Supporting Information

Oka *et al.* 10.1073/pnas.0809420106

SI References

1. Slep KC, *et al.* (2001) Structural determinants for regulation of phosphodiesterase by a G protein at 2.0 A. *Nature* 409:1071–1077.
2. Barren B, Natochin M, Artemyev NO (2006) Mutation R238E in transducin-α yields a GTPase and effector-deficient, but not dominant-negative, G-protein α-subunit. *Mol Vis* 12:492–498.
3. Natochin M, Artemyev NO (1998) Substitution of transducin ser202 by asp abolishes G-protein/RGS interaction. *J Biol Chem* 273:4300–4303.
4. Rarick HM, Artemyev NO, Hamm HE (1992) A site on rod G protein α-subunit that mediates effector activation. *Science* 256:1031–1033.
5. Berlot CH, Bourne HR (1992) Identification of effector-activating residues of Gsα. *Cell* 68:911–922.
6. Chen Z, Singer WD, Sternweis PC, Sprang SR (2005) Structure of the p115RhoGEF rgRGS domain-Gα13/i1 chimera complex suggests convergent evolution of a GTPase activator. *Nat Struct Mol Biol* 12:191–197.
7. Wieland T, Bahtijari N, Zhou XB, Kleuss C, Simon MI (2000) Polarity exchange at the interface of regulators of G protein signaling with G protein α-subunits. *J Biol Chem* 275:28500–28506.
8. Lan KL, *et al.* (1998) A point mutation in Gαo and Gαi1 blocks interaction with regulator of G protein signaling proteins. *J Biol Chem* 273:12794–12797.
9. Kimple RJ, Kimple ME, Betts L, Sondek J, Siderovski DP (2002) Structural determinants for GoLoco-induced inhibition of nucleotide release by Gα subunits. *Nature* 416:878–881.
10. Tesmer VM, Kawano T, Shankaranarayanan A, Kozasa T, Tesmer JJ (2005) Snapshot of activated G proteins at the membrane: the Gαq-GRK2-G$\beta\gamma$complex. *Science* 310:1686–1690.
11. Day PW, *et al.* (2004) Characterization of the GRK2 binding site of Gαq. *J Biol Chem* 279:53643–53652.
12. Venkatakrishnan G, Exton JH (1996) Identification of determinants in the α-subunit of Gq required for phospholipase C activation. *J Biol Chem* 271:5066–5072.
13. Vazquez-Prado J, Miyazaki H, Castellone MD, Teramoto H, Gutkind JS (2004) Chimeric Gαi2/Gα13 proteins reveal the structural requirements for the binding and activation of the RGS-like (RGL)-containing Rho guanine nucleotide exchange factors (GEFs) by G α 13. *J Biol Chem* 279:54283–54290.
14. Beane WS, Voronina E, Wessel GM, McClay DR (2006) Lineage-specific expansions provide genomic complexity among sea urchin GTPases. *Dev Biol* 300:165–179.
15. Suga H, *et al.* (1999) Extensive gene duplication in the early evolution of animals before the parazoan-eumetazoan split demonstrated by G proteins and protein tyrosine kinases from sponge and hydra. *J Mol Evol* 48:646–653.
16. Seack J, Kruse M, Muller WE (1998) Evolutionary analysis of G-proteins in early metazoans: cloning of α- and β-subunits from the sponge Geodia cydonium. *Biochim Biophys Acta* 1401:93–103.

