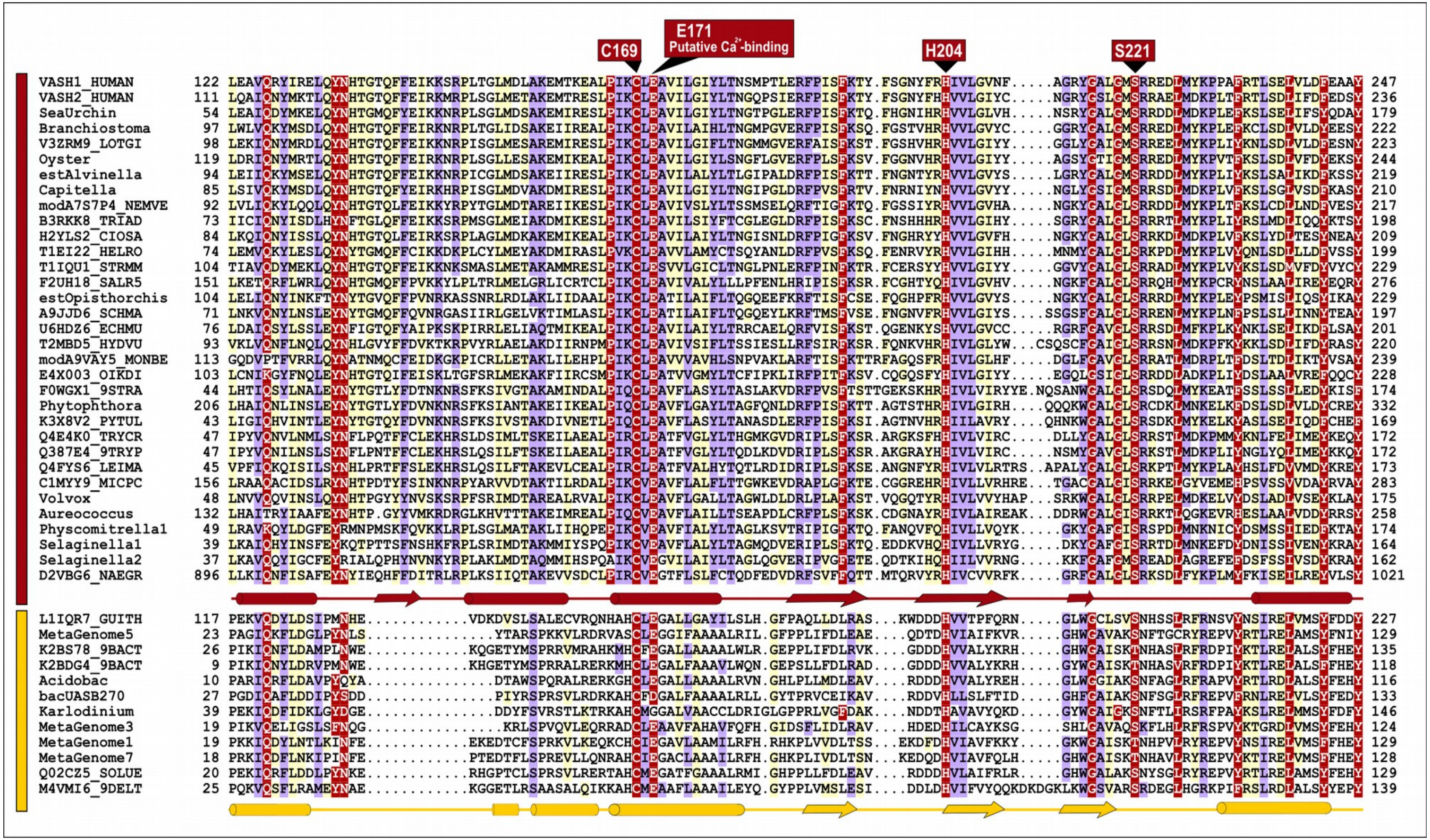
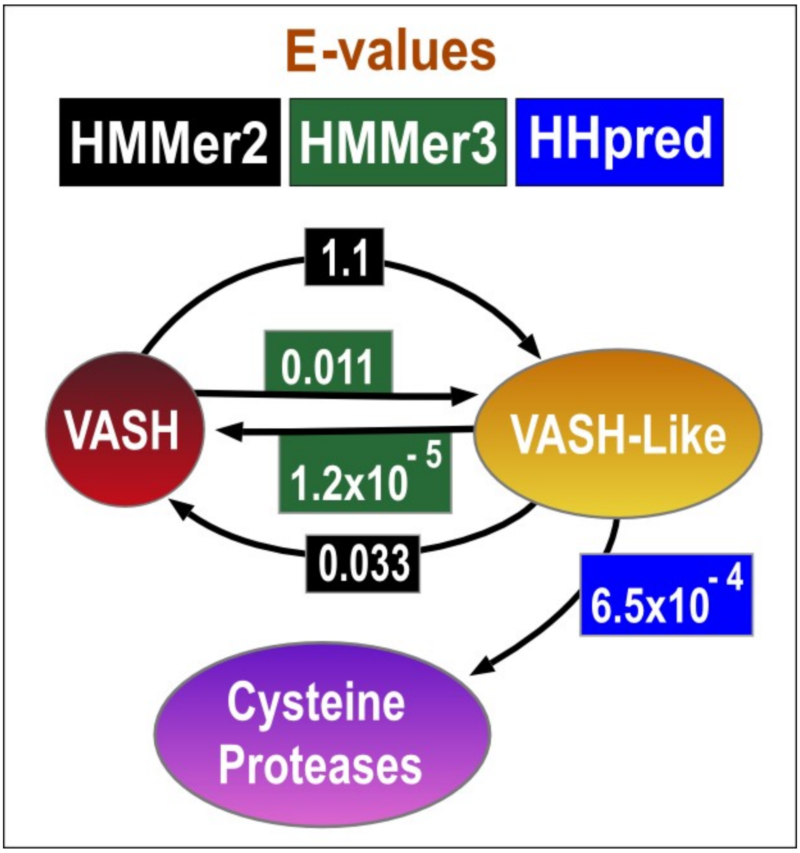

C169
E171
Putative Ca^{2+}-binding
H204
S221
VASH1_HUMAN 122 LEAVQRYIRELQYNHTGTQFFEIKKSRPLTGLMDLAKEMTKEALPIKCLEAVILGIYLTNSMPTLERFPISFKTY.FSGNYFRHIVLGVNF.....AGRYGALGMSRREDLMYKPPAFRTLSELVLDFEAAY 247
VASH2_HUMAN 111 LQAIQNYMKTLQYNHTGTQFFEIRKMRPLSGLMETAKEMTRESLPIKCLEAVILGIYLTNGQPSIERFPISFKTY.FSGNYFHHVVLGIYC.....NGRYGSLGMSRRAELMDKPLTFRTLSDLIFDFEDSY 236
SeaUrchin 54 LEAIQDYMKELQYNHTGMQFYEIKKNRPLSGLMDSAKEMIRESLPIKCLEAVILGIYLTNGTPGLERFPISFKTQ.FHGNIHRHVVLGLVH.....NSRYGALGMSRRDDLMDKPLEFKSLSELIFSYQDAY 179
Branchiostoma 97 LWLVQKYMSDLQYNHTGTQFFEIRKNRPLTGLIDSAKEIIREALPIKCLEAVILAIHLTNGMPGVERFPISFKSQ.FGSTVHRHVVLGVYC.....GGRYGALGMSRREDLMYKPLEFKCLSDLVLDYEESY 222
V3ZRM9_LOTGI 98 LEKIQNYMRDLQYNHTGTQFFEIKKNRPISGLMESAREMIRESLPIKCLEAVILGIFLTNGMMGVERFAISFKST.FGSHVHRHVVLGIYY.....GGLYGAIGMSRREELMYKPLIYKNLSDLVLDFESNY 223
Oyster 119 LDRIQNYMRTLQYNHTGTQFFEIRKNRPLSGLLESAKEMIKEALPIKCLEAVILGIYLSNGFLGVERFPLSFKSV.FGGNVHRHVVLGIYY.....AGSYGTIGMSRREDLMYKPVTFKSLSDLVFDYEKSY 244
estAlvinella 94 LEIIQKYMSELQYNHTGTQFFEIKKNRPICGLMDSAKEIIRESLPIKCLEAVILALYLTLGIPALDRFPISFKTT.FNGNTYRHVVLGVYS.....LGRYGALGMSRRDDLMYKPLIYKSLSALIKDFKSSY 219
Capitella 85 LSIVQKYMSDLQYNHTGTQFYEIRKHRPISGLMDVAKDMIRESLPIKCLEAVILGIYLTNGIPGLDRFPVSFRTV.FNRNIYNHVVLGVYY.....NGLYGSIGMSRRSDLMDKPLVFKSLSGLVSDFKASY 210
modA7S7P4_NEMVE 92 LVLIQKYLQQLQYNHTGTQLFEIKKYRPYTGLMDTAREIIKESLPIKCLEAVIVSLYLTSSMSELQRFTIGFKTQ.FGSSIYRHVVLGVHA.....NGKYGALGLSRRDDLMYKPLTFKSLCDLLNDFVESY 217
B3RKK8_TRIAD 73 IICIQNYISDLHYNFTGLQFFEIKKSRPMSGLMEIAKDMIKESLPIKCLEAVILSIYFTCGLEGLDRFPISFKSC.FNSHHHRHVVLGIHY.....SGRYGALGLSRRRTLMYKPLIYRSLMDLIQQYKTSY 198
H2YLS2_CIOSA 84 LKQIQNYISSLQYNHTGTQLFEIRKSRPLAGLMDKAKEMIKEALPIKCLEAVILAIYLTNGISNLDRFPIGFKSV.FNGHRYYHVVLGVFH.....NGKYGALGLSRRKDLMDKPLVFKSLYDLTESYNEAY 209
