Deciphering the function and evolutionary context of a novel G alpha protein, Gv

Inaugural-Dissertation

Zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Asmaa Abu Obaid

Köln, May 2022

Berichterstatter (Gutachter): Prof. Dr. Sigrun Korsching

Prof. Dr. Niels Gehring

Tag der mündlichen Prüfung: 31.05.2022

Printed with the support of the German Academic Exchange Service

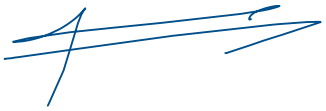
Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.

Teilpublikationen:

Datum, Name und Unterschrift

11.04.2022, Asmaa Abu Obaid



Dedication

To my father's soul, Abdellatif Abu Obaid

To my beloved mother, Muna Abu Obaid

To my soul mate, Shadi Allabadi

To my beloved children, Abdallah, Retal, and Eman

To the person who enlightened me in science, Sigrun Korsching

I affectionately dedicate this work to you

Acknowledgement

Foremost, my distinct thanks and deepest gratitude go to “*the mother of science*” my marvellous supervisor Prof. Sigrun Korsching for the extraordinary supervision and manifold support in the implementation of the entire work of my thesis. I have been honoured to work under your supervision. I will always appreciate that you were extremely supportive, cooperative, insightful to enlighten my scientific career, and you were appreciative to my work all along the way. I am deeply grateful for the very long meetings (about three hours and sometimes ran to seven hours and until 11 pm) you have spent for scientific discussions from which I have learned a lot. I ascertain my gratitude and appreciation for helping me develop my critical thinking and scientific writing skills and for many hours of proofreading which enabled to sculpt my thesis. I will always be grateful for you being a constant source of knowledge, *caring*, and inspiration. Without your sustained scientific and personal support this project would not have been made possible.

It is with warmest affection I express my indebtedness to my soul mate, Shadi Allabadi, for being there for me, being patient and continuously supportive throughout entire my doctorate program. I owe you huge respect and gratefulness for your words of encouragement and push for tenacity; you have been my best cheer leaders. You have given me the drive and discipline to tackle the task with enthusiast and determination. You are always source of motivation, enthusiasm, and generosity of spirit that assisted me surmounts all difficulties during my studying time. For that all, and for the eternal love, taking care of the kids, and believing in me and my work, special thanks to You. Also, I am obligated to extend my gratitude to my kids Abdalla, Retal, and Eman for their understanding, being patient, and responsible during the time I've been busy and away. However, from now on I will compensate all the time I was mentally and physically away from You.

My recognition and gratitude are addressed to those who presided at the heart of this work. To Prof. Niels Gehrings, many thanks for having been a member of my thesis committee, for evaluating this work, and for your valuable insights and support during my doctorate program; also, for allowing me to use your lab resources. I would also like to thank Prof. Tobias Bollenbach for chairing my thesis defense. Also, I am very grateful to Dr. Isabell Witt, the GSfBS graduate school coordinator, who introduced me to the Korsching lab and supported me and my family during my doctoral program.

I am deeply grateful to the people whose assistance was a milestone in the completion of this project: Kim Korsching who kindly provided the script file which was very useful in the bioinformatic analysis; Dr. Thomas Hermanns who kindly generated the

recombinant protein (Gv); Dr. Stephan Optiz who measured the cation levels by atomic absorption spectrometry; the CECAD proteomic facility who performed the MS assays; Dr. Volker Boehm who introduced me to western blotting and kindly assisted in troubleshooting cloning and *in vitro* transcription; Cihan Yilmaz who provided great assistance in qPCR experiments; and Lani Folkenstern who kindly provided the embryos for WISH experiment.

I wish to extend my gratefulness to my lab mate (dude) Dr. Kanika Sharma for the scientific discussions and rich feedback; and also, for the great fun we had throughout the process of this dissertation. Also, thanks to Shahzaib Hassan for the help in lab work and in taking care of the *gnav1*$^{-/-}$ line during my absence. My thanks also go to all the current and former Korsching lab members Mehmet Saltürk, Günes Birdal, Adnan Alsayed, Manish Tomar, and Daniel Kowatschew for their cooporation.

I owe profound gratefulness and sincerest appreciation to my wonderful and supportive family, especially my mother Muna Abu Obaid, for her unconditional love, spiritual support, and encouraging me to always seek for outstanding education. Sincerest appreciation for my brothers Ahmad, Mohammad, Mahmoud, and Hassan Abu Obaid for the considerable help and support in the technical and computer work. My warm gratitude to my sisters Alaa, Maryam, and Maram Abu Obaid for their moral support, enormous love, and for being a source of strength to me.

I would also like to extend my gratefulness to my parents in law Abdallah Ahmad Allabadi and Eman Mohammad Ameen for their caring and continuous support.

It is a great honour to thank those who made this thesis possible. At this point I would like to express my gratitude to all those who were involved and who supported me in the preparation of my doctoral thesis and who may not be mentioned by name here!

Abstract

Heterotrimeric G proteins act as molecular switches inside cells and control a wide range of systematic functions i.e., embryonic development, learning, homeostasis, etc. Loss or gain of function mutations in G alpha proteins result in signalling defects that lead to pleiotropic manifestations; in addition, they are outstanding pharmaceutical targets. Despite the importance of the G alpha subunits in all biological systems, so far there have been no published data about the function of the most recently discovered G alpha family, Gv. Up to now, nothing is known about its role, effector proteins, and signalling pathway, despite Gv being present in zebrafish, a popular vertebrate model organism. Thereby, the purpose of this study is to decipher the so far uncharacterized Gv protein role, particularly in zebrafish. Moreover, the present study was also designed to examine Gv evolutionary origin and dynamics in relation to the other four G alpha families, since only scarce and partially inconsistent information was available for the birth of Gv in the earliest-diverging, unicellular Holozoa, the Teretosporea and no systematic study of its evolutionary dynamic in relation to the other families has been performed.

Firstly, we investigated the ontogenetic and tissue-specific expression of Gv mRNA during embryogenesis and larval development by the histochemical technique wholemount in situ hybridization (WISH). Secondly, we evaluated phenotypic and molecular changes in *gnav1*$^{-/-}$ mutants with an early 13 bp deletion in the mutant *gnav1*. Thirdly, to address the evolutionary origin and dynamics of Gv, also in relation to the other four families, we constructed phylogenetic trees by rigorously searching the five G alpha families in NCBI genomic databases of nearly a thousand species from many major phylogenetic subdivisions in Holozoa.

Our data show for the first time that *gnav1* exhibits broad expression in early embryonic stages starting from gastrulation and becomes confined to specific tissues in larval stages. These expression patterns indicate manifold functions for Gv in embryonic organogenesis including sensory and osmoregulatory organs, cartilaginous tissue and muscle. We also demonstrate altered phenotypes in *gnav1*$^{-/-}$ mutants that include: the number of eggs produced per clutch is markedly decreased; the mutant embryos hatch significantly earlier than the wildtype siblings; three and five dpf mutants show craniofacial defects and overall decrease in alcian blue staining; five and 11 dpf mutant larvae show prominently reduced mineralization in the craniofacial bones. The 5 dpf mutant larvae exhibit highly significant reduction in calcium, magnesium, and potassium levels. The mutant kidneys show significantly increased mRNA levels of NCC and NKA.5 which play a role in calcium and magnesium homeostasis. Taken together, our data suggest novel functions of Gv in cation homeostasis, and subsequently bone and cartilage formation, hatching process, and oviposition.

Our phylogenetic studies reveal the birth of the Gv family in the common ancestor of all Holozoa, i.e., before the divergence into Teretosporea and Filozoa. Moreover, we show that Gv evolutionary dynamics are within the range of those observed for the other four G alpha classes. Furthermore, we have identified two novel G alpha clades, i.e., Gf and Gw, which have not been described before.

Contents

6 List of Figures

7 List of Tables

8 List of Abbreviations

Abbreviations are explained in first occurrence in the text except gene names. In order to avoid distracting the reader they are explained here.

°C	Degrees Celsius
µl	Microgram
µl	Microliter
AB/Tüb	Ab/tübingen
ABC	Ammonium bicarbonate
AC	Adenylyl cyclase
ADGRG	Adhesion g protein-coupled receptor g1
AE1	Anion exchanger 1b
AFC	Actinotrichia forming cells
AGS	Activators of G protein signaling
aLRT	Approximate likelihood-ratio test
AP	Anterior-posterior
ARS	Alizarin red s
ATP	Adenosine triphosphate
Atp1a1a	Atpase Na^+/K^+ transporting subunit alpha 1a, tandem duplicate
b2m	Beta2-microglobulin
bactin	Beta actin
BCIP	5-Bromo-4-chloro-3-indolylphosphate
bp	Base pair
BSA	Bovine serum albumin
CA	Carbonic anhydrase 2-like a
ca2	Carbonic anhydrase 2
CA2 (-15)	Carbonic anhydrase 2-like a (-15a)
$Ca2^+$	Calcium
CaCl2.2H2O	Calcium chloride dihydrate
Calcr	Calcitonin receptor
calcrla	Calcitonin like receptor a
Calcrlb	Calcitonin like receptor b
cAMP	Cyclic adenosine monophosphate
Cas9	CRISPR-associated 9
Cas9-NLS	CRISPR-associated 9-nuclear localization signal
CasR	Calcium sensing receptor

CCL	Ceratohyal cartilage length
cDNA	Complementary DNA
ClC2	Cl^- channel 2c
cm	Centimeter
CNCC	Cranial neural crest cell
CNG	Cyclic nucleotide gated ion-channel
CRISPR	Clustered regularly interspaced short palindromic repeats
DAG	Diacylglycerol
DAPI	4',6-diamidino-2-phenylindole
DE	Distal
DIG	Digoxigenin
DL	Distal late
DMSO	Dimethyl sulfoxide
DNA	Desoxyribonucleic acid
dNTP	Deoxyribose nucleotide triphosphate
dpf	Days post fertilization
DTT	Dithiothreitol
E2	Embryo 2
E3	Embryo 3
ECaC	Epithelium calcium channel
ECM	Extracellular matrix
EDTA	Ethylene diamine tetraacetic acid
eGFP	Enhanced green fluorescent protein
EJC	Exon junction complex
elf1	Eef1a1l1 eukaryotic translation elongation factor 1 alpha 1, like 2
EST	Expressed sequence tag
g	Gram
G Protein	Guanin nucleotide-binding protein
G11	G alpha 11
G12	G alpha 12
G13	G alpha 13
G14	G alpha 14
GABABR1	Gamma-aminobutyric acid B receptor, 1
GAP	Gtpase-activating protein
GC	Guanine-cytosine
GDI	Guanine nucleotide dissociation inhibitors
GDP	Guanosine diphosphate

GEF	Guanine nucleotide exchange factor
GFP	Green fluorescent protein
Gg	G alpha guest
Gi	G alpha I
Gnai11a	G-alpha q
gnav	Guanine nucleotide binding protein (G protein) alpha v1
Go	G alpha o
Golf	G alpha olf
GPA	G protein alpha
GPCR	G-protein coupled receptor
GPKs	GPCR kinases
GPR	G protein-coupled receptor
Gq	G alpha q
GRAFS	Glutamate, Rhodopsin, Adhesion, Frizzled/Taste2, and Secretin
GRCz11	Genome Reference Consortium of zebrafish 11
gRNA	Guidance ribonucleic acid
Gs	G alpha s
Gt	G alpha t
GTP	Guanosine triphosphate
Gv	G alpha v
GXL	G alpha extra long
Gz	G alpha z
H2O	Water
H2O2	Hydrogen peroxide
HA	H^+-atpase-rich ionocyte
HCl	Hydrochloric acid
$HCO3^-$	Bicarbonate
His-GST	Histidine-Glutathione Sepharose Tag
hpf	Hours post fertilization
HR	H^+-atpase-rich ionocyte
ICD	Intercranial distance
ICP-OES	Inductively coupled plasma-optical emission spectroscopy
IF	Immunofluorescence
IP3	Inositol triphosphate
IPTG	Isopropyl β- d-1-thiogalactopyranoside
ISH	In situ hybridization
JAKs	Janus kinases

K^+	Potassium
KC	Keratinocyte
KCl	Potassium chloride
kDa	Kilo dalton
KH2PO4	Monopotassium phosphate
Kir1.1	Renal outer medullary potassium channel
KOH	Potassium hydroxide
KS	K^+-secreting ionocyte
KS strain	Köln strain
L	Liter
LB	Lysogeny broth
LGR	Leucine-rich repeat-containing G-protein coupled receptor
LH	Luteinizing hormone
LiCl	Lithium chloride
LJL	Lower jaw length
LPM	Lateral plate mesoderm
M	Molar
MAFFT	Multiple alignment using fast Fourier transform
MBT	Midblastula transition
MFP	Maruration-promoting factor
MgCl2	Magnesium chloride
MgSO4. 7H2O	Magnesium sulphate heptahydrate
MHB	Midbrain-hindbrain boundary
ml	Milliliter
mM	Millimolar
MRCA	Most recent common ancestor
mRNA	Messenger ribunucleic acid
MS	Mass spectrometry
MSA	Multiple sequence alignment
MyoD	Myoblast determination protein 1
Na^+	Sodium
Na2HPO4	Disodium phosphate
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NaR	Na^+-K^+-atpase rich cell
NBC	Electrogenic Na^+-$HO3^-$ cotransporter 1b
NBT	Nitro blue tetrazolium chloride

NCA.1	Na^+-K^+-atpaseα1-subunitsubtypes(atp1a1a.1)
NCBI	National Center for Biotechnology Information
NCC	Na^+-Cl^- cotransporter-expressing ionocyte
NCX	Na^+/$Ca2^+$exchanger
NCX	Na^+/$Ca2^+$ exchanger 1b
Ncx1b2	Solute carrier family 8-member 1b
ng	Nanogram
$NH4^+$	Ammonium ion
NHE3	Na^+/H^+ exchanger 3b
NKA	Na^+-K^-atpase
nm	Nanometer
NMD	Nonsense mediated decay
NRT	No reverse transcriptase
OD	Optical density
ORF	Open reading frame
P	Pair
PBS	Phosphate buffered saline
PBST	Phosphate-buffered saline tween
PCP	Planar cell polarity
PCR	Polymerase chain reaction
PCT (kidney)	Proximal tubule
PCT (kidney)	Proximal convoluted tubule
PFA	Perfluoralkoxy-polymere
PhyML	Phylogenetic estimation using (Maximum) Likelihood
PIP2	Phosphatidylinositol 4,5-bisphosphate
PKA	Protein kinase A
PKC	Protein kinase C
PLC	Phospholipase C
Plcb1	Phospholipase C-betta 1
Plcg1	Phospholipase C-gamma 1
PMCA	Plasma membrane $Ca2^+$-atpase
PRM	Parallel monitoring reaction
PST	Proximal strait tubule
PTH	Parathyroid peptide hormone
PTU	Phenylthiourea
qPCR	Quantitative polymerase chain reaction
RBP	RNA binding protein

RGS	Regulators of G protein signaling
Rhcg1	Rhesus glycoprotein
RIPA buffer	Radioimmunoprecipitation assay buffer
RNA	Ribonucleic acid
ROMK	An ortholog of the mammalian renal outer medullary K^+ channel (Kir1.1)
rpl37	Ribosomal protein l37
rpl8	Ribosomal protein l 8
rpm	Rounds per minute
RT-PCR	Reverse transcriptase PCR
SD	Standard deviation
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEM	Standard error of the mean
Slc12a3	Solute carrier family 12 (Sodium/Chloride Transporter), subfamily A, Member 3
SLC26	Solute carrier 26-expressing ionocyte
Slc9a3.2	Solute carrier family 9, subfamily A, member 3, tandem duplicate 2
SPR	Subtree pruning and regrafting
SSC	Saline sodium citrate
STATs	Signal transducer and activator of transcription proteins
Stc-1	Stanniocalcin1
T	Tween
TAE	Tris-acetate-EDTA
TBS	Tris buffered saline
TBST	Tris-buffered saline, 0.1% Tween 20
TEAB	Triethylammonium bicarbonate
TEMED	Tetramethyl ethylene diamine
TM	Transmembrane
TMEM165	Transmembrane protein 165
TPM	Transcript per million
TRI	Trizol
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer RNA
TRPV5/6	Transient receptor potential cation channel subfamily5/6
UTR	Untranslated region
UV	Ultraviolet
V	Voltage
Vdra	Vitamin D receptor a

WGS	Whole genome shotgun
WISH	Whole mount in situ hybridization
XL	Extra long
YSL	Yolk syncytial layer
ZGA	Zygotic genome activation

1 Introduction

1.1 G protein-coupled receptors

Receiving information from the extracellular environment and signal transmission within the body are fundamental for all living organisms. The signals are conveyed by many biologically active molecules through receptors. G protein-coupled receptors (GPCRs) are the largest family of membrane receptors in animals and they have a pivotal role in a multitude of biological systems (Fredriksson & Schiöth, 2005). The seven transmembrane (TM) helical domains which transverse through the plasma membrane are structural hallmarks of GPCRs which have been conserved over the course of evolution (Fredriksson & Schiöth, 2005; D. Yang et al., 2021). In contrast, extracellular N-terminus and intracellular C-terminus show high variability (Lagerström & Schiöth, 2008). Various GPCRs can recognize a tremendous variety of signaling molecules. Signaling through GPCRs is initiated by binding the specific ligands to either an extracellular domain (e.g., glutamate receptors) or within transmembrane helices (e.g., rhodopsin receptors), or upon heterodimerization (e.g. the glutamate receptors GABABR1 and GABABR2 subunits) (Lagerström & Schiöth, 2008; Trzaskowski et al., 2012). Some of these ligands are chemokines, peptides, hormones, growth factors, neurotransmitters, tastants and odorants; even light can activate some GPCRs, e.g., rhodopsin (Figure 1) (Trzaskowski et al., 2012).

The size of the GPCR superfamily is extremely huge as they constitute 4% of the entire human protein-coding genome; so far about thousand human GPCRs have been discovered (D. Yang et al., 2021). The superfamily of GPCRs is divided into five main families according to the recent GRAFS classification system which is based on phylogenetic analysis: rhodopsin, secretin, adhesion, glutamate, and frizzled (consisting of frizzled and smoothened) which correspond to class A, B1, B, C, and F, respectively (D. Yang et al., 2021). The largest family and most extensively investigated is the rhodopsin family, which contains the eponymous rhodopsin and a

multitude of neurotransmitter receptors, which show characteristic sequence differences. They regulate a wide range of physiological functions (including various autocrine, paracrine, and endocrine processes) (D. Yang et al., 2021). The adhesion family is the second largest; these receptors have important roles in cell adhesion and migration, and they possess long and diverse N-terminal motifs which determine the specificity of the receptor–ligand binding interactions (D. Yang et al., 2021). The receptors of the glutamate family have been shown to be constitutive dimers and possess large extracellular domains; they perform a variety of different functions including neurotransmission (glutamate receptor) and ionic homeostasis (e.g., calcium-sensing receptor (CaSR)) (D. Yang et al., 2021). Receptors of the secretin family harbor extracellular hormone-binding domain and bind peptide hormones e.g., parathyroid peptide hormone (PTH) (D. Yang et al., 2021). The frizzled receptors share a conserved cysteine-rich domain in the extracellular region; they bind Wnt glycoproteins and other molecules and mediate Wnt signaling in addition to other pathways e.g., planar cell polarity (PCP) (D. Yang et al., 2021). The frizzled receptors are essential for embryonic development and adult organisms as they play a key role in cell proliferation, cell polarity, cell migration, and cell fate determination (Clevers, 2006; Eisenmann, 2005; Gruber et al., 2016).

GPCRs are hugely important since they regulate a myriad of cellular and physiological processes as varied as bone development and remodeling (e.g., PTH receptor), ionic homeostasis (e.g., CaSR), neurotransmission (e.g., dopamine receptor), and immune response (e.g., chemokine receptors) (Hariharan et al., 2020; D. Yang et al., 2021). Furthermore, GPCRs are involved in a plethora of diseases such as bone diseases, male infertility, diabetes, obesity, cancer, asthma, rheumatoid arthritis and Alzheimer disease, and are substantially the target of approximately half (53%) of all modern medicinal drugs (Bassilana et al., 2019; Hauser et al., 2017). Mutation in some GPCRs induces phenotypes in numerous and various tissues. For instance, alteration in the adhesion GPCR, ADGRG6 or GPR126, is associated with adolescent idiopathic scoliosis, arthrogryposis, intellectual disability, diffusing capacity of the lung, urothelial bladder carcinomas, and thinning of the myocardial wall; these diseases affect the musculoskeletal, nervous respiratory, urinary, and circulatory systems in human

(Bassilana et al., 2019; Hauser et al., 2017). Mutation in some GPCRs induces phenotypes in numerous and various tissues. For instance, alteration in the adhesion GPCR, ADGRG6 or GPR126, is associated with adolescent idiopathic scoliosis, arthrogryposis, intellectual disability, diffusing capacity of the lung, urothelial bladder carcinomas, and thinning of the myocardial wall; these diseases affect the musculoskeletal, nervous respiratory, urinary, and circulatory systems in human (Bassilana et al., 2019). Given the importance of GPCRs in both pathological conditions and treatment of disease, it is critical that we comprehensively understand these receptors and their downstream signaling components. Despite the fact that GPCRs have been studied extensively in the past years, many of these GPCRs signaling pathways and their G proteins interactors remain to be addressed.

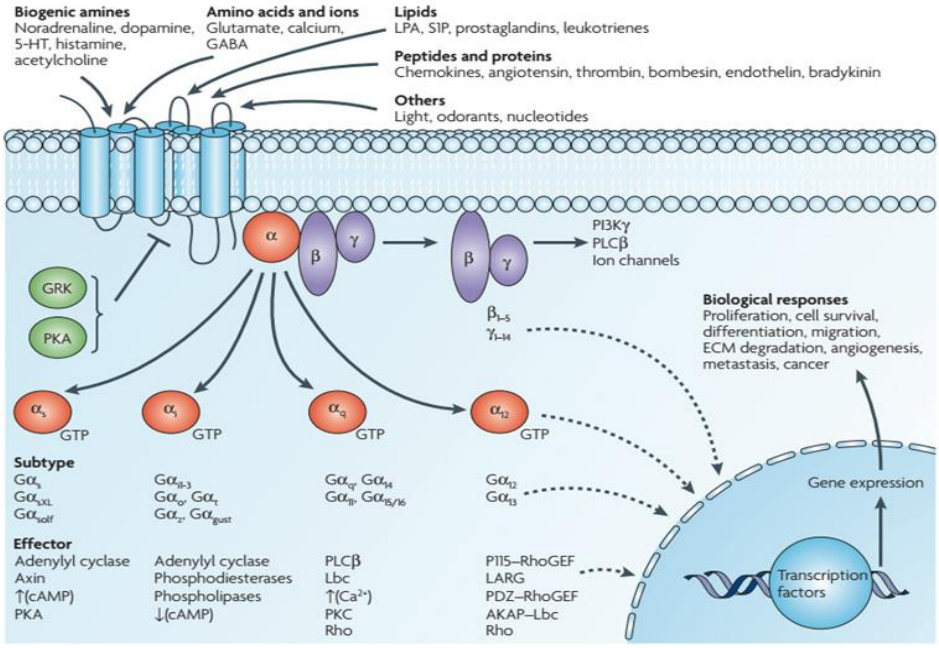


Figure 1: The diverse GPCR signaling cascades within a cell can interact to affect multiplemolecules in the cell, leading to various biological responses. Figure is taken from (Dorsam & Gutkind, 2007).

1.2 GPCR signaling through heterotrimeric G proteins

In the classical paradigm, GPCRs signal transduction involves the activation of downstream heterotrimeric G proteins (Hilger et al., 2018). There are three subunits in heterotrimeric G proteins alpha, beta, and gamma which exist as small subfamilies of one to two dozen members in contrast to the large repertoire of GPCRs, see e.g., (Oka & Korsching, 2011); these genes combine in multiple ways to produce a diverse family of trimeric G proteins.

In the absence of a signal, guanosine diphosphate (GDP) binds the inactive G alpha, and the entire G protein-GDP complex binds to a nearby GPCR. Upon receiving the extracellular signals (diverse ligands and induction mechanisms), conformational changes are induced within the receptor. This is a common feature in all GPCR signaling (Hilger et al., 2018). These changes route the signal to downstream

heterotrimeric G proteins or also other molecules by allowing the receptor to act as a guanine nucleotide exchange factor (GEF) (Weis & Kobilka, 2018). The GEF domain catalyzes the (inactive) G proteins to exchange GDP bound to the G alpha subunit with guanosine triphosphate (GTP) (Figure 2 and Figure 3) (Hilger et al., 2018). The GTP-bound G alpha (active) dissociates from and thereby frees G beta/gamma, resulting in two functional subunits (G alpha and G beta/gamma) (Figure 2) (Hilger et al., 2018). Both signaling components transduce signals by tremendous amounts of different intracellular signaling cascades e.g., adenylyl cyclase (AC), phospholipase C (PLC), or ion channels (Goldsmith & Dhanasekaran, 2007).

