLGPEGTGAFSAYGSTIAHTVADVLSKQLWG
LLGHHLRVKIHEHIIIHGNLTVSKKVKVAGVEVETQEMVDSTEEDAV

>gene66
MFAINPEYMNETVLLDECTPIYLDIGPLWEQVVARVTQWFGVILSLVFLIYYIWNTYKATCGWEELYVCT
VEFCKIIIELYFEYTPPAMIFQTNGQVTPWLRYAEWLLTCPVILIHLSNITGLNDDYSGRTMSLITSDLG
GICMAVTAALSKGWLKALFFVIGCGYGASTFYNAACIYIESYYTMPQGICRRLVLWMAGVFFTSWFMFPG
LFLAGPEGTQALSWAGTTIGHTVADLLSKNAWGMIGHFLRVEIHKHIIIHGDVRRPVTVKALGRQVSVNC
FVDKEEEEEDERI

>gene67
MPPAYLTLFGSVLRQPALLPKNQNTISISQYVRRPLRGCSPRIRTLFRCLVIEGALCNAPRPAGSFSLPS
RRRHPVTRPQPTTRGTTMGNANSFPRPEDDFYTKCGDFHDQNGSLLVQQIGRGLQWVTFVVSWFQLAFYG
QSVKSTLGWEETYVCVVEAIKITFDIFFMYSTPATIYQTNGNTVSWLTYSEWLLTCPVILVQLCNITGLE
SDFSNRTMQVLVSDVGAIVMGTYAAFSAGWLKAVFFLLGCSYGAFTFYHAAKVYIESYYMVPKGLCRNLV
KAMAWCFFTSWVGFPIIFLAGPDGFGHLSIWAEVAAQQICEIFSKNVWGFLGHFLRVKIHEFIIIHGDCR
KLSKVKFAGKELMVMSYVEKDGNDTI

>gene68
MGFQLNPEYLNETILLDDCTPIYLNVGPLWEQKVARGTQWFGVILSLAFLIYYIWITYKATTCGWEELYV
CTIEFCKIVIELYFEFSPPAMIYQTNGEVTPWLRYAEWLLTCPVILIHLSNITGLNDDYSGRTMSLITSD
LGGICMAVTSALSKGWLKWLFFVIGCCYGASTFYHAALIYIESYYTMPHGVCKNMVLAMAAVFFTSWFMF
PGLFLAGPEGTNALSWAGSTIGHTVADLLSKNAWGMIGHFLRLEIHKHIIIHGDVRRPITVNTLGREVTV
SCFVDKEEEDEDERI

>gene69
MSSPQTEPPLAQSLSTCDARGGLDLCSAAQLSEALNACGYWDDPVSLNSAKVGLESAVVALALTLAAFHL
WQLFHGFREWTLFYYSLVVAAQYLISLTATGAVPVALYVPCTAPLNWLRYAAWLLIVPAMVMALCRVTGL
SSVYPRLQDPLLLLSVLMVVLGVSTAAAADITTKAVLFALAGVAGLAAFVYAAIQFIELYYTLNIVVCRR
GIYPLAALWGIVWVGYPVLFIAGPEGFGSLSAPGSAIAYAFIDLAKLSWGLCAWNLHQNCRSYALKYGTT
SKKVALPFGLGTVAVRTFFNENDLEETETGSQTAGAHYLEDEAEALQAA

>gene70
MSLVPRLLAEAAVVFNTSVAPGRPGGRGPQNPSVPLGNGWVGVPEELCYCKGWYNLKGSDSQASFADIMQ
WVVFVLCMLVVVWNAINFFYYSRQTTSWECVFVMTIVCAFYGFSIWHQFDSPVIAYFSTGNISVLLRYGQ
WLFCTPVLLVAVSDMTGRNGSYSREHMLMIMADVFAVTIYVGTAVSPNFWVKFFCTAFSTIAFMYLFFMA
NRCYKKLLDDVPTARCRSLVQLIRVVFFVGWNMHTCSSLANPEVLNFLNALEYQITACFGDLVAKVLFCA
LATYLRVLMHHDAPPEEEDDAIGPNQIAMDDACEGGYEGAARAAAA

>gene71
MKAFKVGSFLFGLVAILSLFYSTIAAGDSSGNYTGYEKYYAEIDVSFGGGKVPQYFEKVQIKFFDKGNEA
EIIAATVFQWLCLVASVGLFGWYTYQLFNGSCGWEVVYVSIVETCKYLLEIFAEPRPPATIYASHGSTYI
WLRYAEWLMTCPVLLIHLSNITGLKEEYSHYTMRLLTSGQGPISMGITSAAATGGVKIFFFCLGLCYGGF
TFYSAAQVYMEAYHMVPRGQCRGIVRMMAWCYFASWPLFPILFLFGPEGFNIISLSGSHIGHAVADLLSK
NLWGWFGWYLRYKIREYVIEKGDIRRVKEIEVLGRKFEVEEYMDEEELAELP

>gene72
MNECRCRPENGRGVASGSGRSPCAAVLRAWRGCRARRRGAPRGGLAPRAHHPGPPPTIQAVFLWIGFASF
VAATVVFVFLFFLHGKREHYVLNFFCVAIAALGYFANANGMGFVNVNSYTVNFARYLDWSITTPLLVIDL
AELGERPFAIKLTLVFADFLMIATGLFCALCPAPMRWGFFAFSCLFFLVLLFILYQGYARTLRLLPTDRD
RNMYRVLIIIFLATWNVYPVIFVIGTEGIRTLNPDVEVIILTFTDLTAKVIFGFIYLKLEDNKYEADIEK
RVTKEVVFFEDENGVKYVRAQDGIALDEQGKIRA

>gene73
MASYNSSQWVHGDDVGYGHWYQDTPNGTLVCSNEDNIAWLKYKGTDAEMLGANVCMWMAFAACLLCLMFY
AYSTWRATCGWEEVYVCLVEMVKVMIEVFHENDSPATLYLSTGNFIMWIRYGEWLLSCPVILIHLSNITG
LQDQYSKRTMQLLVSDLGTITMGVTAALCSNYVKWIFFIAGLCYGVNTYFHAAKVYIESYHIVPKGICRI
CVRVMAWCFFAAWTCYPLLFVFGPEGLGVLSYNASAIGHTIIDIFSKQVWGFVGHYLRIKIHEHIVIHGN
LVKPTKVKVAGMEIDAEEMVEKDEEGAI

>gene74
MSVNLSLWEHGEDAGYGHWYQGTPNGTLVCSHEDNIAWLKNKGTDEEMLGANICMWMAFAACLLCLSFYA
YSTWRATCGWEEVYVCLVEMVKVMIEVFHENDSPATLYLSTGNFIMWIRYGEWLLSCPVILIHLSNITGL
QDQYSKRTMQLLVSDLGTITMGVTAALCGNYVKWIFFILGLCYGVNTYFHAAKVYIESYHIVPKGVCRVC
VRVMAWCFFGAWTCYPLLFVFGPEGLGVLSYNASAIGHTIIDIFSKQVWGFVGHYLRIKIHEHIVIHGNL
VKPTKVKVAGMEIDAEEMVEKDEEGAI

>gene75
MSGSPQPFPGTCGPGYDDDVVCALVVNTTLYDPLFVPLGPPWEQMVSRGTQWFTVILSILMTGWFFYNMY
TGHCGWEVIFVTIIELAKIFIELFWEYETPCMIYSVFGPVTSWLRYIEWLLTCPVILIHLSNLTGLDEEY
SARTMNLLTSDQGTICFGITAALSPHGWIKIVLFLIGLSFGVNTFLTAARVYIESYHQVPKGRCRNLVKY
LCGMFFCSWLLFPLLFIAGPEGTGYLTWSGSTIGHTVADLLSKNIWGLIAHYLQVKIREHILIHGDVRKT
VEKTVAGHTFEIEEFADKDDEDAVV

>gene76
MADLRYGPEPVFGSCADGNPETFCQLVVNTTLYPALLVDLGPVWEQVASRGTQWFTFVLSVLMTGWFVYN
FYTGHCGWEVIFVTTIEFIKIIIELFYEYETPCMIYSVFGPVTSWLRYIEWLLTCPVILIHLSNLTGLDE
EYSARTMGLLTSDQGTICFGITSALSRHGLIKIITFLVALCFGVSTFYSASRVYIESYHQVPKGLCRKLV
RYLTTMFFCSWLMFPLLFLAGPEGMVYLTWSGSTIGHTVADLMSKNIWGLIAHYLQVKIREHIILHGDVR
AKVEKTVAGHTFEVEGFRDEGDDDAEAV

>gene77
MGWKINPLYSDEVAILEICKENEMVFGPLWEQKLARALQWFTVILSAIFLAYYVYSTLRATCGWEELYVC
TVEFTKVVVEVYLEYVPPFMIYQMNGQHTPWLRYMEWLLTCPVILIHLSNITGLNDEYSGRTMSLLTSDL
GGIAFAVLSALAVGWQKGLYFGIGCIYGASTFYHAACIYIESYHTMPAGKCKRLVVAMCAVFFTSWFMFP
ALFLAGPECFDGLTWSGSTIAHTVADLLSKNIWGLIGHFLRVGIHEHILVHGDVRRPIEVTIFGKETSLN
CFVENDDEEDDV

>gene78
MKNSQQIAGGRVDAAHATSLPTMILPLAIGFMILFTSTLAFFVMSMSYETEKRTFHHITMFVTGIAAIAY
LIMAFEGGAAELTLTDGKVRTFFWVRYVDWALTTPLLLLDIGLLANAPLAHIAYVIGCDVLMIVGGLVGP
LLFAANDPAKWLLFYIGCFWFLPILYAMMVEWKTTLTDKARPVYNVIANGTVVLWCAYPVMWAMCEGTGM
LSDETEVILYLVLDMVAKSGLGFILLSSHDAI

>gene79
MPAVGLVFALRPAAERRDGCAGGGGAAAVAGGVRAGGAQLPSSPVAVALLDGMNLRVHEWPAMAYATSFV
VEVLGLVVSSNSSKQRLTFLPMVVVLLPLLSLGLLHTGNSAVMLSQQGRMLLPGRYLTWTCTTVLLLLAT
FISSGLPLRVLLAAVTSDVVMIAAGDLALVSDSLLCFWCGFATSMVAWVWVVWCLWCMISNCQTYCTEME
GTSRALGRLKLLTIVVWTTFPVVWLASALYPDSISAQLEEILWCLCDLFAKLLLGAAVHRGTLCVFEDRL
HSSSSELLRLNIRDLHEQMKCKDRFLAAV

>gene80
MAGMPHARSGATMFIVLACLAIFANLPKADASMSKWKNYIGGENGTIPSTNITYGAEMYRGFYVLDANPD
WVTGPVDDCYCKSWAVSHGTKAEKLGAIVAMWIVFAFCVIVLVFYAIAAWRSTCGWEEVYVCIIELAHVC
IAIFHEIESPSTLYLSTGNQILWLRYAEWLLSCPVILIHLSNLTGMKDDYSKRTMGLLVSDIGTIVFGTT
AAMSPYNYLKIVFWFCGLTYGCVTFFLAAKVYIEAYHTVPKGTCRKIVRVMAWDYFGSWCMFPILFVLGP
EGFGHISAYGSVIAHQVLDITSKNLWSFLGHLLRVKIHEHIIIHGNITKKTKITVAGDPVEVEEYVDVDE
V

>gene84
MKRQQARVMEANNGSASSAVIISNWVCFLCLTGSFLGLLIFSLKYKGPDGKEVYYNGYREQNMLGVFINL
WCAVAYFAKVIQAHGGDDVAFAPLTTVKYIDYITTCPLLTMDLLWNLDAPYKITSGLLVLTCLTFAVASF
LTPKPGAYVWFAVGLSFFTFSYTFILSIVRQRLDFFTECAREIAHVKNCGKAKTSILYLKVALFSYFGIW
IFFPVLWILGERGIQIISDDLNHVLHCILDVIAKSCYGYALLYFRVYFDKKLITSGMNEEEFTKLSKDIV
THSDEV