T1EI22_HELRO 74 LEMVQKYLESLQYNYTGMQFFCIKKDKPLCYLMEYAKDMIRASLPVKCLEAVVLAMYCTSQYANLDRFPVSFKSQ.FENRVYRHVVLGIHH.....MNMYGALGLSRKPDLMYKPLVYQNLSDLLLDFVSSY 199
T1IQU1_STRMM 104 TIAVQDYMEKLQYNHTGTQFFEIKKNKSMASLMETAKAMMRESLPIKCLESVVLGICLTNGLPNLERFPINFKTR.FCERSYYHVVLGIYY.....GGVYGALGLSRRADLMYKPLVYKSLSDMVFDYVYCY 229
F2UH18_SALR5 151 LKETQRFLWRLQYNHTGMQFFPVKKYLPLTRLMELGRLICRTCLPIKCLEAVIVALYLLLPFENLHRIPISFKSR.FCGHTYQHIVLGVHV.....NGKFGALGLSRRQHLMYKPCRYNSLAALIREYEQRY 276
estOpisthorchis 104 LELIQNYINKFTYNYTGVQFFPVNRKASSNRLRDLAKLIIDAALPIKCLEATILAIFLTQGQEEFKRFTISFCSE.FQGHPFRHVVLGVYS.....NGKYGALGLSRRRDLMYKPLEYPSMISLIQSYIKAY 229
A9JJD6_SCHMA 71 LNKVQNYLNSLEYNYTGMQFFQVNRGASIIRLGELVKTIMLASLPIKCLEATILAIFLTQGQEYLKRFTMSFVSE.FNGNVFRHVVLGIYS....SSGSFGALGLSRRENLMYKPLNFPSLSLLINNYTEAY 197
U6HDZ6_ECHMU 76 LDAIQSYLSSLEYNFIGTQFYAIPKSKPIRRLELIAQTMIKEALPIKCLEAVILAIYLTRRCAELQRFVISFKST.QGENKYSHVVLGVCC.....RGRFGAVGLSRRSDLMFKPLKYNKLSELIKDFLSAY 201
T2MBD5_HYDVU 93 VKLVQNFLNQLQYNHLGVYFFDVKTKRPVYRLAELAKDIIRNPMPIKCLEAVIVSIFLTSSIESLLRFSIRFKSR.FQNKVHRHIVLGLYW...CSQSCFGAIGLSRRSNLMDKPFKYKKLSDLIFDYRASY 220
modA9VAY5_MONBE 113 GQDVPTFVRRLQYNATNMQCFEIDKGKPICRLLETAKLILEHPLPIKCLEAVVVAVHLSNPVAKLARFTISFKTTRFAGQSFRHIVLGLHF.....DGLFGAVGLSRRATLMDRPLTFDSLTDLIKTYVSAY 239
E4X003_OIKDI 103 LCNIKGYFNQLEYNHTGTQIFEISKLTGFSRLMEKAKFIIRCSMPIKCLEATVVGMYLTCFIPKLIRFPITFKSV.CQGQSFYHIVLGIYY.....EGQLGSIGLSRRDDLADKPLIYDSLAALVREFQQCY 228
F0WGX1_9STRA 44 LHTIQSYLNALEYNYTGTLYFDTNKNRSFKSIVGTAKAMINDALPIQCLEAVFLASYLTASLAKVDRFPVSFTSTTGEKSKHRHIVLVIRYYE.NQSANWGALGLSRSDQLMYKEATFSSLSSLLEDYKISF 174
Phytophthora 206 LHAIQNLINSLEYNYTGTLYFDVNKNRSFKSIANTAKEIIKEALPIQCLEAVFLGAYLTAGFQNLDRFPISFKTT.AGTSTHRHIVLGIRH....QQQKWGALGLSRCDKLMNKELKFDSLSDLVLDYCREY 332