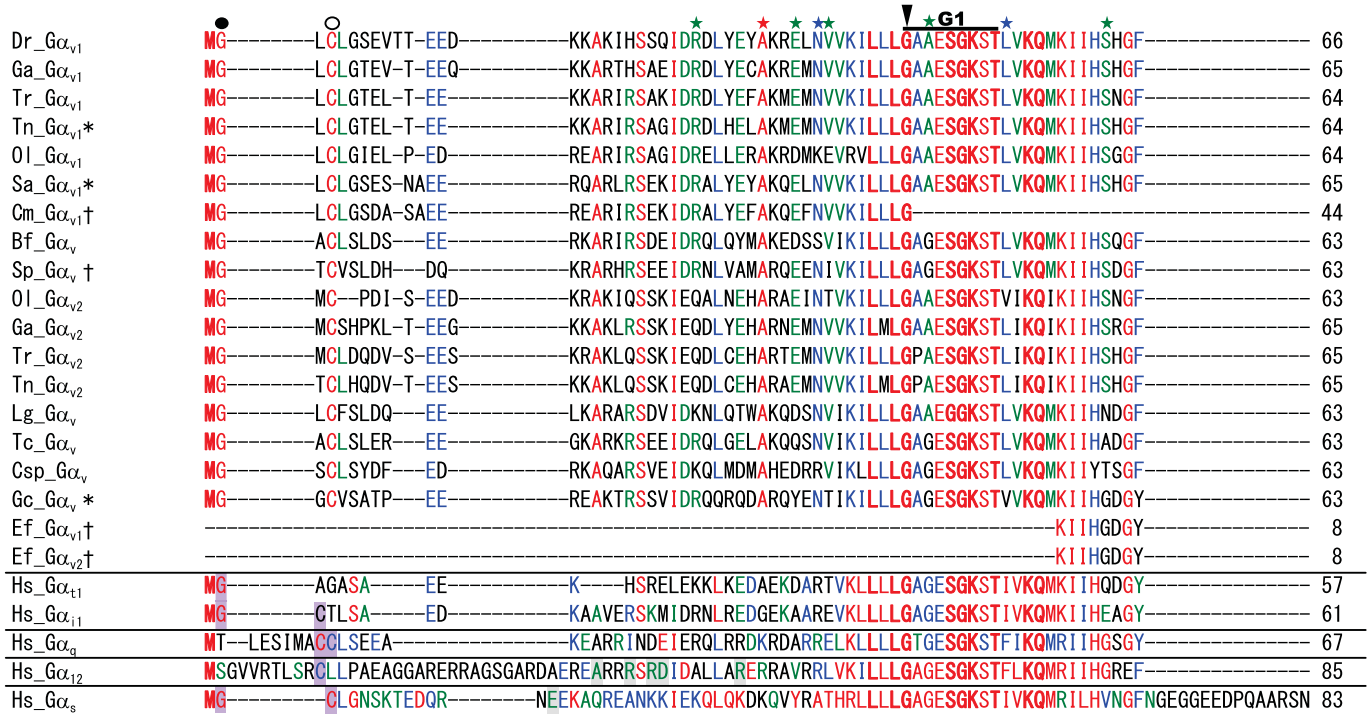


Fig. S1. Amino acid alignment of Gα$_v$ proteins: alignment of deduced amino acid sequences of 19 Gα$_v$ proteins from teleosts, sharks, a cephalochordate (lancelet), a sea urchin, a red flour beetle, an annelid (polychaete worm), a mollusk (limpet), and sponges together with human Gα$_{t1}$, Gα$_{i1}$, Gα$_q$, Gα$_{12}$, and Gα$_s$ sequences as representatives for each class. Conservation was analyzed between all human, mouse, zebrafish, fruit fly, and sponge Gα protein sequences. Residues conserved among all Gα proteins are in red boldface type. Residues in 100%, 80–99%, and 60–79% conservation within a class are in red, blue, and green, respectively. Atypical residues at conserved positions in sequences depicted here are shaded in gray. A comparison of the amino acids conserved in neoteleosts to the corresponding positions in shark and zebrafish sequences shows a 2-fold larger distance from Gα$_{v2}$ compared to Gα$_{v1}$, consistent with the topology of the shark and zebrafish genes as *gnav1* orthologs in the phylogenetic tree using the NJ algorithm (Fig. 1*B*). Residues suggested for contact with regulators of G protein signaling/GTPase-activating proteins, effectors, and both (see refs. 1–13) are shaded in blue, green, and yellow, respectively. In the case of atypical residues at these contact sites the residues are not shaded, but boxed. Known acylation sites in the N-terminal region and Cys residues of Gi proteins, ADP-ribosylated by Pertussis toxin, are shaded in purple. G-boxes (G1–5) are indicated with bars above the alignment. Switch regions (SW1–3) are shaded with hatch marks. Solid and open circles indicate putative N-linked myristoylation and thio-palmitoylation sites for Gα$_v$ proteins, respectively. Solid and open arrowheads are splicing junctions in a codon and between codons for teleost *gnav* genes, respectively. Gv-specific junctions are in red. Stars are Gv-specific motifs color coded for degree of conservation. Dr, zebrafish; Ol, medaka; Ga, stickleback; Tr, fugu; Tn, tetraodon; Sa, dogfish shark; Cm, elephant shark; Bf, lancelet; Sp, sea urchin; Lg, limpet; Tc, red flour beetle; Csp, polychaete worm annelid; Gc, sea water sponge; Ef, fresh water sponge; Hs, human. Asterisks and daggers indicate corrected and partial sequences, respectively. Two single-base ambiguities were resolved as follows: for Sa_Gα$_{v1}$, an ambiguous nucleotide (x) at the position 584 in the sequence DV496403 was replaced with cytosine (X333 to A333); for Gc_Gα$_v$, an ambiguous nucleotide (w) at the position 251 of the sequence Y14247 was replaced with thymine (X68 to L68). A potential pseudogene was rescued as follows: for Tn_Gα$_{v1}$, guanine between L180 and R181, adenine in G188, and guanine between E191 and T192 were removed. The sea urchin predictions Gα$_{oL}$ and Gα$_{oL2}$ (14) exhibit partial overlap in the C-terminal region, with 99% nucleotide identity and no difference in amino acid sequence, but exhibit different exons in the N-terminal region of the sequence. We conclude that Gα$_{oL}$ and Gα$_{oL2}$, which are assigned to the same contig, are different haplotypes of the same gene, a common assembly error in unfinished genomes constructed from two haplotypes of noninbred species. We therefore combined both predictions into a single prediction, Sp_Gα$_v$. Furthermore, we found two lancelet sequences with 97% nucleotide identity, but encoding the identical protein, Bf_Gα$_v$. We concluded that these sequences constitute 2 alleles of a single *gnav* gene locus as well. Tc_Gα$_v$, Ef_Gα$_{v1}$ and Gα$_{v2}$, and Gc_Gα$_v$ were renamed from Gα$_{i2}$, Gα$_6$ and Gα$_7$ (15), and Gα$_o$ (16), respectively.