>gene85
MDTLAWVARELLSTAHDATPATATPSTDHSTPSTDHGSGETFNVTITIGGGHHGGHAGPVDNSIVIGGID
GWIAIPAGDCYCAGWYVSHGSSFEATFAHVCQWSIFAVCILSLLWYAWQYWKATCGWEEVYVCCIELVFI
CFELYHEFDSPCSLYLSTANIVNWLRYSEWLLCCPVILIHLSNVTGLSDDYGRRTMGLLVSDIATIVFGI
TAAMLVSWPKIIFYLLGFTMCCYTFYLAAKVLIESFHQVPKGICRHLVKAMAITYYVGWSFFPLIFLFGQ
SGFKKISPYADVIASSFGDLISKNMFGLLGHFLRVKIHEHILKHGDIRKTTHLRIAGEEKEVETFVEEED
EDTV

>gene86
MLGNGSAIVPIDQCFCLAWTDSLGSDTEQLVANILQWFAFGFSILILMFYAYQTWRATCGWEEVYVCCVE
LTKVIIEFFHEFDDPSMLYLANGHRVQWLRYAEWLLTCPVILIHLSNLTGLKDDYSKRTMRLLVSDVGTI
VWGATSAMSTGYVKVIFFVLGCIYGANTFFHAAKVYIESYHVVPKGRPRTVVRIMAWLFFLSWGMFPVLF
VVGPEGFDAISVYGSTIGHTIIDLMSKNCWGLLGHYLRVLIHQHIIIYGDIRKKTKINVAGEEMEVETMV
DQEDEETV

>gene87
MSRLVAASWLLALLLCGITSTTTASSAPAASSTDGTAAAAVSHYAMNGFDELAKGAVVPEDHFVCGPADK
CYCSAWLHSHGSKEEKTAFTVMQWIVFAVCIISLLFYAYQTWRATCGWEEVYVTIIELVHVCFGLWHEVD
SPCTLYLSTGNMVLWLRYAEWLLTCPVILIHLSNLTGMKNDYNKRTMALLVSDVGCIVWGTTAALSTDFV
KIIFFFLGLLYGFYTFYAAAKIYIEAYHTVPKGICRQLVRLQAYDFFFTWSMFPILFMVGPEGFGKITAY
SSGIAHEVCDLLSKNLWGLMGHFIRVKIHEHILVHGNITKKTKVNVAGDMVELDTYVDQDEEHDEG

>gene88
MAELISSATRSLFAAGGINPWPNPYHHEDMGCGGMTPTGECFSTEWWCDPSYGLSDAGYGYCFVEATGGY
LVVGVEKKQAWLHSRGTPGEKIGAQVCQWIAFSIAIALLTFYGFSAWKATCGWEEVYVCCVEVLFVTLEI
FKEFSSPATVYLSTGNHAYCLRYFEWLLSCPVILIKLSNLSGLKNDYSKRTMGLIVSCVGMIVFGMAAGL
ATDWLKWLLYIVSCIYGGYMYFQAAKCYVEANHSVPKGHCRMVVKLMAYAYFASWGSYPILWAVGPEGLL
KLSPYANSIGHSICDIIAKEFWTFLAHHLRIKIHEHILIHGDIRKTTKMEIGGEEVEVEEFVEEEDEDTV

>gene89
MEPVLGLASTAVRELTAGGSGNPYESYKPPEDPCALTPFGCLTNFWCDPQFGLADAKYDYCYVKAAYGEL
AIVETSRLPWLYSHGSDAEHQGALAMQWMAFALCIICLVFYAYHSWKATTGWEEVYVCVVELVKVLLEIY
KEFESPASIYLPTANAALWLRYGEWLLTCPVILIHLSNITGLKDDYNKRTMQLLVSDIGCVVWGITAAFS
VGWLKWVFFVLGLLYGSNTYFHAAKVYIESYHTVPKGHCRLIVRLMAYCFYVAWTMYPILFILGPEGLGH
MSAYMSTALHGVADMLSKQIWGLLGHHLRVKIFEHILIHGDIRKTTTMQVGGQMVQVEEMVDEEDEDTI

>gene90
METAATMTHAFISAVPSAEATIRGLLSAAAVVTPAADAHGETSNATTAGADHGCFPHINHGTELQHKIAV
GLQWFTVIVAIVQLIFYGWHSFKATTGWEEVYVCVIELVKCFIELFHEVDSPATVYQTNGGAVIWLRYSM
WLLTCPVILIHLSNLTGLHEEYSKRTMTILVTDIGNIVWGITAAFTKGPLKILFFMIGLFYGVTCFFQIA
KVYIESYHTLPKGVCRKICKIMAYVFFCSWLMFPVMFIAGHEGLGLITPYTSGIGHLILDLISKNTWGFL
GHHLRVKIHEHILIHGDIRKTTTINVAGENMEIETFVDEEEEGGV

>gene91
MRGAIAIWAEVGFGGHSCLLSPKAHLPANKMNRILLNAVVDPHATAAATDTNGTEAAASHGECVPNSHEY
GHQWEQDFAMGCLWFCFALSIGILLFYAFENWRATCGWEEVYVCVIELIMVSLEIFKGHSAPATIYQTNG
TRALWLRYAEWLLTCPVILIHLSNITGMHEEYTKRTMTLLVTDIGTIVFGTTAALTKGGLKVLFFFIGLT
YGCVTFFNAAKVYQESFYMVPAGMCRDLVKYMCWIYFISWPMFPVLFVAGPEGFHVISWSGSIIGHTVAD
ILSKNLWGLIGHYLRYKIHEYILIHGDIRKKVHHNILGQDEEIEEFVDEEDDQTTKMSSQG

>gene92
MASLAGLVRGAARHLQSGSTPENPYQTPPNDGCGLTPFGCVEDFWCDPKFGLADAKYGYCFVEAAHGELI
IGPDSNASWLLSRGTDSEKLGMQICQWTAFAIAIFLLGFYAFSAWKATCGWEEVYVCLVEITYVTLEIFN
QFDSPSMIYLSTGNAVYFMRYTEWVLCCPVILIHLSNLSGLKNDYSKRTMRLIVSCLGMLVFGMAGGLST
GWLKWFLFGIGCLYSGQTYFQSAKCYVEAHHCVPKGHCRLVIKLMAYAFFASWGTYPVLWVIGPEGLKHT
SWYTNTIAHTLCDILSKELWPFLGHHLRIKIREHILIHGDIRKKTTIVIAGESLEVEEFVDEEDEDTV

>gene93
MGGIGGGGIQPRDYSYGANGTVCVNPDVCFCLDWQQPFGSNMENNVSQGFQLFTIALSACILMFYAYEWY
KATCGWEEIYVCVVEMSKICIELVHEYDTPFCLYLATGSRVLWLRYAEWLMTCPVILIHLSNITGLGTDY
NKRTMVLLMSDIGCIVFGATAAFANEGYVKCACFLLGMAWGMNTFYNAAKVYYESYVLVPSGICKLLVAV
MAGLYYVSWSLFPILFAIGPEGFGVISLQASTIGHTIADVLSKNMWGLMGHFLRVQIYKHILLHGNIRKP
IKLHMLGEEVEVMALVSEEGEDTV

>gene95
MDDCQCHTWYNSYGTQSENDAEQYVSWFIFGITSLNLLYYIWNIFRATAGWEEIFVCIIEGFAAAVSAWP
QLSNPYTIYLSTGQRVQWIRYFEWLMTCPVILIHLSNLTGMKHHYPLRTMTLIVSDISCIVTGVQCALSN
GYAKIIWFLVSCAFGLVTYFSAFKTYKEAYALVPDLISKRIVLTMATIFFVSWTGYPLLFALGPEGFQQI
SYYAMNIGYCLNDMFSKNIWGLLGHTLRIRAFNNSKDGWVVPRSSSSVGGSGKDMTADQQVWVAG

>gene96
MAPAAPDPHATPDPHATAPPASHSNITYPPAPGGPGGAGPQNPAISLQQALGWSKNGDGEGYVAVLSNQC
YCNGWYDSKGTDDMQVFATIMGWVVFISCLLVILMNFYILKVKGRKAVSWEVMFVMTIVFTFYGFQIFLP
WASPVMVFLNTGNFVVMLRYGQWLFATPVLLVLVADMTGRHANYTFDHMYMVIADVFAVTIFVFGCLSDN
AWVKVFAYTFSLLGFCYLFWQAWRCFGILSDEVPNALTRKLTTVTIIVFYVGWNLHTVFSTIGPEMGMFL
DAVQTEIGATFSDMVAKVAFCLLATVLKLQVFNHYEKHGDPLLAAAEKEQQAAAAAAG

>gene97
MTTISEVCGVWALDNPECIEVSGTNDNVKMAQLCFCMVCVCQILFMASQYPKVGWEAIYLPSCECFLYGL
ASSGNGFIQLYDGRLIPWARYAAWICTCPSILLQINTIHKCKISHFNLNTFIVQADLIMNIMGVTGALTT
NIAFKWIYFAIGCILFIFIVLVVYDIMTSAAKEWKAKGDSKGNLVSTRLILLRWIFIVSWCVYPLLWILS
PQATCAVSEDVISVAHFICDAFAKNMFGFIMWRTLWRDLDGHWDISRHYPQSSYAKDGKEEEQMTAMSQT
DDTEKPHSSQG

>gene98
MSFSSRSMAIMTAVLFLVLPAVKAVEVCKVADLRLVTQTGEFFLAVTFLVMGVSTVGYVLLAFKAIHEKR
KFHFAMCFVTAIATFSYYAMLSGQGWLITPACRQLFYVRYIEWMSTTPLLMLVLGMLADADIAYLLAVMG
GTAMMIFGGLMATISSGHIKWLWFTLSLGVFFALAFVMVRGFKVLVEKNHPSIFELYNKVSTLAAISWAL
YPVVFIFSEGTGDWSPNFEIMLYSVLDILSKVVFGYIVLLSHEGLDRLVGLKGMPAAPTANYGTPTSKSF
AAI

>gene99
MGGQAWLFSVLVAVMLEVVSGVEVCKVADLRLVTQTGEFFLAVTFLGMGVTTVIFVLMALKASHEKRKFY
FTCCYIGMIATFAYYAMLSGQGWIITPSCRQMFYVRYVEWIASTPLLLMLLAWIVGADIALTIAVMGAQL
LMIFGGYMAAISSGHIKWLWFGMALLIFGPIIYIIMSVFKGMVERSDASVAELYNKLSWLTVVTWACYPF
VFLCSEGTGDWSPNFEIMIFGVLDLMSKSVFGFILLLSHQALDRIANLPRFASIGGGDYGTQMADSQRNS
KG

>gene100
MDAVGAVPGAENPFQNPAVIVADWLGFIILGGSSIILCYKLMGFYGSDNKQKYFIGYREEKMLSVYVNTF
AAVVYWARVSAHANGDVGLAAYVHLLKYCDYCFTCPLLTLDLMWSLNLPYKFTYSIWVGICIFSGIGSTK
FQPPAKYMWFAFGMVIMFIVFWHIYSVVIIRLNQIFCGSTKKSANTLQIACAVYFSIWWGFPVLWFLLEF
GVIGHPVTLCMHTVLDVSAKSFYGFMLLSFQLQAEREEWIFLPLQPKIAWNENASETESQADVELGYNSE
YGFTGG

>gene101
MSEGIPTACTTIEEVGSTGVGALTVAFLILVVTTIVFLFKANGSGEQRKYYYLSTYMCGFAAMAYFAMLS
GQGWTAIAGCRQFFYARYADWLVTTVLTVVLLGSVAGASGDVIGGAVGADVIMIFAGYMGSVSVVTTVKW
FWFLIGVMGLIGVVVHFAQTFKAAADSKGGDIATLYGKIAWVTILAWICYPVIWLFSEGFASFSVSFEVC
AYALVDIIAKVVMSFMVMSAHDLLGEGAQTREYAVAYLSHGELVFLSVLHGKHVNTYLANKGVWFRLLVL

>gene102
MVFSQLAGATAAAKSLAPTACTTIEEVGSTGVGALTLGFVLLSVTTVILIYKAFNSDIERRKYFFACTFI
CGISTYAYFAMLSGQGWTAISGCRQFFYARYVDWVLTTPLTILVLGLVAGVELEVIAAVAGADVLMIVCG
YMGAVSVVTTVKWFWYIFGLIMFIPVVYSLARVFRETVLNRKDPEMIQLYGKLAWLTIIIWAFYPVVWLF
SEGFASFSVSFEVVGYTIMDLIAKCLFCFMVVQADPILGPHTHVPTQQQEYVTLDHHVRTRVGVTRLCGI