K3X8V2_PYTUL 43 LIGIQHVINTLEYNYTGTQYFDVNKNRSFKSIVSTAKDIVNETLPIQCLEAVFLASYLTANASDLERFPISFKSI.AGTNVHRHIILAVRY....QHNKWGALGLSRSDKLMYKELKYASLSELIQDFCHEF 169
Q4E4K0_TRYCR 47 IPYVQNVLNMLSYNFLPQTFFCLEKHRSLDSIMLTSKEILAEALPIRCLEATFVGLYLTHGMKGVDRIPLSFKSR.ARGKSFHHIVLVIRC.....DLLYGALGLSRRSTLMDKPMMYKNLFELIMEYKEQY 172
Q387E4_9TRYP 47 IPYVQNILNSLSYNFLPNTFFCLEKHRSLQSILFTSKEILAEALPIRCLEATFVGLYLTQDLKDVDRIPLSFKSR.AKGRAYHHIVLVIRC.....NSMYGAVGLSRKSTLMDKPLIYNGLYQLIMEYKKQY 172
Q4FYS6_LEIMA 45 VPFIQKQISILSYNHLPRTFFSLEKHRSLQSILFTAKEVLCEALPIRCLEATFVALHYTQTLRDIDRIPLSFKSE.ANGNFYRHIVLVLRTRS..APALYGALGLSRKPTLMYKPLAYHSLFDVVMDYKREY 173
C1MYY9_MICPC 156 LRAAQACIDSLRYNHTPDTYFSINKNRPYARVVDTAKTILRDALPIKCVEAVFLALFLTTGWKEVDRAPLGFKTE.CGGREHRHIVLLVRHRE...TGACGALGISRRKELGYVEMEHPSVSSVVDAYRVAY 283
Volvox 48 LNVVQQVINSLQYNHTPGYYYNVSKSRPFSRIMDTAREALRVALPIKCLEAVFLGALLTAGWLDLDRLPLAFKST.VQGQTYRHIVLVVYHAP...SRKWGALGLSRRPELMDKELVYDSLADLVSEYKLAY 175
Aureococcus 132 LHAITRYIAAFEYNHTP.GYYVMKRDRGLKHVTTTAKEIMREALPIQCVEAVFLAILLTSEAPDLCRFPLSFKSK.CDGNAYRHIVLAIREAK...DDRWGALGISRRKTLQGKEVRHESLAALVDDYRRSY 258
Physcomitrella1 49 LRAVKQYLDGFEYRMNPMSKFQVKKLRPLSGLMATAKLIIHQPEPIKCVEAVFIALYLTAGLKSVTRIPIGFKTQ.FANQVFQHIVLLVQYK.....GKYGAFGISRSPDLMNKNICYDSMSSIIEDFKTAY 174
Selaginella1 39 LKAIQHYINSFEYKQTPTTSFNSHKFRPLSRIMDTAKMMIYSPQPIKCVEAVFLALYLTAGMQDVERIPLSFKTQ.EDDKVHQHIVLLVRYG.....DKYGAFGISRRTDLMNKEFDYDNISSIVENYKRAY 164
Selaginella2 37 LKAVQQYIGCFEYRIALQPHYNVNKYRPLAKLMDTAQMMIHSPQAIKCVEGVILAIYLTAGLQGVERIPVGFETE.QDTKIHQHIILVVRNG.....KKFGAFGMSREADLAGREFEFDSFSSIVSDYKRAY 162
D2VBG6_NAEGR 896 LLKIQNFISAFEYNYIEQHFFDITRLRPLKSIIQTAKEVVSDCLPIRCVEGTFLSLFCTQDFEDVDRFSVFFQTT.MTQRVYRHIVCVVRFK.....GRFGALGLSRKSDLFYKPLMYFKISEILREYVLSY 1021