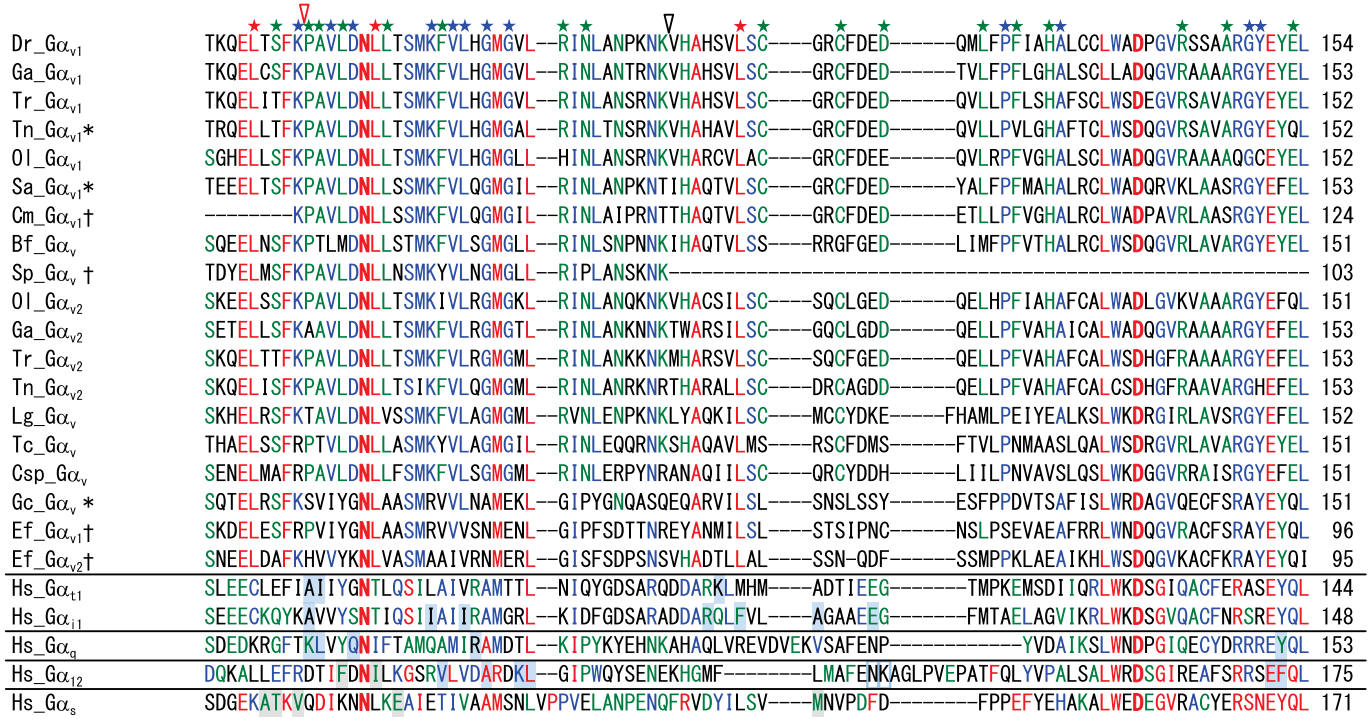


Fig. S1. continued

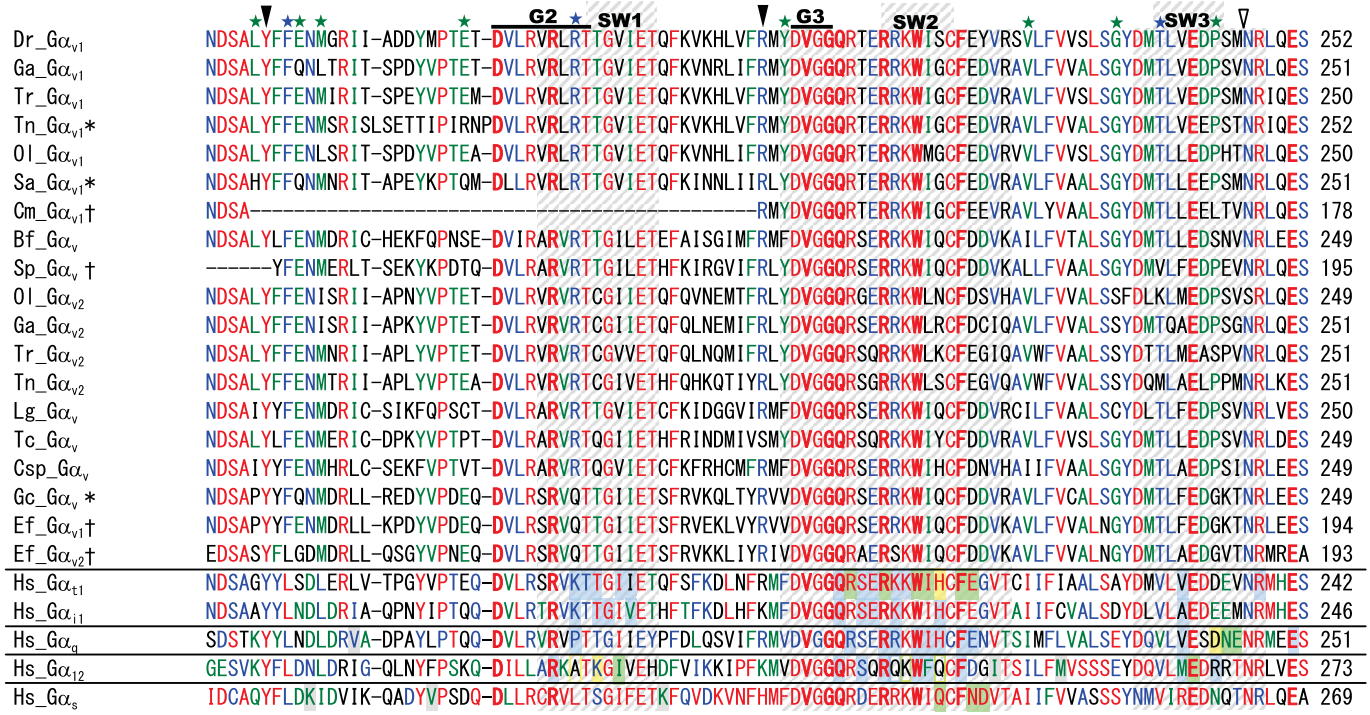
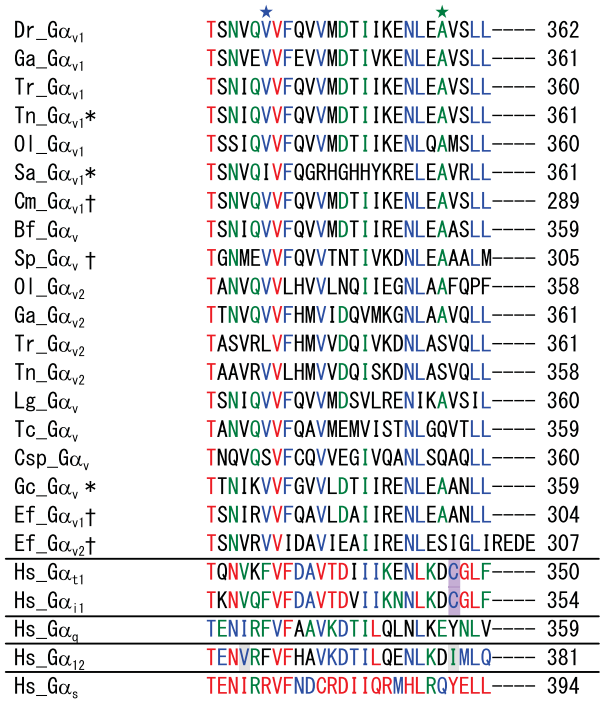


Fig. S1. continued

Fig. S1. continued



Conservation

Red Bold: conserved among all G alpha classes

Red: 100% conserved within a class

Blue: 80-99% conserved within a class

Green: 60-79% conserved within a class

Shaded in grey: atypical residue at the conserved position

★★★: Gv specific conservation (color coded as above)

Modification

Shaded in purple (N-terminal): acylation sites

(C-terminal): ADP-ribosylation site by Pertussis toxin

Contact sites

Shaded in blue: RGS/GAP contact sites

Shaded in green: effector contact sites

Shaded in yellow: both RGS/GAP and effector contact sites

Boxed: atypical residue at contact sites (color coded as above)