>gene103
MPFTQLAGASAAAKALAPTACTTIEEVGSTGVGALTLAFLALTVSTCVMIAKAASSGPEHRKYYFLNVYI
CGISAFAYFAMLSGQGWTAIAGCRQFFYARYVDWFMTTALIILQLGLIAGQDFVSIAAVMGTDMLMVLAG
YMGAVSVVTTVKWFWFILGLSLFMPVIYSLARTFRETVITRKDPEMIELYGKIAWLTIILWAFYPVVWLF
SEGFASFSVSFEVVAYAIMDIISKVVFTFMIVSAHDSLAAPSMAPAQSREYVVPQAKELGSRVFRPQDPG
LESKGPLVCMSVGL

>gene104
MGLTDVASTQAEYTTSTRHFLLIGFLAMFVASVVFCYLGMKKKKDNIPETLVFIITSVAAGSYYIMWSGY
AVAHKTDANDQHRMIFWGRYVDWFITTPLLLLTLSMVAEAEAGMTILLIGIDMLMITSGAIGASVVHPYK
WFWWTLGVIFFIVVMFMLVGLNSTVKENKPDIAIGFRTLIVLTAISWCVYPIVWVLGSEGLAAVHIDVEV
GLICLADLVSKVVFGLYLVFSVNVGDNGGDSAEKTSLVPAFSSEGQSGEDLLRGRVSPWFDLEKCGCRRH
SGLKSHGG

>gene105
MCGILSTLLHVRFTSDGSFDIIVDALAKTQAMRTAPKWRRAHQLCLFRQLFAIYNLHQLLLRSATKCTVL
KTTTANNQVRRKMSIAEYFPLLPGTVNLPGYINFNNFSTYYLQQSLLLQQAENVTQLGLDFFWLAFGLFV
ILTVIFTILTVFRAKETRRLHIVITISVAVEAAAYYAMAMLQGFYPQQSDVDSGYRAFFWARYVAWALSA
PLLLYALSLLGRASTETTLAVIGCAEIFVVTGLFGGLSIDGSRWGWFGFSVAALSLVFLILFGNMSKSAF
RKGRCTGFAFLSLCFFIVGTWSIYPVVWVIGEGTTSVSVNVEILLYAILDVITRGFFGLAILLSEPVAML
YLDHN

>gene106
MGAPINRLLLGLAVLGCLGMAEGVEVCRVADLSLVTQTGEFFLAVTFLGMGVTTIIFVLMMLKAASEKRK
FYFVTCYCCGIATFAYYAMLSGQGWLISPSCRQLFYVRYLDWFFTTPLLLIDLALIAGADWTYVAAVVGG
DMMMIFGGYMASISSGHIKWLWFAISLLIFAPIIYTVLHTFKVLVERSHPAVGELYNKCSWLLVVTWACY
PVVFIFSEGTADWSPNFEVMIYGVLDLMAKCVFGFILLLSHEALDRSTNPPRFASIAGGTATAYGTQMQE
QKPVESYIG

>gene107
MTVARMEGTMEGWWAPNTLSDTAIGAHWLTFLALLACTVFLAYESFAAKGPSGREKFFAGYHEQYNLALY
VNLMASLSYFAKVVSDTHGHNFENVGPFIIGLGNYKYADYMLTCPLLVMDLLFQLRAPYKVTGAVLIFAV
LFCGAITNFYPGKENQQAALAWFGMGVFYYCLSYFFLGYIVSRQYRRLEEMAMGTEAKKALGPLRLAITV
FFTIWVAFPAVWIVSDRGFNVIDASTAEVLHCIADLIAKSLYGFALARFRRYYDKKMFEILENLGYDGEE
AIEELENEMRHMDKEEELEKQAV

>gene108
MTMLEHLEGTMDGWYAENDLGQGAIIAHWVTFFFHMITTFYLGYVSFHSKGPGGKQPYFAGYHEENNIGI
FVNLFAAISYFGKVVSDTHGHNYQNVGPFIIGLGNYRYADYMLTCPLLVMDLLFQLRAPYKITCAMLIFA
VLMIGAVTNFYPGDDMKGPAVAWFCFGCFWYLIAYIFMAHIVSKQYGRLDYLAHGTKAEGALFSLKLAII
TFFAIWVAFPLVWLLSVGTGVLSNEAAEICHCICDVVAKSVYGFALANFREQYDRELYGLLNSIGLDGED
VVQQLEKEMQTNHHKKKSINSPAVG

>gene109
MAAEIKARGYTYNEVVANLDPLLHGRLEGTMQGWYAPANYVGSTMIAHWITFLVLSACTMYLARDSFANR
GPRGNTAYYSGYNEQYNIALYVNLMACVAYYGKVVADTSNHNFSNVGPFIPGLGNYRYADYMLTCPLLVM
DLLFQLRAPFKITSAFLIFVVLLCGVVTDFYPPDVLYGPSVAWFIFGCFWYIIAYIFLYTIITKQYKRLL
EISKETEAKKSLGPLKLAIYTFFSIWLIFPCVWLLTPKGLNMLDEDSAEVLFCICDLLAKSMYGFALSRF
RFYYDKKMYDLFEQLGYDPENIEEEMQKELKL

>gene110
MFVFGFVVMVVAAVFFLVMALRYDGEADTRRHGPMFFYLVFFVTSISAMTYYSMWTETSVLHVEQEGDDE
DKRTIFPARYIDWIVTTPLMLATVSLLGNAPTSSLVAMIGSDLLFFSSLYIGAVQTAGHKWFWWGVGVFF
FIMLVYFMLVELSKVDTGRVKEYDADTLRMLTYFFIICWAFFPLLWLFGQEGTAASSLPVQAAFTTLADL
TAKIAFGLIVVFRSQPDYTEYVPERDPQRVVTSSAFVAPELGGVGMAGLHSRKGAKRSGAEEEEQQQRRR
RCSVG

>Gene111
MDAFTAIASTAVRSLQAGSNPYEDLEPPQNNCVLSPFGCLDNFWCNPALGLADAGYGYCYVKSAYGQLAV
VPTENLAWLTSHGTQAQKVGAKASMWFAFALCIVILLFYAYHSWKATTGWEEVYVCCVELIKVLLEIYHE
FDSPASLYLSTGNWVLWLRYGEWLLTCPVILIHLSNITGLKDDYNKRTMRLLVSDIGCVVWGITAAMTVG
YLKWIFFGLGMLYGSNTYFHAAKVYIESYHTVPKGHCRLVVKAMAFCFYAAWTMFPLLFALGPEGLNVMT
GYTSTICHTVADVLSKQIWGLLGHHLRVKIYEHILIHGDIRKTTTMIVAGESVEVETFVEEEDEDTV

>Gene112
MGTPDPLLSSIPGTDIGLGDWTEYSNYYFLNATNSTHKWVAGPEDDCFCKAWTFNRGSDEESVAAFAIAW
VVFSLSVLQLLYYAYAQWRSTCGWEEVYVGIIELTHICIAIFREFDSPAMLYLSTGNFVVWARYASWLLS
CPVILIHLSNLTGMKGNYSKRTMALLVSDIGTIVWGSTSAMSPHNHVKIIFFFLGLVFGLFTFYAAAKVY
LEAYHTVPKGKCRNIVRFMAWTYYVTWALFPILFILGPEGFGHITYYGSSIGHYVLEIFSKNLWSGTGHY
LRLKIHEHIILHGNLTKKTKINIAGEPLEVEEYVEADDTDEGV

>Gene113
MLNVAHSSDVRLVVISALLVALTWNSYPAAMGIYLGTEDNLAGFMLWISFLSSLVLSAFFCAQIRYGICG
WEVAYLSVIEMMIYVTAIGWESEPPCSAYLISGQIVPWLRYVEWLITCPVILIALSRVGQSEGGYSKRTM
KLLTSDQGTIILGILAASSPDSAGQALFYLCGVAYGFTTFYTAAVVYFEAWQAIPDECRAILKSMIYFFY
SGWLMFPILFIIGPEGAGHITPAGSTIGHVIADLLSKQMWSICEYMMENKLHILNLLAEEEEEEEEEGG

>Gene114
MRRRALLLLLTFICLVALGLASTPAPVVSTSPPPATAASGDAHGAPAAGESGGNHAAPTEHGGEGHSSEN
GGEHGGGHSSEHHSSSDGELYPPTFSEEEAARALQWVCFMISLMLALFYFWQYYIGTSKWEVLYVCFFES
LAYIFDIWYLHKNPISLVFKQGHHAEWLRYMEWLLTCPILLIALSRVGVADGAYSKRTMRLLTSDQGTIL
MGVTAAFAQGYVRLGFYLVGCAFGFNTYYTAAAVYLEAWRNVPDAEKDVVKWMATIYYFSWVMFPVLFAI
GSEGLGYITLSGSVIGHCIADLLSKNVWGLLEWYLEFCVHSRYHAELAALEEEEEEEEGEEGA

>Gene115
MADLRYGPEPVFGSCADGDPESYCRLIVNTTLYPVLVVELGPVWEQLVSRGTQWFTVILSIFMTLWFIYN
FYTGHCGWEVIFVTTIEFIKILIEVFFEYETPCMIYSVFGPVTSWLRYIEWLLTCPVILIHLSNLTGLDE
EYSARTMGLLTSDQGTICFGVTAALSRNGVVKVITFLTALCFGVSTFYSASRVYIESYHQVPKGLCRRLV
RYLTAMFFCSWLMFPLLFLAGPEGLAYLTWSGSTIGHTVADLLSKNIWGLIAHYLQVKIREHIILHGDVR
AKVEKTVAGHTFEVEGFRDEEDDNA

>gene116
MQYTVNYTTMGNDTYCVQGDACFCLPWFDSLGTQSEKTAANALQWTAFGISVAMLIFYAYETWKATTGWE
EVYVCGVEMVKVIIEFFHEFDSPAMLYLSSGERVLWIRYAEWLLTCPVILIPLSNLTGLKDDYNKRTMRL
LVSDIGTIVMGATSAMCTGNWLKILFFCLGLLYASNTFFHAAKIYIESYHIVPKGTCRMLVRLMAWIFFA
SWACFPLLFVLGPEGFGHISWYASTIGHTIIDLMSKNLWGMLGHFLRVQIHKHILLYGDIRKKVMVNVAG
EQMEVETMVEE

>gene117
MGGAARVLLAGQEASLKNGPQNSATRVDNGTLDGWFVPYGDCFCYPWYATRGTTVEYDLAYVMNWIAFAI
SMITLAWYAYEYVVASCGWEEVYVCIIETFNVAFEIFHEFDEPLMLYLSSGVRVPFLRYTEWLMSCPVIL
IHLSNLSGLKDAYNKRTMVLLVSLIGCLVTGVYCAMIPQGWVKWLMYTVSFAYCAITYWQAARVYAETYI
MMPKGGCRRMVAWMAAIFFTCWGIVYPLTFALGPEGLNVISFHLSNFIHCVADLFAKLLWGAMGHHLRFM
IAKHILIHGDLRKTTKVKFGNAELEVETFIEADEASEEGA

>gene118
MSSIGGSGIVTGTTVSPGSTNLVQIDHGGEAMYWIVFGVFSLSALAIAVLTYRKAPALRSHGYCTLAILC
TASIAYFSMATQGGYAFIKVYDETNTARAIYWARYVDWTITTPLLLLDILLMAGLSIGDTLWIVAADLAM
ILTGLFGALLPNRYKWGWFGIGCLFMVFIMWGLLFHGRRAAFLRSKKIGGVYSILSLYLLALWWIYPIAW
GLAEGSNTISSKAEAIFYGCLDLMTKGVFGWMIMLMAEPVVARQHREEEELNGAHPSLLAAPINAPLKGV
WAANGG

>gene119
MSFHADSDLFSDVCYCDAWTIKLGSNSEKHAIAAIDIFAIIVSVLPLMYYMYSSWKSTCGWEEVYVCSIE
LTKCFISLFRSYEAPGVLYLSTGNTLYWLRYGEWLLTCPVILIHLSNITGLKDAYSKRTMKLLVSDIGTI
VFGVSSAASTGPLKIFFFVLGLIYGANTFFHAAKVYVESYHMVPRGTCKLLVRLMAMVYFVSWSMFPIFF
VLGPEGFGVLSHLGSDIAHTWIDLMSKNLWGFIGHLLRQKIHYHILKHGDIRKTTKVNIAGEEVNVETYV
EEEDEETV