L1IQR7_GUITH 117 PEKVQDYLDSIPMNHE............VDKDVSLSALECVRQNHAHCLEGALLGAYILSLH.GFPAQLLDLRAS...KWDDDHVVTPFQRN.....GLWGCLSVSNHSSLRFRNSVYNSIRELVMSYFDDY 227
MetaGenome5 23 PAGIQKFLDGLPYNLS..............YTARSPKKVLRDRVASCLEGGIFAAAALRIL.GFPPLIFDLEAE....QDTDHVIAIFKVR.....GHWGAVAKSNFTGCRYREPVYRTLRELAMSYFNIY 129
K2BS78_9BACT 26 PIKIQNFLDAMPLNWE............KQGETYMSPRRVMRAHKMHCFEGALLAAAALWLR.GEPPLIFDLRVK....GDDDHVVALYKRH.....GHWGAISKTNHASLRFRDPIYKTLRELALSYFHEY 135
K2BDG4_9BACT 9 PIKIQNYLDRVPMNWE............KHGETYMSPRRALRERKMHCLEGALFAAAVLWQN.GEPSLLFDLRAD....GDDDHVVALYKRH.....GYWGAISKTNHASVRFRDPIYKTLRELALSFFHEY 118
Acidobac 10 PARIQRFLDAVPYQYA..............DTAWSPQRALRERKGHCLEGALLAAAALRVN.GHLPLLMDLEAV....RDDDHVVALYREH.....GLWGGIAKSNFAGLRFRAPVYRTLRDLALSYFEHY 116
bacUASB270 27 PGDIQAFLDDIPYSDD..............PIYRSPRSVLRDRKAHCFDGALFAAAALRLL.GYTPRVCEIKAV....RDDVHLLSLFTID.....GHFGAIAKSNFSGLRFREPVFRNLRELVLSYFEDY 133
Karlodinium 39 PEKIQDFIDKLGYDGE..............DDYFSVRSTLKTRKAHCMGGALVAACCLDRIGLGPPRLVGFDAK....NDDTHAVAVYQKD.....GYWGAIGKSNFTLIRSRFPAYKSLRELMMSYFDFY 146
MetaGenome3 19 PIKVQELIGSLSFNQG...............KRLSPVQVLEQRRADCLEAAVFAHAVFQFH.GIDSFLIDLRAV....HDEDHILCAYKSG.....SHLGAVAQSKFLHLRFRSPVYKTGRDLVMSYFEHY 124
MetaGenome1 19 PKKIQDYLNTLKINFE............EKEDTCFSPRKVLKEQKCHCIEGAVLAAMILRFH.RHKPLVVDLTSS...EKDFDHVIAVFKKY.....GKWGAISKTNHPVLRYREPVYNSIRELVMSFFHEY 129
MetaGenome7 18 PRKIQDFLNKIPINFE............PTEDTFLSPREVLLQNRAHCIEGACLAAAILRFH.GHKPLIVDLTSN...KEDQDHVIAVFQLH.....GHWGAISKTNHAVLRYREPVYKTIRELVMSFFHEY 128
Q02CZ5_SOLUE 20 PEKIQRFLDDLPYNKE............RHGPTCLSPRSVLRERTAHCMEGATFGAAALRMI.GHPPLLFDLEAV....RDDDHVLAIFRLR.....GHWGALAKSNYSGLRYREPVYRTLRELAMSYFEHY 129
M4VMI6_9DELT 25 PQKVQSFLRAMEYNAE............KGGETLRSAASALQIKKAHCMEAAFLAAAILEYQ.GYPPLVMSLESI....DDLDHVIFVYQQKDKDGKLKWGSVARSRDEGLHGRKPIFRSLRDLALSYYEPY 139

