Palovcak et al., http://www.jgp.org/cgi/content/full/jgp.201311103/DC1



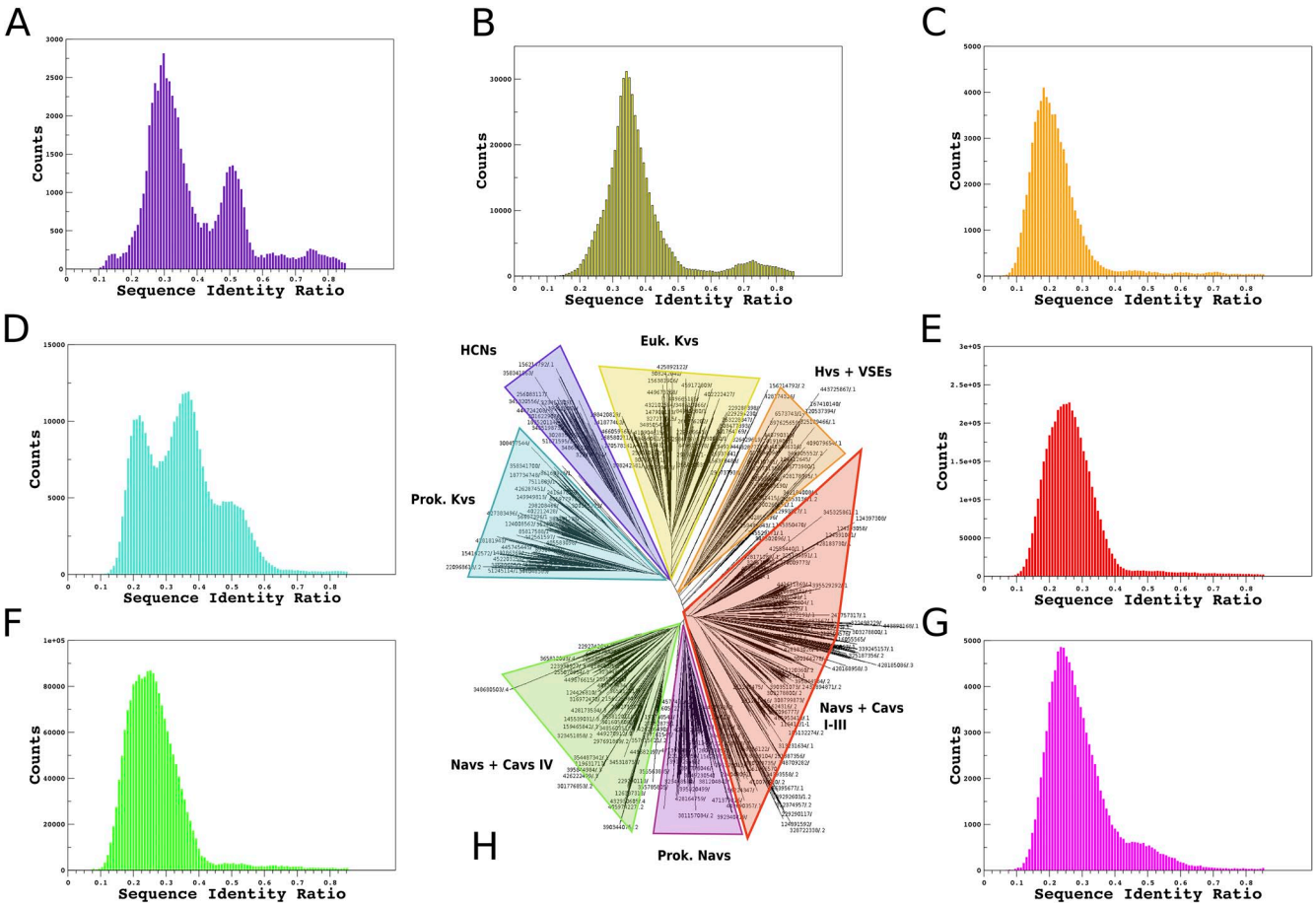


**Figure S1.** Histograms of sequence identity for each of the identified protein families. Specifically: A, HCNs; B, Euk. Kvs; C, Hvs +VSEs; D, Prok. Kvs; E, Navs + Cavs I-III; F, Navs + Cavs IV; G, Prok. Navs. (H) Dendrogram showing how specific branches were assigned to protein families. Sequences outside the highlighted regions were labeled "unclassified." Pairwise identities were calculated by computing the number of matching characters divided by the length of the MSA (including the gap symbol) for every pair of sequences in each family MSA.