Splice sites

▽: junction between two codons

▼: junction within a codon

▽: Gv specific junction

Fig. S1. continued

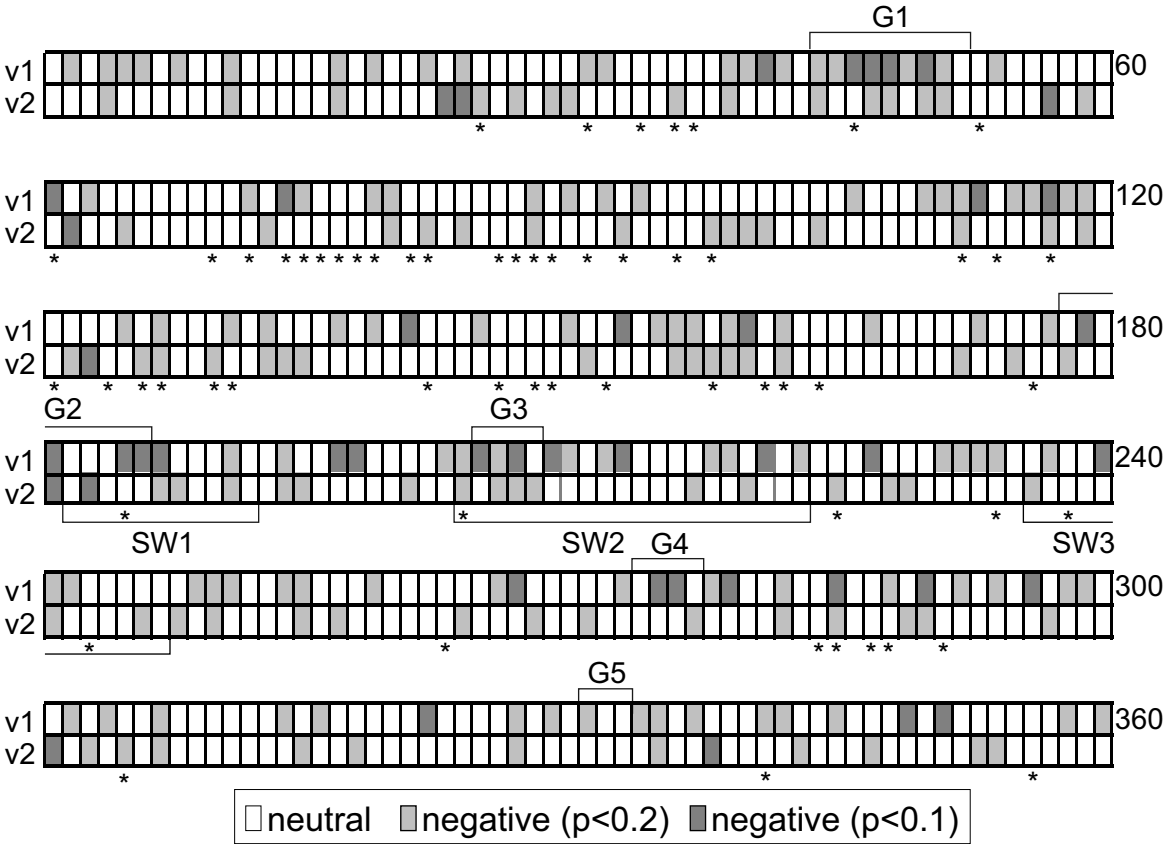


Fig. S2. Single-site d_N/d_S analysis of teleost *gnav* genes: sequence plot of neutral (open boxes) and negatively selected (shaded boxes) sites (codons) in each *gnav* family. Dark and light shading corresponds to significance levels of $P < 0.1$ and $P < 0.2$, respectively. Amino acid numbers are shown to the right. G-boxes and switch regions are indicated. Asterisks indicate Gv-specific motifs.