Supplemental Fig. S1.

Representative multiple sequence alignment of the vasohibin and vasohibin-like families.

The alignment shows in Figure 1 corresponds to the currently defined core of the transglutaminase-like cysteine protease family (Ginalski *et al.*, 2004; Makarova *et al.*, 1999; Makarova *et al.*, 2000; Chatterjee *et al.* 2012), but we were able to extend the alignment between vasohibin and vasohibin-like families further toward their C-termini, including a predicted alpha-helix. Putative catalytic residues are shown in red. A putative calcium-binding residue E171 in human VASH1 is labelled. Families are indicated by coloured bars to the left of the alignment: vasohibin and vasohibin-like sequences are indicated in red and yellow, respectively. Secondary structure predictions (Jones, 1999) were performed independently for the vasohibin and vasohibin-like families. The limits of the protein sequences included in the alignment are indicated by flanking residue positions. Alignments were produced with T-Coffee (Notredame *et al.*, 2000) using default parameters and slightly refined manually. The alignment was presented with the program Belvu using a colouring scheme indicating the average BLOSUM62 scores (which are correlated with amino acid conservation) of each alignment column: red (>3), violet (between 3 and 1.5) and light yellow (between 1.5 and 0.5) (Sonnhammer and Hollich, 2005). Sequences were obtained from UniProt, GenBank and JGI (Joint Genome Institute) databases, but were supplemented by manually assembled ESTs and FGENESH+ predicted gene models. The "mod" prefix identifies UniProt sequences corrected by gene prediction software FGENESH+ (Solovyev *et al.