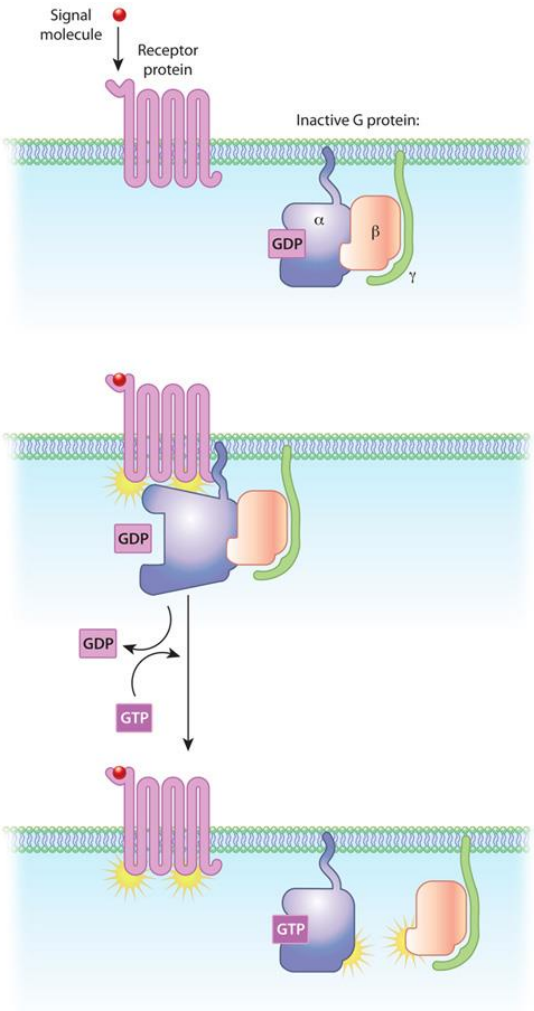


Figure 2: Schematic representation shows GPCR activation by signal molecule and transducing the signal through activating heterotrimeric G proteins. Figure is taken from (O'Connor & Adams, 2010).

It has been shown that different receptors couple to different alpha, beta, and gamma subfamily members, and these in turn activate different effector cascades; some GPCRs activate only one family of G proteins, whereas other receptors may activate multiple families (Dorsam & Gutkind, 2007; Okashah et al., 2019). These pathways interact with each other to form a network that regulates multiple components of the cellular machinery controlling a wide range of cellular processes, which in turn regulate systemic functions such as embryonic development (Figure 4) (Dorsam & Gutkind, 2007; O'Connor & Adams, 2010). In addition to GPCR-mediated activation, cytoplasmic proteins can also activate G protein signaling in GPCR-independent manner; such as activators of G protein signaling (AGS, guanine nucleotide

dissociation inhibitors (GDI)) that block nucleotide exchange (Figure 3) (Blumer et al., 2012). Recent study showed that non-GPCR proteins which have GEF activity also trigger G protein signaling in cells, and they proposed that signaling is activated by efficient release of G beta/gamma subunits rather than generation of G alpha-GTP (Garcia-Marcos, 2021). The discovery of non-GPCR activators for G proteins led to perceiving that G alpha and G beta/gamma may exist associated with alternative binding partners independent of the classical heterotrimer, thus exacerbating the networks of G protein signaling pathways (Blumer et al., 2012; Garcia-Marcos, 2021).

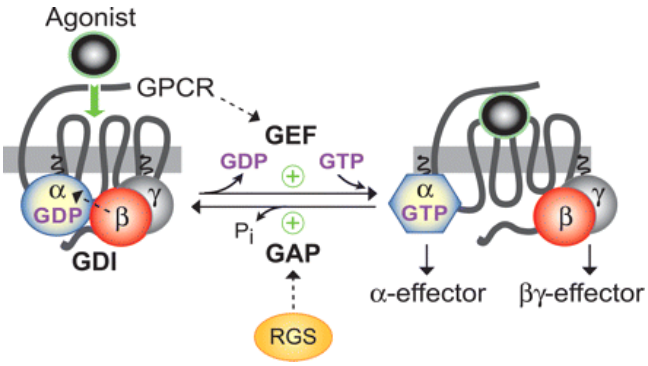


Figure 3: Schematic representation depicts the guanine nucleotide cycle of heterotrimeric G proteins. Figure is taken from (Johnston & Siderovski, 2007) .

Termination of the GPCR signaling is crucial to maintain normal body physiology. Several mechanisms are involved in signal termination including G protein inactivation and receptor desensitization. G proteins are inactivated when the GTP attached to the catalytic G alpha subunit will eventually hydrolyze to GDP, allowing its re-association with the G beta/gamma subunits and thus terminating the G protein signaling cascade (Hilger et al., 2018). Regulator of G protein signaling (RGSs) act as GTPase-activating proteins (GAPs) and accelerate the hydrolysis of GTP to GDP in the G alpha subunit; also, the effector itself may possess intrinsic GAP activity thus terminating the transduced signal e.g., PLCβ and AC (Hilger et al., 2018). Furthermore, the signal can be terminated by phosphorylating the GPCR receptor itself (desensitization) which uncouples its associated G protein. This process is mediated by Kinases such as GPCR kinases (GRKs) or protein kinase A (PKA) and protein kinase C (PKC) that are activated by receptor-triggered second messenger signaling (Figure 1) (Dorsam & Gutkind, 2007; O'Connor & Adams, 2010). Upon GPCR phosphorylation, the receptor

affinity to β-arrestin increases thereby sterically inhibiting G protein coupling as well as promoting internalization of the active receptor by endocytosis or ubiquitination (Hilger et al., 2018).

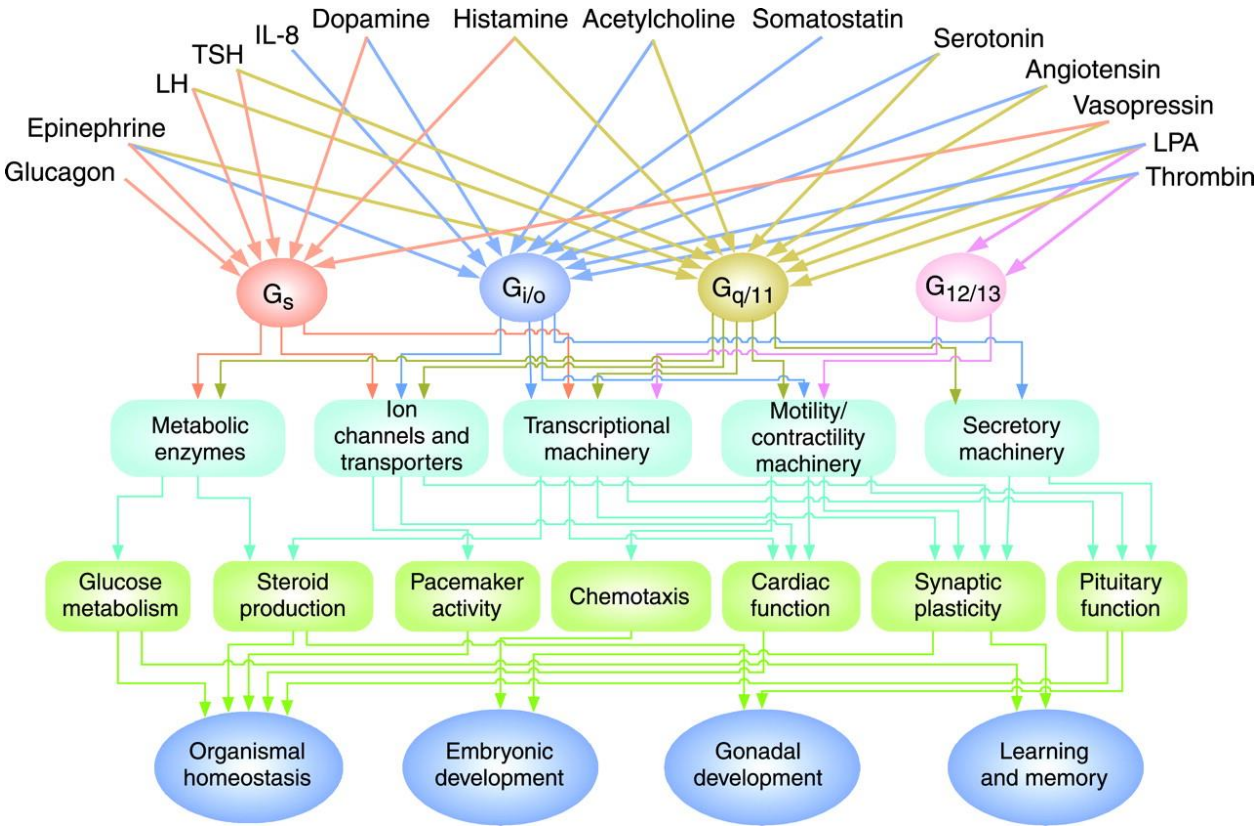


Figure 4: G protein signaling networks. G proteins regulate systemic functions by receiving signals from many extracellular agents thus regulating cellular machineries. Figure is taken from (Neves et al., 2002).

Heterotrimeric G proteins are typically referred to by their G alpha subunits and major research efforts have been focused on signaling via G alpha subunits. The G alpha proteins act as molecular switches determined by the bound nucleotide inside cells. They control a wide range of systemic functions such as embryonic development, learning, memory, and homeostasis through the regulation of metabolic enzymes, ion channels, transporter proteins and other parts of the cell machinery, controlling transcription, motility, contractility and secretion (Figure 4) (Neves et al., 2002). Loss or gain of function mutations in G alpha proteins result in signaling defects that lead to pleiotropic manifestations (Weinstein, 2001).

G alpha subunits are composed of a conserved GTPase domain, a Ras-like (Ras), and a helical domain (Figure 5). The GTPase domain promotes GTP hydrolysis to GDP and facilitates the binding of G beta/gamma, GPCRs, and effectors; it also includes flexible loops (switches) I, II, and III (Figure 5) (Hilger et al., 2018; Oldham & Hamm, 2008). The helical domain is positioned before switch region I and is specific to G alpha proteins, it forms a lid over the pocket containing the bound GTP/GDP, burying it in the core of the protein (Figure 5) (Hilger et al., 2018; Oldham & Hamm, 2008). All G alpha proteins are modified at or near their N-termini by linking of the fatty acids myristate (N-Myristoylation) and/or palmitate (thio-Palmitoylation); several alpha subunits contain both myristate and palmitate (Wedegaertner et al., 1995). Thio-Palmitoylation is a post-translational modification in which palmitate is covalently attached to cysteine residues through thioester linkage in all G alpha proteins, except Gt (Oldham & Hamm, 2008). N-Myristoylation is a co-translational modification, wherein the initiating methionine is removed followed by the addition of myristate irreversibly on the next glycine residue through amide linkage; all members of the Gi family (Gi1, Gi2, Gi3, Go, Gz, and Gt) are myristoylated (Oldham & Hamm, 2008). Both thio-Palmitoylation and N-Myristoylation serve to anchor the G alpha protein in the membrane thereby increasing the efficient concentration for interaction with the GPCRs (Rocks et al., 2005; Webb et al., 2000; Wedegaertner & Bourne, 1994). Also, they play important roles in regulating protein stability (Xia et al., 2021) and modulating protein-protein interactions (Linder et al., 1991).

1.3 The G alpha protein families

Vertebrate G alpha subunits can be divided into five families based on sequence similarity: Gs (contains Gs, Golf, GXL); Gi (contains Gi1, Gi2, Gi3, Gt, Go, Gg, and Gz); Gq (contains Gq, G11, G14, and G15/16); G12 (encompasses G12 and G13); and the newly described Gv family, which has two members, Gv1 and Gv2 (Malbon, 2005; Oka et al., 2009).

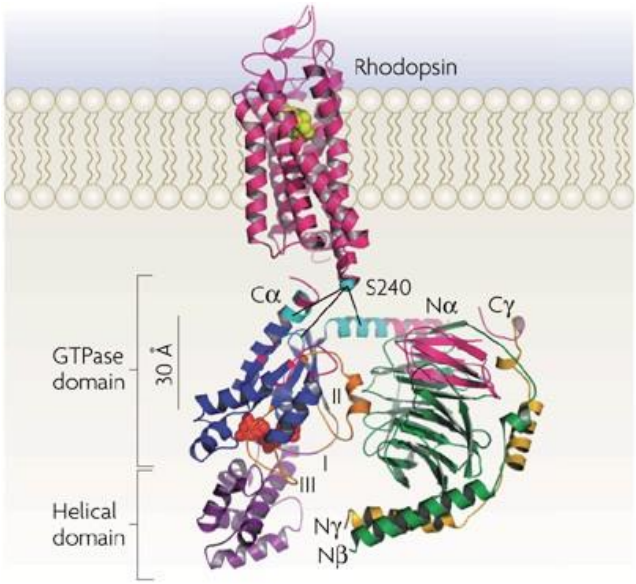


Figure 5: Schematic representation of the common structure for heterotrimeric G proteins interacting with GPCRs. Different colors refer to different subunits. receptor contact sites on the G protein are indicated by pink and cyan colors; specific point-to point interactions between the GPCR and heterotrimeric G proteins is between Ser240 on intracellular loop 3 of the receptor (cyan sphere) with specific regions of the G alpha N terminus (Nα), C terminus (Cα) and α4–β6 loop (cyan). The G alpha N terminus encompass the lipid modification sites. Figure is taken from (Oldham & Hamm, 2008).

The Gs family (named after its stimulatory role) is one of the first G proteins to be identified, they are activated by a large group of GPCRs and the Gs protein is expressed ubiquitously (Figure 6) (Malbon, 2005; Syrovatkina et al., 2016). The other members of Gs family are Golf (olf stands for olfaction) which is predominantly expressed in the olfactory sensory neurons, and the GXL which is an extra-long (XL) version of Gs that possesses an XL exon and is found in the neuroendocrine tissues (Figure 6) (Malbon, 2005; Syrovatkina et al., 2016). The general function of Gs proteins is to *stimulate* intracellular signaling pathways by activating AC which synthesizes the second messenger cyclic adenosine 3', 5' -monophosphate (cAMP) from ATP. cAMP in turn activates the cAMP-dependent PKA (Figure 1) (Goldsmith & Dhanasekaran, 2007). PKA phosphorylates myriad downstream targets hence regulating enormous number of physiological processes such as glycogen metabolism, synthesis of hormones as estrogen, progesterone, aldosterone, and cortisol, calcium and fluid homeostasis, among others (Goldsmith & Dhanasekaran, 2007; Syrovatkina et al., 2016).

The Gi family (named after its inhibitory function) is the largest and most diverse among G alpha families. Gi proteins play diverse roles in development and mediate aspects of Wnt-frizzled signaling (Goldsmith & Dhanasekaran, 2007; Malbon, 2005). Gi proteins are broadly expressed but some members exhibit tissue specialization: Go is predominantly expressed in the neurons, Gt (t stands for transducin) is expressed in the eye, Gg (g stands for gustducin) is expressed in the taste receptor cells, Gz is expressed in the platelets and sensory tissues (Figure 6) (Malbon, 2005; Syrovatkina et al., 2016). Gi activity counteracts Gs function by inhibiting AC thereby lowering cAMP levels, thereby the GPCRs coupled to Gi counteract the activity of the GPCRs coupled to Gs (Figure 1) (Malbon, 2005). Consequently, the levels of cAMP control the activity of different ion channels and PKA proteins which in turn modulates many physiological processes as vision, taste, vomeronasal function, muscle contractility, cell migration, developmental processes, lipid metabolism, regulation of immune cells, renal function, and platelet activation, etc. (Goldsmith & Dhanasekaran, 2007; Malbon, 2005).

Within the Gq family, Gq and G11 are ubiquitously expressed, while the expression of G14 and G15/16 is more confined. G14 is found in the kidney, liver and lung whereas G15/16 is specifically found in the hematopoietic cells (Figure 6) (Malbon, 2005; Syrovatkina et al., 2016). The Gq proteins activate the well-known Gq effector, PLCβ, which induces the cleavage of membrane-bound phosphatidylinositol 4,5-bisphosphate (PIP2) into diacylglycerol (DAG) and inositol triphosphate (IP3), both server as second messengers (Figure 1) (Goldsmith & Dhanasekaran, 2007; Syrovatkina et al., 2016). IP3 activates IP3 receptors (ligand-gated calcium channel) thus inducing calcium release from the endoplasmic reticulum into the cytosol. The increase in calcium levels in the cytosol results in a cascade of intracellular changes and activity through calcium binding proteins and calcium-sensitive processes. For instance, the IP3-dependent pathway is a signal transduction pathway for many hormones (Goldsmith & Dhanasekaran, 2007). DAG activates PKC proteins in the plasma membrane thus stimulating downstream cellular signaling cascades including those responsible for cytoskeleton regulation (Goldsmith & Dhanasekaran, 2007). It has been shown that the Gq family plays several physiological roles and is involved in

muscle contractility, platelet activation, hormone secretion in pituitary gland, neurotransmission, and leukocyte migration (Syrovatkina et al., 2016).

The G12 family is the last one discovered before Gv, therefore less knowledge is available in the literature compared to Gs, Gi, and Gq. Both G12 and G13 are ubiquitously expressed and are essential for normal embryonic development (Figure 6) (Malbon, 2005). It has been reported that the G12/13 family is involved in activation of the Rho family of small GTPases by direct interaction with GEF and recruiting RhoGEF which results in a variety of effects that include platelet activation, muscle contraction, and modulating cell cytoskeleton thus controlling cell migration (Syrovatkina et al., 2016).

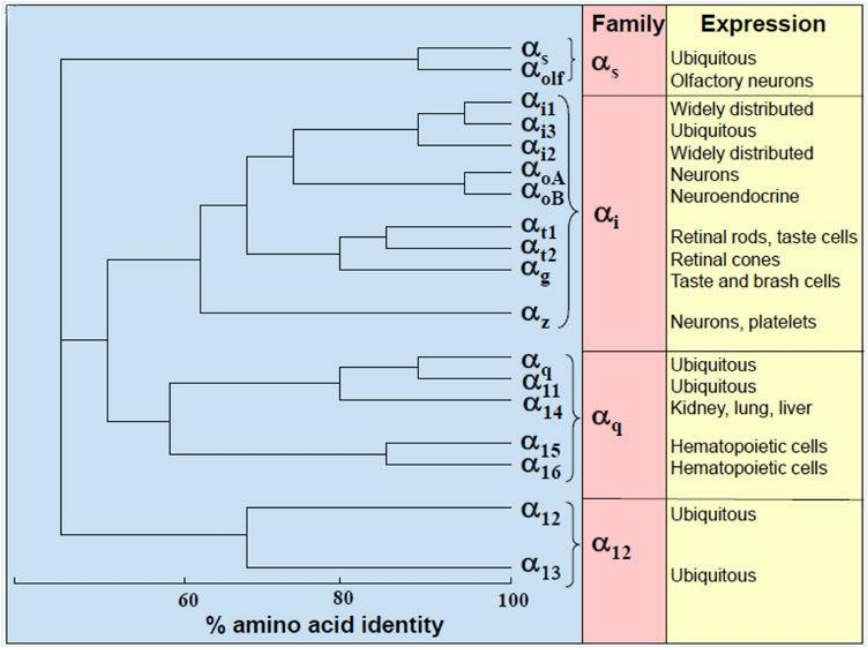


Figure 6: G proteins localization in human and mouse tissue illustrated the phylogenetic tree of four G alpha families Figure is taken from (Syrovatkina et al., 2016).

The Gv proteins are the most recently discovered members of G alpha proteins (by the Korsching lab); they encompass two members, Gv1 and Gv2 (Oka et al., 2009). In spite of the fact that the Gv family is phylogenetically as ancient as the other four families and that G alpha proteins in general are extensively studied, it is astonishing that the Gv family was discovered so late. This could be due to the fact that Gv is absent in many of the common animal models used to study G alpha proteins including

tetrapod vertebrates, *Drosophila melanogaster*, and *Caenorhabditis elegans* (Oka et al., 2009).

1.4 The most recently discovered G alpha protein, Gv

Oka and colleagues described Gv proteins (encoded by *gnav* genes) as an independent novel class of G alpha subunits (Oka et al., 2009). Despite the importance of the G alpha subunits in all biological systems, so far there has been no published data about the Gv role, its effector proteins, and signaling pathway. This is presumably because it was only identified a few years ago in addition to its absence in higher vertebrates and several invertebrate model organisms. However, Gv has orthologues in sponges, one of the earliest-diverging animal phyla, hence it is as old as the other four classes of G alpha proteins (Oka et al., 2009).

Gv protein sequence is highly conserved, e.g., Gv has 70% amino acid sequence homology among vertebrates (Oka et al., 2009). Many teleost lineages including the prominent model system zebrafish possess *gnav* genes, on the other hand Gv is absent in several different lineages (Oka et al., 2009). Oka and colleagues showed that there has been a single gene duplication early in the vertebrate lineage, resulting in 1–2 *gnav* genes per species. They also showed that the Gv has undergone frequent gene loss events throughout animal evolution and they could not find Gv orthologs in mammals, chicken, reptiles, amphibians, jawless fish, ascidians, fruit fly, mosquitoes, bee, moth, several nematodes, leech, and cnidarians (Oka et al., 2009). In the meantime, very few phylogenetic studies investigated the evolution of Gv family in addition to the other four families. Those studies reported the presence of *gnav* genes in many species and also in a very early diverging single-celled eukaryote, *Capsaspora owczarzaki*, which belongs to the Filasterea class, a sister clade to animals and choanoflagellates (de Mendoza et al., 2014; Krishnan et al., 2015; Lokits et al., 2018). This seemed to suggest that gnav genes originated in the MRCA of the Filozoa, which encompass the Filasterea and all Choanozoa (animals and choanoflagellates). However, all three studies concluded that Gv originated in the common ancestor of all Holozoa (comprising both Filozoa and Teretosporea) even

though they either did not investigate Teretosporea (Krishnan et al., 2015; Lokits et al., 2018) or did not detect Gv in Teretosporea (de Mendoza et al., 2014).

Gv protein possesses all the expected motifs for G alpha proteins: helical domain, three switch regions, 5 G-box motifs (which play a role in GDP/GTP binding), two conserved residues essential for GTPase activity, and acylation sites (palmitoylation and myristoylation) which are essential in G protein signaling (Figure 7) (Oka et al., 2009; Oldham & Hamm, 2008). Gv protein harbors sixty amino acids as Gv-specific motifs within the helical domain (Figure 7) (Oka et al., 2009). The functions of the helical domain are not completely clear so far, however, it seems to affect GTPase activity and may be involved in the interaction with GPCRs, regulator and effector proteins (Chung, 2013; Oldham & Hamm, 2008). Thus, the Gv-specific motifs of the helical domain might suggest a set of Gv-specific interaction partners. Furthermore, in comparison to the four other G alpha classes, some residues essential for interaction with RGS are conserved in Gv proteins, as well as a larger motif interacting with phosphodiesterase γ (Oka et al., 2009), suggesting potential coupling to these molecules. However, this has not been empirically examined so far, hence the biological role of Gv, i.e., its activation mechanisms with respect to its GPCR and effectors activated by it are still unknown.

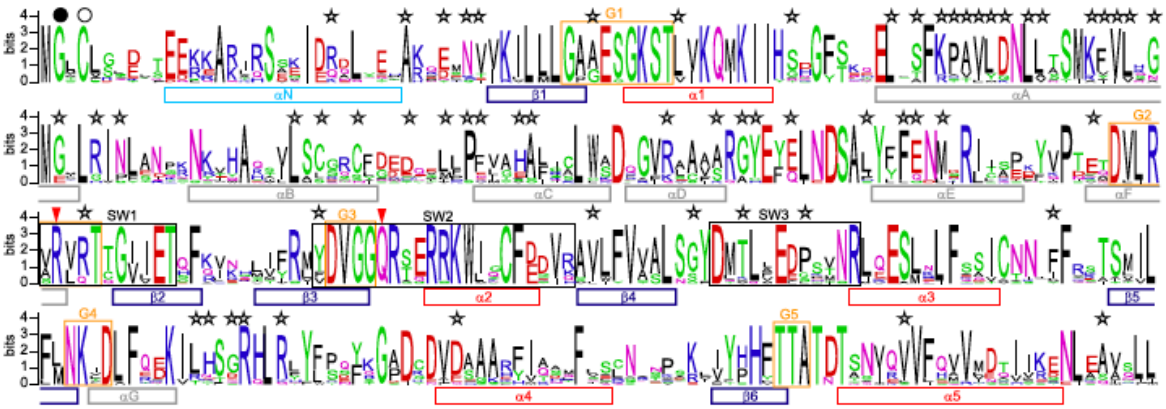


Figure 7: Web logo for Gv proteins illustrates conserved features of 19 Gv proteins. Secondary structures are indicated under the sequence inside the bars. Light blue bar: N-terminal helix; gray bars: helices within helical domain; red bars: helices within GTPase domain; dark blue: beta-sheet. G-boxes and switch regions are denoted with orange and black boxes, respectively. Black and white circles above the logo marks the putative sites for N-linked

myristoylation and thio-palmitoylation, respectively. Gv conserved motifs are indicated by stars, and red arrowheads denote residues critical for GTPase activity. Figure and figure legend are taken from (Oka et al., 2009).