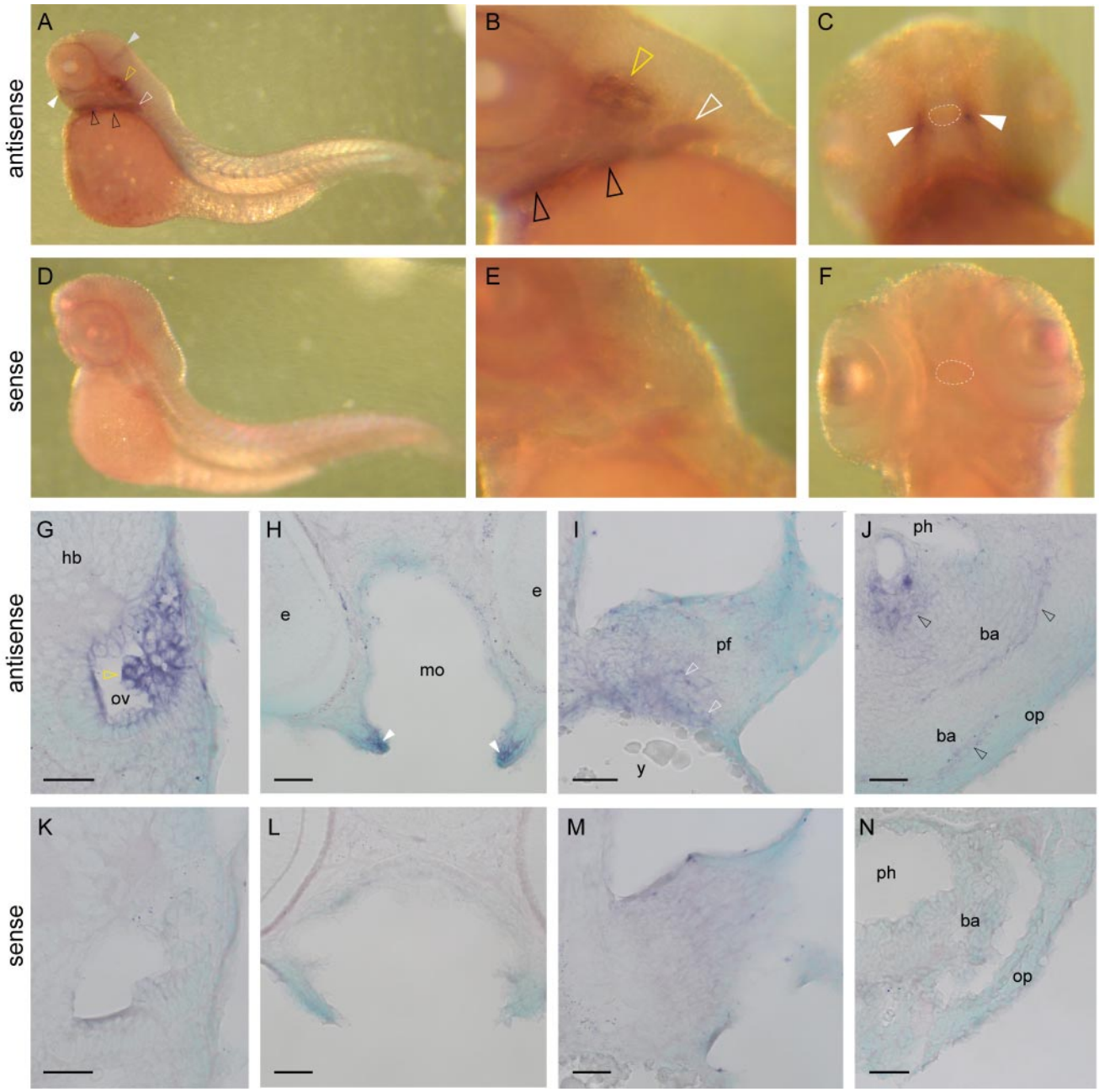


Fig. S3. Expression pattern of *gnav1* in zebrafish larvae: whole-mount in situ hybridization of *gnav1* probe (*A–C* and *G–J*) and negative control (sense probe, *D–F* and *K–N*) with 3 dpf zebrafish larvae. Embryos were grown in 0.0045% of phenylthiourea (Sigma) from 12 h postfertilization until used to prevent pigmentation. Probes were synthesized with T3 RNA polymerase (Roche) by following manufacturer's instructions. Templates for in vitro transcription were amplified using primers listed in Table S3. Probes were hybridized at 65 °C overnight. Specific probe hybridization was detected with anti-DIG antibody conjugated with alkaline phosphatase (1:5,000; Roche), using Nitroblue tetrazolium chloride/5-bromo-4-chloro-3-indolyl-phosphate (Roche) as substrates. Stained larvae were observed and photographed with a Nikon SMZ-U binocular and an attached Nikon CoolPix 950 digital camera. Stained larvae were cryo-sectioned (8 μm), counterstained in 0.001% methyl green, mounted with VectaMount (Vector), and documented on a Zeiss Axioplan microscope and an attached AxioCamMRc5. (*A* and *D*) Lateral view of whole larvae. (*B* and *E*) lateral view around the developing inner ear. (*C* and *F*) Ventral view of head region. Dotted circle, mouth. (*G–N*) Cross-sections of stained larvae at the levels of developing inner ear (*G* and *K*), mouth (*H* and *L*), pectoral fin (*I* and *M*), and branchial arches (*J* and *N*). Dorsal is to the top. White and gray arrowheads indicate the cell clusters near the lip and midbrain–hindbrain boundary, respectively. White, yellow, and black arrowheads point to labeled cells within pectoral fin (pf), otic vesicle (ov), and branchial arches (ba), respectively. e, eye; hb, hindbrain; mo, mouth cavity; op, operculum; ph, pharynx; y, yolk. (Scale bars: 50 μm.)