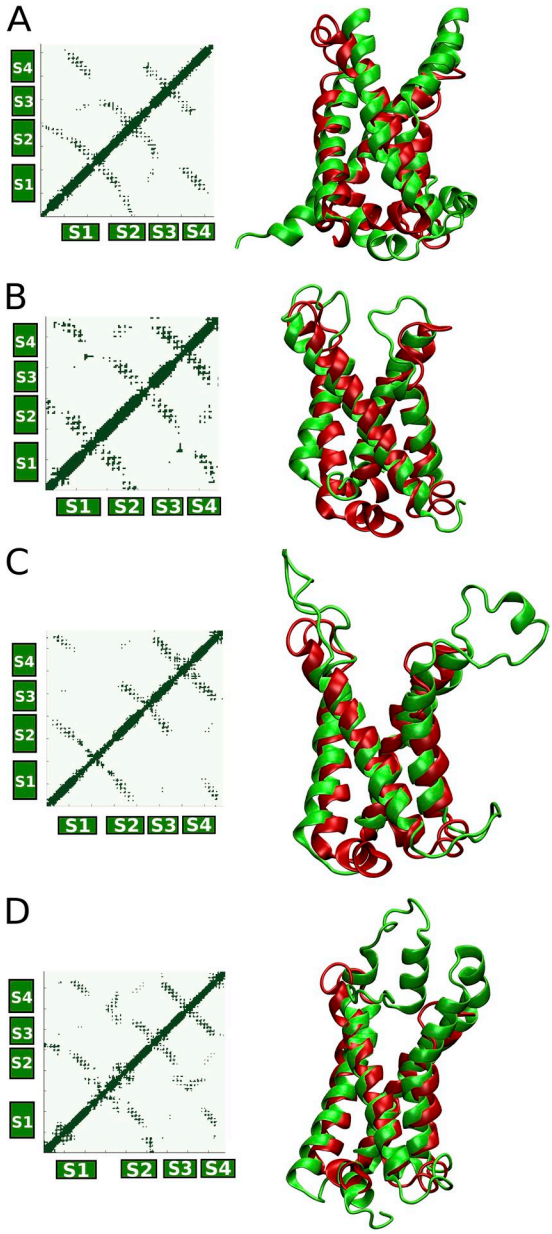


**Figure S2.** Contact maps for VSD homologues are highly similar. (A; left) Contact map for the solution structure of the archaean voltage-gated potassium channel KvAP VSD (Protein Data Bank accession no. 2KYH). (Right) Superposition of the KvAP VSD structure (green) onto the NavAb VSD (red). (B; left) Contact map for the bacterial sodium channel NavRh VSD (Protein Data Bank accession no. 4DXW). (Right) Superposition of the NavRh VSD (green) on the NavAb VSD (red). (C; left) Contact map for the mammalian Kv1.2 VSD (Protein Data Bank accession no. 3LUT). (Right) Superposition of the Kv1.2 VSD (green) on the NavAb VSD (red). (D; left) Contact map for the Kv1.2/2.1 paddle chimera VSD (Protein Data Bank accession no. 2R9R). (Right) Superposition of the paddle chimera VSD (green) on the NavAb VSD (red). All superpositions were performed with scripts within VMD.

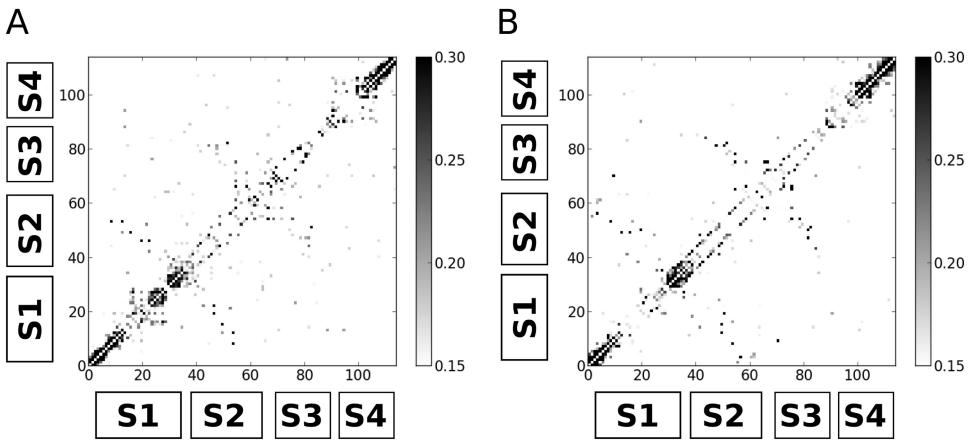


**Figure S3.** EC maps of Kv and Nav/Cav channels are similar. (A) EC score matrix for a partition of the VSD MSA containing only Kv sequences. Sequences were partitioned by combining the sequences identified in Fig. S1 (B and D). Total MSA contained 3,391 unique sequences, with 1,796 "effective sequences" at a 90% sequence identity threshold. (B) EC score matrix for a partition of the VSD MSA containing only Cav or Nav sequences. Sequences were partitioned by combining the sequences identified in Fig. S1 (E and F). Total MSA contained 1,832 sequences, with 1,071 "effective sequences" at a 90% sequence identity threshold.

