Supplemental Methods

Data Collection

Human or mouse sequences were used as queries to search for orthologs of NALCN and *Saccharomyces cerevisiae* queries were used to search for orthologs of fungal calcium channels. We used BLASTp to search NCBI's non-redundant protein database, the Joint Genome Institute's genomes, or the Origins of Multicellularity protein database (Altschul et al. 1997). Putative orthologs were reciprocally BLASTed into the genome from which the original query came to verify strict orthology between subject and query.

Alignment

We used the GUIDANCE server (with the MAFFT option) to make alignments and to prune the alignments of the most unreliable columns, leaving 50% of the columns (Katoh et al. 2005; Penn et al. 2010). The alignment of one sequence from the apusozoan *Thecamonas trahens* was found by GUIDANCE to be unstable, and was discarded. Since GUIDANCE often leaves areas with a high proportion of indels, we also removed columns that were more than 50% gapped using the Gap-Streeze server (Los Alamos HIV Sequence Database: http://www.hiv.lanl.gov/content/sequence/GAPSTREEZE/gap.html). For ion channels, this combined strategy produced alignments that consisted mainly of the trans-membrane regions, pore loops, and the intra-cellular linker between domains III and IV.

Phylogenetics

We used both maximum likelihood (ML) and Bayesian methods to estimate phylogenies. The Whelan and Goldman model with a class of invariant sites (+I), 4 gamma distributed rate categories (+G), and estimated amino acid frequencies (+F) was chosen by the Akaike information criterion in Prottest as the best model and was used for ML inference (Whelan and Goldman 2001; Abascal, Zardoya, and Posada 2005). ML and bootstrap trees were estimated in Garli (Zwickl 2006) with the final ML tree being the best of four independent replicates. The bootstrap proportions are out of 100 pseudo-replicates. Bayesian estimation was done using PhyloBayes 3.3 under default 'automatic stopping-rule' conditions (Lartillot and Philippe 2004). The authors of PhyloBayes recommend the CAT-GTR or CAT-Pois models for datasets larger than 1,000 aligned columns. Both of our datasets are over this threshold, so we chose the default CAT-Pois model. The mean numbers of site-classes assigned by the CAT model were averaged over the posterior distributions of both chains pooled together. The mean of the data sets were 84.84 and 101.2 for the data sets in Figures 1 and Supp. Fig. 3, respectively.

Data Submission

Data sets, alignments and trees used for phylogenetic analysis were submitted to TreeBase (accession URL: http://purl.org/phylo/treebase/phylows/study/TB2:S12662)

References

Abascal F, Zardoya R, Posada D. ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21:2104 –2105.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. Nucl. Acids Res. 25:3389–3402.

Katoh K, Kuma K, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Research 33:511 –518.

Lartillot N, Philippe H. 2004. A Bayesian Mixture Model for Across-Site Heterogeneities in the Amino-Acid Replacement Process. Molecular Biology and Evolution 21:1095 –1109.

Penn O, Privman E, Landan G, Graur D, Pupko T. 2010. An Alignment Confidence Score Capturing Robustness to Guide Tree Uncertainty. Molecular Biology and Evolution 27:1759–1767.

Whelan S, Goldman N. 2001. A General Empirical Model of Protein Evolution Derived from Multiple Protein Families Using a Maximum-Likelihood Approach. Mol Biol Evol 18:691–699.