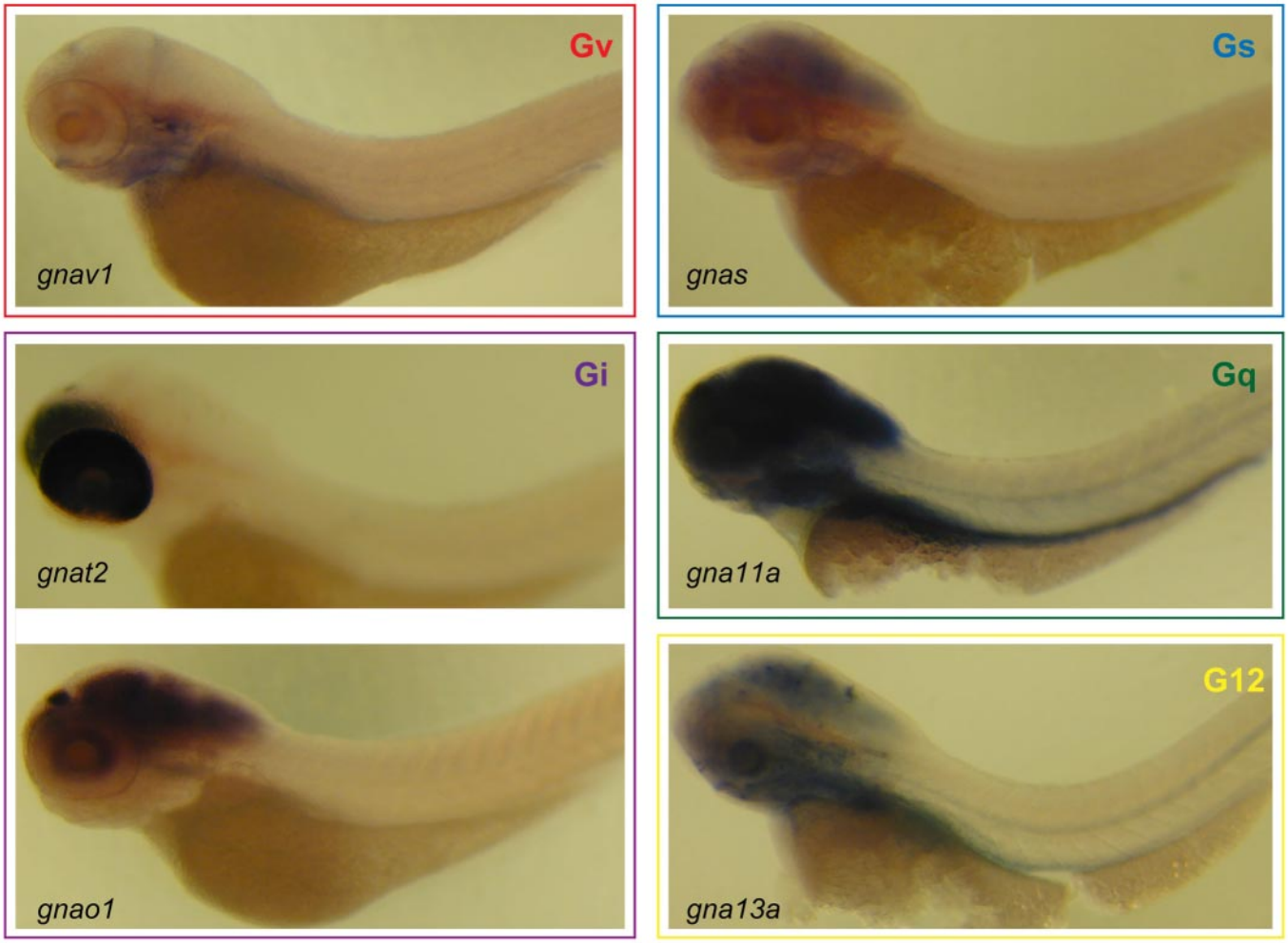


Fig. S4. Expression pattern of *gnav* compared to *gna* genes of the other 4 classes. Whole-mount in situ hybridization of *gnav*, *gnas*, *gnat2*, *gnao1*, *gna11a*, and *gna13a* probes with 3 dpf zebrafish larvae was performed as described in Fig. S3. All images are lateral views, and anterior is to the left. Note that expression patterns are characteristically different and none from the other 4 classes is similar to that of *gnav1*. Colored frames enclose genes from the same class (red, Gv; blue, Gs; purple, Gi; green, Gq; yellow, G12). Primers used to clone *gna* genes are as follows: gnas-fw, 5′-aagactgaggaccagcgaaa-3′; gnas-rv, 5′-gctggacaggctaactggac-3′; gnat2-fw, 5′-ctggtgaagctgccacagta-3′; gnat2-rv, 5′-gcttctctacaagcgccatt-3′; gnao1-fw, 5′-ccagtccaacgctgtctttt-3′; gnao1-rv, 5′-cgctccttgtctccgtactc-3′; gna11a-fw, 5′-cgatcaggttctggtggaat-3′; gna11a-rv, 5′-tgaaaggcgagttggagtct-3′; gna13a-fw, 5′-agaaactgcacatcccttgg-3′; and gna13a-rv, 5′-ttttggctgggcaagtagtc-3′.

Table S1. List of teleost *gnav* genes

Gene	Protein	Ensembl	Chromosome	Start	End	Orientation
Dr_gnav1	Dr_Gα_{v1}	ENSDARG00000043006†	22	9,130,267	9,186,364	−
Ol_gnav1	Ol_Gα_{v1}	ENSORLG00000012122	1	31,556,132	31,583,202	−
Ol_gnav2	Ol_Gα_{v2}	ENSORLG00000007929	8	10,625,829	10,632,451	+
Ga_gnav1	Ga_Gα_{v1}	ENSGACG00000019015†	Group IX	15,176,864	15,193,316	−
Ga_gnav2	Ga_Gα_{v2}	ENSGACG00000011158†	Group XI	9,238,874	9,244,473	+
Tr_gnav1	Tr_Gα_{v1}	SINFRUG00000129994‡	Scaffold 189	438,137	448,743	−
Tr_gnav2	Tr_Gα_{v2}	SINFRUG00000135273‡	Scaffold 115	694,830	698,835	+
*Tn_gnav1**	Tn_Gα_{v1}	GSTENG00026255001‡	18	2,116,654	2,124,011	−
Tn_gnav2	Tn_Gα_{v2}	GSTENG00018570001‡	3	10,269,269	10,272,711	+