Expression of Gv can be taken as the first indication of possible functional roles. So far, Gv expression is studied only in zebrafish; the Gv mRNA expression was analyzed by RT-PCR for adult tissues and WISH for 3 days post fertilization (dpf) larvae (Figure 8) (Oka et al., 2009). The expression analysis by RT-PCR showed that *gnav1* is broadly, but not ubiquitously expressed in many adult tissues as gill, kidney, olfactory epithelium, stomach, eye, brain, liver, spleen, and skin (Figure 8) (Oka et al., 2009). Also, Gv mRNA is expressed in many larval tissues that include the inner ear, lower lip, branchial arches, pectoral fins, and midbrain-hindbrain boundary (MHB) region (Figure 8) (Oka et al., 2009). Furthermore, we could find Gv in two transcriptome databases for zebrafish: (Pasquier et al., 2016) showed that Gv mRNA is present (at different levels) in unfertilized egg, adult zebrafish ovaries, testes, brain, gills, heart, muscles, liver, head kidney, bones, and intestine. Gv mRNA is also expressed in many different developmental stages at varying levels (Papatheodorou et al., 2018). The presence of Gv in several osmoregulatory tissues such as kidney, gills, and intestine, either by RT-PCR (Oka et al., 2009) or RNA-seq data (Pasquier et al., 2016), could suggest a role for Gv in ionic homeostasis. Interestingly, all but one species reported by (Oka et al., 2009) to possess the gnav gene are living in an aqueous environment which can pose challenges to keep their ionic and osmotic homeostasis. This would be consistent with Gv having a role in ion homeostasis.

In order to examine the function of Gv, a previous thesis work attempted to knockout *gnav1* in zebrafish by CRISPR/Cas9 system (Ivandic, 2015). Ivandic could introduce a frameshift mutation in *gnav1*. Preliminary analysis of the mutant phenotype indicated that it might be involved in calcium homeostasis (Ivandic, 2015).

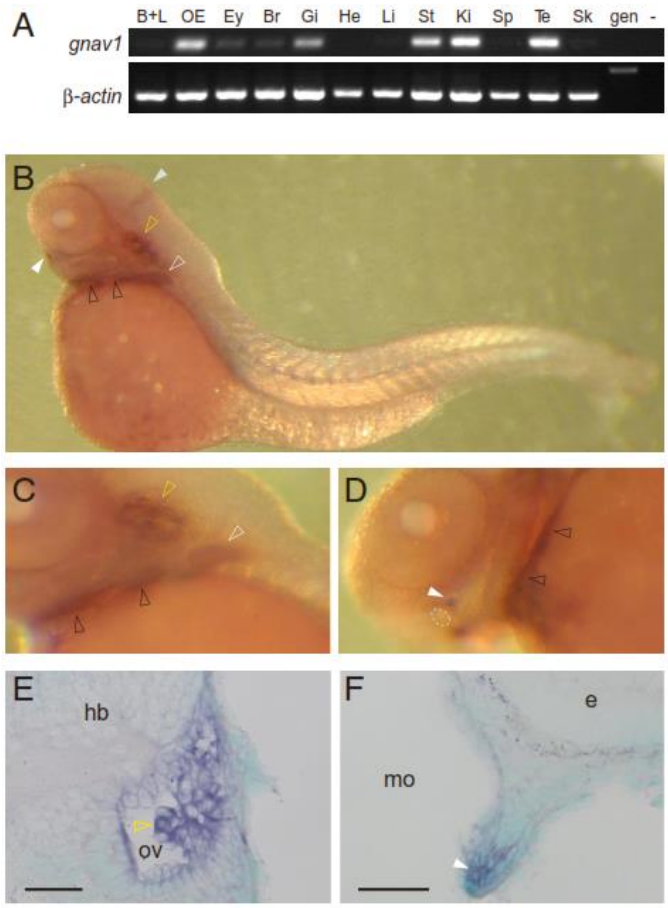


Figure 8: Expression pattern of *gnav1* gene in adult zebrafish tissues and 3 dpf zebrafish larvae. (A) Expression of *gnav1* in adult zebrafish organs determined by RT-PCR. B+L: barbels and lips; OE: olfactory epithelium; Ey: eye; Br: brain; Gi: gill; He: heart; Li: liver; St: stomach; Ki: kidney; Sp: spleen; Te: testis; Sk: skin; gen: genomic DNA; -: negative control. (B–F) wholemount in situ hybridization for 3 dpf larvae. Dotted circle: mouth. (E and F) *gnav1* expression in the developing inner ear (E) and lower lip (F). White and gray solid arrowheads indicate the cell clusters next to the lower lip and the MHB, respectively. White, yellow, and black open arrowheads point to labeled cells within pectoral fins, otic vesicle (ov), and branchial arches, respectively. E: eye; hb: hindbrain; mo: mouth. Figure and Figure legend are taken from (Oka et al., 2009).

1.5 Ionic homeostasis regulation in zebrafish

Regulation of ionic homeostasis and maintaining the ionic concentrations and pH at normal levels are crucial to enable normal physiological processes. Aquatic vertebrates such as fish need the capacity to survive under diverse environmental conditions in terms of ionic contents and pH levels. For instance, zebrafish is challenged by the hypotonic freshwater environment which contains low ionic concentrations and fluctuating pH, hence the fish should be able to tightly regulate the

ion levels, e.g., by minimizing passive ion loss and osmotic water gain. In teleosts such as zebrafish, specialized ionocytes are present in the adult gills and embryonic/larval skin; these cells are the major extra-renal organs that are involved in active absorption of electrolytes from the aquatic environment (P.-P. Hwang & Chou, 2013). Moreover, the intestinal (gut) epithelia play a role in ionic homeostasis by absorbing the electrolytes contained in the food (Whittamore, 2012). Importantly, the zebrafish kidney regulates the ionic homeostasis by reabsorbing ions through active transcellular and passive paracellular transport (Ying Jey et al., 2015). Zebrafish kidney generates large amounts of low osmotic prourine in order to dispose of the excess water gained from the surrounding; the high urine flow rates show the exceptional functional ability of zebrafish kidney (Flik et al., 2009). In terms of electrolyte transportation, the ionocytes (that are present in gills/embryonic skin) are similar to renal tubular cells, and many ion channels and transporters are coexpressed in both tissues (Figure 9) (P.-P. Hwang & Chou, 2013). Therefore, the ionocytes and kidney share many functions with respect to ion transport in zebrafish, however, their relative contribution to ionic homeostasis depends on the type of ion that is being regulated.

At least five types of ionocytes expressing different sets of ion transporters have been identified so far: H+-ATPase-rich (HR), Na^+/K^+-ATPase-rich (NaR), Na^+/Cl^- cotransporter (NCC), K^+-secreting (KS), and solute carrier family 26 (SLC26)-expressing cells (Figure 9) (Ying Jey et al., 2015). These ionocytes perform trans-epithelial H^+ secretion/Na^+ uptake/$NH4^+$ excretion, Ca^{2+} uptake, Na^+/Cl^- uptake, K^+ secretion, and Cl^- uptake/$HCO3^-$ secretion, respectively, by distinct sets of ion transporters (Figure 9) (Ying Jey et al., 2015). In addition to the ion transporters, several isoforms of the Na^+/K^+-ATPase (NKA) pump are present at the basolateral membrane and function to maintain the electrochemical gradient during ion transportation; the pump also provides the driving force for active transportation.

Similar ion transporters and pumps are expressed in adult and larval zebrafish kidneys. The embryonic or larval kidney is called pronephros; it is a flat organ that begins to filtrate the blood at 2 dpf. The pronephros consists of one glomerulus at the anterior midline of the body, which is connected to bilateral pronephric tubules

(containing proximal and distal tubule segments), which connect to the pronephric duct, which ends in the cloaca (Figure 10) (Drummond et al., 1998). In the juvenile zebrafish kidney this pronephros has developed into a structure called mesonephros, which is composed of hundreds of glomeruli and their associated nephrons (the tubules) to maintain the increasing osmoregulatory demand (Figure 11) (G. F. Gerlach & Wingert, 2013). In contrast to mammalian kidney development, no further maturation into metanephros occurs in zebrafish. The mesonephros just increases in size and ramifications; eventually the adult kidney is a flattened organ located along the dorsal body wall and divided into tail, trunk and head regions (Figure 11 and Figure 10) (McCampbell et al., 2015).

Each functional unit (nephron) consists of a glomerulus which is responsible for blood filtration, followed by proximal convoluted tubule (PCT), proximal straight tubule (PST), distal early (DE), and distal late (DL) tubule segments which fuse with the pre-existing distal tubule segment of the pronephric tubules (Figure 10) (G. F. Gerlach & Wingert, 2013). The proximal tubule exhibits brush borders and carries out ionic reabsorption, also it is the predominant site for acid-base balance in the zebrafish nephron. The proximal tubule contains Na^+/H^+ exchanger which is responsible for hydrogen ion extrusion by the carbonic anhydrase activity (Kersten & Arjona, 2017). In contrast to the proximal tubule, the distal tubule lacks the brush borders, and it has a low permeability to water thus it carries out the osmodilution of the prourine and potassium extrusion toward the lumen (Kersten & Arjona, 2017). In addition, the reabsorption of sodium and chloride takes place in the distal tubule by the Na+/K+/Cl− cotransporter and the Na+/Cl− cotransporter (NCC). The prourine is drained out of the body through the collecting duct (formerly named pronephric duct) (Kersten & Arjona, 2017).

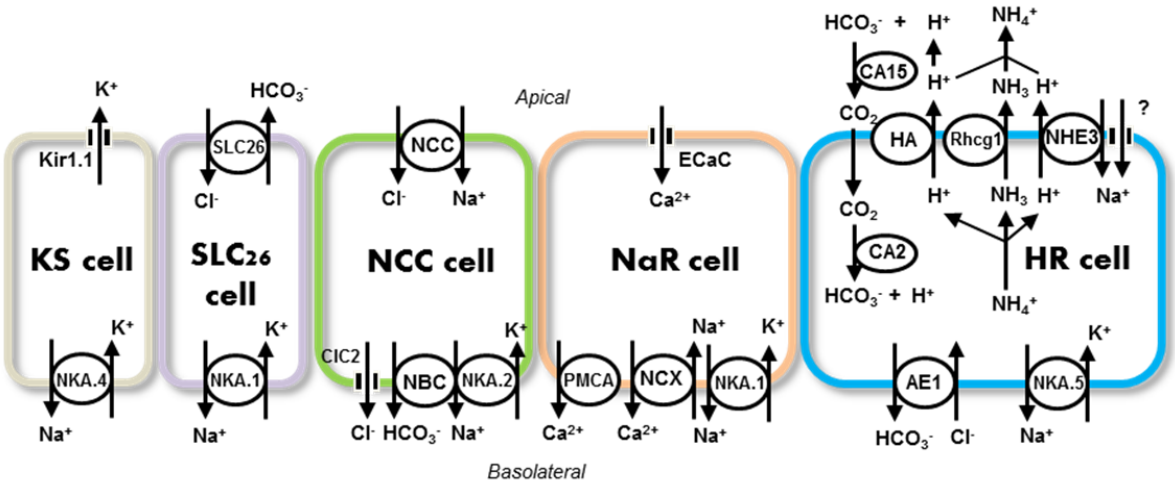


Figure 9: the five types of ionocytes identified so far in zebrafish. The ionocytes expressing different sets of ion transporters are as follows: H^+-ATPase-rich (HR), Na^+/K^+-ATPase-rich (NaR), $Na^+/Cl-$ cotransporter (NCC), K^+-secreting (KS), and solute carrier family 26 (SLC26)-expressing cells. Question mark indicates unidentified ion transporter pathway. For abbreviations the reader is referred to the list of abbreviations. Figure is taken from (Ying Jey et al., 2015).

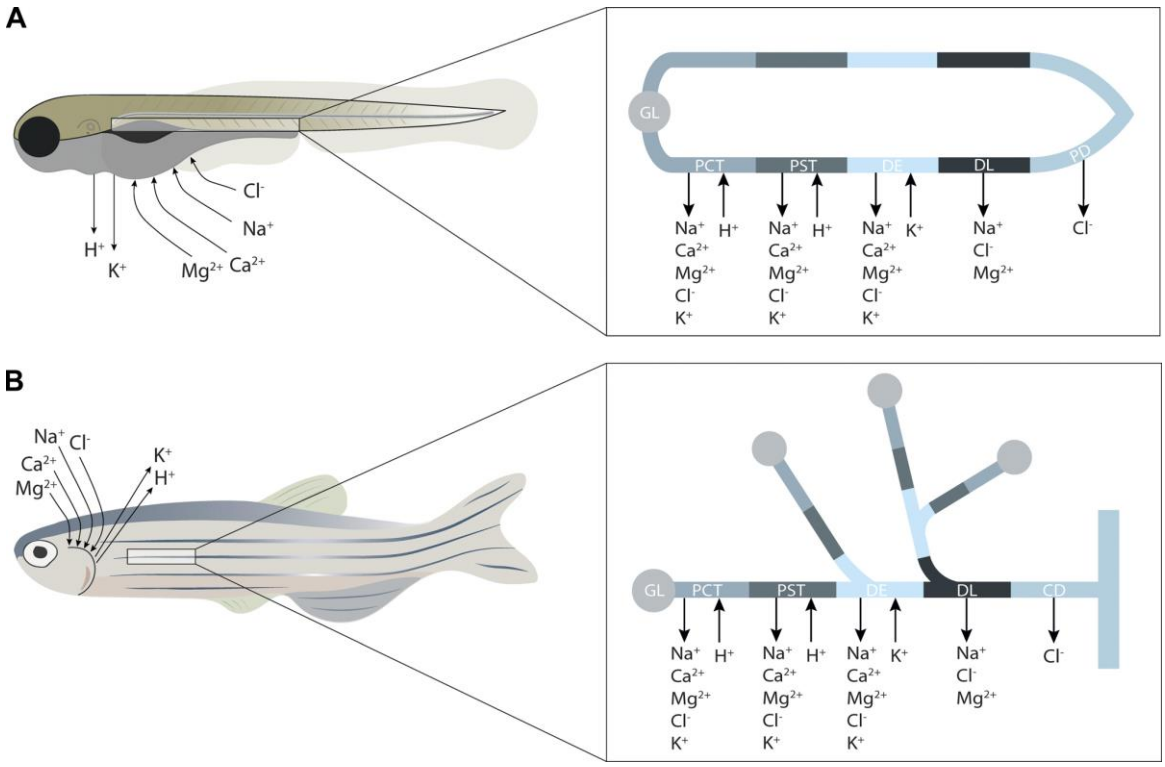


Figure 10: Schematic drawing shows the fluxes of the sodium, chloride, calcium, magnesium, and potassium, and protons, transported in the gills/embryonic skin (A) and nephrons of larval

(A) and adult zebrafish (B). In larvae (A) skin ionocytes and the kidney work together to minimize ion losses. Active transcellular absorption of magnesium, calcium, sodium, and chloride are performed by skin ionocytes, while potassium and hydrogen ions are mostly excreted by the kidney. Larvae secrete and reabsorb ions through the ion channels, transporters, and claudins in the pronephros. The Larval pronephros is composed of bilateral nephrons that fuse at a single central glomerulus. The main segments of the nephron are the proximal convoluted tubule (PCT), the proximal straight tubule (PST), the distal early tubule (DE), the distal late tubule (DL), and the pronephric duct (PD). In adult zebrafish the flux of indicated ions is regulated in the gills and mesonephros (adult kidney). Figure is taken from (Kersten & Arjona, 2017).

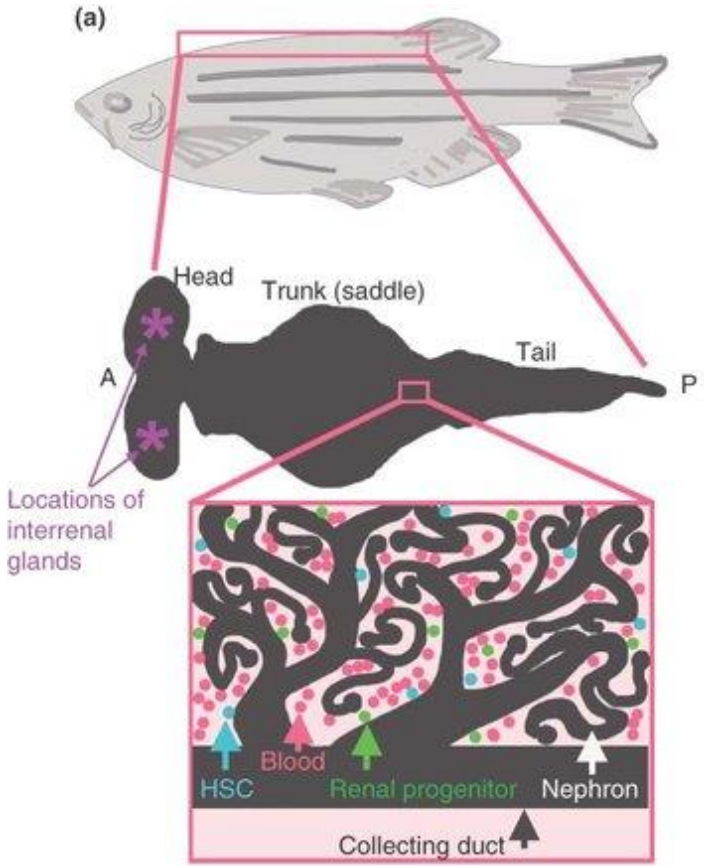


Figure 11: Schematic drawing illustrates structures of the adult zebrafish kidney (mesonephros). (a) The adult kidney is located at the dorsal body wall indicated by the pink box. (A) the main parts of the adult kidney, i.e., head, trunk, and tail. The magnified pink box represents the tubular network in mesonephros and the other structures are indicated in the figure, modified from (G. F. Gerlach & Wingert, 2013)

Calcium is both a structural element (bones) and a central regulator of manifold physiological activities. The calcium is mainly absorbed from the aquatic environment either by passive diffusion or actively transported through calcium transporters (Ying

Jey et al., 2015). Calcium is predominantly regulated by gills/embryonic skin, whereas magnesium is more efficiently handled in the kidney (Figure 10) (Kersten & Arjona, 2017). On the other hand, sodium (absorption), potassium (extrusion), and chloride (absorption) are regulated cooperatively by both organs (Figure 10) (Kersten & Arjona, 2017). With respect to acid-base homeostasis and nitrogenous metabolic waste, gills/embryonic skin are the major organs which handle the hydrogen ions secretion and ammonia excretion; while the kidney appears to have a minor role in these processes (Ying Jey et al., 2015).

Ionic homeostasis is a fundamental cellular phenomenon; malfunctions of the ion channels, transporter, or pumps can lead to impaired absorption of ions, and eventually alter the ionic homeostasis in the kidney (Kersten & Arjona, 2017). Since ions play a significant role in physiology and many different biological processes, such altered ionic balance of the blood and tissues of the body can have far-reaching physiological consequences (Kersten & Arjona, 2017). Therefore, it is critical to comprehensively understand the ionic homeostasis, and in recent years zebrafish increasingly have been used as an animal model to study the mechanism of ionic homeostasis in body fluids and kidney diseases (P.-P. Hwang & Chou, 2013; Lieschke & Currie, 2007).

1.6 Zebrafish as a model organism

Zebrafish (*Danio rerio*) is a freshwater tropical fish native to southeast Asia. Zebrafish was introduced by George Streisinger as an animal model to the scientific community in 1981 (Streisinger et al., 1981). In the following decades, zebrafish has been positioned at the frontline of biomedical research owing to the remarkable experimental advantages of this model (Lieschke & Currie, 2007). This powerful model provides multiple attractive features over mammalian vertebrate model organisms such as mice. The maintenance of the fish in the laboratory is inexpensive, easy, and efficient. Zebrafish are small in size, i.e., the length of larvae and adult fish ranges from about 3.5 to 30 mm, respectively (Lieschke & Currie, 2007; Strähle et al., 2012). Another attractive feature is that its fertilization is external thus it develops externally

(in contrast to internal development in mice); the eggs are large compared to other fish (around 0.7 mm); in addition, zebrafish larvae are optically translucent in contrast to the opaque mice embryos (Lieschke & Currie, 2007). The optical clarity of zebrafish larvae, coupled with the abundance of available fluorescent transgenic lines enables the researchers to monitor zebrafish in vivo through all developmental stages under dissecting microscope (Streisinger et al., 1981). Furthermore, zebrafish produces several hundred eggs in a single mating while mammals distinctly produce fewer offspring; the zebrafish embryo develops rapidly, i.e., the 24 hpf embryo already has many organs; and zebrafish have a short generation time enabling large-scale use, improving the statistical power and accuracy of the experiments (Spence & Smith, 2006). Importantly, in terms of anatomy and histology, zebrafish show a high degree of similarity with mammalian organs, tissues, and cellular systems together with their associated physiological functions (Lieschke & Currie, 2007). Therefore, using zebrafish as a model for vertebrates overcomes many limitations in using the mammalian animal models.

The groundbreaking CRISPR/Cas9 technology for gene editing has quickly become a method of choice for genetic manipulation in zebrafish, due to its efficacy and versatility. Moreover, many genetic resources including whole genome sequencing and assembly project are available (W. Y. Hwang et al., 2013). Therefore, the simplicity and versatility of genetic manipulation and in vivo assays of zebrafish enabled the researchers to significantly utilize them for therapeutic drug discovery, and also in fundamental research such as evolutionary science, genetics, neurobiology, and developmental biology (Lieschke & Currie, 2007; Strähle et al., 2012).

The early development of zebrafish has been described in detail by Kimmel and colleagues. They classified the zebrafish developmental stages from fertilization to hatching into seven periods: zygote, cleavage, blastula, gastrula, segmentation, pharyngula, and hatching periods (Kimmel et al., 1995).

Unfertilized eggs are enclosed by an acellular chorion and contain cortical granules (Figure 12) (Rojas et al., 2021). Upon fertilization, the egg is called **zygote** and at this

stage calcium waves are triggered in many species including zebrafish. For mammals it has been reported that this calcium signaling is elicited by PLC that is released from the sperm into the oocyte (Figure 12). The activation of PLC results in the generation of DAG and IP3; the latter binds to the IP3 receptor resulting in the release of calcium from the intracellular stores, which in turn contribute to egg activation progression including meiosis resumption (Figure 12) (Rojas et al., 2021). When this rapid and transient increase in intracellular calcium levels propagates throughout the egg, enzymes are released from the cortical granules and start a cortical reaction resulting in the separation of the oocyte cell membrane from the chorion (Rojas et al., 2021). The chorion swells and lifts away from the zygote thus establishing the perivitelline space (Figure 12) (Hart & Fluck, 1996). The wave of calcium release also causes the actin microfilaments to sequester the non-yolk cytoplasm towards one extremity of the zygote thus it streams to what will be the animal pole; the yolk-rich side opposite the animal pole is called the vegetal pole (Kimmel et al., 1995). The accumulation of non-yolk cytoplasm at the animal pole forms the blastodisc and its formation changes the shape of the zygote from spherical to pear shaped (one cell stage in Figure 13) (Hart & Fluck, 1996).

In zebrafish the **cleavage** period begins shortly after fertilization; the cell divisions are meroblastic (incomplete) and confined to the blastodisc at the animal pole. At this stage, the cells (blastomeres) divide synchronously and regularly at about 15-minute intervals (Kimmel et al., 1995). The maternal mRNA contents contained in the oocyte during oogenesis is sufficient to drive all early steps of embryogenesis before the zygotic genome activation (ZGA) (Solnica-Krezel, 2020).

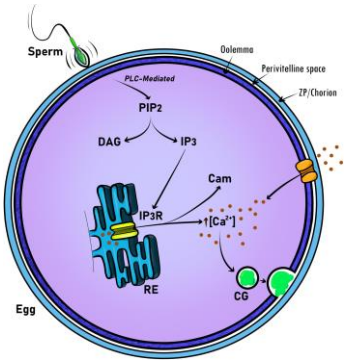


Figure 12: Schematic representation of the common signaling pathway that involve intracellular calcium signaling at oocyte activation and fertilization. PIP2: 4,5-bisphosphate; DAG :1,2-diacylglycerol; IP3: inositol 1,4,5-triphosphate; IP3R: IP3 receptor; PLC: phospholipase C; cam: calmodulin; CG: cortical granule. Figure is taken from (Rojas et al., 2021).