>gene120
MAAGLEGLVSSASRGLHASIPENPYHSDGHHLPCGLTPFGCMDDFWCNPEYGMSYAGYTYCFSELAFGKL
VMVPEADAGWLHSHGTQAEFVAATACQYTALSLALLLLSFYAYSAWKATCGWEEGYVCCVEVLFVTLEIS
NEFNSPATLYLSTGNYCYFLRYGEWLLSCPVILIHLSNLSGLKNDYSMRTMRLLVSCIGMLITGMAGGLG
VGWVKWTLYFVSCAYSAQTYLQAAKCYVEVYATVPKGYCRTVVKLMAYAFFTAWGAYPILWAIGPEGLKY
ISGYSNTIAHTFCDILAKEIWTFLGHHLRIKIHEHILIHGDIRKKVQVRVAGELMNVEELMEEEGEDTV

>gene121
MTISDTNIGQPLYYGHTCQCDAWEIPYGTDMEETAGNVMQIVTIAFSVLPLFFYAYSAWKATCGWEEVYV
CIVELICMVLSYEFERQSPATLYLSTGNQVVLLRYSEWLFTCPVILIHLSNITGLKDNYTKRTMALIVSD
IGCIVCGVLAASSQGPIKIVFFLMGLTYGCMTYYQAGRIYIEAYHMVPKGRCKLYVRLMAWIYFITWLCF
PIWFLCGPEGYGVFSFYGSSVAHSISDIWSKMVWGFIGHNLRIEIHKHILKHGDIRKKVNVNIAGEEVNV
ETYVEEEDEETV

>gene122
MNNTCDCLDWHQDLGGHAENIVNNVVQWLCFGLSIANLAYFAWSTFRATCGWEEVYVCSVELTKISIEMW
HEYSNPFTIYTSTGRWVTWLRYAEWLMTCPVILIHLSNLSGLTHDYSARTMGLLVSDIGTIVLGISAAFA
TGPLKAIFFCFACCYGATTMFHAAKVFYDSHNEVPKGACQTEVKWLAGLYFFSWNMYPVIWLLSQEGYQV
LTWHEAGILYNIADILAKNIWTLIGHDLRLRIHKFVAAGGKLNKASAHNTVEADNESKKSSP

>gene123
MADLFSLDGSGATRQLLAQSIAGAVNNTDCYCDAWYTSLGDDWLFTFGNVMTWIVFALSILVLVFYVWHS
FHATTGWEECYVAVIELGMACLELYTEFYTPFTLFLSNGNIVAWLRYAEWLMSCPVILIHMANLTGLNNS
YSPRTMSLLISDVGCIVLGSTAAMATGWVKWVFFCMALSFGLTTYYNAARVYMEAFHTVPKGICRQLVMA
LAWIYFTSWSCFPILFILGPEGLDHIPFMASSVAHNFLDLISKNLWGMLAHMLRVKIHEHILIHGDIRTK
TKVSVAGQDVEVETFVEPGTVEDAV

>gene124
MHKVKYEIPQASRHHRLAQPQGQTDFRPLHRKAPTTILCTSLKQLQAIILNFFRQAIQIRRKRFLTPRMD
HPVARSLIVGKVYSDLSNGSLVIPSNDCFCLKWLKSRGSTIEQKIANALQWFAFGLSVLILIYYAYATWK
TTCGWEEVYVCCVELTKVVIEFFHEFDAPSMLYLSNGNRVLWLRYAEWLLTCPVILIHLSNLTGLKDDYN
KRTMRLLVSDIGTIVWGTTAAMSTGYVKVIFFLMGVIYGANTFFHAAKVYIEAYHTVPKGLCRQVVRVMA
WLFFASWGMFPVLFLLGPEGFGHIGLYGSTIGHTVIDLLSKNCWGLLGNFLRVKIHEHILLYGDIRKVQK
IKVAGQELEVETMMTEEATDTI

>gene125
MRTPLLAVLAIMAFLATTAVCEEAREVVTLDDGTELQMHGDEDLSGPKRQLMKKALFGDPVPIPEYDVCV
PTMFGCSNDFWCNPEYNLTDAGFGFCQITWPEQERAFAVGLFKKLSWQIPHGSETEKKVLRGFQWVAFAF
SLGVLAYYAYNSWKATTGWEEVYVCIVELIKVCLEIWHEPDHPATLYLSAGNFILWLRYGEWLLTCPVIL
IHLSNLTGLKDNYSKRTMMLLISDIGCIVWGVTSALTTLTWLKWFFFFIGLCYGLSTFFHAGRVYIESYM
MLPKGQCRKILRIMAGVYYFSWSAFPILFILGPEGFGHISAYASTIGHTINDMLSKQLWTLLGHHLRNLI
HIHIIKHGNLTRKTKMTFLGQDVEVEEMVEEEGDDTV

Methods

TRACE

The main goal of TRACE is to use partial sequence information from transcriptome sequencing to generate cDNA with fully preserved 5' ends. This requires total-RNA or mRNA sample and also transcriptome sequencing results to generate gene specific primers. For plant and hard tissue samples it is generally recommended to use trizol (phenol-cholorform) extraction with a tissue homogenizer. The 5' RNA ligation and RT-PCR reactions are detailed below. See Mandl et. al. 1991 Biotechniques and Maruyama K, Sugano S. 1994 Gene for detail on capped RNA selection (i.e. non-degraded or truncated RNA). This protocol is similar to the Ambion RLM-RACE kit, which can be bought commercially. Although at the time of debugging the protocol we were not aware of the Ambion kit. We also tested alternative kits that relied on reverse-transcriptase with terminal transferase activity to generate 5' overhang for amplification of cDNA ends, however we did not obtain full length 5' ends.

11-17-2010

Algae TRACE

extracted the following samples, loaded ~30-40 mL culture and used 4-5 mL trizol

CACC 51, 53

UTEX 232, 1293, 1644, 2377, 2565, 2601,

see word doc. for conc. <1 µg total RNA discarded samples

combined and concentrated samples in Zymo RNA-5 columns, split CACC 19 sample into 2 columns

sample #	amt loaded [µg]
19	8
51	1.5
55	7
53	6
1644	3
1293	1.5
232	4.5

eluted in 15 µL total H_2O

SAP rxn

15	total RNA
2	10x SAP buffer
2	SAP enzyme
1	RNase Inh.

Incubate @ 37°C for 1 hour

dilute to 100 µL, purify Zymo RNA-5

elute → normal 7 µL H_2O

↘ small 6 µL 1x TAP buffer

tube	sample	
1	19	normal rxn
2	53	
3	55	
4	232	
5	51	small rxn
6	1293	
7	1644	low amt of starting RNA

TAP rxn

normal	small	
6	5	RNA
1	—	10x TAP buffer
2	1	TAP enzyme
1	—	RNase Inh.

incubate @ 37°C for 1 hour

ran out of TAP enzyme → reorder overnight

tube 1-4 : transferred RNA to per tubes added 1 µL RNase Inh.

tube 5 : performed TAP rxn

tube 6-7 : transferred RNA to per tubes

11-18-2010

Algae TRACE

continue rxn from 11-17, wait for enzyme for tubes 1-4, 6-7

RNA ligation

normal	small	
2-6	6	RNA
1	1	5' RACE adapter (10μM)
1	1	10x RNA ligase buffer
2	2	T4 RNA ligase
4-0	—	H_2O

use small rxn for tube 5 with RNA adapter 2

incubate @ 37°C for 1 hour

Reverse Transcription

use Superscript III RT kit, prepare the following mix

1	oligo-dT primer (50μM)
1	dNTP mix (10mM each)
3	ligated RNA
8	H_2O

setup two RT rxn for tube 5

loaded 3μL ligated → "L" tube

6μL RNA → "H" tube

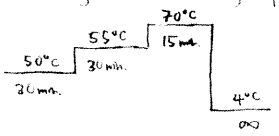
used DNA adapter 1 - oligo dT primer

heat to 65°C for 5 min, incubate on ice immediately for > 1 min.

add to tube

4	5x first-strand buffer
1	0.1M DTT
1	RNase Inh.
1	RT enzyme

incubate using the following programm



add 1μL of RNase H and incubate for 20 min @ 37°C

①

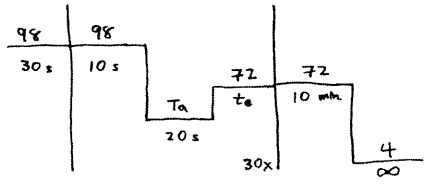
11-18-2010

Algae TRACE

pcr rxn

4	5x phu HF buffer
1	dNTP mix (10mM each)
1	primer A (10μM)
1	primer B (10μM)
X	template → 2μL for transcripture / cDNA
add to 20μL	H_2O 50-100ng for genome
0.2	phu polymerase

use the following protocol for pcr



T_a = annealing temperature

t_e = extension time, roughly $\frac{15\text{-}30 \text{ sec}}{1 \text{ kb}}$

CACC 19 pcr

tube	template	primer A	primer B	Tq	te
19-1	1μL CACC 19 genome (80 ng/μL)	CACC19 GSP1	CACC 19 GSP2	62	1 min.
~~19-2~~		~~CACC19 GSP3~~	~~CACC19 GSP2~~	~~55~~	~~1 min.~~

too low don't run rxn

both rxn are for Scherffelia dubia opsin 3

based on mRNA transcript region between GSP1, GSP2 should be 500 bp

but there could be large introns in between so a longer extension time is used

CACC 51 pcr

tube	template	prime A	primer B	Ta	te
51-1L	1μL CACC51 cDNA "L"	DNA adapter 2	446915	58	30 s
51-1H	" "H"	"	"	"	"

both rxn are for Tetraselmis cordiformis opsin 1, should be ≥ 100bp, missing helix 1 & 2 in ref. seq.

②

Virus production

Lentiviral protocol
Amy Chuong
Modified 2011-01-29

Materials

- Low passage HEK293FT cells (less than 15)
- Sterile D10
- Sterile virus production media
  - 500 mL ultraculture (Lonza 12-725F)
  - 5 mL pen/strep
  - 5 mL sodium pyruvate
  - 5 mL sodium butyrate (100x 0.5 M stock solution made from powder)
- Sterile 20% sucrose (200g/L)
  - 10g sucrose
  - PBS to 50 mL
- Trypsin-EDTA
- Fugene 6 / Mirus TransIT
- Ultracentrifuge tubes
- T175 plate (BD Falcon 353112)
- 100 mm plate (BD Falcon 353003)
- 140 mm plate (BD Falcon 353025)
- Viral plasmids, packaging plasmids

Day 0:

1. Passage cells such that they will be almost 100% confluent on day 1.

Day 1:

2. Transfect with Fugene 6 when cells are almost 100% confluent. For a T175 flask:

DNA mix	22 ug	lentiviral plasmid of interest
	15 ug	pΔ 8.74 (ampR)
	7 ug	pMD2.G (vsvg coat protein, ampR)
Fugene 6	132 uL	
DMEM	Final volume 4.5 mL	

   - Place Fugene into DMEM without touching sides of plastic tube. Mix with tapping, then let sit 5 min at room temperature.
   - Simultaneously mix DNA (carrier + pDelta + vsvg) in another tube.
   - Add DNA to Fugene + DMEM mix while tapping tube lightly to mix.
   - Let mixture rest 20-30 min at room temperature.
   - Simultaneously replace HEK cell media with 16 mL fresh D10.
   - Add 4.5 mL Fugene+DMEM+DNA to T175 flask, distributing evenly with rocking.