The gene list of teleost *gnav* orthologs is shown. Gene identification numbers (IDs) and locations given here are derived from Ensembl release 48 (December 2007). Dr, zebrafish; Ol, medaka; Ga, stickleback; Tr, fugu; Tn, tetraodon. Note that a rare alternative start codon is used for Tn_Gα_{v2} and that a noncanonical splicing donor sequence (GC) instead of the canonical one (GT) is used for Tr_Gα_{v2} in intron 3 and for Tn_Gα_{v2} in introns 3 and 8. Other sequences used for phylogenetic analysis (Fig. 1) are as follows: XM_968117 (Tc_Gα_{v}, previously predicted as Gα_{i2}); AB006548 (Ef_Gα_{v1}, previously named Gα_{6}); AB006549 (Ef_Gα_{v2}, previously named Gα_{7}); Y14247 (Gc_Gα_{v}, previously reported as Gα_{o}); EE049147 and DV496403 (Sa_Gα_{v1}); AAVX01181130, AAVX01051239, AAVX01180564, AAVX01051240, AAVX01416005, AAVX01144285, and AAVX01077290 (Cm_Gα_{v1}); JGI protein ID 57428 (Bf_Gα_{v}); SPU_024792 and SPU_024793 (Sp_Gα_{v}, combined from previously named Gα_{oL} and Gα_{oL2}); JGI protein IDs 185322 and 227716 (Lg_ Gα_{v} and Csp_Gα_{v}); XP-685500 (Dr_Gα_{s}); AAS92627 (Dr_Gα_{olf1}); ENSDARG00000045415 (Dr_Gα_{olf2}); ENSDARG00000044199 (Dr_Gα_{t1}); ENSDARG00000042529 (Dr_Gα_{t2}); ENSDARG00000016676 (Dr_Gα_{o1}); ENSDARG00000036058 (Dr_Gα_{o2}); ENSDARG00000021647 (Dr_Gα_{i1a}); ENSDARG00000044760 (Dr_Gα_{i1b}); ENSDARG00000017294 (Dr_Gα_{i2a}); NP_001001818 (Dr_Gα_{i2b}); ENSDARG00000030644 (Dr_Gα_{i3}); ENSDARG00000069358 (Dr_Gα_{z}); ENSDARG00000011487 (Dr_Gα_{q}); ENSDARG00000053326 (Dr_Gα_{11a}); ENSDARG00000010002 (Dr_Gα_{11b}); ENSDARG00000025013 (Dr_Gα_{14a}); BC077106 (Dr_Gα_{14c}); BC077141 (Dr_Gα_{15a}); ENSDARG00000056654 (Dr_Gα_{15b}); ENSDARG00000031543 (Dr_Gα_{15c}); ENSDARG00000063231 (Dr_Gα_{15d}); ENSDARG00000025826 (Dr_Gα_{12}); NP_001012243 (Dr_Gα_{13a}); ENSDARG00000037924 (Dr_Gα_{13b}); NP_000507 (Hs_Gα_{s}); NP_002062 (Hs_Gα_{olf}); NP_000163 (Hs_Gα_{t1}); NP_005263 (Hs_Gα_{t2}); NP_001095856 (Hs_Gα_{t3}); NP_066268 (Hs_Gα_{o}); NP_002060 (Hs_Gα_{i1}); NP_002061 (Hs_Gα_{i2}); NP_006487 (Hs_Gα_{i3}); NP_002064 (Hs_Gα_{z}); NP_002063 (Hs_Gα_{q}); NP_002058 (Hs_Gα_{11}); NP_004288 (Hs_Gα_{14}); NP_002059 (Hs_Gα_{15}); NP_031379 (Hs_Gα_{12}); NP_006563 (Hs_Gα_{13}); NP_477502 (Dm_G-iα65A); NP_523684 (Dm_G-oα47A); NP_725191 (Dm_dgq); NP_477506 (Dm_G-sα60A); NP_524118 (Dm_Gfα); NP_001036421 (Dm_cta); CG30054 (Dm_CG30054); CG17760 (Dm_CG17760); CG40005 (Dm_CG40005); GCY14248 (Gc_Gα_{q}); GCY14249 (Gc_Gα_{s}); AB006543 (Ef_Gα_{1}); AB006544 (Ef_Gα_{2}); AB006545 (Ef_Gα_{3}); AB006546 (Ef_Gα_{4}); AB006547 (Ef_Gα_{5}); AB006550 (Ef_Gα_{8}); and AB006551 (Ef_Gα_{9}).

*A possible pseudogene (see Fig. S1).

†The predictions that lack a small part of the protein sequence (mostly N-terminal).

‡The predictions that lack ≥ 1 exons.

Table S2. List of EST clones for teleost *gnav* genes

Gene	EST accession	Length (bp)	Coverage of *gnav* CDS (%)	Nucleotide identity (%)
Dr_gnav1	CT736382.2	817	74	99
	EE310328.1	848	73	99
	EB899666.1	650	58	99
	AL916853.1	539	18	98
	DT073850.1	853	14	100
	CD595967.1	650	19	90
	BI983956.1	569	13	100
	BG799669.1	444	10	98
	CK674763.1	447	4	100
Ol_gnav1	AM344465.1	815	68	99
	AM304638.1	770	64	99
	AM303462.1	758	61	99
	AM317248.1	737	59	99
	AM336808.1	735	58	99
	AM326993.1	689	55	99
	AM353138.1	700	54	99
	AM315079.1	510	47	99
	AM367625.1	577	43	100
	AM370671.1	350	26	100
	AM322694.1	369	27	98
	AM383108.1	402	26	96
Ol_gnav2	AM149874.1	494	43	99
Ga_gnav1	DN665837.1	1323	81	93
	DN662698.1	1371	77	94
Tr_gnav2	CA846482.1	795	32	99

The list of EST clones for teleost *gnav* genes identified from the NCBI database is shown. Dr, zebrafish; Ol, medaka; Ga, stickleback; Tr, fugu.

Table S3. List of primers used for RT-PCR, cloning for full-length *gnav1*, and amplification of in vitro transcription templates

Primer	Sequence	Use
Dr-gnav1-fw	5′-gtgtggccgttgttttgatga-3′	RT-PCR
Dr-gnav1-rv	5′-catgtcgtatccactcagag-3′	RT-PCR
Dr-b-actin-fw	5′-ccccattgagcacggtatt-3′	RT-PCR
Dr-b-actin-rv	5′-agcggttcccatctcctg-3′	RT-PCR
Dr-gnav1–5′UTR-fw	5′-ccaactggacctttagctcttc-3′	Full-length cloning
Dr-gnav1–3′UTR-rv	5′-ccagcatatgcttggctcatc-3′	Full-length cloning
gnav1-N-fw	5′-atgggtctgtgtttgggctc-3′	Template amplification
gnav1-N-rv	5′-catcaaaacaacggccacac-3′	Template amplification
gnav1-M-fw	5′-ggtgtgcgttcgtctgcggc-3′	Template amplification
gnav1-M-rv	5′-catgtcgtatccactcagag-3′	Template amplification
T3-promoter site	5′-tattaaccctcactaaagggaa-3′	Template amplification

Reverse (rv) or forward (fw) primers were attached with the T3 RNA polymerase promoter site at their 5′ end for the anti-sense or sense probes, respectively.