Support The Information of the I Martinson et al. 10.1073/pnas.1321716111

SI Methods

Identification, Sequence Assembly, and Cloning of Erg Channel Orthologs. Trichoplax adhaerens, Daphnia pulex, and Caenorhabiditis briggsae Ether-a-go-go related gene (Erg) channels were identified using TBLASTN searches of genome drafts (1–4), and existing gene predictions were used in sequence and phylogenetic analyses. A combination of mammalian, Drosophila, and nematode Erg channel amino acid sequences were used for queries. Searches routinely identified Erg, Elk, Ether-a-go-go (Eag), HCN, and CNG channels. However, only true Erg family orthologs were reciprocal best matches to previously cloned Erg channels in BLASTP queries against REFSEQ.

Total Nematostella RNA was prepared from 1- to 2-mo-old polyps using the RNAeasy Miniprep Kit (Qiagen). Before lysis, whole polyps were cut into small cubes, preserved in RNAlater (Qiagen), frozen at –80 °C, and ground with a mortar and pestle. cDNA was reverse transcribed from total RNA using oligo dT priming and SuperScript III reverse transcriptase (Life Technologies). Nematostella vectensis Erg1 (NvErg1) and NvErg4 were amplified in fragments, and full-length clones were assembled using overlap extension. Three independent clones of each gene were sequenced and compared with the genome sequence to generate a consensus coding sequence reported in [Dataset S1](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1321716111/-/DCSupplemental/sd01.xlsx). Sequences have been submitted to GenBank (NvErg1, KF877721; NvErg4, KF877722). No splice variants or nonsynonymous mutations were identified. Down and out (DAO), Drosophila Erg (DmErg), Caenorhabditis elegans Erg ortholog (CeErg), and Anopheles gambiae Erg (AgErg) were cloned by RT-PCR from total RNA samples prepared from adult Canton-S flies, mixed stage worms, and adult mosquitos. The 3′ end of AgErg coding was determined by RACE PCR, and there are several small splicing differences in the clone we amplified and informatics gene predictions. The experimentally determined sequence is included in [Dataset S1](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1321716111/-/DCSupplemental/sd01.xlsx). Mouse Erg3 (MmErg3) was cloned by RT-PCR from total RNA isolated from mouse cortex. All clones were sequence verified.

Electrophysiology: Recording and Data Analysis. All chemicals described below were obtained from Sigma Aldrich. Xenopus laevis ovaries were obtained from eNASCO, and mature de-folliculated oocytes were prepared by digestion with Type II Collagenase diluted to $1-2$ mg/mL in Ca^{2+} -free ND98 (98 mM NaCl, 2 mM KCl , 1 mM $MgCl_2$, 5 mM Hepes, pH 7.2). Following digestion, oocytes were rinsed extensively in ND98 culture solution (98 mM NaCl, 2 mM KCl, 1 mM MgCl₂, 2 mM CaCl₂, 2.5 mM Na-Pyruvate, 100 U/mL penicillin, 100 μg/mL streptomycin, 5 mM Hepes, pH 7.2). Oocytes were maintained in this culture solution until recording.

Channel coding sequences were cloned into the pOX expression vector (5), and transcripts were prepared by runoff transcription using the T3 mMessage mMachine kit (Life Technologies). Transcripts were precipitated in 2.5 M LiCl, and pellets were extensively rinsed in cold 70% (vol/vol) ethanol and resuspended in nuclease-free water supplemented with RNase Inhibitor (Superasin, Life Technologies). RNA was diluted to empirically determine optimal concentrations $(2-50 \text{ ng/µL})$ and injected into oocytes in a 50 nL volume with a Nanoject II injection system (Drummond Scientific). Oocytes were incubated at 18 °C for 1–3 d before recording.

Recordings were carried out under constant perfusion of a low Cl– recording solution (see Methods). Glass capillary electrodes were filled with 3 M KCl and had tip resistances or 0.4–1 MΩ.