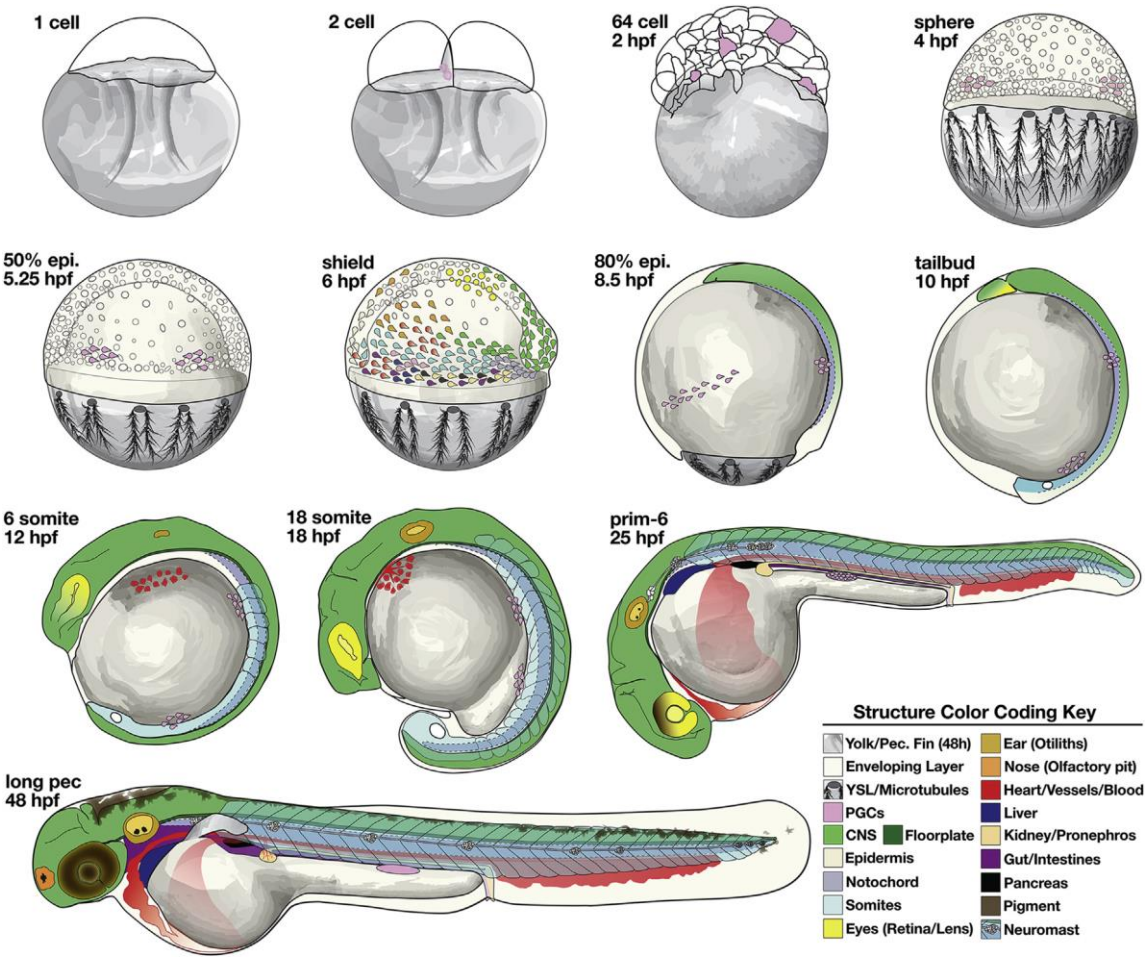


Figure 13: Schematic overview represents zebrafish embryo at different developmental stages beginning with one cell stage over gastrula, early somatogenesis stage, and pharyngula stage

to hatching stage. Color code for tissues is indicated. Lateral view anterior to left. The reader is referred to the text for details. YSL: yolk syncytial layer; PGC: primordial germ cells. Figure is taken from (Pathak & Barresi, 2020).

The **blastula** period (approximately 2-5 hours post fertilization (hpf)) begins at the 128-cell stage during which a high number of blastomeres appears as a solid half ball perched on the yolk cell (see 64 cell stage in Figure 13) (Hart & Fluck, 1996). After 10 cell divisions (512 cells) during the blastula period, the midblastula transition (MBT) commences wherein significant changes occur: the cell cycle lengthens and becomes asynchronous; the cells become motile; the yolk syncytial layer (YSL) forms and spreads beneath the blastodisc; and zygotic transcription is initiated (Solnica-Krezel, 2020). Downregulation of maternal RNA, which has already started after fertilization, is increasingly enhanced at the onset of zygotic transcription in MBT to facilitate the rapid transition of developmental control from maternal RNA to the zygotic genome (Figure 14) (Solnica-Krezel, 2020). The contribution of maternal RNA to early embryogenesis, in part gastrulation (the next stage), varies among species, however it is quite significant in zebrafish (Solnica-Krezel, 2020).

Late in the blastula period, the epiboly begins which involves the movement of YSL towards the vegetal pole and drives the blastoderm (cells of the embryo proper) in the same direction, thus the blastoderm thins and extends over the yolk sac (Solnica-Krezel, 2020). Thereafter, the development of the embryo is measured as percent epiboly. In the **gastrula** period (about 5-10 hpf), epiboly reaches 50% and continues until the end of gastrulation (morphogenetic cell movements include involution, convergence, and extension) (Solnica-Krezel, 2020). In the shield stage (6 hpf) during the gastrula period, the cells move towards the blastoderm margin, involute, and also converge towards the animal pole; they accumulate at the front edge forming a small bulge that is called the embryonic shield (see 6 hpf stage in Figure 13) (Solnica-Krezel, 2020).

The embryonic shield comprises two layers, the outer layer epiblast and the inner layer hypoblast separated by a fissure; they eventually correspond to ectoderm and mesendoderm germ layers, respectively (Kimmel et al., 1995). The ectoderm derivatives include tissues such as epidermis, central nervous system, neural crest,

and sensory placodes; while the mesendoderm (mesoderm and endoderm) derivatives include many tissues such as the epithelial lining of the internal organs, somites, cartilage, bone, notochord, kidneys, and several other organs (Kimmel et al., 1995). Fate maps have been constructed for the cells of embryos at the onset of gastrulation; they depict which cell or region gives rise to a particular tissue (Figure 15 and Figure 13). At 10 hpf, 100% epiboly is reached and gastrulation ends; at this stage the entire yolk sac is engulfed by the YSL, and the tail bud is formed (Figure 13) (Solnica-Krezel, 2020). Gastrulation results in narrowing the blastoderm along the mediolateral axis and elongating it from head to tail thus initiating the body axes (dorsal-ventral and anterior-posterior) (Figure 13) (Solnica-Krezel, 2020).

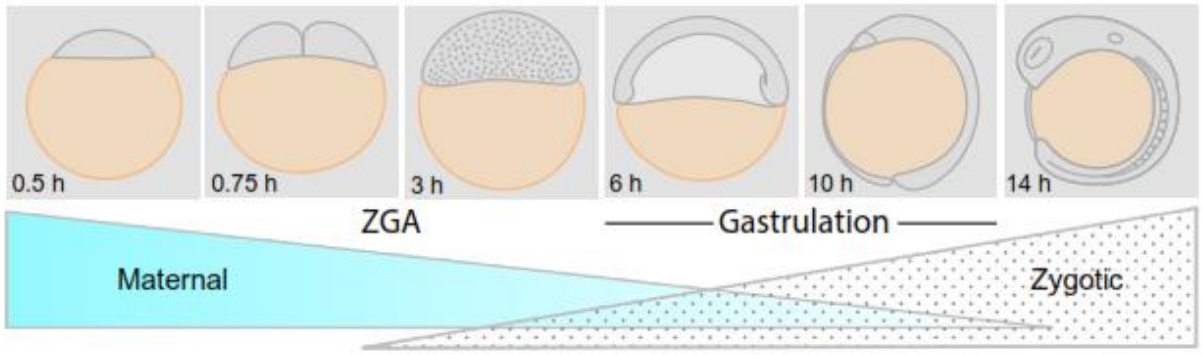


Figure 14: The dynamic changes in the level of maternal and embryonic RNA product during early embryogenesis. Figure is taken from (Solnica-Krezel, 2020).

Next, the **segmentation** period begins at 10.5 hpf and continues until 24 hpf during which the somites, segmented precursors of myotomes and sclerotomes, become visible. Somatogenesis starts at 10.5 hpf and continues by notochord-derived signals (Stickney et al., 2000). During this period the embryo elongates and other distinct morphogenesis events occur; e.g. brain primordium can be determined by the thickened neural keel; the optic primordium can be observed; and tail bud extends within which Kupffer's vesicle can be seen (Figure 15c and 10-12 hpf stage in Figure 13) (Kimmel et al., 1995).

In the **pharyngula** period spectacular morphogenetic changes occur (about 24-48 hpf); the embryo visibly elongates; it possesses a sculptured head and symmetrical bilateral body; the trunk begins to lift off the yolk and the tail lengthen; the heart begins

to beat, more somites develop, and the embryo shows spontaneous movements (Figure 13) (Kimmel et al., 1995). Also, the primordia of some organs become visible such as the brain, otic vesicle, and pronephros. Moreover, the yolk sac size is reduced; the membranous fin fold expands; and the tail is curved (Figure 13) (Kimmel et al., 1995). During this period the primordia of the pharyngeal arches also become visible which eventually gives rise to the craniofacial skeleton (Kimmel et al., 1995). Also, the neural keel (which was formed at 12 hpf) undergoes in-folding at the midline, thereafter it rounds into the cylindrical neural rod (see 18 hpf embryo in Figure 13) (Kimmel et al., 1995). At 1 dpf stage, the brain is prominently sculptured; the main parts can be distinguished, particularly the forebrain (contains the diencephalon and telencephalon), midbrain (subdivides into dorsal optic tectum and ventral tegmentum at this stage), and hindbrain. The cerebellum forms prominently at the posterior side of the MHB (Kimmel et al., 1995). Overall, in this period the animal possesses the classic vertebrate blueprint.

The **hatching** period extends from 48-72 hpf. Even when raising embryos at optimal conditions, the embryos from the same clutch hatch intermittently during the whole third day of development, and occasionally later (Nüsslein-Volhard & Dahm, 2002). During the hatching period morphogenesis of several organ primordia is already rather complete; the pronephric filtration begins; the yolk is depleted and the growth of the head is pronounced (Figure 13). The primordia of the pectoral fins, the jaws, and the gills are rapidly developing at this stage; otic vesicle walls are thickened and can be easily distinguished (Figure 13) (Kimmel et al., 1995). After the embryos hatch, they are called larvae, but the embryos that have hatched are not necessarily more developmentally advanced than the unhatched ones (Kimmel et al., 1995).

Most of the organogenesis is complete in the 3 dpf stage. The name given to this stage (protruding-mouth) refers to the evident morphogenetic changes occurring in the mouth; as it widely opens and protrudes anteriorly (Kimmel et al., 1995). Other featured morphogenesis includes: partitioning of the otic vesicle and formation of the semicircular canals; extension of the pectoral fin; cartilage condensation continues in the anterior pharyngeal arches (mandibular and hyoid arches) and begins in the posterior arches (ceratobranchial arches); gill filament begins to develop; the gut

becomes visible; the proximal convoluted tubule (PCT) of the pronephros starts to coil (Kimmel et al., 1995). At 4 dpf, the larvae develop digestive organs to an extent sufficient for food processing, thereafter the 5 dpf larvae show independent feeding and transition to free-swimming behavior. All organs are already developed and functioning at 5 dpf (Ng et al., 2005; Strähle et al., 2012).

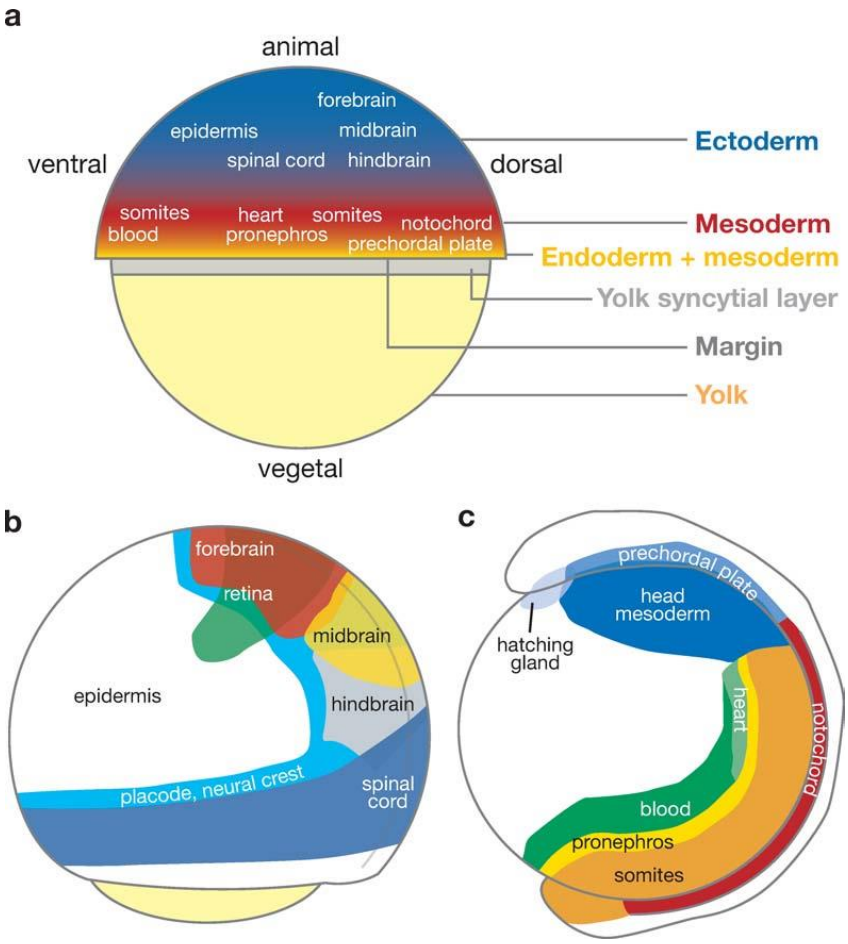


Figure 15: Fate maps of zebrafish embryos at three different stages. (a) fate map for the embryo in the gastrula stage 50% epiboly. (b) fate map of embryo at 90% epiboly. (c) fate map of mesoderm for embryo at early somite stage; lateral view anterior to the left. Color code and corresponding tissues are indicated. Figure is taken from (Schier & Talbot, 2005).

1.7 Aims

G proteins are extremely important since they are involved in many cellular processes and play a huge role in nearly all biological systems; in addition, they are outstanding pharmaceutical targets. As mentioned before, the Gv family was the most recent discovery of all the G alpha families and its function is not known yet. Thus, discovery of its role will answer a very fundamental question.

Herein, this dissertation seeks to perform more comprehensive analysis of Gv protein and unearth its role, using zebrafish (*Danio rerio*) as a model system, using a three-pronged approach.

The first aim is to attain a thorough analysis for Gv ontogenetic and tissue-specific expression during embryogenesis as a basis for functional consideration. The second aim is to validate, then investigate, and characterize the phenotype of the available Gv knockout on the functional and molecular level. The third aim is to delineate the evolutionary origin of Gv, to analyse Gv evolutionary dynamics in relation to the other G alpha families in a large data set of nearly thousand species and to mine the results for correlation to ecological context.

By studying the Gv protein from these different aspects we expect to gain a better understanding of Gv evolution and function. This final puzzle piece will bring us closer to completing the whole picture on G protein signalling, one of the most essential signal transduction pathways in the animal kingdom.

2 Materials and Methods

2.1 Materials

2.1.1 Media, buffers, and other solutions

In this study, all the solutions were prepared using milliQ water from a Millipore Milli Q Synthesis water purification system. The pH was adjusted whenever it is required either with 37% HCl (VWR) or dissolved sodium hydroxide (VWR). Solutions were sterilized by autoclaving either once or twice, particularly those used for *in situ* hybridization experiments. All the solutions were stored at room temperature, unless otherwise mentioned.

10x artificial fresh water stock solution: 344 mg $CaSO_4.2H_2O$, 394.4 mg $MgSO_4.7H_2O$, 292.2 mg NaCl, 174.2 mg K_2HPO_4, and 136.1 mg KH_2PO_4 were dissolved in milliQ water and filled up to 1L.

1x artificial fresh water: 100 ml of 10x artificial fresh water stock solution was filled upt to 1L with milliQ water.

60x E3 Premix Medium: 34.4 g of NaCl, 1.5 g of KCl, 5.8 g of $CaCl_2$ x 2 H_2O and 9,8 g $MgSO_4$ x 7 H_2O were dissolved in 2L milliQ water.

1x E3 Medium: 16.66 ml of the 60x E3 Premix Medium were filled up to 1L milliQ water.

1x E3 $^+$ PTU (0.04%): 40 mg of PTU were dissolved in 1L 1x E3 medium overnight.

LB-medium 1 liter: 20 g of LB Broth (Sigma) were suspended in 1 L of distilled water, autoclaved, and stored at 4 °C.

Ampicillin: stock solution is 100 mg/ml and the working solution is 100 µg/ml.

10x Pronase stock: 100 mg Pronase powder was dissolved in 10 ml 1x E3 medium and stored at -20 °C.

1x Pronase: 100 µl 10x Pronase stock was added to 900 µl 1x E3 Medium.

10x Phosphate buffered saline (PBS); pH 7.4-7.6: 80 g of NaCl, 2 g of KCl, 14.4 g of Na_2HPO_4 and 2.5 g of KH_2PO_4 filled up to 1L with milliQ water and stirred until dissolved. The pH was adjusted with NaOH to 7.4-7.6.

1x PBS: 100 ml 10x PBS is filled up to 1 L with milliQ water.

1xPBST: 100 ml of 10x PBS were filled up with autoclaved milliQ water to 1L and 100 µl of Tween-20 were added.

Ringer solution: 677 mg NaCL, 21 mg KCl, and 119.15 mg HEPES were added to 100 ml milliQ water, pH was adjusted to 7.2, and stored at 4 °C.

4% PFA: it was always prepared freshly. 6 g of PFA was dissolved in 150 mL of 1x PBS at 65 °C.

Hybridization solution: 11 ml of formamide, 5.5 ml of 20x SSC, 1100 µl of tRNA, 22 µl of Heparin (50 ng/ml), 202.4 µl of citric acid to pH 6.0, and 2.2 µl of Tween 20 were mixed together. The volume was adjusted with autoclaved milliQ water to 22 ml.

1 M Magnesium chloride: 20.33 g $MgCl_2$ was dissolved in 100 ml autoclaved milliQ water.

20X Saline-sodium citrate (SSC) buffer pH 7.0: 175.32 g NaCl, 88.23 g tri-Sodium citrate dihydrate (0.3 M $C_6H_5Na_3O_7.2H_2O$) were dissolved in autoclaved milliQ water filled up to 1L. pH was adjusted to 7.

5 M Sodium chloride: 292 g NaCl was dissolved in autoclaved milliQ water filled up to 1L.

Detection buffer for cryosection ISH pH 9.5: It was always freshly prepared. 100 ml 1M Tris-HCl pH 9.5, 20 ml of 5 M NaCl, and 10 ml 1M $MgCl_2$ were added to 870 ml autoclaved milliQ water.

60 µl of Tween 20 were mixed together and the volume was adjusted to 60 ml using twice autoclaved water.

Staining solution: It was always freshly prepared. 70 µl of BCIP and 46 µl of NBT were added to 20 ml of staining buffer and vortexed. The buffer was covered with aluminum foil to protect it from light.

staining buffer for WISH: It was always freshly prepared by adding 3 ml 1M Tris-HCl, 1.5 ml 1M $MgCl_2$, 0.6ml 5M NaCl, and 30 µl Tween20 24.6ml autoclaved milliQ water and filled up to 30 ml.

DNA extraction buffer (for swab method): 20 ml of 1 M Tris pH 7.5, 25 ml of 100 mM EDTA, 5 ml of 5 M NaCl, and 0.5 g of SDS were mixed together. The volume was adjusted to 100 ml with twice autoclaved water.

"Quick" DNA extraction buffer: 10 mM Tris pH 8.2, 2 mM EDTA, 0.2 % Triton, 200 µg/ml Proteinase K, stored at -20º C.

RIPA buffer: 3ml of 5 M NaCl, 1 ml NP-40 (IGEPAL® CA-630 from Sigma), 5ml of 10 % Sodium deoxycholate, 1ml of 10 % SDS, 5ml of 1 M Tris, pH 8.0, 1ml of 0.5 M EDTA, pH 8.0 were filled up to 100 ml with milliQ water.

SDS-PAGE running buffer: 30g Tris, 144g Glycine, and 10g SDS were dissolved in milliQ water filled up to 1 L.

10% ammonium persulfate: 100 mg ammonium persulfate was dissolved in 1 ml milliQ water.

10x TBS: 60.57g Tris and 87.66 g NaCl were dissolved in 1L milliQ water. The pH was adjusted to 7.4 with HCl.

1x TBS: 100 ml 10x TBS were filled up to 1L with milliQ water.

1xTBS-T: 100 ml of 10x TBS were filled up to 1L with milliQ water and 1 ml of Tween-20 were added.

10x fast transfer buffer: 58.15g Tris, 47.66g HEPES, 0.95g DMF, 2.93g EDTA, 3.656g (37%) $NaHSO_3$ filled up to 1L with milliQ water.

1x fast transfer Buffer: 50 ml 10x transfer buffer and 100 ml methanol were filled up to 1L with milliQ water.

1 M Tris-HCl; pH 8.8, 6.8, 9.5 and pH 7.5: 121.1g Tris was filled up to one liter and stirred until dissolved. pH was adjusted as required.

2.1.2 Enzymes, Kits, antibodies, and vectors

The following Kits, enzymes, vectors and antibodies were used in this study according to the manufacturer's instructions, unless otherwise mentioned:

Enzymes	**Manufacturer/Supplier**
FastDigest DpnI	ThermoFisher Scientific
HindIII-HF	New England Biolabs
MfeI	New England Biolabs
Pronase	Sigma
Proteinase K	Roche
T3-RNA Polymerase	Roche
TURBO Dnase	ThermoFisher Scientific
Kits	
KOD Hot Start Master Mix	Toyobo
My-Budget 5x PCR-Mastermix Ready-to-Load	Bio-Budget
OneTaq DNA Polymerase	New England Biolabs
primaQUANT 2x qPCR-CYBR-Green-Blue-MasterMix ohne ROX	Steinbrenner-laborsysteme
Q5 High-Fidelity DNA Polymerase	New England Biolabs
Amersham ECL Western Blotting Detection Reagent	GE Healthcare
EnGen® Spy Cas9 NLS	New England Biolabs
HiScribe T7	New England Biolabs
Nucleospin extract II	Macherey-Nagel
NucleoSpin Plasmid kit	Macherey-Nagel

Protector RNase Inhibitor	Roche
SuperScript IV Reverse Transcriptase	ThermoFisher Scientific
Antibodies	
Alexa Fluor 488 goat α-rabbit IgG	Invitrogen
Anti-DIG-AP IgG from sheep, polyclonal	Sigma-Aldrich
Anti-Gv rabbit IgG, polyclonal	Ivandic, 2015
Anti-Acetylated mouse Tubulin	Sigma
Vectors	
pCS2P-k虚lc19e-linker-eGFP-mut- donor	Hisano et al., 2015
pGEM-T Easy Vector System I	Promega
pOPINK	Addgene

2.1.3 Chemicals and reagents

All the chemicals and reagents used in this study were obtained from VWR, unless otherwise mentioned, at molecular biology grade. Chemicals and reagents obtained from other sources are listed below.

Description	**Manufacturer/Supplier**
Acetic anhydride	Sigma
BCIP (5-bromo-4-chloro-3-indolyl-phosphate)	Roche
Blocking Reagent	Roche
BSA (Bovine Serum Albumin)	Sigma
di-Sodium hydrogene phosphate	Merck
EDTA	Sigma
Gel Loading Buffer II	ThermoFisher Scientific
Glycogen (5 mg/ml)	ThermoFisher Scientific
Heparin	Sigma-Aldrich
Midori green advance DNA stain	NIPPON Genetics
NBT (4-Nitro blue tetrazolium chloride)	Roche
Nuclease free water	Gibco

PageRuler prestained protein standard	Life Technologies
Phenol Red	Fluka
Phenylthiocarbamide	Sigma
Pierce Lane Marker Reducing Sample Buffer	ThermoFisher Scientific
Protease Inhibitor Cocktail	Sigma
Sheep Serum	Sigma
TAE buffer 50x (Rotiphorese)	Roth
TAE buffer 50x (Rotiphorese)	Roth
TissueTek cryo-embedding medium	Sakura
Triethanolamine	Merck
Triton X-100	Merck
Tween20	Sigma
Vecta mount permanent mounting medium	Vector Labs
Yeast tRNA	Thermo Fischer Scientific

2.1.4 List of primers

Primer3Plus web tool (Untergasser et al., 2007) was used to design primers. The primers were produced by ThermoFisher Scientific as desalted standard DNA oligos at the scale of 25 nmoles. Upon delivery, the primers were dissolved in nuclease-free water to a final concentration of 100 µM and stored at -20 ºC. The primers were further diluted to 10 µM upon use, unless otherwise mentioned.