*, 2006). The "est" prefix identifies consensus sequences manually reconstructed by assembling highly similar NCBI-GenBank expressed sequence tags from identical species (conceptual translations). Sequences are named according to their UniProt identification or common name. Database of origin, accession numbers, and species name: UniProt:VASH1_HUMAN, *Homo sapiens*; UniProt:VASH2_HUMAN, *Homo sapiens*; SeaUrchin, UniProt:UPI0000E46146, *Strongylocentrotus purpuratus*; Branchiostoma, JGI genome FGENESH+ predicted gene and NCBI_Est BW730620, *Branchiostoma floridae*; UniProt:V3ZRM9_LOTGI, *Lottia gigantea*; Oyster, UniProt:UPI0005C3CA01, *Crassostrea gigas*; estAlvinella, NCBI_Ests: GO188185 and GO140334, *Alvinella pompejana*; Capitella, JGI genome FGENESH+ predicted gene, *Capitella sp.*; UniProt:A7S7P4_NEMVE, *Nematostella vectensis*; UniProt:B3RKK8_TRIAD, *Trichoplax adhaerens*; UniProt:H2YLS2_CIOSA, *Ciona savignyi*; UniProt:T1EI22_HELRO, *Helobdella robusta*; UniProt:T1IQU1_STRMM, *Strigamia maritima*; UniProt:F2UH18_SALR5, *Salpingoeca rosetta*; estOpisthorchis, NCBI_Est:ES417530, *Opisthorchis viverrini*; UniProt:A9JJD6_SCHMA, *Schistosoma mansoni*; UniProt:U6HDZ6_ECHMU, *Echinococcus multilocularis*; UniProt:T2MBD5_HYDVU, *Hydra vulgaris*; UniProt:A9VAY5_MONBE, *Monosiga brevicollis*; UniProt:E4X003_OIKDI, *Oikopleura dioica*; UniProt:F0WGX1_9STRA, *Albugo laibachii*; Phytophthora, JGI genome FGENESH+ predicted gene, *Phytophthora ramorum*; UniProt:K3X8V2_PYTUL, *Pythium ultimum*; UniProt:Q4E4K0_TRYCR, *Trypanosoma cruzi*; UniProt:Q387E4_9TRYP, *Trypanosoma brucei brucei*; UniProt:Q4FYS6_LEIMA, *Leishmania major*; UniProt:C1MYY9_MICPC, *Micromonas pusilla*; Volvox, JGI genome FGENESH+ predicted gene, *Volvox