Bath electrodes were placed in 1 M NaCl and isolated with an agar bridge. A Dagan CA-1B was used in two electrode voltage clamp mode for recordings, and data were collected with the pClamp 10 Acquisition Suite (Molecular Devices). Currents were sampled at 10 KHz and filtered at 2 KHz. Currents were not subtracted for leak, but capacitive transients were partially compensated. Voltage activation data and exponential fits of inactivation and deactivation time courses were generated in Clampfit (Molecular Devices). Single exponentials were used to fit inactivation and some tail currents (deactivation) using the equation $I_t = I_i + Ae^{-t/\tau}$, where (I_t) is current at time t, I_i is initial current level, A is amplitude, and τ is the time constant. A subset of tail currents fit more accurately with a double exponential function $(I_t = I_i + A_1 e^{-t/\tau} + A_2 e^{-t/\tau}$. In these cases, the fastest exponential accounted for greater than 65% of the current amplitude and is reported here. Boltzmann fits of activation data were performed in Origin (Originlab). Data from individual cells were fit with a single Boltzmann distribution, $f(V) = (A_1 - A_2)/(1 + e^{(V - V50)/s}) + A_2$, where V is the voltage, V_{50} is the midpoint voltage, s is the slope factor, and A_1 and A_2 are the minimum and maximum asymptotes, respectively. Data were normalized before averaging, and normalized Boltzmann fits shown in figures were generated with the arithmetic means of individual V_{50} and s measurements.

HsErg1 and DmErg Gating Models. Gating models for human Erg1 (HsErg1) and DmErg were based on a gating scheme developed to describe Shaker channel activation (6). In the model, each of the four voltage sensors undergoes two independent activation steps before a final concerted channel opening. We added an inactivated state accessible only from the open state for the HsErg1 model. Rate constant parameters were estimated by fitting the gating models to families of outward currents in IChMascot (7). Currents were recorded in response to 1 s depolarizations ranging in 10 mV steps from -110 to $+50$ mV, with 400 ms tail currents recorded at –70 mV. Once fitting had captured the salient features of each gating phenotype, fitting was stopped and rate constant parameters were recorded. The purpose of this exercise was solely to develop gating models that could demonstrate the disparate responses of I_{Kr} -like (HsErg1) and delayed rectifier-like (DmErg) channels to plateau depolarization. No further attempts were made to optimize fit, and we did not attempt to constrain individual gating parameters with additional electrophysiological measurement. Current responses of the HsErg1 and DmErg gating models to a simplified plateau action potential were then calculated in IonChannelLab (7).

Sequence Analysis. Multiple sequence alignments of the eag domain (22 taxa from Eag, Elk, and Erg subfamilies) and cyclic nucleotide-binding homology domain (CNBHD) (26 taxa from Eag, Elk, and Erg subfamilies) were generated in Jalview v2.8 (8) by T-Coffee (9) and Muscle (10), respectively, under default settings and manually edited. Pairwise estimates of evolutionary divergence between sequences were calculated in MEGA5 (11) using the number of amino acid differences per site. Ambiguous positions were removed for each sequence pair. A distance matrix was constructed for the entire eag domain using 192 positions. Separate distance matrices were constructed for 27 predicted eaginteracting CNBHD residues (labeled in Fig. S4) and noninteracting CNBHD residues (118 positions), as defined in a cocrystal structure of the Eag1 eag and CNBHD domains (12). Based on these sequence distance matrices, we quantified the mean \pm SEM for multiple pairwise comparisons between or within various

sequence groups. In-group or between-group comparisons were statistically evaluated with an unpaired two-tailed t test.

Secondary Structure Prediction Per–Arndt–Sim Domain. The amino acid secondary structure was predicted using the algorithm JNet implemented by Jalview v2.8 (8, 13). JNet uses several different neural networks and decides on the most likely prediction via a jury network. Alpha helical and beta sheet regions present in the consensus prediction are marked in Fig. 5A.