Type of experiment/Primer name	**Oligonucleotide**
CRISPR/cas9	
Core-gRNAOligo1	AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC
GvExon1gRNA	GCGTAATACGACTCACTATAGGGCTCAGAGGTGACAACAGGTTTTAGAGCTAGAAATAGC
GvgRNA-CH	GCGTAATACGACTCACTATAGGTGATGGACACCATTATCAGTTTTAGAGCTAGAAATAGC
eGFP-gRNA	GCGTAATACGACTCACTATAGGCGAGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGC

krtt1c19e-gRNA	GCGTAATACGACTCACTATAGCTTTACTTAACAAGGGACGGTTTTAGAGCTAGAAATAGC
GvEx1-40bplinker	AGAGGTGACAACAGAGGAGGACAAGAAGGCGAAGATCCACGGAGGAGGTGGTTCAGGTGG
eGFPtailRC	AAGCTTCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTT
Hind3GvEx9CH-40bpRC	TAATGGTGTCCATCACCACCTGGAAAACCACCTGGACGTTAAGCTTCACACCTCCCCCTG
gRNAeGFPF-Ex1-23bpF	GGCGAGGGCGATGCCACCTACGGAGAGGTGACAACAGAGGAGGACA
gRNAeGFP-Ex9-20bpR	GGCGAGGGCGATGCCACCTACGGTAATGGTGTCCATCACCACC
f1ori_R3	TTAATGCGCCGCTACAGGGC
reverse-2	GCGGATAACAATTTCACACAGG
qPCR primers	
atp1a1a.5 F	GTG GTT CTG TGC CTT CCC AT
atp1a1a.5 R	CGGGTGTTCATTTTGATGTT
ca2-F	AGGACGCAGTTGATAAGCCT
ca2-R	TGGACTTGATAGCATCCATA
calc-F1	TTCTTGTTGCCTACGCTCTG
calc-R1	TTCGATCTGGTGACGATTCC
calcr-F	CATTCTGGTGCCCTTATTGG
calcr-R	CCACTGCCTTTTCAAAGCTC
calcrla-F	AAAGTGGGCAGTGGTTCTTG
calcrla-R	ATTGCGGTCCTTCTTCCTTC
calcrlb-F	AGTTCATCAGCGAGCACATC
calcrlb-R	GTTCCGTTGTTTGGCTCTTG
calcr-R	CCACTGCCTTTTCAAAGCTC
casr F	TGGGTCTTTTCTACATCCCT
casr R	GTCTTTAAAGCCGGGTATATG
casr-F2	TGTACAGAATGCTCGGATGG
casr-R2	TTGGACCAGGAGTTGTTTGG
ECaC F	TCCTTTCCCATCACCCTCT
ECaC R	GCACTGTGGCAACTTTCGT
GAPDH F	GTGGAGTCTACTGGTGTCTTC
GAPDH R	GTGCAGGAGGCATTGCTTACA
Gna11a F	ACGTCACATCTATCATGTTCCT
Gna11a R	CCATGCGATTCTCGTTGTC
Gv-QPCR-Ex1-4 F	TCC ATC TAT GAA TCG CCT GCA G

Gv-QPCR-Ex1-4 R	GAA CAT GTC GGC GAT GAA GC
Gv-QPCR-Ex6-9 R	CCG AGT GAG CGT GAA CCT T
Gv-QPCR-Ex6-9F	CGA GTA CGC CAA GAG AGA GTT
HRatp1a1a.5 F	CGGGTTCTTTACCTACTTTGTGA
HRatp1a1a.5 R	CAGCCAATCCGTAATCCCAC
HRslc9a3.2-F	GCGAAACCCACCCTGGCAAAC
HRslc9a3.2-R	GGCGAAGGAGTCTGTGGAGCG
KSatp1a1a.4 F	TTCTGCCACTTCTGCCTTCC
KSatp1a1a.4 R	CACCTTGATTCCAGCACTCC
NarAtp1a1a.1 F	GCCCTGAGCAATTAGACGATG
NarAtp1a1a.1 R	TACGGCTACAATGGCACCCT
NCCatp1a1a.2 F	TCTACCTTTGGGCACCGTCAC
NCCatp1a1a.2 R	TGCTTGGATCATCCCGATTT
NCCslc12a3-F	CGATGATGGCGGTTTGACAC
NCCslc12a3-R	TGAAACCCAGACGGAACCTG
ncx1bF1	ATTGAGGCCATTACGGTCAG
ncx1bR1	GATTGACACCACGAAACACG
Plcb3F	AGCAGGCTAAAATGGCTGAA
Plcb3R	GCCCTGTGGTGATGTTTCTT
Plcg1 F	AAAGCGGTGGACATTTCAG
Plcg1 R	TAACTTCTTGTGCTTGATGAGG
pth 2 F	ATACGTTGTTTGGAGAAAGCC
pth 2 R	CATTGTGCATCAGCTGAACTT
pth1 F	GTTTCCATCAACGGGAATTT
pth1 R	CATCAGCTGCACTTCATTCA
rpl37 F	ATGACGAAGGGTACGTCG
rpl37 R	TTATGAGGAGCTGGACGC
slc12a3 F	CGATGATGGCGGTTTGACA C
slc12a3 R	TGAAACCCAGACGGAACCTG
slc12a10.2 F	GACCCAAGGTGGAGAGGACG
slc12a10.2 R	CAGTTGATACCGATACTCAGC
stc1-F1	CTCCGCGTACGAACTGGAT
stc1-R1	TGTCACACGTTGAGTTTTCG
vdra-F	TCATCATGCTGAGGTCCAAC
vdra-R	CCAACCTGAAACTTGACCAG
b2m F	GCC TTC ACC CCA GAG AAA GG
b2m R	GCG GTT GGG ATT TAC ATG TTG

elf1 F	CTT CTC AGG CTG ACT GTG C
elf1 R	CCG CTA GCA TTA CCC TCC
Slc26a4-F	TTCATTCGGATAGGGGTCAG
Slc26a4-R	AAGAGAATGGCGTCGTGAAC
bactin -F	GCC TTC CTT CCT GGG TAT GG
bactin-R	CAG ACG GAG TAT TTA CGC TCA G
rpl8-F	CCGAGACCAAGAAATCCAGAG
rpl8-R	CCAGCAACAACACCAACAAC
Genotyping	
Ex1GvF	ACAAATGTCCGCGTCTTCTG
Ex1GvR	CGACCAATGTAAACATCAAAATC
Cloning	
gnav1–EcoR1-F	TAAGCAGAATTCAATGGGTCTGTGTTTGGGCTC
gnav1-HindIII-R	TGCTTAAAGCTTGAGCAGAGAAACAGCCTCCAG
pOPINK GNAV1 1-F	AAGTTCTGTTTCAGGGCCCGATGGGTCTGTGTTTGGGCTCAGAG GTGACAAC
pOPINK GNAV1 362-R	ATGGTCTAGAAAGCTTTAGAGCAGAGAAACAGCCTCCAGGTTCT CCTTG
zf.GvHindIII-F	TAAGCAAAGCTTATGGGTCTGTGTTTGGGCTCA
zf.GvSacI-R	TGCTTAGAGCTCTCAGAGCAGAGAAACAGC
***In situ* hypridization primers**	
gnav1-T3F	TATTAACCCTCACTAAAGGGAATCACTCGGTTCTCTCGTGTG
gnav1-T3R	TATTAACCCTCACTAAAGGGAAAATGGTGTCCATCACCACCT
gnav1-F	TCACTCGGTTCTCTCGTGTG
gnav1-R	AATGGTGTCCATCACCACCT
gnav1-F3	CTGTGTTTGGGCTCAGAGGT
gnav1-R3	TGTGGTTCTCAGTCGCACTC
gnav1-T3F3	TATTAACCCTCACTAAAGGGAACTGTGTTTGGGCTCAGAGGT
gnav1-T3R3	TATTAACCCTCACTAAAGGGAATGTGGTTCTCAGTCGCACTC
gnav1-n-f	ATGGGTCTGTGTTTGGGCTC
gnav1-n-r	CATCAAAACAACGGCCACAC
gnav1-N-T3f	TATTAACCCTCACTAAAGGGAAATGGGTCTGTGTTTGGGCTC
gnav1-N-T3r	TATTAACCCTCACTAAAGGGAACATCAAAACAACGGCCACAC

2.1.5 Technical equipment and disposables

Type and model	Manufacturer
Cryostat CM1900	Leica
Electrophoresis Chamber	Bio-Rad
Electrophoresis Powersupply Power Pac 300	Bio-Rad
Fluorescence Microscope BZ-9000	Keyence
Fusion FX-6 Edge system RPN2134	Vilber Lourmat
Gel Documentation System GelDoc XR	Bio-Rad
Halogen Lamp iLux 150NL	Visitool
Heating Plate MEDAX	StörkTronic
Incubator (dry)	Memmert
Incubator (wet)	GFL KMF
Laser Scanning Microscope LSM 510	Zeiss
Laser Scanning Microscope SP8	Leica
Microcentrifuge (cooling) Z233	MK-2 Hermle
Microcentrifuge Micro Star 17	VWR
Micropipette Puller P-97	Sutter Instruments
Microwave	900 & Grill Severin
Mini Gel Tank and Blot Module Set	life technologies
Peristaltic Pump	Pharmacia Fine Chemicals
pH-Meter	Calimatic 766 Knick
Pipets (10/20/200/1000 µl) Research Plus	Eppendorf
Pneumatic Pico Pump PV 830	World Precision Instruments

Precision scale LA 120 S	Sartorius
Rotary Mixer R1	Pelco
Scale universal	Sartorius
Shaker Bio Shaker 3D	BIOSAN
SLR Camera D5100	Nikon
Spectrophotometer NanoDrop One	Thermo Fisher
Stereomicroscope Stemi 2000	Zeiss
Thermal mixer Thermomixer comfort	Eppendorf
Thermcycler LifeEco	BIOER
Vortex Vortex Genie 2	Bender & H

Disposables	**Manufacturer/Supplier**
96-well PCR plates	Bio-Rad
Borosilicate Capillary GB100F-10	Science Products
Conical Tubes 15 ml, 50 ml	Falcon
Coverslips 24x60 mm	VWR
Latex Gloves powderfree	VWR
Microcentrifuge Tubes 1.5 ml, 2 ml	Axygen
Microloader Tips 20 µl	Eppendorf
Microscope Slides Standard	VWR
Microscope Slides Superfrost	VWR
Microseal 'B' PCR Plate Sealing Film	Bio-Rad
Pasteur Pipet	VWR
PCR-Tubes	VWR
Petri Dish Ø 60mm, 100 mm	VWR
Pipet Tips 10/200/1000 µl	VWR
Pipet Tips with Filter	SurPhob Biozym
PVDF membranes	Millipore

2.2 Fish husbandry and breeding

The adult zebrafish were kept in the fish facility with a constantly circulating and filtering water system with day/night cycle of 14 hours light and 10 hours darkness at 27–28 °C. The water system contains a mixture of equal parts of deionized water and tap water, and it is kept at pH 7.0 – 7.4. The zebrafish strains used in this study were Ab/Tüb^{wildtype} (Tübingen), and KS$^{gnav1-/-}$ mutant line that has been generated in the Korsching laboratory from the KS strain (Cologne origin). The mutants were crossed out several times with the Ab/Tüb^{wildtype} strain, hence the resulting strain of the (KS/Ab/Tüb)$^{gnav1-/-}$ was used in all the experiments conducted in this study. The wildtype and mutant fish were raised in the same conditions in all experiments conducted.

In the early evening, single-pairs were put into mating tanks each possessing a sieve and removable divider, at least three pairs from each genotype were prepared for mating. Next morning, the dividers were removed; the spawned eggs that fell through the sieve at the bottom of the tank were collected, washed, and treated according to the respective experiment.

For the fish fecundity analysis, the eggs were washed and kept in E3 medium; the total number of fertilized eggs were counted after 90 minutes of mating. For analyzing the egg viability, the eggs were washed and kept in E3 medium; the live and dead eggs were counted in the first five days post fertilization. The majority of the mortality was observed in the first 24 hpf, both for wildtype and mutant, so in later experiments we evaluated the survival rate only at 1 dpf embryonic stage. With respect to analyzing the hatching rates, the spawned eggs were raised in either E3 medium or artificial freshwater. The hatched and unhatched eggs were counted in the first three days post fertilization.

For fish maintenance, the wildtype and mutants (eggs, embryos, larvae, and adults) were raised in the same conditions until adulthood. The eggs were raised in Petri dishes containing E3 medium until 5 dpf larval stage. On day 5, the larvae were transferred to cages containing about 1L of E2 medium without waterflow, and the

larvae were fed twice a day with paramecia and dry food, respectively. Thereafter, 20 dpf larvae were transferred to large water tanks that are connected to the circulated water system, and the fish were fed with live brine shrimp (Arthemia) and flakes until adulthood.

All animal experiments were performed in accordance with the German Animal Welfare Act as well as with the General Administrative Directive for the Execution of the Protection of Animals Act and were approved by the governmental animal care and use office (Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen, Protocol No. 8.87-51.05.20.10.217).

2.3 Molecular biology techniques

2.3.1 Preparation of genomic DNA from adult zebrafish and embryos

Fin biopsy and swab methods were performed as described in (Bedell et al., 2012 and Breacker et al., 2017), respectively, in order to obtain the DNA from the adult fish (grown to approximately 2.5–3 months) with some modifications on the fin biopsy technique. In the latter method, the caudal fin was lysed in 50µl “Quick" DNA extraction buffer freshly supplemented with 200 ng/µl Proteinase with shaking in a thermomixer at 350 rpm and 50-55 °C for two hours. Then, 300 µL of nuclease-free water was added to the lysed fin, vortexed, and only 2-5 µl were used for the PCR reaction. To obtain DNA from the injected embryos (5 dpf), they were sacrificed on ice and then lysed and processed as described in the swab method (Bedell et al., 2012). DNA samples were stored at 4°C until they were used in PCR.

The obtained results from fin biopsy and swab methods were similar, therefore, swab-based genotyping method was adopted, as it has many advantages over fin biopsy-based genotyping method. Swabbing represents a less invasive and reliable alternative, requiring less time and effort. In particular, there is no need to anesthetize the fish thus negating its side effect on the fish behavior and biological system. On the other hand, fin biopsy is time-consuming and has some drawbacks, as It has potential effect on the fish behavior, especially that the fin is used in swimming and reproductive

behavior (Spence et al., 2008). Unlike swabbing-based method, using anesthesia in fin biopsy can elicit cortisol responses, thereby leading to inadvertent consequences on the results of studying behavior, homeostasis, and immune responses (Carter et al., 2011; Sladky et al., 2001).

2.3.2 DNA amplification by PCR and colony PCR

PCR reaction was conducted for different purposes in this study including: amplification of templates for *in vitro* transcription, genotyping of zebrafish in order to identify genetic modification, subcloning of specific DNA fragment, and endpoint RT-PCR for cDNA fragments. Primers were designed using Primer3Plus web tool (Untergasser et al., 2007), and they are listed in 2.1.4. Q5 high-fidelity DNA polymerase (New England Biolabs) was used to generate DNA templates for *in vitro* transcription, transformation into vectors, and producing gRNAs acoording to the manufacturer' instructions. While GoTaq green master mix (Bio-Budget) was used for all other PCR reactions. 10-100 ng of DNA template was used for each reaction; for colony PCR, a bacterial colony was picked from the plate with a pipette tip and mixed with the PCR reaction solution, the tip was then used to inoculate 4 ml LB medium for bacterial culture.

The standard PCR conditions for Taq polymerase are as following: initial denaturation at 95 °C for 5 minutes, (denaturation at 95 °C for 30 seconds, primer annealing at 55-60 °C for 45 seconds, and extension at 72 °C for 1 minute) for 40 cycles, and final extension at 72 °C for 5 minutes; then the samples were kept at 4 °C until further use. The optimal annealing temperature was calculated using the Tm-calculator provided on the ThermoFisher website and the range of the optimal annealing was 55-60 °C.

2-5 µl from PCR products were resolved by electrophoresis on 1% agarose Tris-acetate-EDTA (TAE) gels stained with Midori green (NIPPON Genetics), and visualized by trans-UV illumination using the Gel Doc XR (Bio-Rad).

Then the PCR product was purified from either a total reaction solution or from a gel band by NucleoSpin Gel and PCR Clean-up kit (Macherey–Nagel) according to the manufacturer's instructions. The total DNA concentration was measured with a

NanoDrop photometer (ThermoFisher Scientific, Germany) and the samples were stored at -20 ºC until further use.

2.3.3 DNA sequencing

Sequencing was carried out for different DNA fragments in this study including: amplified DNA template used for *in vitro* transcription to generate RNA probes, amplified genomic DNA fragments used for identifying genomic modification, DNA fragments used for subcloning, and plasmids. Sequencing was performed by Microsynth (Germany) sequencing services. DNA samples were prepared as described by the company and usually contained 4-30 ng/µl of DNA in PCR products or 40-100 ng/µl plasmid.

2.3.4 RNA extraction, RT-PCR, and qPCR

In order to analyze the differential expression for many genes in the wildtype and mutants (adult kidney, embryos, and larvae) all the samples were treated equally in all the following steps. The freshly-dissected kidneys and pools of 20 embryos/larvae (from 6 hpf to 5 dpf raised in E3 medium) (see 2.5.2 and 2.5.3 for sample collection) were stored in 500 µl of TRI Reagent (Sigma-Aldrich) RNA extraction buffer at - 80 °C until being extracted. On the day of extraction, the samples were thawed on ice and then homogenized for 10 seconds by ultra soundwaves with Sonicator (Bioruptor, Diagenode) at 30 W. The samples were then incubated at room temperature for 5 minutes and the RNA extraction was carried out according to the manufacturer's instructions. Eventually, the RNA pellet was dissolved in 30 µl of nuclease-free 1 mM Tris-HCl, 0.1 mM EDTA, and pH 8.5 solution. Thereafter, the extracted RNA was cleaned up from DNA contamination by DNA-free DNA Removal Kit (Invitrogen) according to the manufacturer's instructions. The total RNA concentration was determined with a NanoDrop photometer (ThermoFisher Scientific, Germany); only samples with OD 260/280 ratio ≥ 1.8 were processed further and stored at - 80 °C.

Next, the SuperScript IV First-Strand cDNA Synthesis System (ThermoFisher Scientific) was used to generate cDNA from 1 µg of total RNA in a 20 µl reaction with

2.5 µM Oligo d(T)20 primer according to the manufacturer's instructions. We included negative control which contained no reverse transcriptase (NRT) for each sample in this assay. Ultimately, all the cDNA samples were diluted one to ten folds with nuclease-free water, and then aliquoted in sterile PCR tubes and stored at -20 °C until further use. Before conduction Quantitative RT-PCR (qPCR), 10 ng of cDNA was used as template in (endpoint) PCRs using My-Budget PCR master mix (Bio-Budget) for 35 PCR cycles, and 0.5 mM final concentration of primers (see 2.1.4). The samples were resolved by electrophoresis on 1% agarose TAE gels stained with Midori green (NIPPON Genetics), and visualized by trans-UV illumination using the Gel Doc XR (Bio-Rad).

We considered several criteria when we designed the primers for the qPCR assay including the followings: the amplified product should have a size of 80–200 bp; the amplified region and the primers should have a GC content of 50–60%; the amplified region should span two exons or at least one primer spans exon-exon junction; the primers annealing temperature should be 59-60 °C; and the primers were designed to target the 3'UTR region as much as possible.

qPCR experiment was performed with the CFX96 Touch Real-Time PCR Detection System (Bio-Rad). The qPCR reaction was carried out using primaQUANT qPCR-SYBR-Green master mix (Steinbrenner- Laborsysteme) with 10 ng of cDNA, 0.5 µM final concentration of primers in 5 µl reactions. The reaction mixes were prepared in 96 well qPCR plate on ice, the plate was sealed, spun for ten seconds at 1000 rpm, and put into the thermal cycler. Two negative control samples were included in each assay: no template control (NTC) in order to monitor contamination and primer-dimer formation; and a pool of NRT to monitor genomic DNA contamination. As a starting point, in order to evaluate the efficiency of the primers, we performed a 5-fold dilution series from pooled cDNA samples (wildtype and mutants) of 6 steps for each primer. A graph was generated with the results and the slope of the data series plot was calculated which allowed to calculate the primer efficiency; a value between 85% and 110% was considered optimal. Furthermore, the geNorm algorithm in the software package qbase^{+} was used in order to determine the optimal reference gene that exhibits stable expression levels among different samples and under different

experimental conditions (Hellemans et al., 2007). We evaluated the stability ranking of six candidates (*gapdh*, *elf1*, *b2m*, *bactin*, *rpl8*, and *rpl37*) to determine the optimal number and most stable reference genes for being used in two different assays; one with wildtype and mutant kidney samples, and another one with embryos/larvae from different genotypes and developmental stages. As a result, *b2m* and *bactin* were optimal for the kidney assays, while *rpl8* and *rpl37* were optimal for the developmental assays. However, we used *elf1*, *b2m*, *bactin*, and *rpl8* in the kidney assays for more robust evaluation, in particular for siblings, unless otherwise mentioned.

qPCR was conducted with 2 step cycler settings, as follows: initial denaturation at 95 °C for 2 minutes, (denaturation at 95 °C for 5 seconds, primer annealing at 60 °C for 30 seconds) for 39 cycles, including real-time fluorescence measurement at the end of elongation, followed by Melt Curve analysis ranging from 65 °C to 95 °C in steps of 0.5°C.

All data were analyzed using the Bio-Rad CFX Manager software (Bio-Rad). qPCR reactions were performed in replicates, i.e., at least three biological replicates and two technical replicates. Dispersion of technical replicates Cq values from the mean (standard deviation (SD)) was calculated; and the samples were excluded from the analysis if the SD value was ≥ 0.5 cycle. We also evaluated the melting curves for each sample, and we excluded the samples that show a secondary peak. We considered including all types of samples (genotypes) being tested for the respective gene (amplicon) in the same qPCR run. All mRNA levels were calculated as fold expression relative to respective reference genes using the Pfaffl method, the fold changes of the mRNA levels of the same sample relative to different reference genes were averaged. Finally, the significance was determined with two tailed unpaired t-test using GraphPad Prism version 9 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com. Error bars on the plots always indicate SEM.

2.4 CRISPR/Cas9-based genome editing in zebrafish

2.4.1 Guide RNA template assembly and preparation of gRNA

The guide RNA (gRNA) target sequence in exon 1 of *gnav1* was used from (Ivandic, 2015). In order to determine the best gRNA target sequence in exon 9 of *gnav1* in the zebrafish (*Danio rerio*) genome assembly GRCz11, we used the web tool CHOPCHOP. We chose the target site with the best predicted genome editing efficiency and with no potential off-targets (Labun et al., 2019). The eGFP gRNA and krtt1c19e gRNA sequences were used from (Hisano et al., 2015). The gRNA oligos consist of the T7 promoter, 18-20 bp target sequences, and 20 bp sequence that overlap to a generic gRNA template (Core-gRNA). All the gRNA oligos used in this assay to amplify the DNA templates are listed in 2.1.4, i.e., Core-gRNAOligo1, GvExon1gRNA, GvgRNA-CH, eGFP-gRNA, and krtt1c19e-gRNA.

The DNA templates of the gRNAs were generated by a cloning free method. the reaction was assembled as follows: 4µl 5X Buffer, 0.4µl 10mM dNTP, 1µl 100 µM Oligo1 (Core-gRNA), 1µl 100 µM Oligo2 (gRNA of interest), 0.2 µl Q5 High-Fidelity DNA polymerase (New England Biolabs), and nuclease-free water up to 20µl were added in a PCR tube. The assembled reaction was denatured at 98°C for 2 minutes, then annealed at 50°C 10 minutes, and finally extended at 72°C for 10 minutes in the thermal cycler. 1µl of the product mixed with Gel Loading Buffer II (ThermoFisher Scientific) was run on 2 % agarose gel (for 10-30 minutes at 90 voltage) to evaluate the presence of a 120 bp band. Thereafter, the DNA template solution was purified by NucleoSpin Gel and PCR Clean-up kit (Macherey–Nagel) according to the manufacturer's instructions.

Next, the purified DNA templates of the gRNAs were used for in vitro transcription with the HiScribe T7 kit (New England Biolabs). The reaction was assembled as follows: 1-2 µg of DNA templates, 2 µl 10x reaction buffer, 2 µl 10 mM NTP mix, 2 µl T7 polymerase, 0.5 µl Protector RNase Inhibitor (Roche), and nuclease-free water up to 20 µl were added and the mix was incubated overnight at 37 °C. The next day the DNA template was digested by incubating the transcription mix with 1 µl TURBO

DNase (2 U/µL) (ThermoFisher scientific) for 15 minutes at 37°C. The digestion was stopped by adding 5 µl 100mM EDTA of pH 8. The RNA was then precipitated by adding 0.1 volume of 3M sodium acetate, 2.5 volume of pre-chilled 100 % ethanol, 1µl 5mg/ml glycogen, and then incubated at - 20 °C overnight. RNA was pelleted by spinning for 30 minutes at 15,000 rpm at 4°C. Pellet was washed with 70% pre-chilled ethanol and centrifuged again for 5 minutes. The supernatant was removed and the pellet was dried and re-suspended in 30 µl nuclease-free water. 1 µl of the gRNA was run with Gel Loading Buffer II (ThermoFisher Scientific) on 2 % agarose gel (for 10-30 minutes with 90 voltage) to evaluate the presence of a distinct band. RNA concentration was measured with a NanoDrop photometer (ThermoFisher Scientific, Germany); only samples with OD 260/280 ratio ≥ 1.8 were stored at -80 °C until processed further.