Day 2:

3. Gently remove transfection media 24 hours post-transfection and replace with 20 mL virus production media.

Day 3 & 4:

4. Spray ultracentrifuge tubes with ethanol in hood, then leave to air dry under UV.
5. Collect virus supernatant from each plate into 50 mL conical and replace with 20 mL fresh virus production media. This can be done 2x, 3x usually doesn't yield good virus.
6. Spin at 1000 rpm for 5 min to pellet cellular debris and filter through virus production media prewetted **0.45 (NOT 0.22) micron** filter flask.
7. Store at 4 C, or spin down.
8. Transfer 20 mL of supernatant to ultracentrifuge tubes and gently pipette 2 mL of 20% sucrose + PBS solution to the bottom of the supernatant to form sucrose cushion (virus will pellet out, light debris will not be at bottom).
9. Balance all 6 tubes with sterile PBS to a milligram within the holder. If there are less than 6 flasks of virus, balance 6 tube rotor with dummy tubes. Make sure tubes are full enough or they may collapse. Seal lid with vacuum grease.
10. Spin in SW-28 rotor in pre-chilled Beckman ultracentrifuge at 25,000 rpm, 4 C, 2 hours. Don't forget vacuum.
11. Handle ultracentrifuge tubes gently to prevent disturbing pellet. Aspirate supernatant. Place centrifuge tube upside down on kimwipe and use Pasteur pipette to remove any additional media on side of tubes – don't want virus production media in resuspension.
12. Resuspend pelleted virus in 25 uL cold PBS (per tube). Let PBS sit on pellet for 1-2 hrs at 4 C, then gently pipette PBS up and down without producing bubbles and combine resuspensions as appropriate.
13. Aliquot 2.5-5 uL/tube (1 uL per animal injection, try to minimize freeze-thaw cycles. Freeze at -80 C for up to a year. Use mammalian cell-freezing box (1 C/min) for optimal titer preservation.) For actual use, thaw virus on ice and centrifuge at 5,000 rpm for 5 min to pellet out any clumps. Keep virus cold in preparation for surgery.

Neuron culture

This protocol is modified from Richard Tsien and Yingxi Lin's lab protocols. The most important factors in consistent high viability neuron cultures are: fast dissection (<20 minutes for 5 pups), even matrigel coating, and media at the right pH. Generally hippocampal dissection from 5 pups should yield 0.5 to 1.6 million neurons total, and I plate at 20,000 neurons per coverglass. I typically prepare matrigel one or two days ahead of dissection and make new matrigel every two weeks. Most neuron/glia aggregation (microexplants) is caused by poor coverglass surface treatment. Matrigel should be diluted at 1:40 ratio from stock. I add ~75uL to coverglass in 24 well plate and incubate at 37C for 1-2 hours to polymerize. All liquid matrigel must be removed before plating neurons, there should be a clear white film on the coverglass. For hippocampal neuron culture, you should see processes growing within 24 hours. Imbalanced pH seems to affect neuron processes growth and also transfection efficiency.

1. Prepare surgery instruments, sterilize work area, cool Dissection/Dissociation Medium (DM) in small 5cm petridishes
2. Prepare plates, coat coverslips with Matrigel (75ul per coverslip)
3. Prepare Digestion Solution (DS), Inhibitor Solution (IS), balance with 1M NaOH, leave in 37degC bath
4. add Papain to Digestion Solution, put back into 37degC bath and wait a few minutes for the Papain to dissolve completely
5. Filter sterilize Digestion and Inhibitor Solution
6. get rat pups, put on ice, dissect, keep on ice at all time
7. Transfer hippocampus / cortex tissue into 40ml Falcon tubes, remove media by twice adding 1ml of Digestion Solution and then aspirate, finally add 3ml of Digestion Solution, incubate 6-8 minutes at 37degC
8. While incubating, remove Matrigel, dry in hood with cover off
9. remove Digestion Solution, replace with Inhibitor Solution by twice adding 1ml of Inhibitor Solution and then aspirate, finally add 3ml of Inhibitor Solution, put back at 37degC for another 4 minutes
10. remove Inhibitor Solution, wash three times with 1ml Plating Medium, transfer to fresh 50m Falcon tube and bring volume to 2ml
11. titrurate (1ml pipette tip, 10 times up/down), rest 1 minute, take off 1ml supernatant aliquot
12. add 1ml Plating Medium and repeat previous step
13. count cells in hematocytometer
14. plate roughly 20,000-50,000 cells per well (in 60-75ul), will adhere in 2-3 hours, then add 1ml warm Plating Medium
15. check glia in1-2 days, add 4AraC when glia density is 50-70%

Hanks' Balanced Salt Solution (HBSS)

1	L	H_2O
1	bottle	Hanks' Balanced Salts
350	mg	$NaHCO_3$
2.38	g	HEPES (final conc. 10mM)

adjust pH to 7.3-7.4, sterilize with 0.2μm filter and make 40mL aliquots
it should not contain Ca^{2+} or Mg^{2+}

Ky/Mg Stock

200	mL	HBSS
378	mg	Kynurenic acid (final conc. 10mM)
1.9	g	$MgCl_2$ (final conc. 100mM)

mix vigorously to dissolve, adjust pH to 7.4 with NaOH
sterilize with 0.2μm filter and make 50mL aliquots

Dissection/Dissociation Medium (DM)

500	mL	HBSS
3.47	g	D-Glucose (final conc. 35mM)
50	mL	Ky/Mg Stock

sterilize with 0.2μm filter and make 50mL aliquots
can store at 4°C for 1 month

Digestion Solution (DS)

1.6	mg	L-Cysteine hydrochloride
100	U	Papain
5	mL	Dissection/Dissociation Medium

adjust pH to 7.4 with NaOH (by eye), sterilize with 0.2μm filter
exact concentration of L-Cysteine hydrochloride is not critical, up to 4mg is suitable for enzyme activity

Inhibitor Solution (IS)

50	mg	Albumin from bovine serum (BSA)
50	mg	Ovomucoid trypsin inhibitor
5	mL	Dissection/Dissociation Medium

adjust pH to 7.4 with NaOH (by eye), sterilize with 0.2μm filter

Base Medium

500	mL	MEM
2.5	g	D-Glucose
50	mg	transferrin
1.19	g	HEPES (final conc. 10mM)

this is used for making Plating and AraC medium

Insulin Solution

250	mg	Insulin
20	mL	HCl (10mM stock)

make 1mL aliquots, final concentration is 12.5mg/mL

Plating Medium (PM)

500	mL	Base Medium
5	mL	L-Glutamine (200mM stock)
1	mL	Insulin Solution (12.5mg/mL stock)
50	mL	Heat Inactivated Fetal Bovine Serum (HI FBS)
10	mL	B27 Supplement

adjust pH to 7.3-7.4, sterilize with 0.2μm filter and make 40mL aliquots

AraC

500	mL	Base Medium
1.25	mL	L-Glutamine (200mM stock)
10	mL	B27 Supplement
500	uL	AraC (4mM stock)
25	mL	Heat Inactivated Fetal Bovine Serum (HI FBS)

adjust pH to 7.3-7.4, sterilize with 0.2μm filter and make 40mL aliquots

Matrigel

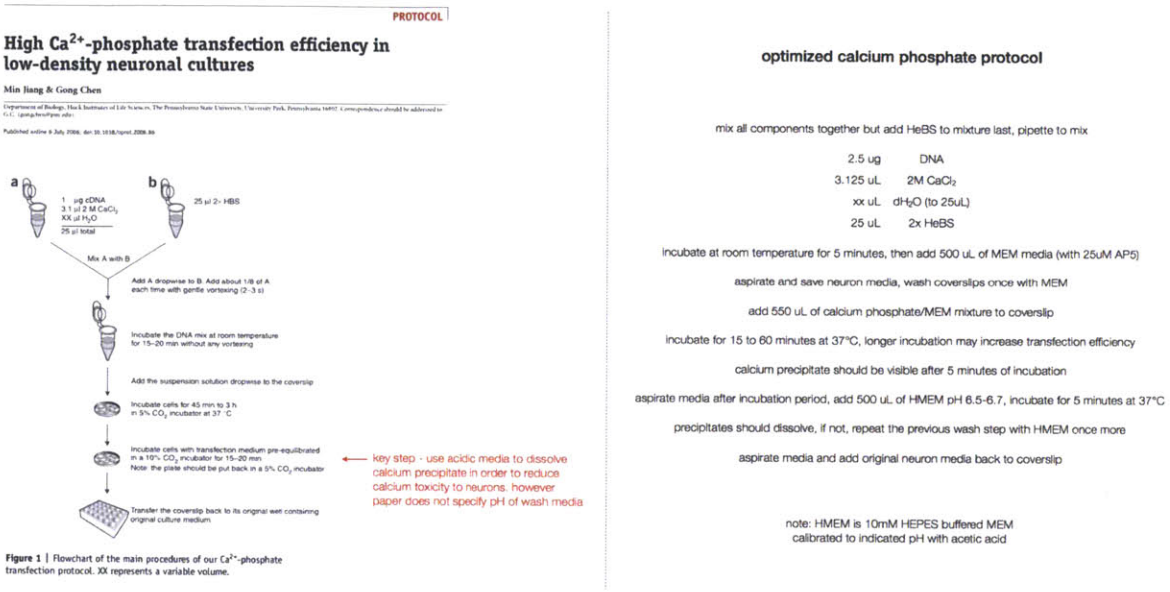
250	uL	Matrigel
12	mL	DMEM

always keep on ice, will turn into hard gel at room temperature
make new solution every 3 weeks

all solutions should be stored at -20°C unless otherwise stated

Calcium phosphate transfection

This low toxicity calcium phosphate protocol is modified from Min Jiang and Gong Chen 2006 Nature Protocols. We control the pH of both incubation and wash media to guarantee fine calcium phosphate precipitates and full resuspension of precipitates. We generally use incubation time around 20 minutes, although 60 minutes can also be used to increase transfection efficiency (upwards of 50-70% transfection). However cell health seems to deteriorate with long incubation even with full precipitate removal.



PROTOCOL

High Ca^{2+}-phosphate transfection efficiency in low-density neuronal cultures

Min Jiang & Gong Chen

Department of Biology, Huck Institutes of Life Sciences, The Pennsylvania State University, University Park, Pennsylvania 16802. Correspondence should be addressed to G.C. (gongchen@psu.edu).

Published online 9 July 2006; doi:10.1038/nprot.2006.90

a: 1 μg cDNA, 3.1 μl 2 M $CaCl_2$, XX μl H_2O, 25 μl total

b: 25 μl 2× HBS

Mix A with B

Add A dropwise to B. Add about 1/8 of A each time with gentle vortexing (2–3 s)

Incubate the DNA mix at room temperature for 15–20 min without any vortexing

Add the suspension solution dropwise to the coverslip

Incubate cells for 45 min to 3 h in 5% CO_2 incubator at 37 °C

Incubate cells with transfection medium pre-equilibrated in a 10% CO_2 incubator for 15–20 min. Note: the plate should be put back in a 5% CO_2 incubator

← key step - use acidic media to dissolve calcium precipitate in order to reduce calcium toxicity to neurons. however paper does not specify pH of wash media

Transfer the coverslip back to its original well containing original culture medium

Figure 1 | Flowchart of the main procedures of our Ca^{2+}-phosphate transfection protocol. XX represents a variable volume.

optimized calcium phosphate protocol

mix all components together but add HeBS to mixture last, pipette to mix

2.5 ug	DNA
3.125 uL	2M $CaCl_2$
xx uL	dH_2O (to 25uL)
25 uL	2x HeBS

incubate at room temperature for 5 minutes, then add 500 uL of MEM media (with 25uM AP5)

aspirate and save neuron media, wash coverslips once with MEM

add 550 uL of calcium phosphate/MEM mixture to coverslip

incubate for 15 to 60 minutes at 37°C, longer incubation may increase transfection efficiency

calcium precipitate should be visible after 5 minutes of incubation

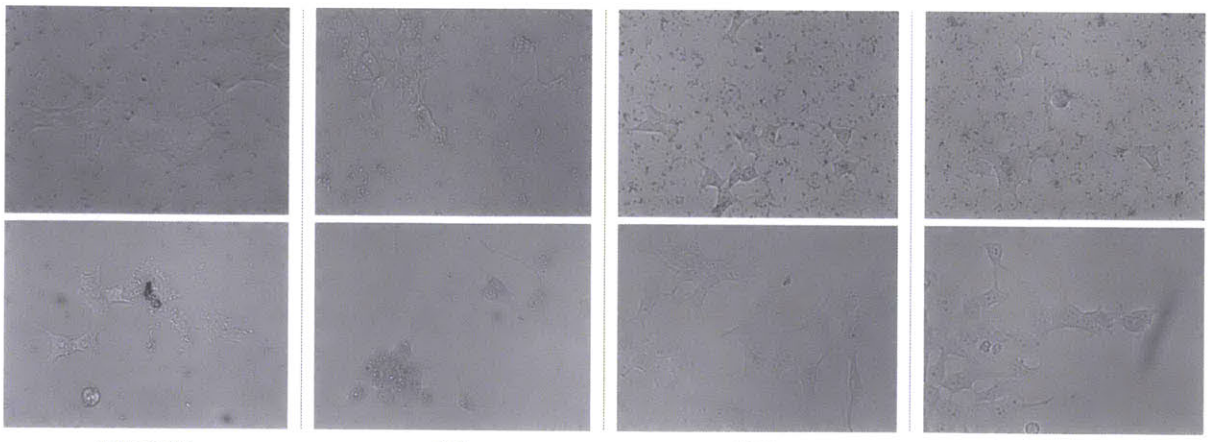
aspirate media after incubation period, add 500 uL of HMEM pH 6.5-6.7, incubate for 5 minutes at 37°C

precipitates should dissolve, if not, repeat the previous wash step with HMEM once more

aspirate media and add original neuron media back to coverslip

note: HMEM is 10mM HEPES buffered MEM
calibrated to indicated pH with acetic acid