Phylogeny Construction. Phylogenetic reconstruction was limited to the core transmembrane K^+ channel motif because we wanted to determine the position of Erg channels that have

- 1. Holt RA, et al. (2002) The genome sequence of the malaria mosquito Anopheles gambiae. Science 298(5591):129–149.
- 2. Stein LD, et al. (2003) The genome sequence of Caenorhabditis briggsae: A platform for comparative genomics. PLoS Biol 1(2):E45.
- 3. Colbourne JK, et al. (2011) The ecoresponsive genome of Daphnia pulex. Science 331(6017):555–561.
- 4. Srivastava M, et al. (2008) The Trichoplax genome and the nature of placozoans. Nature 454(7207):955–960.
- 5. Jegla T, Salkoff L (1997) A novel subunit for shal K+ channels radically alters activation and inactivation. J Neurosci 17(1):32–44.
- 6. Ledwell JL, Aldrich RW (1999) Mutations in the S4 region isolate the final voltagedependent cooperative step in potassium channel activation. J Gen Physiol 113(3): 389–414.
- 7. Santiago-Castillo JA, Covarrubias M, Sánchez-Rodríguez JE, Perez-Cornejo P, Arreola J (2010) Simulating complex ion channel kinetics with IonChannelLab. Channels (Austin) 4(5):422–428.

lost the eag domain independently of the domain itself. Because the eag domain interacts with the channel through the CNBHD domain, both domains were removed to avoid bias. Sequence regions used are provided in [Dataset S3](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1321716111/-/DCSupplemental/sd03.rtf). Sequences were aligned in MEGA5 using CLUSTALW and trimmed to remove areas of length polymorphism. MrBayes (14) was used to construct a Bayesian inference phylogeny under a mixed model using the following conditions: two independent runs of four chains were run for 1,000,000 generations. Trees were sampled every 500 generations (2,000 trees per run), and the first 25% were discarded. The phylogeny is based on a consensus of the two runs.

- 8. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ (2009) Jalview Version 2-A multiple sequence alignment editor and analysis workbench. Bioinformatics 25(9):1189–1191.
- 9. Notredame C, Higgins DG, Heringa J (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol 302(1):205–217.
- 10. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32(5):1792–1797.
- 11. Tamura K, et al. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28(10):2731–2739.
- 12. Haitin Y, Carlson AE, Zagotta WN (2013) The structural mechanism of KCNH-channel regulation by the eag domain. Nature 501(7467):444–448.
- 13. Cole C, Barber JD, Barton GJ (2008) The Jpred 3 secondary structure prediction server. Nucleic Acids Res 36(Web Server issue):W197–201.
- 14. Ronquist F, et al. (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61(3):539–542.

Fig. S1. DmErg and HsErg1 gating models. (A) Gating model showing transitions from a resting closed channel (upper left) to open and inactivated channels (lower right). The center circle represents the pore domain in closed (solid red), open (green outline), and inactivated (solid black with green outline) conformations. Four independently moving voltage sensors (one from each subunit) are depicted as circles surrounding the pore in resting (black), intermediate (blue), and activated (green) conformations. Following independent activation of all four voltage sensors, the pore opens. An inactivated state, accessible only from the open state, was included for HsErg1, but not DmErg. Labels at the right margin highlight closed, open, and inactivated channel conformations. Rates for transitions are labeled and rate constants for the DmErg and HsErg1 models are listed in the table inset. Forward rate constants are given for the following transitions: resting to intermediate state of the voltage sensor (α), intermediate to activated state of the voltage sensor (γ), channel opening (O), and channel inactivation (I). The reverse conformational transitions are recovery from inactivation (R), channel closing (C), activated to intermediate state of the voltage sensor (δ), and return from intermediate to resting voltage sensor (β). (B) Simulated K⁺ currents elicited in response to 1 s voltage steps ranging from –100 mV to +40 mV in 10 mV increments for the DmErg and HsErg1 models. The holding potential was –100 mV, and 400 ms tails were elicited at –70 mV. The models capture the key gating features of both channels: subthreshold activation and rapid deactivation for DmErg and I_{Kr}-like rectification and current rebound for HsErg1.

Fig. S2. MmErg3 (Left) and NvErg1 (Right) currents are not altered by coexpression with DAO. Control currents are shown in black, and +DAO currents are shown in red. Currents were recorded in response to a 1 s +50 mV step followed by a 400 ms step to –70 mV and are normalized in amplitude for comparison. The holding voltage was –100 mV, and we used the same amount of DAO RNA that facilitated functional expression of DmErg and CeErg. DAO did not affect the I_{Kr} phenotype and did not significantly alter the current amplitude ($n = 4$ for each).