2.4.2 Construction of donor vector

We used the plasmid pCS2P-krtt1c19e-linker-eGFP-mut-donor from (Hisano et al., 2015) and made some modifications. Our target was to obtain a donor insert that harbors the two Gv 40 bp homology arms surrounding the linker-eGFP mutated sequence, and eGFP-gRNA target sequences at each side of the donor insert. First, we digested the plasmid to remove the krtt1c19e homology arms and obtain the linker-eGFP mutated sequence separately. The plasmid was digested with MfeI and BtgI (New England Biolabs) according to the manufacturer's instructions. The digested product was run on a gel and the band of interest (957 bp) was purified and eluted with 25 µl nuclease-free water using NucleoSpin Gel and PCR Clean-up kit (Macherey–Nagel) according to the manufacturer's instructions. The purified donor insert (template 1) was used in the next PCR reaction. In order to amplify the donor sequence that harbors the Gv 40 bp homology arms surrounding the linker-eGFP mutated sequence, we performed two consecutive PCR reactions (template 2 and 3) (see below). These were followed by a third PCR reaction to introduce the eGFP-gRNA target sequences at each side of the donor insert. Eventually, the final product would be the donor insert that has Gv homology arms, linker-eGFP mutated sequence,

and eGFP-gRNA target sequences in addition to two cut sites (for BtgI and HindIII) to enable easier modifications on the homology arms in the future.

In more details, the first PCR was set up to add Gv exon 1 homology arm and also a HindIII cut site (introduced at the 3' of the donor insert) to the linker-eGFP mutated sequence. The PCR amplification was conducted using Q5 High-Fidelity DNA polymerase (New England Biolabs) according to the manufacturer's instructions. Template 1 was used in addition to the forward (GvEx1-40bplinker) and reverse (eGFPtailRC) primers in 2-step PCR reaction (initial denaturation step at 98 °C for 30 seconds, followed by 35 cycles from 2 steps (denaturation at 98 °C for 10 seconds, extension at 72 °C for 90 seconds), and final elongation at 72 °C for 2 minutes. The resulting template was digested with DpnI (ThermoFisher Scientific) to clean up the reaction from the methylated donor insert template according to the manufacturer's instructions; and then it was purified as previously described in 2.3.2. The latter purified product (template 2) which already contains Gv exon 1 homology arm and the linker-eGFP mutated sequence was used as a template in the second PCR reaction. Q5 High-Fidelity DNA polymerase was used along with template 2, forward (GvEx1-40bplinker), and reverse (Hind3GvEx9CH-40bpRC) primers in 2-step PCR reaction similar to the first PCR conditions. The second PCR reaction resulted in a donor insert that has both Gv homology arms and linker-eGFP mutated sequence. The purified product from the second PCR reaction was used as a template (template 3) in the third and last PCR reaction.

The last PCR reaction was carried out using KOD Hot Start Master Mix (Toyobo) using template 3, forward (gRNAeGFPF-Ex1-23bpF) and reverse (gRNAeGFP-Ex9-20bpR) primers in 2-step PCR reaction. Thereafter, 1 µl of One Taq DNA polymerase (New England Biolabs) was added to the PCR reaction mix followed by an additional thermal cycle at 72 °C for 10 minutes in order to generate 3'A-tailed insert (a requirement for ligation into pGEM-t vector). The final product (donor insert) was run on a gel and the 1126 bp band was purified and eluted with 20 µl nuclease-free water using NucleoSpin Gel and PCR Clean-up kit (Macherey–Nagel) according to the manufacturer's instructions. The sequence of donor insert was confirmed by DNA sequencing.

2.4.3 Cloning the donor insert (cassette integration)

The donor insert was inserted into pGEM-T Easy Vector (Promega). Before carrying out the ligation reaction, the doner insert was incubated at 65 °C for 10 minutes and then kept on ice for 3 minutes. Next, a ligation reaction was set up using 5 µl 2X Rapid Ligation Buffer which contains T4 DNA Ligase, 1µl pGEM-T Easy Vector (50ng), 160 ng donor insert, and nuclease-free water to a final volume of 10 µl. The ligation reaction was incubated overnight at 4 °C.

Thereafter, the ligated product was transformed into chemical competent cells (DH5α cells); 2 µl of the ligation mix was added to 50 µl chemical competent cells and incubated for 30 minutes on ice. After incubation, the cells were heat shocked at 42 °C for 2 minutes; they were then returned to ice for 2 minutes. 600 µl LB media was added to the cells followed by incubation for 45 minutes at 37 °C with shaking. The cells were briefly pelleted and resuspended in 100 µl fresh LB media. 100 µl of the transformation culture was plated on LB/ampicillin/IPTG/X-Gal plates and incubated overnight at 37 °C. The next day, only the blue colonies were picked as this mostly indicates successful insertion into the vector. A colony PCR was performed as described in 2.3.2 with simultaneous inoculation into 4 ml LB/ampicillin media. The cultures were incubated overnight at 37 °C. Plasmids from positively tested clones were extracted from LB cultures using the NucleoSpin Plasmid kit (Macherey-Nagel) according to the manufacturer's instructions. Plasmid (donor vectors) was then sequenced with either M13 forward or SP6 primers to confirm the correct insertion and intact donor sequence.

2.4.4 CRISPR/Cas9 microinjections

Needles generated from borosilicate glass capillaries (1.0 mm O.D. x 0.58 mm I.D, Harvard Apparatus) were used for microinjections after being pulled with a Flaming/Brown micropipette puller (Sutter Instruments). The tip of the capillaries was broken by a sharp forceps to open the capillary and create a sharp tip; the drop volume to be injected by the needle was adjusted to 1-2 nL by using an object micrometer. One cell stage wildtype embryo was placed in agarose microinjection molds, oriented

and injected with the cas9/gRNA mix via Pneumatic Pico Pump PV 830 (World Precision Instruments). The instrument was calibrated to an eject pressure of ~50 psi while the hold-pressure was adjusted as such to avoid the unwanted in- or outflow.

In order to check the efficiency of the gRNA against exon 9, microinjections were performed using 300 pg of Cas9-NLS (New England Biolabs) and 25 pg of Gvexon9-gRNA mixture. One cell stage wildetype embryos were injected with 1 nl of the Cas9/gRNA mix in the cell mass or in the yolk sac immediately after fertilization.

After injection, the embryos were raised in E3 medium for 5 days at 28°C. 5 dpf larvae were sacrificed on ice and placed in separate tubes; their genomic DNA was examined for potential mutations as described in 2.3.1 to 2.3.3.

Unfortunately, it was not possible to perform the subsequent steps for plasmid Co-injection with both gRNAs and Cas9 protein and raising of injected eggs within the time frame of this study.

2.5 Histological methods

2.5.1 Generation of Digoxigenin labeled RNA probes

First, DNA templates for the sense and antisense *gnav1* strands were amplified by PCR using Q5 High-Fidelity DNA polymerase (New England Biolabs) according to the manufacturer's instructions. The primers used to generate the DNA templates of the probes harbor T3 promoter sequence and they are listed in 2.1.4. cDNA of 5 dpf larvae was used as a template in the PCR reaction. The DNA templates were then purified and sequenced as described in 2.3.2 and 2.3.3, respectively. After evaluating the sequence of the DNA templates, the DIG-labeled RNA probes were produced from these templates by in vitro transcription using T3 RNA polymerase (Roche) according to the manufacturer's instructions. The generated DIG-labeled RNA probes were then purified as follows: the in vitro transcription product was precipitated with 75 µl of pre-chilled 100 % ethanol, 2µl 0.2M EDTA, and 2.5 µl of 4M Lithium chloride for 2 hours at - 80 °C. Next, the RNA probes were pelleted by spinning for 30 minutes at 12000

rpm at 4°C. Pellets were washed with 70% pre-chilled ethanol and centrifuged again for 5 minutes. The supernatant was discarded and the RNA pellets were dried and re-suspended in 30 µl nuclease-free 1 mM Tris-HCl, 0.1 mM EDTA, and pH 8.5 solution. RNA concentrations were measured with a NanoDrop photometer (ThermoFischer Scientific, Germany); only samples with OD 260/280 ratio ≥ 1.8 were stored at - 80 °C until further use.

2.5.2 Wholemount *in situ* hybridization for 6 hpf embryos to 5 dpf larvae

The zebrafish embryos were raised until 12 hpf in E3 medium, and then the E3 medium was supplemented with 0.0045% (w/v) Phenylthiourea to block pigmentation. After the embryos reached the desired stage, they were sacrificed on ice and fixed in 4% paraformaldehyde (PFA) in 1xPBS overnight at 4 °C. Next day, the fixed samples were washed with 1xPBS and dehydrated gradually in methanol, then stored in 100% methanol at -20 °C until further use. WISH for zebrafish embryos (6, 12, 24, and 48 hpf) and larvae (3 and 5 dpf) were carried out as described in (Thisse & Thisse, 2008) with some modifications. The samples were rehydrated gradually in 75/50/25 % methanol in 1xPBS for 5 minutes at room temperature, and then washed with 1xPBS/0.1%Tween 20 (PBST) extensively. Thereafter, we permeabilized only 24-120 hpf embryos/larvae by digestion with proteinase K (10 µg/ml); 24 hpf embryos were digested for 15 minutes, and 48-150 hpf were digested for 30 minutes). The embryos/larvae were re-fixed in 4% PFA in 1xPBS for 20 minutes at room temperature, followed by four washes in PBST. The embryos/larvae were pre-hybridized in a hybridization solution for 1 hour at 65°C. Hybridization was carried out with 200 ng/ml DIG-labeled RNA probe in hybridization solution overnight at 65°C water bath.

The following day, the embryos/larvae were washed with 50% formamide/5x SSC and then gradually changed to 2x SSC (25/50/75/100% 2x SSC) for 15 minutes at 65°C each step. The embryos/larvae were washed twice for 30 minutes in 0.2x SSC at 65°C to quench unspecific signals and then stepwise changed to PBST (25/50/75/100% PBST) for 10 minutes at room temperature for each step. The embryos/larvae were blocked in 2% sheep serum/ 0.2% BSA/ 0.5 mg/ml levamisole in PBST for 60 minutes

at room temperature and then incubated in a 1:5000 dilution of α-DIG antibody conjugated to alkaline phosphatase (Roche) in the blocking solution overnight at 4°C.

The next day, embryos/larvae were washed with PBST seven times for 15 minutes at room temperature. They were then washed with the staining buffer three times for 5 minutes followed by incubation with NBT/BCIP in the staining buffer to start the color reaction. The embryos/larvae were checked regularly in the first 2 hours, if no staining was observed, they were further incubated at room temperature overnight. The staining reaction was stopped by immersion into 1xPBS. The embryos/larvae were then stored in PBST until being photographed with Olympus SZX16 stereo microscope (Olympus, Japan) using the cellSens Standard Imaging Software (Olympus).

2.5.3 *In situ* hybridization for kidney cryosections

First, after the adult zebrafish were sacrificed on ice, the head was removed and then the fish were opened at the ventral side from anterior to posterior. Swim bladder, spleen, intestine, stomach, heart, and adipose tissue were removed, so that the kidney was exposed. The opened fish was attached to the stage and immersed in 4% PFA, and fixed overnight at 4 °C. Next day, the zebrafish Kidneys were removed from the dorsal body wall from posterior to anterior and washed with 1xPBS followed by incubation with 30% sucrose overnight at 4°C to cryoprotect the samples. After cryoprotection they were embedded in TissueTek (Sakura) and frozen. 10 µm thick cryosections of the kidney were produced with a cryostat (Leica CM1900) and dried at 37 °C for 2 hours. The kidney sections on the slides were fixed in 4 % PFA in 1xPBS for 5 minutes at room temperature. Fixation was followed by three washing steps in 1x PBS for 5 minutes, the tissues were then permeabilized with proteinase K (1 µg/ml) in 1x PBS for 10 minutes at room temperature. The slides were then washed three times with 1x PBS for 5 minutes at room temperature followed by fixation in 4% PFA in 1x PBS for 5 minutes at room temperature and another three washes with 1x PBS for 3 minutes each. Thereafter, the slides were incubated for 10 minutes at room temperature with 5 mM acetic anhydride/0.1% triethanolamine/HCL followed by a wash with 1x PBS for 5 minutes. Hybridization was carried out with 100-200 ng/ml

DIG-labeled RNA probe in hybridization solution supplemented with 5x Denhardt's reagent overnight at 60°C in a wet chamber.

Next, the sections were shortly washed with 5x SSC followed by incubation in 50 % formamide/2x SSC for 30 minutes and subsequently in 0.2x SSC for 60 minutes at 65°C. Then, the sections were incubated for 15 minutes in 0.2x SSC, and washed for 5 minutes in 1x PBS at room temperature. The sections were blocked for 60 minutes in 0.5 % blocking reagent (Roche) in 1x PBS at room temperature. Then, they were incubated with a 1:500 dilution of α-DIG antibody conjugated to alkaline phosphatase (Roche) in 0.5 % blocking reagent for 2 hours at room temperature. Thereafter, the sections were washed with 1xPBS six times for 10 minutes each, followed by three washes with the detection buffer for 5 minutes. The sections were then incubated with NBT/BCIP in the detection buffer to start the color reaction. Slides were checked regularly, if no signal was observed within the first 2 hours at room temperature, the slides were then incubated with the detection solution overnight at 4°C. The color reaction was stopped by rinsing the slides in 1x PBS for 5 minutes followed by brief wash in water. Slides were mounted with VectaMount (Vector Laboratories) and photographed with the Keyence BZ-9000 microscope (Keyence, Japan).

2.5.4 wholemount immunofluorescence staining for larvae and adult kidney

The 3 and 5 dpf larvae and adult kidneys were prepared for wholemount immunofluorescence staining as explained in 2.5.2 and 2.5.3, respectively. After fixation, the samples were washed six times with 1xPBST for 10 minutes at room temperature. They were then stored in 100% methanol at - 20 °C until further use.

The samples were rehydrated in a PBST dilution series in methanol 75/50/25% for 5 minute each, washed six times for 10 minutes in PBST, and blocked with PBST supplemented with 5% sheep serum, 1.5% BSA, and 1% DMSO overnight at 4 °C. Thereafter, the samples were incubated in primary antibody solution (1:100) in the blocking buffer for 3 days at 4 °C.

Next, the samples were washed six times for 10 minutes in PBST, and then incubated in a secondary antibody solution (Alexa 488, 1:200) in PBST for 2 days at 4 °C. The samples were then washed extensively in PBST (six times for 60 minutes each). Only the kidneys were stained with DAPI for 2 hours at room temperature and then washed with PBST twice for 10 minutes. All the samples were then embedded in 100% glycerol and imaged with a Leica SP8 confocal microscope (Leica, Germany). The ImageJ (NIH) distribution FIJI was used for data analysis; images were enhanced by linear adjustment of the image brightness and contrast and slight background subtraction.

2.5.5 Cartilage and bone staining

Acid-free double staining protocol (alcian blue/ Alizarin Red S (ARS)) was carried out in order to stain the bones and cartilages of zebrafish larvae. Alcian blue is a positively charged dye stains the cartilage through an electrostatic interaction with the negatively charged acidic mucopolysaccharides thereby it binds the proteoglycans of epithelial and connective tissue (Adolphe et al., 1997). ARS staining binds to the mineralized bone matrix (mainly calcium) thus it stains the bones (Puchtler et al., 1969). The wildtype and mutant zebrafish larvae were raised in artificial freshwater and fed, 5-11 dpf, with paramecia until the desired age (3, 5, and 11 dpf). The larvae were sacrificed on ice and then fixed in 4% PFA in PBS overnight at room temperature with horizontal shaking. The samples were then stored in the fridge at 4 °C in 4% PFA solution. Thereafter, the acid-free double staining protocol was performed for all the fixed samples as described in (Walker & Kimmel, 2007). Following the alcian blue/ARS staining, the embryos/larvae were stored in 50% glycerol/0.1% KOH at 4 °C until being photographed with Olympus SZX16 stereo microscope (Olympus, Japan) using the cellSens Standard Imaging Software (Olympus). The cartilage parameters were measured as described by (Hillegass et al., 2008) using The ImageJ (NIH) distribution FIJI software as follows: intracranial distance (ICD) was determined by measuring the distance between the medial edge of each eye. The ceratohyal (angular) cartilage length (CCL) was determined by measuring the distance from a reference line drawn between the posterior edges of the hyosymplectic cartilage to anterior edge of ceratohyal cartilage. The lower jaw length (LJL) was determined by measuring the

distance from a reference line drawn between the posterior edges of the hyosymplectic cartilage to the anterior edge of the Meckel's cartilage.

2.6 Whole body ion content measurement

Here, we measured the whole-body ion contents (calcium, magnesium, sodium, and potassium) of 5 dpf larvae raised in artificial freshwater using atomic absorption spectrometry method. Twenty-five wildtype and *gnav1*$^{-/-}$ larvae were pooled as one sample sacrificed on ice. Then, they were rinsed briefly in 1 ml deionized water three times. The water was removed and the (wet) weight of the samples were measured by precision scale (Sartorius). The samples were dried out at 50 °C overnight and the (dry) weight was measured. Thereafter, 100 µ L HNO3 (65%) was added to the samples for digestion at 65 °C overnight. The digested solutions were diluted with deionized water to a volume of 5 mL and then filtered through a 0.2 µm filter into a sterile conical tube. The samples were kept at - 20 °C until analysis. The ion contents were measured using inductively coupled plasma-optical emission spectroscopy (ICP-OES) spectrogreen (Spectro Ametek) in the core facility of the geographical institute of Cologne university as described by (Bruder et al., 2021).

2.7 Protein analysis methods

2.7.1 Sample preparation and protein extraction

Adult zebrafish were sacrificed on ice; the adult organs (kidney and testes) were dissected as described in 2.5.3 and washed with 1x PBS. A single organ or 3-4 each were pooled in 1.5 mL centrifuge tubes, and then lysed in 30 µl RIPA buffer with 1x Protease inhibitor cocktail (Sigma Aldrich).

Zebrafish embryos were raised in E3 medium until the desired age. Embryos less than 3 dpf were dechorionated by 1 mg/ml pronase digestion (Sigma Aldrich) for 5 minutes, followed by a wash with 1x PBS. A pool of 50 larvae were collected in 1.5 mL centrifuge tube, and then washed twice with Ringer solution. The embryos/larvae were then deyolked as the high proportion of yolk proteins distorts protein migration on SDS-

PAGE (Link et al., 2006). Deyolking was carried out by adding 10 mM EDTA/1x protease inhibitors/Ringer solution and pipetting with a narrow tip to disrupt the yolk cells, followed by twice washes with Ringer solution. Next, the embryos/larvae were lysed in 90 µl RIPA buffer mixed with 10 µl 10x protease inhibitor cocktail (Sigma Aldrich).

All the samples (adult organs, embryos, and larvae) were homogenized with a probe sonicator for 10 seconds with 30 W on ice. The samples were incubated on ice for 20 minutes, followed by spinning for 30 minutes at 13000 rpm (4 °C). The supernatant was transferred into a fresh clean 1.5 mL centrifuge tube. Protein concentration was measured using standard Bradford assay (Sigma Aldrich) according to the manufacturer's instructions.

2.7.2 SDS-PAGE and Western blotting

The protein samples were incubated with a reducing sample buffer (ThermoFisher Scientific) at 55 °C for 5 minutes and then kept on ice. SDS-PAGE and Western blot were performed using Mini Gel Tank and Blot Module Set (Life Technologies) according to the manufacturer's instructions.

Briefly, denaturing SDS-PAGE polyacrylamide gels (i.e., resolving and stacking gel) were prepared as follows: 2 ml 40% Acrylamide, 2 ml Tris-HCl 1.5 M pH 8.8 resolving buffer, 3.820 ml distilled water, 80 µl 10 % ammonium persulfate, and 80 µl SDS were mixed to prepare 10 % resolving gel; and for preparing 4 % stacking gel, 0.3 ml 40% Acrylamide, 0.75 ml Tris-HCl 0,5 M pH 6,8 buffer, 1.890 ml distilled water, and 30 µl 10 % ammonium persulfate were mixed. TEMED and 10% ammonium persulfate were added immediately before pouring the gel in the molds and the resolving gel was overlaid with 1 mL isopropanol. After the gels has polymerized entirely, they were placed inside of the vertical SDS-PAGE electrophoresis apparatus (Life Technologies), fixed and immersed 1x SDS-PAGE running buffer (250 mM Tris-HCl, 192 mM Glycine, 1% SDS). Thereafter, mixture of 40 µg of total protein along with five µL of the PageRuler prestained protein standard (Life Technologies) marker were loaded and subsequently separated on the 10% resolving gel. SDS-PAGE was

conducted with an initial run using voltage 90 for 30 minutes, followed by a run using 110 voltage for 1 hour. The gels were then washed and the proteins were transferred to an activated PVDF membranes (Millipore) and subjected to immunoblotting for 70 minutes at 10 volts in 1x fast transfer buffer.

Next, the membranes were blocked in 5% milk in 1xTBS-T (50mM Tris-HCl, 150mM NaCl, and 0.1% Tween-20, pH 7.4) for one hour at room temperature. The membranes were then incubated with primary antibodies diluted in the blocking buffer overnight at 4 °C (1:200 for anti-Gv serum and the purified anti-Gv antibody). Next day, the membranes were washed three times with 1xTBS-T for 10 minutes and incubated for 2 hours with the secondary antibody (1:3000 horseradish peroxidase anti-rabbit antibody (Dako) in blocking buffer) at room temperature. The membranes were then washed again three times in TBS-T, each for 10 minutes. Amersham ECL Western Blotting Detection Reagent was used for signal detection and Images were recorded using Fusion FX-6 Edge system (Vilber Lourmat) documentation system according to the manufacturer's instructions. Band intensity and protein molecular mass were determined using Image Lab software 6.0.1 (Bio-Rad). The pre-stained protein ladders were included at each side of the gel for more accurate measurements. The intensity ratio of each mutant band to its corresponding wildtype band was measured and compared across all the respective blots. The increase or decrease of the band intensity compared to its corresponding wildtype band was considered distinct when it is more than two folds or less than 0.5-fold, respectively.

2.7.3 Recombinant protein production

Gv recombinant protein was produced kindly by Dr. Thomas Hermanns from Hofmann's laboratory as follows: Zebrafish *gnav1* was ligated into a pOPIN-K vector (Addgene) which possesses an N-terminal 6His-GST-tag using the primers listed in 2.1.4 and as described by (Bird, 2012). The *gnav1*-pOPIN-K construct was transformed into *Escherichia coli* (Strain: Rosetta (DE3) pLysS). The cells were then allowed to grow in 12 L LB media at 37 °C until the OD600 of 0.8 was reached. The bacterial cultures were cooled down to 18 °C; and Gv protein expression was induced by adding 0.2 mM isopropyl β-d-1-thiogalactopyranoside (IPTG); bacteria were then

grown for 16 hours. the cultures were harvested by spinning at 5000 × g for 15 minutes followed by freeze thaw cycle. The pellets were resuspended in binding buffer (300 mM NaCl, 20 mM Tris-HCl pH 7.5 20 mM imidazole, 2 mM β-mercaptoethanol) containing DNase and Lysozyme. Thereafter, they were homogenized by sonication using 10s pulses with 50 W for a total time of 10 minutes. Lysates were clarified by centrifugation at 50,000 × g for 1 hour at 4 °C and the supernatant was used for affinity purification on HisTrap FF columns (GE Healthcare) according to the manufacturer's instructions. For all constructs, the 6His-GST tag was removed by incubation with 3C protease and concurrent dialysis in the binding buffer. The liberated affinity-tag and the His-tagged 3C protease were removed by a second round of affinity purification with HisTrap FF columns (GE Healthcare). Subsequently, Gv protein was purified with a final size exclusion chromatography (HiLoad 16/600 Superdex 75) in 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 2 mM dithiothreitol (DTT), concentrated using VIVASPIN 20 Columns (Sartorius). The recombinant protein was then flash frozen in liquid nitrogen, and stored at − 80 °C until analysis.