Calcium Phosphate Neuron Transfection



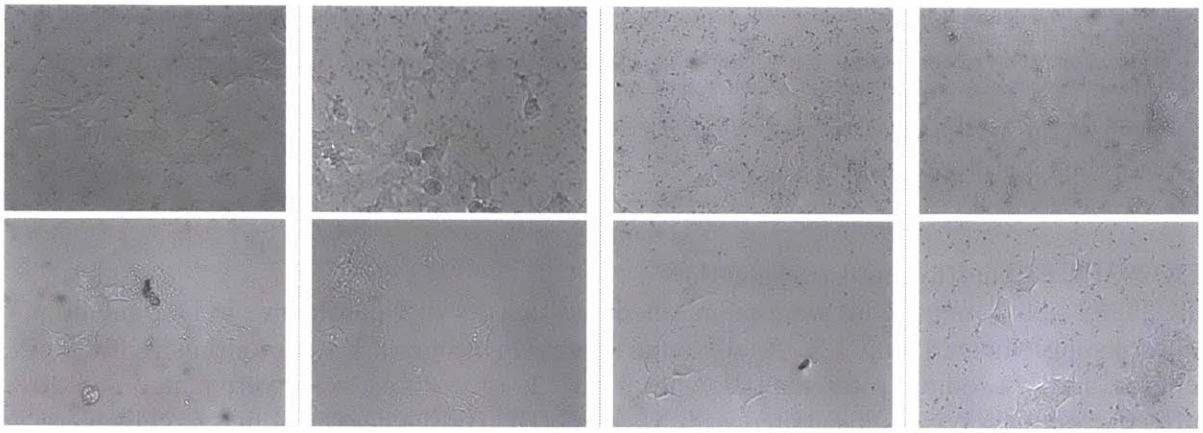
MEM unmodified
pH 7.54, osm 295
residual precipitate

HMEM
pH 4.87, osm 300
no precipitate, but dead cells

HMEM
pH 6.01
no precipitate, healthy cells

HMEM
pH 6.51
no precipitate, healthy cells

HMEM Wash pH Titration in HEK293FT
top row: calcium phosphate precipitation before wash; bottom row: after wash with indicated media



HMEM Wash pH Titration in HEK293FT
top row: calcium phosphate precipitation before wash; bottom row: after wash with indicated media

	[DNA] (ug/uL)	DNA/well (ug)	1X (1/24well)
DNA	0.155	2.0	12.9
dH_2O (total 15ul)			8.975
$CaCl_2$ (2M) (ul)			3.125
2x HeBS (ul)			25

Protocol

1 mix DNA, H2O, and CaCl2 in tube

2 remove neuron culture media from well and save it in 50 mL conical

3 wash neuron coverslips once with MEM (pH 7.1), then leave 500 uL of MEM (pH 7.1) in well

4 add HeBS to the DNA/CaCl2 tube and mix by pipetting, wait 30 seconds for precipitates to form

5 add CaPhos mix droplet wise to well and return plate to incubator for 20-60 minutes. Precipitates should appear within 5 minutes

6 after incubation period, remove media and wash once with acidic MEM (pH 6.8), then leave 1 mL of acidic MEM (pH 6.8) in well and return to incubator for 5 minutes

7 after incubation period, check to make sure precipitates have been dissolved. If not, repeat previous step, otherwise continue onto next step

8 remove media and wash once with acidic MEM (pH 6.8) to remove residual calcium, then add back the original neuron culture media to well

Note:

recipe is for 24 well plate, adjust volume accordingly for other plates
incubation time can vary from 20-60 minutes, depending on labeling density desired and cell viability. Generally 20-30 min. should be sufficient
incubation MEM pH can be between 7.1-7.3; acidic wash MEM pH should be between 6.7 to 6.9, generally more acidic media will yield faster resuspension of CaPhos precipitates

HEPES (or other buffers) should be added to stabilize MEM pH

generally 1-3 ug of DNA is enough for each well

HEK 293FT cell culture and transfection

HEK 293FT cells (Invitrogen) were maintained between 10-70% confluence in D10 medium (Cellgro) supplemented with 10% fetal bovine serum (Invitrogen), 1% penicillin/streptomycin (Cellgro), and 1% sodium pyruvate (Biowhittaker)). For recording, cells were plated at 5-20% confluence on glass coverslips coated with Matrigel (BD Biosciences). Adherent cells were transfected approximately 24 hours post-plating either with TransLT 293 lipofectamine transfection kits (Mirus) or with calcium phosphate transfection kits (Invitrogen), and recorded via whole-cell patch clamp between 36-72 hours post-transfection.

In vitro electrophysiology

Whole cell patch clamp recordings were made using a Axopatch 200B amplifier, a Digidata 1440 digitizer, and a PC running pClamp (Molecular Devices). Neurons were bathed in room temperature Tyrode containing 125 mM NaCl, 2 mM KCl, 3 mM CaCl2, 1 mM MgCl2, 10 mM HEPES, 30 mM glucose, 0.01 mM NBQX and 0.01 mM GABAzine. The Tyrode pH was adjusted to 7.3 with NaOH and the osmolarity was adjusted to 300 mOsm with sucrose. HEK cells were bathed in a Tyrode bath solution identical to that for neurons, but lacking GABAzine and NBQX. Borosilicate glass pipettes (Warner Instruments) with an outer diameter of 1.2 mm and a wall thickness of 0.255 mm were pulled to a resistance of 3-9 MΩ with a P-97 Flaming/Brown micropipette puller (Sutter Instruments) and filled with a solution containing 125 mM K-gluconate, 8 mM NaCl, 0.1 mM CaCl2, 0.6 mM MgCl2, 1 mM EGTA, 10 mM HEPES, 4 mM Mg-ATP, and 0.4 mM Na-GTP. The pipette solution pH was adjusted to 7.3 with KOH and the osmolarity was adjusted to 298 mOsm with sucrose. Access resistance was 5-30 MΩ, monitored throughout the voltage-clamp recording. Data was analyzed using Clampfit (Molecular Devices) and MATLAB (Mathworks, Inc.)

Ion Conductance Recording

Whole-cell patch clamp recordings were performed in isolated HEK293FT cells to accurately measure parameters from single cells. All recordings were performed using an Axopatch 200B amplifier and Digidata 1440 digitizer (Molecular Devices) at room temperature. In order to allow isolated cell recording, cells were plated at a lower density of 15,000 cells per well in 24-well plates that contained round glass coverslips (0.15 mm thick, 25 mm in diameter, coated with 2 % Growth Factor Reduced Matrigel in DMEM for 1 h at 37 ^{o}C). For most recordings, Tyrode was used as the extracellular solution, and the intracellular solution consisted of (in mM) 125 K-Gluconate, 8 NaCl, 0.1 $CaCl_2$, 0.6 $MgCl_2$, 1 EGTA, 10 HEPES, 4 MgATP, 0.4 NaGTP, pH 7.3 (KOH adjusted), with 295-300 mOsm (sucrose adjusted). Extracellular and intracellular solutions used for testing ion permeability is listed in Table S1.

Table S1. Compositions of solutions used in ion permeability experiments

Solution	[Na] (mM)	[K] (mM)	[Ca] (mM)	[H] (mM)	pH	Other
Intracellular	0	140	0	5.10E-05	7.4	5 mM EGTA, 2 mM MgCl2, 10 mM HEPES
145 mM NaCl	145	5	1	5.10E-05	7.4	10 mM HEPES, 5 mM glucose, 2 mM MgCl2
145 mM KCl	0	145	1	5.10E-05	7.4	10 mM HEPES, 5 mM glucose, 2 mM MgCl2
90 mM CaCl2	0	5	91	5.10E-05	7.4	10 mM HEPES, 5 mM glucose, 2 mM MgCl2
5 mM NaCl	5	5	1	5.10E-04	6.4	135 mM NMDG, 10 mM HEPES, 5 mM glucose, 2 mM MgCl2

Liquid junction potentials were measured as previously described [16] to be 5.8 mV for the 90 mM $CaCl_2$ and 4.9 mV for the 5 mM NaCl extracellular solutions, which were corrected during recording; the others were < 1 mV in junction potential.

In all patch clamp recordings, a stringent cutoff of access resistance less than 25 MΩ and holding current less than ± 50 pA was applied in order to ensure accurate measurement. Typical membrane resistance was between 500 MΩ - 2 GΩ and pipette resistance was between 4 - 10 MΩ.

Photostimulation of patch clamped cells was conducted by a 470 nm LED (Thorlabs) at 10 mW/mm^2 unless otherwise stated. For most experiments, 1s illumination was delivered to measure transient and steady-state photocurrents.

In utero electroporation

All procedures were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and approved by the Massachusetts Institute of Technology Committee on Animal Care. C57BL/6J E16-timed pregnant mice were used for electroporation. Surgery was done under ketamine-xylazine anesthesia and buprenorphine analgesia, DNA solution containing plasmids of interest were injected into lateral ventricle of each embryo using a pulled capillary tube. Five square pulses (50ms width, 1Hz, 35V) were applied using tweezer electrode for electroporation.

Slice preparation

P20–P40 mice were used for slice preparation. Mice were anesthetized with isofluorane and transcardialy perfused with ice-cold cutting solution containing 110 mM choline chloride, 25 mM $NaHCO_3$, 25 mM D-glucose, 11.6 mM sodium ascorbate, 7 mM $MgCl_2$, 3.1 mM sodium pyruvate, 2.5 mM KCl, 1.25 mM NaH_2PO_4 and 0.5 mM $CaCl_2$. The brain was then carefully removed and mounted in vibrating blade microtome (Leica VT1000S). 300-mm thick coronal slices of the visual cortex were cut with a vibrating metal blade at 90 Hz and 0.1 mm/s cutting speed. Sectioned slices were incubated in 37C cutting solution for 30-45 minutes before transfer to room temperature oxygenated artificial cerebrospinal fluid (ACSF) for recording. ACSF

contained 127 mM NaCl, 2.5 mM KCl, 25 mM $NaHCO_3$, 1.25 mM NaH_2PO_4, 12 mM D-glucose, 0.4 mM sodium ascorbate, 2 mM $CaCl_2$, and 1 mM $MgCl_2$.

Optics

For in vitro culture experiments all illuminations were done with (Thorlabs) LEDs, but separate (Thorlabs) LEDs were used for imaging versus opsin excitation. For slice experiments (Thorlabs) LEDs were used for opsin illumination but we used mercury lamp for visualizing opsin expression. It was necessary to separate opsin excitation light source from imaging light source in order to prevent opsin excitation LED power rundown. We did not observe any power drifts over the course of at least six months.

Illumination intensity was modulated by analog voltage input to the LED controllers. Due to the two log illumination intensity dynamic range that must be spanned and the low illumination intensity used for blue wavelengths, the analog voltage input often ranged from 0.05 to 0.5 volts (V). For these low voltage inputs, it was necessary to increase the input pulse length in order to obtain the desired output light pulse width. For example, to obtain 5 ms light pulse from 0.1 V input, we had to increase input pulse width to 6.5 ms and there is correspondingly a 1.5 ms delay in light onset. These are idiosyncrasies of the (Thorlabs) LED controllers and we adjusted timing values appropriately in all datasets shown.