*Table S1*
*Seed MSA of various VSDs*

| Sequence | VSD |
|---|---|
| Nav1.4_III: | --RRACFKIVEHNWFETFIVFMILLSSGALAFED------------------------ |
| Shaker: | -------SSQAARVVAIISVFVILLSIVIFCLETLPEFKHYKVF----------NTTTNG |
| rat_Kv1.2: | -------SSGPARIIAIVSVMVILISIVSFCLETLPIFRDENEDMHGGGVT---FHTYSQ |
| Nav1.4_I: | --RRGAIKVLIHALFSMFIMITILTNCVFMTMSDP------------------------ |
| VSP: | RVQFRVRAVIDHLGMRVFGVFLIFLDIILMIIDLS------------------------ |
| Nav1.4_II: | KFKNIIHLIVMDPFVDLGITICIVLNTLFMAMEHY------------------------ |
| Nav1.4_IV: | KIQGMVYDLVTKQAFDITIMILICLNMVTMMVETD------------------------ |
| Cav1.2_III: | RFRLQCHRIVNDTIFTNLILFFILLSSISLAAEDPVQHTSFRNH---------ILFYFDI |
| Shaw: | -------SSNAAKTIGVVSVFFICISILSFCLKTHPDMRVPIVRNITVKTANG-----SN |
| Cav1.2_II: | --------AVKSNVFYWLVIFLVFLNTLTIASEHY------------------------ |
| NavAb: | -MYLRITNIVESSFFTKFIIYLIVLNGITMGLETS------------------------ |
| Shab: | -------TSFAARVIAVISILFIVLSTIALTLNTLPQ----------------------- |
| KvAP: | GRVRNIGDVMEHPLVELGVSYAALLSVIVVVVEYT------------------------ |
| Kv7.1: | VYNFLERPTGWKCFVYHFAVFLIVLVCLIFSVLST------------------------ |
| Cav1.2_I: | ---RACISIVEWKPFEIIILLTIFANCVALAIYIP------------------------ |
| Kv11.1: | HRWTILHYSPFKAVWDWLILLLVIYTAVFTPYSAAFLLKE------------------TE |
| Hv1: | DFRGMLRKLFSSHRFQVIIICLVVLDALLVLAELILDL----------------------K |
| Cav1.2_IV: | KIQGMVYDLVTKQAFDITIMILICLNMVTMMVETD------------------------ |
| | llllllllllltttttttttttttttuuuu |
| | \|------------S1-------------\| |
| Nav1.4_III: | --IYIEQRRVIRTILEYADKVFTYIFIMEMLLKWVAYGF--------------KVYFTNA |
| Shaker: | TKIEEDEVPDITDPFFLIETLCIIWFTFELTVRFLACPN-------------KLNFCRDV |
| rat_Kv1.2: | STIGYQQSTSFTDPFFIVETLCIIWFSFEFLVRFFACPS-------------KAGFFTNI |
| Nav1.4_I: | -----------PPWSKNVEYTFTGIYTFESLIKILARGFCV----------DDFTFLRDP |
| VSP: | ---LPGKSESSQSFYDGMALALSCYFMLDLGLRIFAYGP--------------KNFFTNP |
| Nav1.4_II: | -----PMTEHFDNVLTVGNLVFTGIFTAEMVLKLIAMDP--------------YEYFQQG |
| Nav1.4_IV: | -----NQSQLKVDILYNINMIFIIIFTGECVLKMLALR---------------QYYFTVG |
| Cav1.2_III: | VFTTIFTIEIALKILGNADYVFTSIFTLEIILKMTAYGA--------------FLHKG-S |
| Shaw: | GWFLDKTQTNAHIAFFYIECVCNAWFTFEILVRFISSPN-------------KWEFIKSS |
| Cav1.2_II: | -----NQPNWLTEVQDTANKALLALFTAEMLLKMYSLGL--------------QAYFVSL |
| NavAb: | ----KTFMQSFGVYTTLFNQIVITIFTIEIILRIYVHR---------------ISFFKDP |
| Shab: | LQHIDNGTPQDNPQLAMVEAVCITWFTLEYILRFSASPD-------------KWKFFKGG |
| KvAP: | ----MQLSGEYLVRLYLVDLILVIILWADYAYRAYKSGD---------------PAGYVK |
| Kv7.1: | ---IEQYAALATGTLFWMEIVLVVFFGTEYVVRLWSAGCRSKYV----GLWGRLRFARKP |
| Cav1.2_I: | --FPEDDSNATNSNLERVEYLFLIIFTVEAFLKVIAYGL--------------LFHPN-A |
| Kv11.1: | EGPPATECGYACQPLAVVDLIVDIMFIVDILINFRTTYVNANEEVVSHPGRIAVHYFK-G |
| Hv1: | IIQ-PDKNNYAAMVFHYMSITILVFFMMEIIFKLFVFR---------------LEFFHHK |
| Cav1.2_IV: | -----NQSQLKVDILYNINMIFIIIFTGECVLKMLALR---------------QYYFTVG |
| | uuuuuuuuuuuuuttttttttttttttttttlll lllllll |
| | \|------------------S2------------\| \|------ |
| Nav1.4_III: | WCWLDFLIVDVSIIS--LVANWLGYSELG-------------------PIKSLRTLRALR |
| Shaker: | MNVIDIIAIIPYFITLATVVAEEEDTNLNLPKAPVSPQDKSSNQAMSLAILRVIRLVRVFR |
| rat_Kv1.2: | MNIIDIVAIIPYFITLGTELAEKPEDAQQGQQA-----------MSLAILRVIRLVRVFR |
| Nav1.4_I: | WNWLDFSVIMMAYLT--------EFVDLG-------------------NISALRTFRVLR |
| VSP: | WEVADGLIIVVTFVVTIFYTVLDEYVQETGAD------------GLGRLVVLARLLRVVR |
| Nav1.4_II: | WNIFDSIIVT--------LSLVELGLANV------------------QGLSVLRSFRLLR |
| Nav1.4_IV: | WNIFDFVVVILSIVGLALSDLIQKYFVSP------------------TLFRVIRLARIGR |
| Cav1.2_III: | FCRNYFNILDLLVVSVSLISFGIQSSAIN-------------------VVKILRVLRVLR |
| Shaw: | VNIIDYIATLSFYIDLVLQRFASHLENA-------------------DILEFFSIIRIMR |
| Cav1.2_II: | FNRFDCFVVCGGIL-----ETILVETKIMSP-------------------LGISVLRCVR |
| NavAb: | WSLFDFFVVAISLVP----------TSSG--------------------FEILRVLRVLR |
| Shab: | LNIIDLLAILPYFVSLFLLETNKNATDQFQD-------------VRRVVQVFRIMRILR |
| KvAP: | KTLYEIPALVPAGLLALIEGHLAGL------------------------GLFRLVRLLR |
| Kv7.1: | ISIIDLIVVVASMVVLCVGSKGQVFATSA-------------------IRGIRFLQILR |