2.7.4 Protein quantification using mass spectrometry and targeted parallel reaction monitoring

For quantitative proteomics, the samples were collected and proteins were extracted as described in 2.7.1. Protein extraction was followed by acetone precipitation as follows: proteins were precipitated by adding 4x volume of ice-cold acetone; incubated for 2 hours at - 80 °C, followed by centrifugation for 15 minutes at 16.000 × g at 4 °C. The pellet was washed twice with ice-cold acetone and then resuspended in 8M urea/50 mM TEAB buffer with 1x protease inhibitors. The concentrations were measured as previously described in 2.7.1 and 50 µg of total proteins was used for mass spectrometric quantifications.

Two methods were performed for protein digestions; in-solution digestion for total proteins quantification in the samples; in-gel digestion for identification of gel-separated and Coomassie-stained proteins.

For in solution digestion, samples were treated with 5 mM DTT for 1 hour at room temperature to reduce disulfide bonds. Carbamidomethylation (alkylation) was then performed by incubation with 40 mM chloroacetamide for 30 minutes at room temperature in the dark. The samples were digested with Lys-C protease at an enzyme:substrate ratio of 1:75 for 4 hours at room temperature. Thereafter, the urea concentration was diluted to 2 M with 50 mM TEAB followed by Trypsin digestion at an enzyme:substrate ratio of 1:75 overnight at room temperature. Next day, the samples were acidified to stop the enzymatic digestion with 1% formic acid and centrifuged 3,500g for 5 minutes to remove the precipitate. Finally, the peptides in the samples were separated and cleaned up using C18 Stage Tips (Thermo Fisher Scientific) for label-free proteomics of total protein levels.

For in-gel digestion, an equal amount of total protein lysate (50 ug) from wildtype and mutant samples were separated by SDS-PAGE as described in 2.7.2 along with the Gv recombinant protein. The gel was stained with Coomassie blue, destained, and washed with water; the resulting protein smear corresponding to the Gv band was cut from the gel for which the approximate molecular weight ranges from 40 to 50 kDa, followed by in-gel digestion. The gel was submerged with 50 mM ABC/50% acetonitrile, vortexed and incubated for 20 minutes at room temperature; the supernatant was discarded and this step was repeated once again. The gel was treated with 5 mM DTT for 30 minutes at 56 °C, then washed with acetonitrile, vortexed for 1 minute, and incubated for 15 minutes at room temperature. The sample was alkylated with 55 mM chloroacetamide for 30 minutes at room temperature in the dark and the solution was discarded. The sample was then washed, again, twice with acetonitrile. The gel was dried in the Speedvac for approximately 5 minutes; it was then destained with 50 mM ABC for 15 minutes at room temperature, washed with acetonitrile, and dried again. Thereafter, the digestion reaction was performed by a digestion solution of 10 ng/µL of 90% Trypsin and 10% LysC in 50 mM ABC for 30minutes at 4 °C. The excess digestion solution was removed and the gel was covered with 50 mM ABC and incubated overnight at 37 °C while shaking at 750 rpm.

Next day, the supernatant of the gel pieces was transferred into new 1.5 mL centrifuge tubes and the digested peptides were extracted from the gel as follows: the gel was

covered with 30% acetonitrile/3% TFA, vortexed, and incubated for 20 minutes at room temperature; the extract was combined it with the supernatant from previous step; the gel was covered with 100% acetonitrile, vortexed, and incubated for 20 minutes at room temperature; and the extract was combined it with the supernatant from previous steps. The collected supernatants were dried down in the Speedvac to remove the organic solvent. Finally, the peptides in the samples were separated and cleaned up using C18 Stage Tips (Thermo Fisher Scientific) for targeted parallel reaction monitoring (PRM) analysis.

Once the samples were processed for proteomics, the mass spectrometry (MS) and targeted PRM assay were performed by the CECAD Proteomics Facility as described by (Bourmaud et al., 2016). The samples were analyzed on a Orbitrap Exploris 480 (Thermo Scientific) mass spectrometer that is coupled to Ultimate 3000 HPLC System (Thermo Scientific). All mass spectrometric raw data were processed with MaxQuant version 1.5.3.8 software by the CECAD Proteomics Facility.

2.8 Quantification and statistical analysis

All statistical analyses and the corresponding graphs in this dissertation were performed using GraphPad Prism version 9 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com. The significance was determined with two tailed unpaired t-test, unless otherwise mentioned, and p values < 0.05 were considered significant. Significance is indicated in all figures as follows: ns = not significant, *p = 0.01 to 0.05, **p = 0.001 to 0.01, ***p 0.0001-0.001, ****<0.0001. Error bars on the plots always indicate SEM. Some data are presented as box-and-whisker plot: the horizontal lines of the box represent the 25 percentile, median, and 75 percentile, the whiskers represent the total range. + is the mean.

2.9 Phylogenetic analysis for identifying the G alpha genes

2.9.1 Data mining

In order to delineate the genes of the five G alpha families, all annotated G alpha protein sequences of *Drosophila melanogaster*, together with *Tribolium castaneum* Gv and G12, and *Bombyx mori* G12 protein sequences served as queries for tblastn algorithm (Altschul et al., 1990). The representative queries were used to search in all the Whole Genome Shotgun (WGS) databases available at NCBI for selected groups. The search was maintained at the phylum, subphylum, (sub) class, infraclass, or order levels, since for some phyla, classes, or other classification it was not possible to obtain results because of the big database size, therefore the searches were maintained at the possibly searchable levels.

The G alpha orthologs were identified through tblastn search in the WGS of Holozoa in the following groups (with different scientific classification levels): Ichthyosporea, Corallothytrea, Filasterea, Choanoflagellatea, Porifera, Placozoa, Ctenophora, Cnidaria, Echinodermata, Hemichordata, Tardigrada, Nematoda, Cephalopods, Arachnida, Myriapoda, Crustacea, Collembola, Diplura, Palaeoptera, Blattodea, Hemiptera, Thysanoptera, Hymenoptera, Coleoptera, Diptera, Trichoptera, and Lepidoptera. The Archaeognatha genomes displayed only a partial genomic coverage, therefore it was excluded. The plants G proteins sequences from (Oka et al., 2009) served as an outgroup for rerooting the trees (see below).

The candidate G alpha protein sequences were retrieved from the blastn search as excel sheets contain the information about the predicted sequences such as the expectation-values (E-values), sequence coverage, length, and position of each candidate sequence in the WGS. E-values of 0.001 and a minimal sequence length of 100 amino acids were used as cutoff values to extract the candidates of G alpha sequences, regardless of query coverage. The genomic sequences of all candidate genes were extended up to 10 kb in the 5’ and 3’ direction to obtain the complete amino acid sequence prediction. A written script, that was created and kindly provided by Kim Korsching, allowed to automatically find the amino acid sequence of each gene

(according the previous criteria) by matching to corresponding orthologous G protein sequences using GenWise web tool (www.ebi.ac.uk/Wise2/, last accessed February 22, 2022). That was carried out using the full gene sequence and the trimmed gene sequence where the undefined nucleotides (N) were removed. Prior to tree calculation, the resulting amino acid sequences from full length or trimmed nucleotide sequences were compared using (https://text-compare.com/, last accessed February 22, 2022); in order to retain the nucleotide sequence that gives a larger intact amino acid sequence. Furthermore, we reduced redundancy within the dataset for each single species using (https://www.bioinformatics.nl/cgi-bin/emboss/skipredundant, last accessed February 22, 2022) with a 98% identity redundancy threshold.

2.9.2 Phylogenetic tree construction

To construct the phylogenetic trees, the sequences were aligned with the online multiple alignment tool MAFFT version 7 (Katoh & Standley, 2013) using the E-INS-I strategy with the default parameters. A huge number of candidate sequences were obtained from the searches even after removing the redundant sequences, however, the maximum number of sequences to be aligned in MAFFT is approximately 900 sequences. Thereby, initial trees were calculated for each group using the representative queries (see above), zebrafish G alpha sequences from (Oka et al., 2009), and the outgroup plant's G proteins. The multiple sequence alignments (MSA) were gap-striped with a 90% tolerance to remove the regions with gaps in over 90% of sequences using Gap Strip/Squeeze v2.1.0 (https://www.hiv.lanl.gov/content/sequence/ GAPSTREEZE/gap.html, last accessed February 22, 2022).

The phylogenetic trees were calculated using a Maximum likelihood algorithm (PhyML), PhyML-aLRT with SPR setting for tree optimization and chi square-based aLRT for branch support on Phylemon server available online (Guindon et al., 2010; Sánchez et al., 2011). Trees were constructed using FigTree software (http://tree.bio.ed.ac.uk/software/figtree/, last accessed February 22, 2022).

The newly predicted genes were named according to previously named orthologs or closest paralogs from other species, beginning with more basal genes. After annotating the new genes from the initial analysis, a specific set of G alpha sequences were chosen from each group (representative genes for each group) and used as reference genes in the (representative) reference tree (200-300 genes). Thereafter, the reference genes were used in recalculating trees for each group and also combined groups, in order to verify the annotation and position of each species gene relative to the species genes from other groups. The nomenclature we use is as follows: the term family refers to the top hierarchical level (Gs, Gi, Gq, G12, Gv), whereas the subfamilies are designated as class (e.g., the Gi family encompass Gi and Go classes).

3 Results

Heterotrimeric G proteins act as molecular switches inside cells and control wide range of systematic functions i.e., embryonic development, learning, homeostasis, etc (Neves et al., 2002). Major research efforts have been focused on signaling via G apha subunits. Vertebrate G alpha proteins comprise five main families, four of which - Gs, Gi, Gq, G12 - have been studied a lot. However, for the fifth and newly described family, Gv, so far nothing is known about the interaction partners and the functional role. This family has 1-2 members per species, and only one in zebrafish, a popular vertebrate model organism.

The initial characterization of Gv expression in 3 days post fertilization (dpf) larvae suggested possible functions in development of sensory organs and ion homeostasis (Oka et al., 2009). To further examine the function of Gv, an earlier study attempted to knockout *gnav1* in zebrafish by CRISPR/Cas9 system (Ivandic, 2015). Ivandic obtained a 13 base per (bp) deletion in exon 1, which resulted in a premature stop codon (nonsense mutation) in the first exon, see (Figure 16). Preliminary analysis of the mutant phenotype indicates that it might be involved in calcium homeostasis (Ivandic, 2015). Herein, this dissertation set out to perform more comprehensive analysis and unravel Gv role, in particular ionic homeostasis. We also attempt to achieve whole gene knockout and simultaneous knockin of a reporter gene to obtain an unambiguous null mutant for Gv and to facilitate analysis of Gv expression localization.

As a starting point, we began with an investigation for the ontogenetic and tissue-specific expression by optimizing WISH at several embryonic and larval stages. Concomitantly we implemented qPCR to quantify Gv mRNA levels at the same and additional developmental stages. From these data we would gain deeper insight into *gnav1* tissue specialization and attempt to explicate its role during embryogenesis.

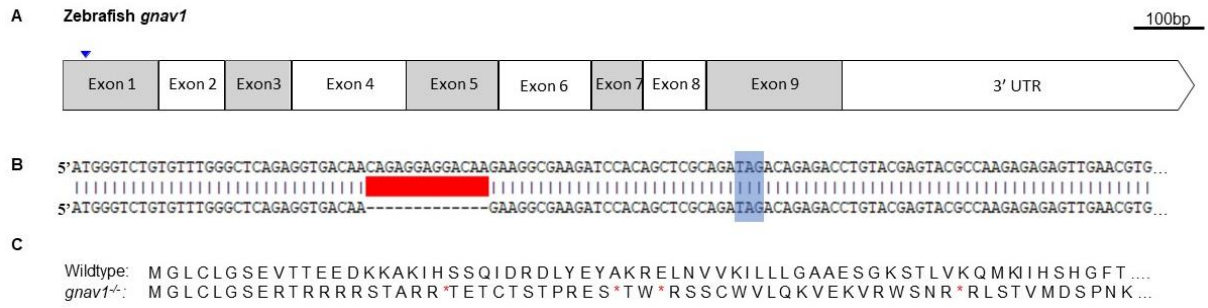


Figure 16: Zebrafish *gnav1* mutation at the DNA and protein levels**. (**A) Scheme shows the coding sequence of *gnav1* and the mutation site (blue arrow head). (B) Alignment of the wildtype sequence with the mutant sequence indicating the deletion of 13bp (red highlight) and the premature stop codon (blue highlight). (C) illustration for the consequence of the frameshift mutation; the corresponding predicted amino acid sequences of wild type and mutants aligned to show the presence of a premature stop codon (*) and the blue arrowhead points to the deletion site.

3.1 Studying Gv mRNA expression in wildtype during embryogenesis and in adult kidney

So far, *gnav1* expression pattern was only investigated at 3 dpf, in the protruding-mouth stage of zebrafish larvae (Oka et al., 2009). Here, for the first time we elucidate the ontogenetic time course of expression of *gnav1* from 6 hpf, gastrula through 5 dpf larval stage using WISH as well as qPCR. Furthermore, we investigated *gnav1* expression in adult kidney, since a previous thesis had found Gv-like immunoreactivity in adult kidney (Ivandic, 2015).

3.1.1 Gv mRNA is broadly expressed in the gastrulation and early segmentation periods

We selected embryonic stages where a variety of morphogenetic changes occur: 6 hours post fertilization (hpf), 12 hpf, 24 hpf, 2 dpf, 3 dpf, and 5 dpf stages. The embryos/larvae were staged prior to collection according to criteria outlined by Kimmel et al., 1995. We conducted WISH using two different non-overlapping antisense probes. Results obtained were indistinguishable between probes, so only results for one of the probes are shown. Sense probes were included as controls to evaluate the

specificity of the signal, and no signal was obtained with sense probes for all of the analyzed stages.

At the gastrula period of development (6 hpf, shield stage) we observe that Gv mRNA is interestingly expressed broadly throughout the blastoderm (Figure 17). Zebrafish embryo enters the shield stage at 6 hpf; at this stage it is called blastoderm and appears as a cap over the yolk sac, likewise *gnav1* expression in Figure 17A appears as a stained cap over the yolk. Gastrulation at this stage is defined by generation of the germ ring that results from the morphogenetic cell movements that include involution, convergence, and extension (Kimmel et al., 1995). Here, we show that *gnav1* is also expressed in the embryonic shield which results from the folding at the marginal region of the germ ring (Figure 17B**,** C).

The embryonic shield encompasses two layers, the epiblast and hypoblast separated by a fissure and migrate in different directions (Kimmel et al., 1995). At gastrulation, the cell's plasticity is lost, so the cell fate is restricted to either epiblast which corresponds to the primitive ectoderm; or hypoblast that gives rise to the definitive mesoderm and endoderm. Fate maps showing future organ or tissue contribution have been constructed for cells at the onset of gastrulation (Figure 15) (Schier & Talbot, 2005). During the gastrula stage Gv mRNA appears to be expressed broadly in both epiblast and hypoblast (Figure 17B), suggesting that Gv plays a role in the development of many different organs and tissues.

The origin of the Gv mRNA found in the gastrula could be due at least partially to maternal mRNA, since it has been shown that in zebrafish some maternal transcripts are still present at 6 hpf and declines further afterward (Mathavan et al., 2005). Indeed, in another fish species, Japanese flounder, transcriptome data show significant levels of *gnav1* in the unfertilized eggs, i.e. before the onset of zygotic transcription, thus maternal RNA (Fu et al., 2017).

Thereafter, *gnav1* expression was analyzed at 12 hpf (early segmentation stage) and found to be clearly less intense compared to the gastrula stage (Figure 18). This supports the maternal contribution in 6 hpf since most of the maternal RNA is expected

to be degraded rapidly in the first few hours of development (Figure 14) (Solnica-Krezel, 2020).

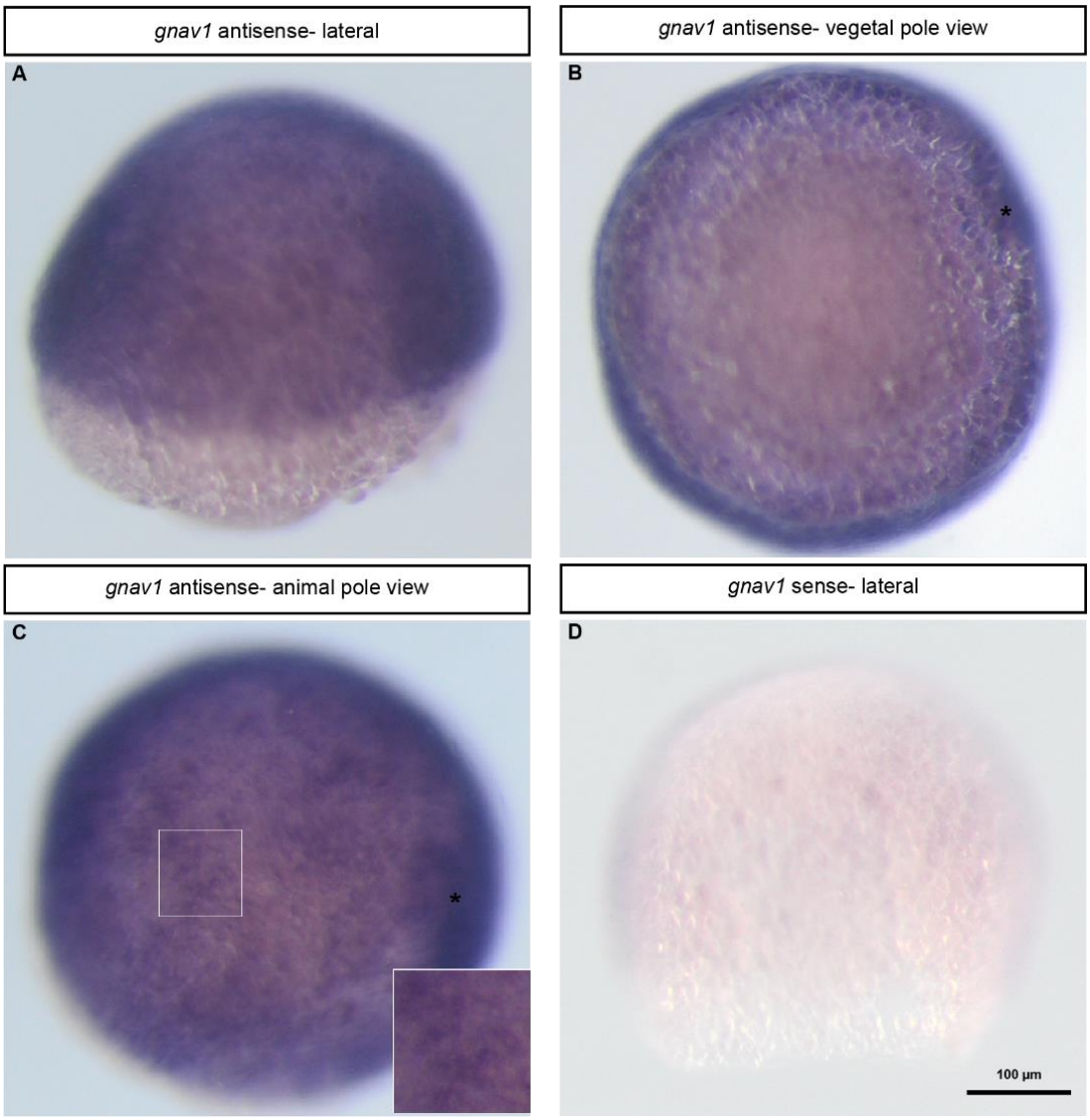


Figure 17: Spatiotemporal expression pattern of zebrafish *gnav1* in 6 hpf wildtype embryo. WISH analysis of *gnav1* expression in 6 hpf embryos using antisense (A-C) and sense (D) probes. (A) lateral view shows that *gnav1* is broadly expressed in the blastoderm including the ventral and dorsal regions. (B) vegetal pole view shows the broad expression of *gnav1* particularly in the shield area. (C) *gnav1* expression at the animal pole, in part in the shield at the eventual dorsal region; the white box magnification shows the broad, but not ubiquitous, expression of *gnav1* in many cells. (D) no background signal is observed in the negative control (sense probe). Asterisks: embryonic shield.

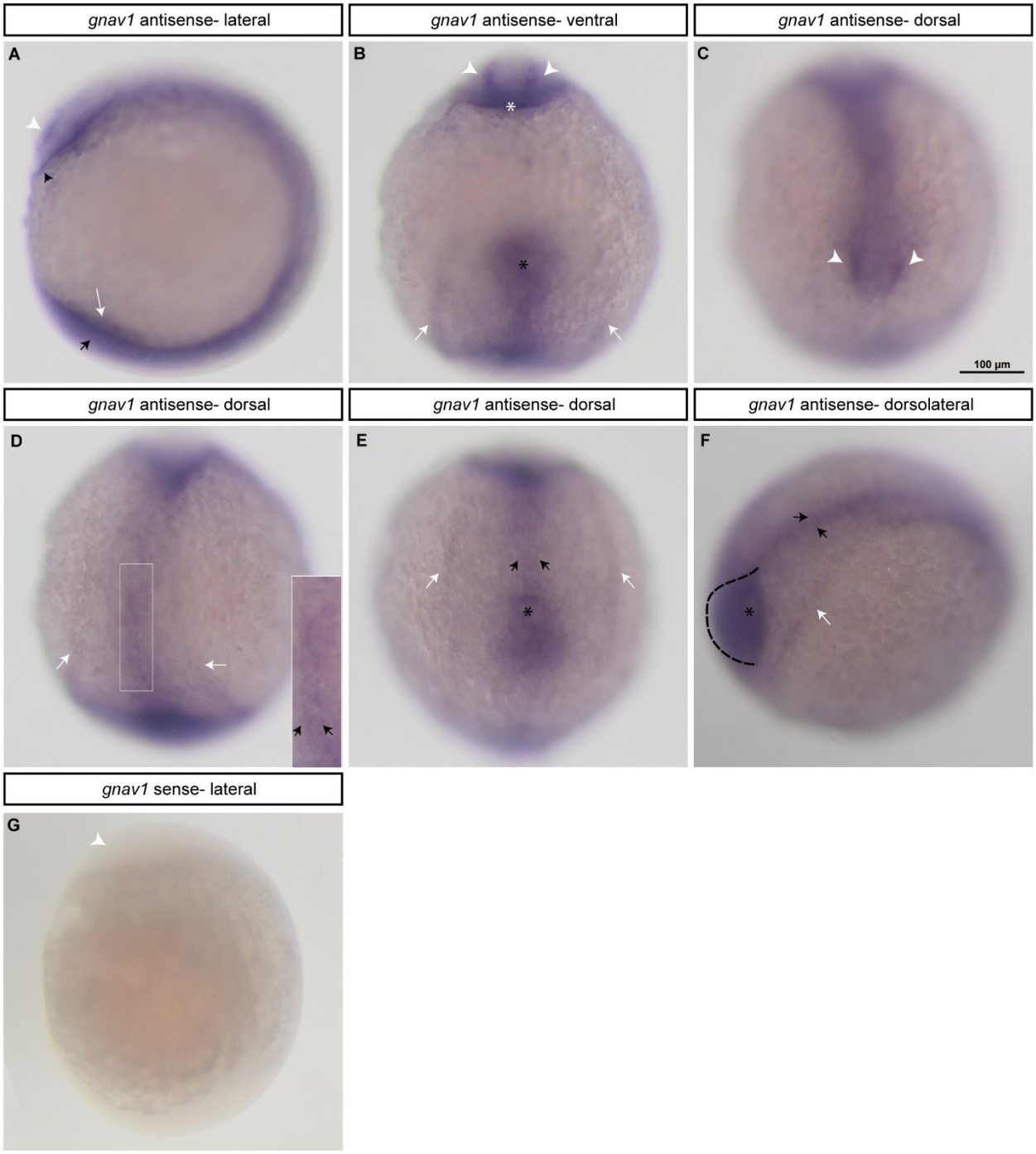


Figure 18: Expression pattern of zebrafish *gnav1* in 12 hpf wildtype embryo. WISH analysis of *gnav1* expression in 12 hpf embryos using antisense (A-F) and sense (G). (A) lateral view, *gnav1* is broadly expressed, and the signal is distinct in the eye primordium, notochord, tail bud, and presumably intermediate mesoderm (pronephros progenitor cell field). (B) ventral view, predominant expression of *gnav1* in YSL in the axial mesoderm, the eye primordium, the tissue surrounds Kupffer's vesicle, and bilateral stripes in the presumably intermediate mesoderm. (C) *gnav1* expression along the body axis, in particular the brain and eye primordium. (D) dorsal view anterior is up, *gnav1* expression in the axial mesoderm, notochord; magnification shows *gnav1* expression in the notochord. (E) the bilateral stripes of

the presumptive intermediate mesoderm originate from the middle region of the embryo and extend into the tail bud. (F) dorsolateral view anterior to the right, expression is present in the tail bud and notochord. (G) no background signal is observed in the negative control (sense probe). White arrowheads: eye primordium; white arrow: presumably intermediate mesoderm (pronephros progenitor cells); black asterisks: Kupffer's vesicle; white asterisks: axial mesoderm; black arrows: notochord; black arrowhead: hatching gland progenitors; dashed line: tail bud.