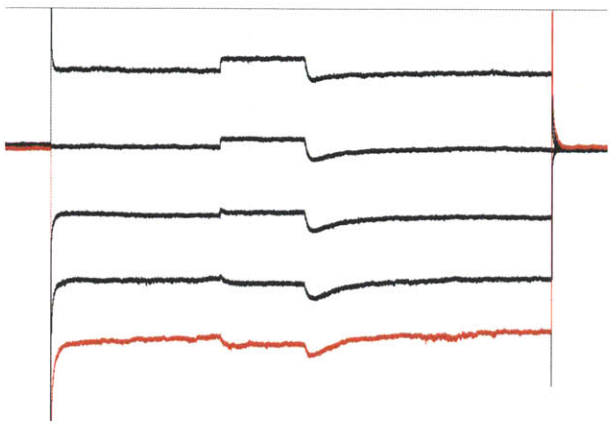
We measured illumination spot size by focusing on a microscope slide coated with appropriate wavelength Alexa dye and photobleaching it for 10 minutes under full intensity illumination. We imaged the photobleached microscope slides along with a micrometer calibration slide to determine the pixel to micrometer conversion. We drew a circle around the clearly photobleached region in an image editor to approximate the illumination spot size.

Action spectrum data was taken with a monochromator (Till-Photonics Polychrome IV). Due to brief white light leakage during wavelength transition (undocumented monochromator design fault), we had to use a separate (Uniblitz) shutter to gate the light pulses. We used approximately equal photon flux across wavelengths for illumination (at most 12% difference from the target photon flux) by using the monochromator's built-in intensity adjustment. We wrote a (QuickMacros) script to automate the wavelength and intensity selections and synchronize with electrophysiology recording.

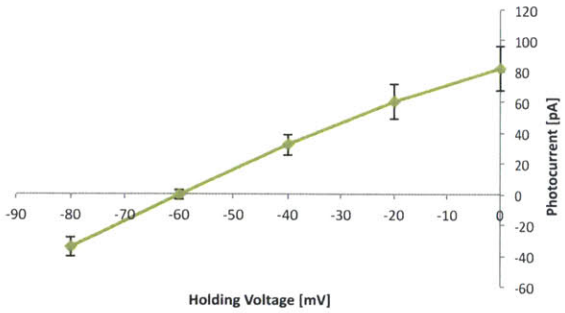
Miscellaneous data

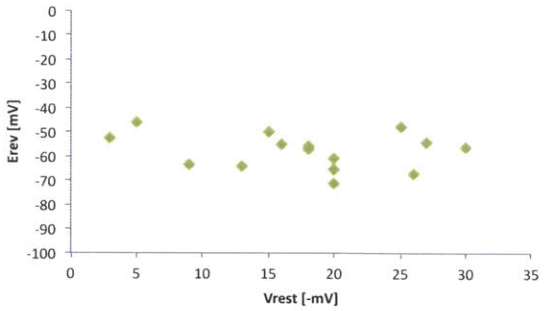
Bizarro

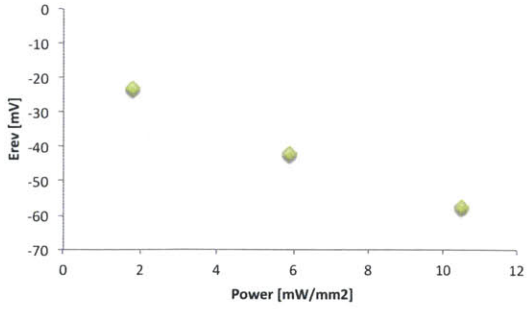
Bizzaro is the W96Y mutation of ArchT that shows bidirectional activity. An example HEK293FT cell voltage clamped trace from -120 to 0 mV in +30 mV increment is shown below.

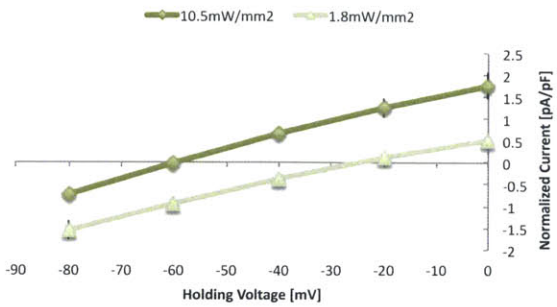


The steady state response is consistent but the transient on and off responses can vary greatly. We can tune the effective steady state reversal potential from -20 to -60 mV by modulating the illumination light density and the reversal is independent of the resting potential of the cell. The data shown below are for HEK293FT cell under green light (543 nm) illumination, but this effect can occur under all wavelength from UV to red although the absolute reversal versus light density relationship is shifted.









We have also done all ion replacement measurements (except for protons) and this reversal shift effect is independent of all ions. Ernst Bamberg's lab has previously demonstrated

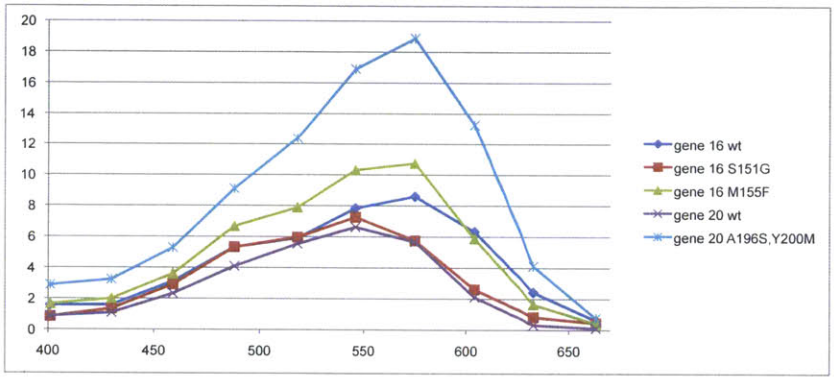
proteorhodopsins can have variable vectoriality, but only under different pH conditions (Lorinczi E et al 2009 J Mol Biol., Friedrich T. et. Al. 2002 J Mol Biol.). Thus bizarro is the first example of an opsin that may have variable vectoriality under physiological conditions and can be tuned to have different effective reversal potentials. One of the curious features of bizarro is that the light off transient current is always depolarizing with a long current decay constant (significantly longer than membrane time constants), similar to a channelrhodopsin. Previously Bamberg's lab has also shown that ChR2 may be a leaky proton pump (Feldbauer K. et. Al. 2009 PNAS). Thus bizarro may also exhibit leaky pump behavior.

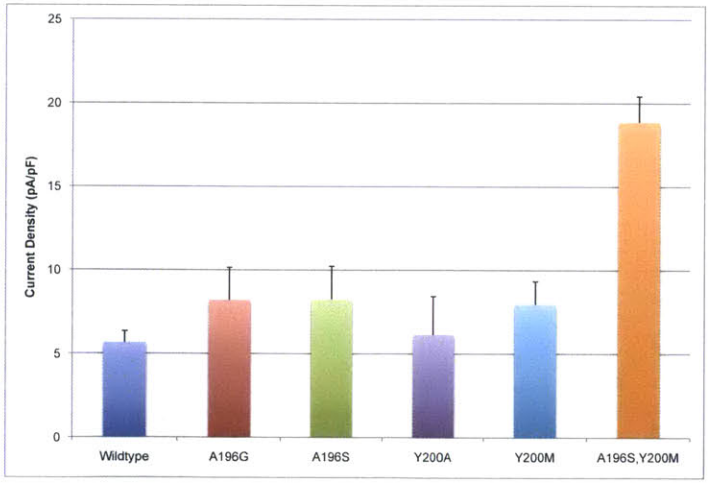
We tried many other W96 mutants and found that even W96F does not exhibit any bidirectional behavior. The difference between phenylalanine and tyrosine is only an hydroxyl group on the aromatic ring, so it appears the hydroxyl group is somehow facilitating in inward pump or channel activity. We also tested the W96Y mutation on other archaerhodopsins such as Arch, AR1, and AR2 (NCBI ID P96787, P69051, P29563), and they all exhibited similar bidirectional behavior although the light density versus effective reversal relationship was shifted. We tested the W96Y analog mutation on other bacteriorhodopsins and found that none of them exhibited this bidirectional behavior.

Bizarro is not very practical to use due to its light density dependence, but its unique biophysical properties is worth further investigation. It may shed light on the evolution of opsin pumps and channels.

Mac

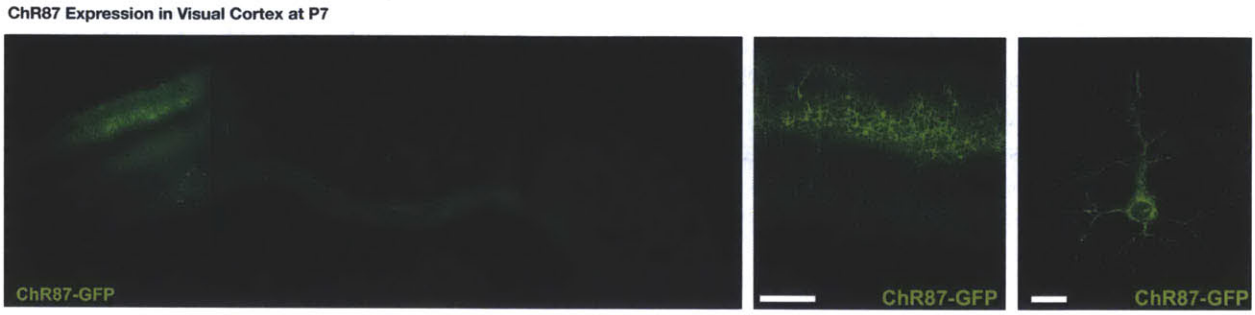
We screened many mutants of Mac based on existing literature (primarily work done by Har Gobind Khorana's lab) and found couple residues that showed minor but statistically insignificant improvements. However when we tried double mutants the current density improved almost four fold, which is significantly higher than the wildtype and is nonlinear combination of the individual mutants. The double mutant is also spectrally red-shifted to a peak of around 570 nm. The action spectra are not equal photon flux in this case, but merely the max light power from the monochromator at each wavelength. We numerically adjusted for the light power differences by scaling the resulting current response, although this may be greatly biased as scaling factors can be as high as 700%. So shifts should only be interpreted as relative. We performed limited testing in neuron culture, but the double mutant did not show significantly higher current density than wildtype due to trafficking.





Gene87 Histology

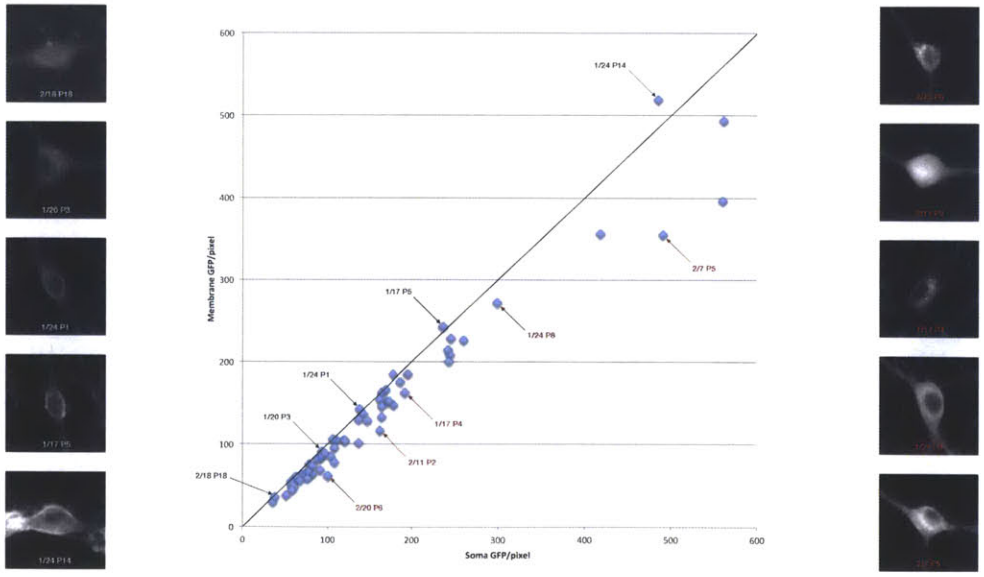
Gene87 is the best membrane trafficking channelrhodopsin we have found. Below is an histology showing clear corpus callosum expression of Gene87 in P7 mice. Although gene87 does not possess unique biophysical properties, its excellent trafficking (and low toxicity based on neuron resting potential) makes it an ideal candidate for applications where early expression is needed (e.g. development studies) or where low toxicity is required (e.g. stem cells).



in utero electroporation of pCAG-ChR87-GFP in left hemisphere at E15, imaged at P7; scale bar = 200 um (middle) and 20 um (right)

Quantifying fluorescence

We took quantitative images of opsin-GFP fluorescence on a conventional epifluorescence microscope. Despite the lack of z-resolution, it was possible to visually discern membrane trafficking. Below is an analysis of averaged soma intensity versus averaged membrane intensity. Off-diagonal cells generally exhibited poor membrane localization as expected. We did not include the membrane GFP analysis in the main dataset though since many opsins did not have good membrane expression.