VAS.

V
A

Fig. S3. Sequence alignment of the eag domain of 12 Erg family, six Elk family, and four Eag family K⁺ channels. Channel names are given at the left margin, and amino acid positions are indicated after the slash. Species prefix abbreviations are as follows: Ag, A. gambiae; Dm, Drosophila melanogaster; Hs, Homo sapiens; Mm, Mus musculus; Nv, N. vectensis; Ta, T. adhaerens. Residue positions reaching 65% identity and similarity (BLOSUM62) are boxed shaded in blue or yellow, respectively. Gaps are indicated with dashes, and an alignment ruler is provided at the top margin. Note two insertions and reduced conservation in the DmErg sequence. CeErg, CbErg, and NvErgs2–5 produced no significant alignment and were omitted.

Fig. S4. Distance matrix for the eag domain alignment. Numbers indicate amino acid substitutions per site for the indicated pairwise sequence comparisons. Color codes highlight degree of conservation (min = 0, red; max = 0.9, blue; average= 0.49). Taxon labels and numbers are given in the left columns, and taxon numbers are repeated in the top row. Note the reduced conservation in DmErg.

PNAS PNAS

Fig. S5. Sequence alignment of the CNBHD of 18 Erg family, six Elk family, and four Eag family K⁺ channels. Channel names and amino acid ranges are given at the left margin, and a ruler is included above the alignment blocks. Species prefix abbreviations are as follows: Ag, A. gambiae; Cb, C. briggsae; Ce, C. elegans; Dm, D. melanogaster; Hs, H. sapiens; Mm, M. musculus; Nv, N. vectensis; Ta, T. adhaerens. Residue shading indicates >65% identity (blue) or similarity (yellow) (BLOSUM62). Dashes indicate gaps. Red dots below the alignment indicate residues that interface with the eag domain in a crystal structure of Eag1 (1) .

1. Haitin Y, Carlson AE, Zagotta WN (2013) The structural mechanism of KCNH-channel regulation by the eag domain. Nature 501(7467):444-448.

Fig. S6. Distance matrices for the CNBHD alignment. Separate matrices are provided for 27 CNBHD positions predicted to interface with the eag domain (red dots in Fig. S5) (A) and 118 noninterface positions (B). Numbers indicate amino acid substitutions per site in the indicated pairwise comparisons. Taxon names and numbers are given in the left columns, and numbers are repeated in the top row. Sequence distance is color coded from red (max) to blue (min): interface matrix, max = 0, min = 0.85, average = 0.48; noninterface matrix, max = 0, min = 0.82, average = 0.52.

NAS PNAS

Movie S1. A 46-s video of wave contractions in a several-wk-old Nematostella polyp. Frames were taken every 0.45 s, and the video is set to run at 10x speed. The video frame is 3.24×3.8 mm.

[Movie S1](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1321716111/-/DCSupplemental/sm01.avi)

N
A
M

 $\tilde{\mathbf{A}}$

Dataset S1. The Excel table provides the amino acid sequence of each EAG superfamily channel examined in this study. Channels are grouped by species, and columns (in order) provide the name used in this manuscript, gene ID, protein ID, and the actual amino acid sequence used. In some cases, the sequence used here differs slightly from the protein ID due to empirical evidence or refined prediction

[Dataset S1](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1321716111/-/DCSupplemental/sd01.xlsx)

Dataset S2. Verified DNA coding sequences are given for NvErg1, NvErg4, and AgErg. GenBank accession numbers are KF877721, KF877722, and KJ493818, respectively

[Dataset S2](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1321716111/-/DCSupplemental/sd02.rtf)

Dataset S3. FASTA format file of sequences used in the phylogeny presented in Fig. 6. The phylogeny was built on the core K+ channel motif, which includes the voltage sensor and pore. Names of sequences are the same as given in Fig. 6

[Dataset S3](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1321716111/-/DCSupplemental/sd03.rtf)