*Table S1*
*Seed MSA of various VSDs (Continued)*

| Sequence | VSD |
|---|---|
| Cav1.2_I: | YLRNGWNLLDFIIVVVGLFSAILEQATKADGANAL------GGKGAGFDVKALRAFRVLR |
| Kv11.1: | WFLIDMVAAIPFDLLIFGSGSEE------------------------LIGLLKTARLLR |
| Hv1: | FEILDAVVVVVSFILDIVLLFQEHQFEAL------------------GLLILLRLWRVAR |
| Cav1.2_IV: | WNIFDFVVVILSIVGLALSDLIQKYFVSP------------------TLFRVIRLARIGR |
| | ltttttttttttttt uuuu uuuttttttttt |
| | --S3----------\| \|---------- |
| Nav1.4_III: | PLRALSRFEGMRVVVNALLGAIPSI |
| Shaker: | IFKLSRHSKGLQILGRTLKASMREL |
| rat_Kv1.2: | IFKLSRHSKGLQILGQTLKASMREL |
| Nav1.4_I: | ALKTITVIPGLKTIVGALIQSVKKL |
| VSP: | LARIFYSHQQMKASSRRTISQNKRR |
| Nav1.4_II: | VFKLAKSWPTLNMLIKIIGNSVGAL |
| Nav1.4_IV: | VLRLIRGAKGIRTLLFALMMSLPAL |
| Cav1.2_III: | PLRAINRAKGLKHVVQCVFVAI--- |
| Shaw: | LFKLTRHSSGLKILIQTFRASAKE- |
| Cav1.2_II: | LLRIFKITRYWNSLSNLVA------ |
| NavAb: | LFRLVTAVPQMRKIVSALISV---- |
| Shab: | VLKLARHSTGLQSLGFTLRNSYKEL |
| KvAP: | FLRILLIISRGSKFLSAIADAADKL |
| Kv7.1: | MLHVDRQGGTWRLL---------- |
| Cav1.2_I: | PLRLVSGVPSLQVVLNSI------- |
| Kv11.1: | LVRVARKLDRYSEYGAAVLF----- |
| Hv1: | IINGIIISVKTRSERQLLR------ |
| Cav1.2_IV: | VLRLIRGAKGIRTLLFALMMSLPAL |
| | ttttttttllllllllllll |
| | --S4----------------\| |