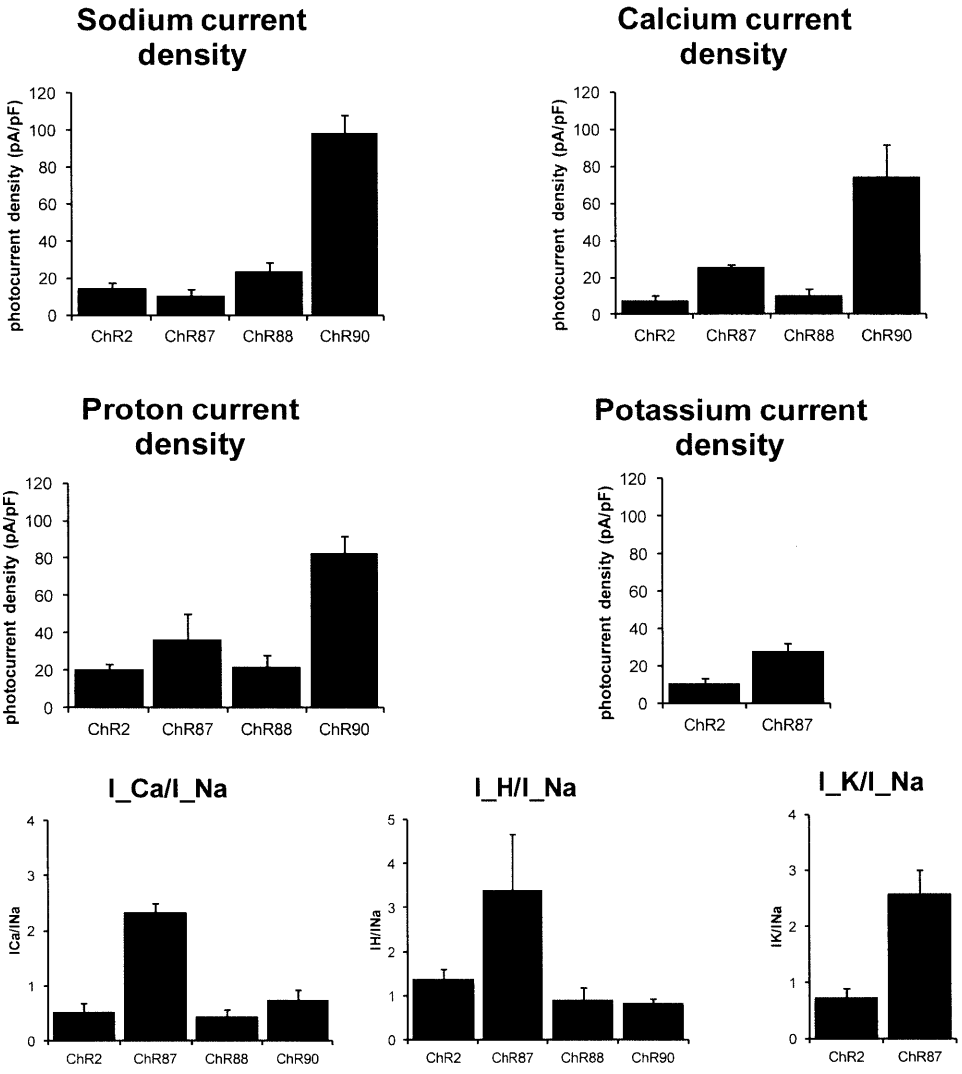


GFP along and above the diagonal typically have good membrane expression or dim uniform expression. GFP below the diagonal have intracellular aggregates or significant cytosolic expression. note: image intensity scale is different in each picture for ease of visualization

Soma vs Membrane GFP

Ion selectivity

Data taken in HEK293FT cells courtesy of Yongku Cho. See methods for detail. In general we did not find any channelrhodopsins that were truly selective for any ion and all of them had a significant proton current component.



Bibliography

1. Oesterhelt, D. & Stoeckenius, W. Functions of a new photoreceptor membrane. *Proceedings of the National Academy of Sciences of the United States of America* **70**, 2853-2857 (1973).
2. Beja, O. et al. Bacterial rhodopsin: Evidence for a new type of phototrophy in the sea. *Science* **289**, 1902-1906 (2000).
3. Brown, L.S., Dioumaev, A.K., Lanyi, J.K., Spudich, E.N. & Spudich, J.L. Photochemical reaction cycle and proton transfers in neurospora rhodopsin. *Journal of Biological Chemistry* **276**, 32495-32505 (2001).
4. Harz, H. & Hegemann, P. Rhodopsin-Regulated Calcium Currents in Chlamydomonas. *Nature* **351**, 489-491 (1991).
5. Danon, A. & Stoeckenius, W. Photophosphorylation in Halobacterium halobium. *Proceedings of the National Academy of Sciences of the United States of America* **71**, 1234-1238 (1974).
6. Sineshchekov, O.A., Jung, K.H. & Spudich, J.L. Two rhodopsins mediate phototaxis to low- and high-intensity light in Chlamydomonas reinhardtii. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 8689-8694 (2002).
7. Lanyi, J.K. Bacteriorhodopsin. *Annual review of physiology* **66**, 665-688 (2004).
8. Kuhlbrandt, W. Bacteriorhodopsin--the movie. *Nature* **406**, 569-570 (2000).
9. Nagel, G. et al. Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 13940-13945 (2003).
10. Bamann, C., Kirsch, T., Nagel, G. & Bamberg, E. Spectral characteristics of the photocycle of channelrhodopsin-2 and its implication for channel function. *Journal of molecular biology* **375**, 686-694 (2008).
11. Kato, H.E. et al. Crystal structure of the channelrhodopsin light-gated cation channel. *Nature* **482**, 369-374 (2012).
12. Gunaydin, L.A. et al. Ultrafast optogenetic control. *Nature neuroscience* **13**, 387-392 (2010).
13. Yizhar, O. et al. Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature* **477**, 171-178 (2011).
14. Berndt, A., Yizhar, O., Gunaydin, L.A., Hegemann, P. & Deisseroth, K. Bi-stable neural state switches. *Nature neuroscience* **12**, 229-234 (2009).
15. Mattis, J. et al. Principles for applying optogenetic tools derived from direct comparative analysis of microbial opsins. *Nature methods* **9**, 159-172 (2012).
16. Berndt, A. et al. High-efficiency channelrhodopsins for fast neuronal stimulation at low light levels. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 7595-7600 (2011).
17. Kleinlogel, S. et al. Ultra light-sensitive and fast neuronal activation with the Ca(2)+-permeable channelrhodopsin CatCh. *Nature neuroscience* **14**, 513-518 (2011).
18. Nagel, G. et al. Light activation of channelrhodopsin-2 in excitable cells of Caenorhabditis elegans triggers rapid behavioral responses. *Current biology : CB* **15**, 2279-2284 (2005).
19. Nagel, G. et al. Channelrhodopsin-1: a light-gated proton channel in green algae. *Science* **296**, 2395-2398 (2002).

20. Boyden, E.S., Zhang, F., Bamberg, E., Nagel, G. & Deisseroth, K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nature neuroscience* **8**, 1263-1268 (2005).
21. Zhang, F. et al. Multimodal fast optical interrogation of neural circuitry. *Nature* **446**, 633-639 (2007).
22. Sineshchekov, O.A., Govorunova, E.G., Wang, J. & Spudich, J.L. Enhancement of the long-wavelength sensitivity of optogenetic microbial rhodopsins by 3,4-dehydroretinal. *Biochemistry* **51**, 4499-4506 (2012).
23. Zhang, Y.P. & Oertner, T.G. Optical induction of synaptic plasticity using a light-sensitive channel. *Nature methods* **4**, 139-141 (2007).
24. Chow, B.Y. et al. High-performance genetically targetable optical neural silencing by light-driven proton pumps. *Nature* **463**, 98-102 (2010).
25. Gradinaru, V. et al. Molecular and cellular approaches for diversifying and extending optogenetics. *Cell* **141**, 154-165 (2010).
26. Madisen, L. et al. A toolbox of Cre-dependent optogenetic transgenic mice for light-induced activation and silencing. *Nature neuroscience* **15**, 793-802 (2012).
27. Ma, D. & Jan, L.Y. ER transport signals and trafficking of potassium channels and receptors. *Current opinion in neurobiology* **12**, 287-292 (2002).
28. Stockklausner, C., Ludwig, J., Ruppersberg, J.P. & Klocker, N. A sequence motif responsible for ER export and surface expression of Kir2.0 inward rectifier K(+) channels. *FEBS letters* **493**, 129-133 (2001).
29. Han, X., Qian, X., Stern, P., Chuong, A.S. & Boyden, E.S. Informational lesions: optical perturbation of spike timing and neural synchrony via microbial opsin gene fusions. *Frontiers in molecular neuroscience* **2**, 12 (2009).
30. Kleinlogel, S. et al. A gene-fusion strategy for stoichiometric and co-localized expression of light-gated membrane proteins. *Nature methods* **8**, 1083-1088 (2011).
31. Tang, W. et al. Faithful expression of multiple proteins via 2A-peptide self-processing: a versatile and reliable method for manipulating brain circuits. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **29**, 8621-8629 (2009).
32. Szymczak, A.L. et al. Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nature biotechnology* **22**, 589-594 (2004).
33. Lin, J.Y., Lin, M.Z., Steinbach, P. & Tsien, R.Y. Characterization of engineered channelrhodopsin variants with improved properties and kinetics. *Biophysical journal* **96**, 1803-1814 (2009).
34. Zhang, F. et al. Red-shifted optogenetic excitation: a tool for fast neural control derived from Volvox carteri. *Nature neuroscience* **11**, 631-633 (2008).
35. Prigge, M. et al. Color-tuned channelrhodopsins for multiwavelength optogenetics. *The Journal of biological chemistry* **287**, 31804-31812 (2012).
36. Greenberg, K.P., Pham, A. & Werblin, F.S. Differential targeting of optical neuromodulators to ganglion cell soma and dendrites allows dynamic control of center-surround antagonism. *Neuron* **69**, 713-720 (2011).
37. Grubb, M.S. & Burrone, J. Channelrhodopsin-2 localised to the axon initial segment. *PloS one* **5**, e13761 (2010).
38. Lewis, T.L., Jr., Mao, T., Svoboda, K. & Arnold, D.B. Myosin-dependent targeting of transmembrane proteins to neuronal dendrites. *Nature neuroscience* **12**, 568-576 (2009).

39. Zhao, S. et al. Cell type-specific channelrhodopsin-2 transgenic mice for optogenetic dissection of neural circuitry function. *Nature methods* **8**, 745-752 (2011).
40. Petreanu, L., Huber, D., Sobczyk, A. & Svoboda, K. Channelrhodopsin-2-assisted circuit mapping of long-range callosal projections. *Nature neuroscience* **10**, 663-668 (2007).
41. Petreanu, L., Mao, T., Sternson, S.M. & Svoboda, K. The subcellular organization of neocortical excitatory connections. *Nature* **457**, 1142-1145 (2009).
42. Froemke, R.C. & Dan, Y. Spike-timing-dependent synaptic modification induced by natural spike trains. *Nature* **416**, 433-438 (2002).
43. Zorzos, A.N., Scholvin, J., Boyden, E.S. & Fonstad, C.G. Three-dimensional multiwaveguide probe array for light delivery to distributed brain circuits. *Optics letters* **37**, 4841-4843 (2012).