carteri*; Aureococcus, JGI genome FGENESH+ predicted gene, *Aureococcus anophagefferens*; Physcomitrella1, JGI genome FGENESH+ predicted gene, *Physcomitrella patens*; Selaginella1, JGI genome FGENESH+ predicted gene, *Selaginella moellendorffii*; Selaginella2, JGI genome FGENESH+ predicted gene, *Selaginella moellendorffii*; UniProt:D2VBG6_NAEGR, *Naegleria gruberi*; UniProt:L1IQR7_GUITH, *Guillardia theta*; MetaGenome5, NCBI_gi:689677707, uncultured bacterium; UniProt:K2BS78_9BACT, uncultured bacterium; UniProt:K2BDG4_9BACT, uncultured bacterium; Acidobac, UniProt:UPI00036EC8CC, *Acidobacteriaceae bacterium*; bacUASB270, NCBI_gi:672404298, *bacterium UASB270*; Karlodinium, NCBI_Ests: EC161919 EC160142, *Karlodinium micrum*; MetaGenome3, NCBI_gi:735013117, *archaeon GW2011_AR4*; MetaGenome1, NCBI_gi:735021449, *archaeon GW2011_AR19*; MetaGenome7, NCBI_gi:735018696, *archaeon GW2011_AR17*; UniProt:Q02CZ5_SOLUE, *Solibacter usitatus*; UniProt:M4VMI6_9DELT, *Bdellovibrio exovorus.*



Supplemental Fig. S2.

HMMer and HHpred comparison E-values.

Numbers correspond to HMMer (versions 2 and 3) profile-versus-sequence (black and green boxes) and HHpred profile-versus-profile (blue box) comparison E-values from global profile search results (Eddy, 1996; Finn *et al.*, 2015; Soding *et al.*, 2005). Arrows indicate the profile search direction. Both versions of HMMer (used with default parameters) yielded highly concordant results. Profile-to-sequence (HMMer) and Profile-to-profile (HHpred) matches were evaluated in terms of an E-value, which is the expected number of non-homologous proteins with a score higher than that obtained for the database match. An E-value much lower than one indicates statistical significance.