For $D_{KL}$ calculation: I, inner membrane; u, outer membrane; t, transmembrane region. In the study, only positions in which NavAb did not contain a gap were considered. This eliminated positions that were not present or alignable across VSD subfamilies.

*Table S2*
*Residue correspondences for high $D_{KL}$ positions between NavAb and Shaker*

| NavAb bacterial voltage-gated $Na^+$ channel | *Shaker* voltage-gated $K^+$ channel |
|---|---|
| S11 | A226 |
| F14 | V229 |
| I22 | I237 |
| N25 | S240 |
| E32 | E247 |
| N49 | E283 |
| F56 | F290 |
| E59 | E293 |
| R63 | R297 |
| F71 | F307 |
| D74 | D310 |
| W76 | M312 |
| S77 | N313 |
| D80 | D316 |
| E96 | R362 |
| R99 | R365 |
| R102 | R368 |
| R105 | R371 |
| R108 | K374 |
| T111 | R377 |
| R117 | Q383 |

*Table S3*
*Residue correspondences for top evolutionarily coupled residue pairs, ordered by strength*

| NavAb bacterial voltage-gated $Na^+$ channel | *Shaker* voltage-gated $K^+$ channel |
|---|---|
| T57-L88[a] | T291-F324 |
| I18-I55[a] | S233-W289 |
| I60-F81[a] | L294-I317 |
| V9-L62[a] | Q224-V296 |
| I64-F72[a] | F298-C308 |
| I60-F71[a] | L294-F307 |
| I53-S87[a] | I287-Y323 |
| I60-V84[a] | L294-I320 |
| F37-G42[a] | E271-T276 |
| L21-I55[a] | V236-W289 |
| R63-S77[a] | R297-N313 |
| L24-V52[a] | L239-C286 |
| T28-V52[a] | I243-C286 |
| T15-F71 | A230-F307 |
| L21-V52[a] | V236-C286 |
| I64-F81[a] | F294-I317 |
| F56-S87[a] | F290-Y323 |
| V9-F14[a] | Q224-V229 |
| L21-F56[a] | V236-F290 |
| N49-E96 | E283-R362 |
| F72-L78[a] | C308-V314 |
| N25-E96 | S240-R362 |
| N25-V52[a] | S240-C286 |
| T15-W76 | A230-M312 |

The three coevolving residue pairs that are not in contact are not shown in Fig. 3 A but are mentioned in the text.
[a]Positions in contact in the NavAb structure.

*Table S4*

*Reference amino acid frequencies for three transmembrane helix topological regions*

| Amino Acid | Inner membrane interface | Transmembrane interface | Outer membrane interface |
|---|---|---|---|
| A | 0.08968 | 0.11462 | 0.09011 |
| C | 0.00672 | 0.01176 | 0.00712 |
| D | 0.02660 | 0.00906 | 0.02878 |
| E | 0.03459 | 0.01192 | 0.03137 |
| F | 0.06874 | 0.09043 | 0.07244 |
| G | 0.07105 | 0.08694 | 0.09580 |
| H | 0.02270 | 0.01184 | 0.02479 |
| I | 0.07352 | 0.10535 | 0.06813 |
| K | 0.04709 | 0.00848 | 0.02191 |
| L | 0.12547 | 0.16591 | 0.11537 |
| M | 0.03059 | 0.03699 | 0.02923 |
| N | 0.02962 | 0.01559 | 0.03599 |
| P | 0.03782 | 0.02663 | 0.05335 |
| Q | 0.02466 | 0.01093 | 0.02662 |
| R | 0.06413 | 0.01359 | 0.02636 |
| S | 0.05438 | 0.04615 | 0.05681 |
| T | 0.05173 | 0.05925 | 0.05797 |
| V | 0.07178 | 0.10574 | 0.07086 |
| W | 0.02840 | 0.03439 | 0.03637 |
| Y | 0.04060 | 0.03432 | 0.05054 |