At 12 hpf the first six somites, segmented precursors of myotomes and sclerotomes, become visible. Somatogenesis starts at 10.5 hpf and continues by notochord-derived signals (Stickney et al., 2000). During this stage the embryo elongates and other distinct morphogenesis events occur; e.g., a brain primordium can be determined by the thickened neural keel; the optic primordium can be observed; and tail bud develops within which Kupffer's vesicle can be seen (Kimmel et al., 1995). We observe broad *gnav1* expression distributed along the embryonic axis at 12 hpf; with enhanced expression in the **head (**axial mesoderm, eye and brain primordia**)**, **notochord**, the presumably lateral plate mesoderm (**LPM**), and **tail bud** (Figure 18).

In the **head** region, the expression is distinctly observed in both the optic and brain primordia (Figure 18A, C). We also observe a distinct staining marking the boundaries between the yolk syncytial layer (YSL) surface in the axial mesoderm and the neural keel at the anterior polster where both mesoderm and neural keel indent the yolk cells (midsagittal groove in anterior neural keel), also where the hatching gland forms ventral to the head (Figure 18B) (Kimmel et al., 1995). Additionally, *gnav1* is expressed in the **notochord** primordium dorsally along the AP axis (Figure 18D, E, F, black arrows). The notochord is a rod-like structure surrounded by an epithelial sheath which is encased in a thick extracellular matrix (ECM), it provides mechanical support to the developing embryo and instructs patterning of the surrounding tissues (Ellis, Bagwell, et al., 2013; Ellis, Hoffman, et al., 2013).

At the posterior-most dorsal area, a prominent Gv mRNA signal can be also observed as bilateral stripes, which originate from the middle region of the embryo and extend into the tail bud (Figure 18A-F, white arrows). These stripes presumably correspond to the **LPM** (Prummel et al., 2019). A similar expression pattern has been observed for *cdh17,* which encodes cadherin-17 protein that functions in calcium-dependent cell adhesion and is proposed to be involved in the normal development of the zebrafish

kidney (Horsfield et al., 2002). Therefore, we propose that the bilateral signal most likely corresponds to the pronephros (kidney) progenitor cell field, yet hematopoietic and gut primordia, which are closely adjacent, cannot be excluded (Brown et al., 2000; Gays et al., 2017; Gering et al., 2003, p. 2).

Moreover, we observe an expression in the **tail bud** and tissues surrounding the Kupffer's vesicle, but not the Kupffer's vesicle itself (Figure 18B, E, F; asterisk points to Kupffer's vesicle). Kupffer's vesicle is a transient structure unique for teleost embryos, and it is required for proper formation of internal organ laterality (Essner et al., 2005). The cells lining Kupffer's vesicle develop into tail mesodermal derivatives, including notochord and muscle (Kimmel et al., 1995; Warga & Kane, 2018), suggesting that Gv is expressed already in the primordia of these tissues.

Due to the generally broad staining along the AP axis, we expect expression of *gnav1* also in the primordia of other organs, but because of the broad staining we could not clearly delineate Gv expression in e.g., primordia for pharyngeal arches and otic vesicles. These structures will become more clearly visible in the next pharyngula period.

3.1.2 Gv mRNA is localized in several organs' primordia during the pharyngula and hatching periods (24 hpf, "prim-5 stage" and 48 hpf, "long-pec stage", respectively)

At one day post fertilization (24 hpf, "prim-5 stage"), Gv mRNA expression is broadly observed throughout the embryo like the earlier developmental stages, additionally the expression now extends to the YSL in yolk sac and its extension. Distinct expression for *gnav1* can be observed in several organs' primordia: the **sensory** organs such as eyes, brain, and otic vesicles; the **osmoregulatory** organs as the pronephros and YSL; primordia of **cartilaginous** structures as the pharyngeal arches region and pectoral fin bud; the **somites**; and the **hatching gland**.

Gv mRNA is clearly visible in the brain and eyes (**sensory** organs) at 1 dpf stage. Particularly, Gv mRNA is present in all main brain subdivisions, forebrain, midbrain,

and hindbrain (Figure 19). This appears to be a continuation of the head staining observed at 12 hpf. The signal in the brain shows an interesting pattern among the different regions. For instance, the signal in the **forebrain** is more prominent in the ventricular zone, in addition to a signal in the diencephalon and at the boundaries between its wall (from which the eyes develop) and the eyes (Figure 19E, F). Similar to the forebrain, *gnav1* is also expressed in the **midbrain** and **hindbrain** ventricular zones (Figure 19F). Also, a prominent staining is observed in the dorsal midbrain (tectum), and the posterior midbrain (Figure 19F); the signal intensifies at the horizontal furrow that separates the tectum and the tegmentum (Figure 19A, red arrowhead). A distinct expression is also observed at the MHB (Figure 19A, F). In the **hindbrain**, *gnav1* expression includes the cerebellum and rhombomeres (Figure 19A). There is also a conspicuous expression in the **floor plate** of the developing spinal cord - which lies immediately dorsal to the notochord - along the midline up to the forebrain (Figure 19H). Moreover, a prominent signal is evident in the **eyes**, which is more intense in the lens (Figure 19E), while low expression can be observed in the **otic vesicles** primordia (Figure 19I).

Furthermore, *gnav1* appears to be distinctly expressed in organs that regulate **ionic and osmotic homeostasis**. Confirming our interpretation at 12 hpf (see 3.1.1), clear expression is observed in the **pronephros** of 1 dpf embryo (Figure 19A, D, C). The zebrafish pronephros is a flat organ that develops bilaterally, and at this stage comprises the pronephric tubule and pronephric duct; the onset of glomerular filtration commences at 2 dpf (Drummond et al., 1998). Another vital organ regulating the ionic homeostasis is the yolk; *gnav1* expression is evident at the **YSL** of the most-posterior region and appears more intense at the yolk extension (Figure 19A, D) (Langenbacher et al., 2011).

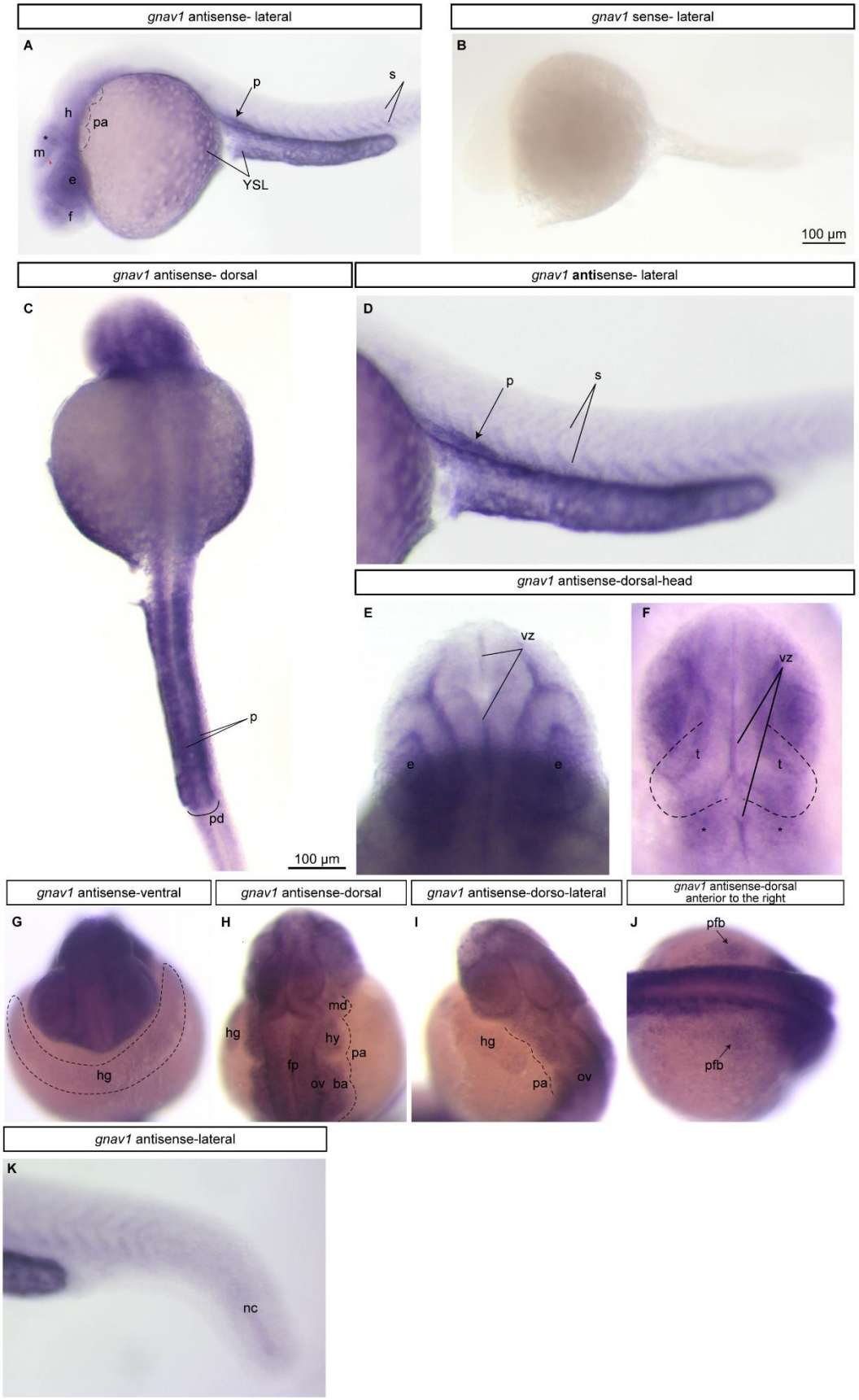

gnav1 antisense- lateral
A
p
s
h
pa
m
e
f
YSL
gnav1 sense- lateral
B
100 µm
gnav1 antisense- dorsal
C
p
pd
100 µm
gnav1 antisense- lateral
D
p
s
gnav1 antisense-dorsal-head
E
vz
e
e
F
vz
t
t
gnav1 antisense-ventral
G
hg
gnav1 antisense-dorsal
H
md
hg
hy
pa
fp
ov
ba
gnav1 antisense-dorso-lateral
I
hg
pa
ov
gnav1 antisense-dorsal anterior to the right
J
pfb
pfb
gnav1 antisense-lateral
K
nc

Figure 19: Expression pattern of zebrafish *gnav1* in 24 hpf wildtype embryo. WISH analysis of *gnav1* expression in 24 hpf embryos using antisense (A, C-K) and sense (B). (A) lateral view anterior to the left, *gnav1* is broadly expressed throughout the embryo; the signal is distinct in the sensory organs (eyes, brain, MHB, and otic vesicles), the osmoregulatory organs (the pronephros and YSL), primordia of cartilaginous structures (the pharyngeal arches region and pectoral fin bud), the somites, and the hatching gland. *gnav1* expression is also evident at the YSL of the most-posterior region and appears more intense at the yolk extension. (B) lateral view, no background signal is observed in the negative control (sense probe). (C-D) dorsal view, *gnav1* expression is distributed along the body axis, in particular distinct expression in the pronephros bilaterally extends on the posterior trunk to the tail dorsal to the yolk extension; the latter is distinctly stained. Weak expression is observed in the somites. (E-F) Dorsal view at the head region, *gnav1* expression in the brain that is more prominent in the ventricular zones and MHB, in addition to an expression at the boundaries between diencephalon's wall and the (lens) eyes. (G) ventral view anterior to front, weak expression for *gnav1* in the hatching gland. (H-I) Dorsolateral view anterior is top, weak expression for *gnav1* in the otic vesicles and the pharyngeal arches region, in part in the mandibular, hyoid, and branchial arches. Expression is also present in the floor plate. (J) Dorsal view anterior to the right, weak expression is present in the pectoral fin bud. (K) lateral view anterior to the left, weak expression for *gnav1* in the tail, somites, and notochord. Distinct expression is observed in the YSL at the yolk extension. ba: branchial arches region; e: eye; f: forebrain; fp: floor plate; h: hindbrain; hg: hatching gland; hy: hyoid region; m: midbrain; md: mandibular region; nc: notochord; ov: otic vesicles; pa: pharyngeal arches region; p: pronephros; pd: pronephric duct; pfb: pectoral fin bud; s: somites; vz: ventricular zone; YSL: yolk syncytial layer. Asterisk: midbrain-hindbrain boundaries (MBH).

Prominent Gv expression could be observed in the somites of 24 hpf embryo (Figure 19A, D), in contrast to the 12 hpf stage, where the first six somites become visible, but do not yet express Gv. By 1 dpf the segmentation is completed, and the embryo has a complete set of somites which vary from 30-34 pairs (Stickney et al., 2000). Somites form in anterior-posterior fashion around the notochord; the furrow initiates prior to the formation of morphologically distinct somites (Kimmel et al., 1995). The somite expression of *gnav1* appears more prominent in the caudal region than the trunk, which has more mature somites, suggesting a transient role for Gv in the maturation of somites. Interestingly, *gnav1* expression is not homogenous within the somites, but appears as v-shaped lateral stripes in all segments at 24 hpf, reminiscent of the v-shaped muscle fibers (Kimmel et al., 1995) and the observed lateral v-shaped pattern of *myoD* expression (a muscle differentiation factor) (Weinberg et al., 1996). *myoD* gene expression also occurs in an anterior-to-posterior wave, but somewhat earlier than that of Gv, with an onset just prior to somite formation and decrease at the onset of somite maturation (Holley et al., 2000; Weinberg et al., 1996). Taken together, this suggests that Gv could be involved in the maturation of myotomes, but at a later stage than *myoD*.

Expression in the notochord has strongly decreased in the 24 hpf stage, suggesting a role for Gv only in the earlier stages of notochord development, with the possible exception of the tail, which retains weak Gv expression at 24 hpf (Figure 19K). In any case, the expression of *gnav1* in tail notochord and tail somites in 1 dpf embryo is consistent with its expression in the tissues surrounding the Kuppfer's vesicle at 12 hpf stage (cf. Figure 18), because these surrounding tissues give rise to tail notochord and somites.

Further weak expression was observed in several tissues such as the **hatching gland**; the primordia of **cartilaginous** tissue including the pharyngeal arches region and pectoral fin bud primordia. The expression only could be identified by increasing the chromogenic reaction time in the WISH experiment. The **hatching gland** appears at this stage like a collar on the pericardium area over the anterior yolk sac ventral to the head. The hatching gland is considered a conspicuous feature at the pericardial region through the pharyngula period (Figure 19G) (Kimmel et al., 1995). **The pharyngeal arches** regions give rise to the craniofacial skeleton. In the following period (hatching), the anterior arches give rise to the jaw; the first (mandibular) arch gives rise to Meckel's cartilage and palatoquadrate; the second (hyoid) pharyngeal arch develops to hyosymplectic and ceratohyal. whereas the posterior arches ultimately develop to the ceratobranchial arches 1-5 (Figure 19H). We next evaluated the expression of *gnav1* in these and other tissues in the second day of development.

In the hatching period, the overall expression of *gnav1* decreases considerably in **2 dpf** embryo (long-pec stage) and appears more restricted compared to 1 dpf (Figure 20). Gv mRNA decreases in the brain, hatching gland, floor plate, the YSL in yolk sac, brain ventricular zones, and MHB, and it vanishes in the eye and somites (Figure 20A, D, E). On the other hand, the expression increases drastically in the otic vesicle, pectoral fin, and the pharyngeal arches region. In pronephros and yolk extension the expression of Gv appears unaltered (Figure 20A, C, I). Intriguingly, in contrast to all previously described patterns, which were bilaterally symmetrical, a specific expression restricted to the left side of the embryo is evident in the gut primordium region (Figure 20I). This finding would be consistent with an expression of *gnav1* in the gut primordium, although it is difficult to be certain, due to the deep position of the

gut and its rather incomplete development at this stage. In fact, a break of bilateral symmetry constitutes a robust sign of specific expression in internal organs such as gut and liver.

As observed in the 1 dpf stage, at 2 dpf *gnav1* is also expressed in **cartilaginous** organs; the pharyngeal arches region and pectoral fin. Prominent morphological changes occur in the pharyngeal region at this stage; the anteriormost pharyngeal arches are well-defined and distinct from the posterior arches, but *gnav1* expression is similar in both the more mature (anterior) and less mature posterior arches (Figure 20F-H). With some difficulty one can visualize early precartilage condensations in the developing mandibular and hyoid arches, but not in the branchial arches; these condensations express *gnav1* (Figure 20G). The **pectoral fin** bud is elongated and curved compared to 1 dpf stage; the innermost precartilage core starts at this stage to differentiate within the mesenchyme and that is the first cartilage differentiate in the embryo (Figure 20J) (Kimmel et al., 1995). Similarly, the *gnav1* staining appears rather stronger in the mesenchyme in the center of the pectoral fin bud, and appears linked to the differentiation process within the fin bud as well as the anterior arches at this stage.

The next stage analyzed was 3 dpf. At this stage hatching is complete, but the larvae are not yet feeding.

3.1.3 Pronounced Gv expression in the gut and the proximal convoluted tubule of the 3-5 dpf larvae

The transition from 2 dpf to 3 dpf stage involves the transition from the embryo enclosed with chorion to the hatched larvae, but not yet feeding (Strähle et al., 2012). Most of the organogenesis is complete in the 3 dpf larvae and the mouth widely opens and protrudes anteriorly. By 96 hpf (4 dpf) the larvae possess developed digestive organs to an extent sufficient for food processing (Ng et al., 2005). At 5 dpf stage the larvae show independent feeding and transition in free-swimming behavior (Strähle et al., 2012).

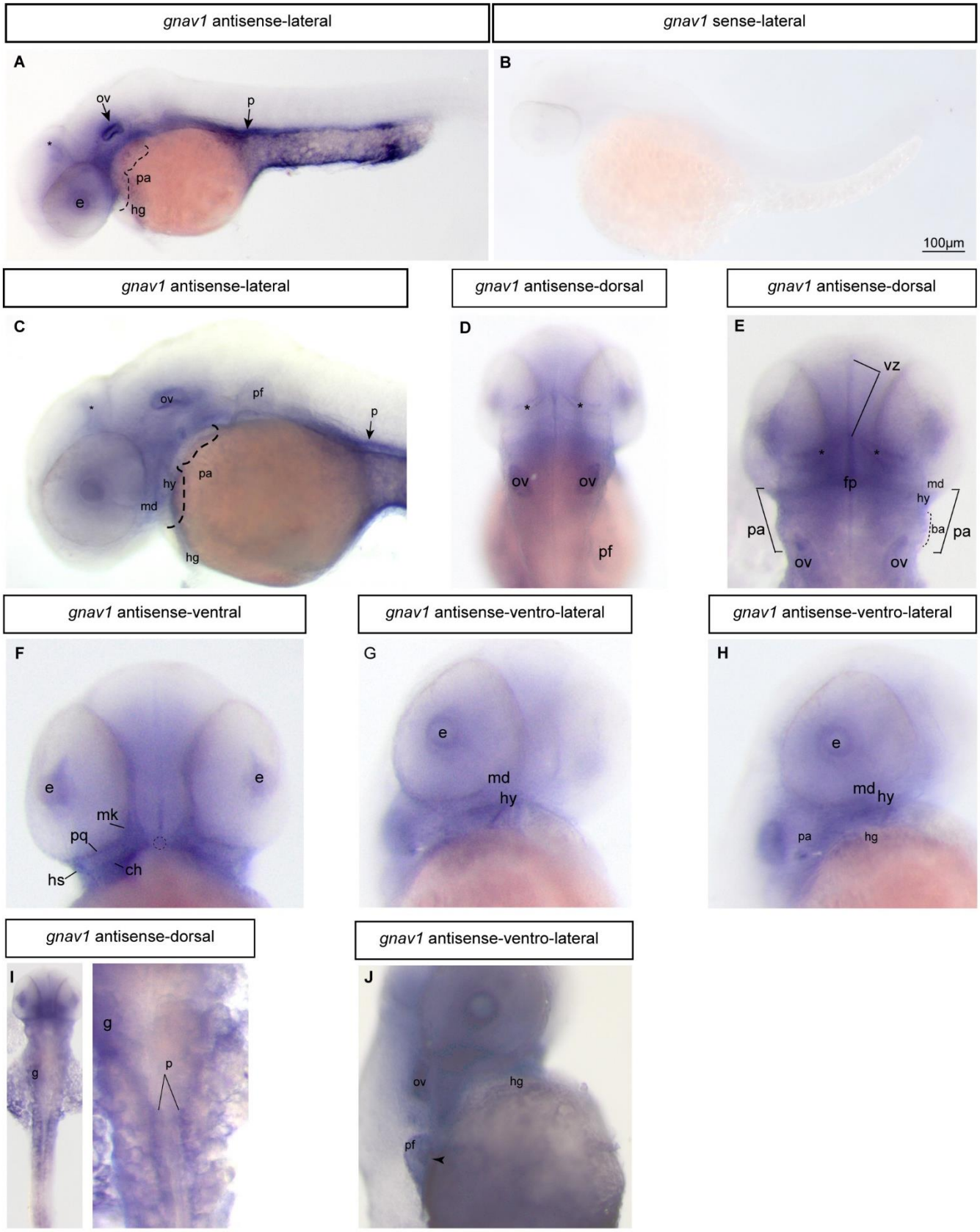

gnav1 antisense-lateral
gnav1 sense-lateral
A
ov
p
pa
e
hg
B
100µm
gnav1 antisense-lateral
gnav1 antisense-dorsal
gnav1 antisense-dorsal
C
ov
pf
p
hy
pa
md
hg
D
ov
ov
pf
E
vz
fp
md
hy
pa
ba
pa
ov
ov
gnav1 antisense-ventral
gnav1 antisense-ventro-lateral
gnav1 antisense-ventro-lateral
F
e
e
mk
pq
ch
hs
G
e
md
hy
H
e
md
hy
pa
hg
gnav1 antisense-dorsal
gnav1 antisense-ventro-lateral
I
g
g
p
J
ov
hg
pf

Figure 20: Expression pattern of zebrafish *gnav1* in 2 dpf wildtype embryo. WISH analysis of *gnav1* expression in 2 dpf embryos using antisense (A, C-J) and sense (B) probes. (A-E) lateral view anterior to the left, *gnav1* expression is more restricted compared to 1 dpf; weak *gnav1* expression in the brain, hatching gland, floor plate, the YSL in yolk sac, brain ventricular zones, and MHB; whereas it is distinct in the otic vesicle, pectoral fin, the pharyngeal arches region, pronephros, and yolk extension. (B) lateral view, no background signal is observed in the negative control (sense probe). (F-H) ventral and ventrolateral views with different focus to show specific structures clearer, there is prominent staining for *gnav1* in the pharyngeal arches' region; (F) it is distinct in the mandibular (Meckel's cartilage and palatoquadrates) and the hyoid cartilages (ceratohyal and hyosymplectic); in (G) one can visualize the chondrocytes' condensation with some difficulty, and in (H) the immature branchial arches. (I) dorsal view anterior is top, *gnav1* expression restricted to the left is evident in the gut primordium region; also, the bilateral staining of the pronephros is prominent. (J) ventrolateral view, the expression is distinct in the pectoral fin, in particular the innermost fin core. ba: branchial arches region;ch: ceratohyal; e: eye; f: forebrain; fp: floor plate; h: hindbrain; hg: hatching gland; hs: hyosymplectic; hy: hyoid; g: gut; md: mandibular region; mk: Meckel's cartilage; ov: otic vesicles; pa: pharyngeal arches region; p: pronephros; pd: pronephric duct; pf: pectoral fin; pq: palatoquadrate; vz: ventricular zone; YSL: yolk syncytial layer. Asterisk: midbrain-hindbrain boundaries (MBH); dashed circle: mouth; black arrowhead: innermost fin core.

Gv expression was not analyzed in subsequent stages for technical reasons, because probe access to inner tissues becomes impaired with continuing growth. However no major changes in Gv expression pattern may be expected in later stages, as all organs are already developed and functioning at 5 dpf.

Gv mRNA expression in 3 dpf larvae continues and appears more distinct, compared to 2 dpf, in the **otic vesicle** (includes the sensory epithelium); **osmoregulatory organs**: pronephros, gut, and yolk; **cartilaginous** tissues: pectoral fin and pharyngeal arch (Figure 21A). On the other hand, in the brain of 3 dpf larvae expression becomes more restricted, and only a weak signal remains in the MHB (Figure 21A).

In the **otic vesicle**, strong staining is evident in substructures, presumably the cristae and possibly the anterior macula (Figure 21F, red asterisks), i.e., in sensory epithelium. Additionally, weak expression is observed in the newly formed semicircular canal, a non-sensory structure (Figure 21F, black arrowhead).

The **pronephros** expression of *gnav1* becomes more restricted compared to the 2 dpf stage. The PCT, which now looks like a hook, is clearly stained, while faint signal is observed in the remaining parts of the pronephros (Figure 21A, G).

The expression **in the gut** is consistent with the signal observed in 2 dpf embryo, in both stages the signal is on the left side of the 3 dpf larvae (Figure 20G). The signal does not change in spatial distribution, but becomes much stronger in 3 dpf (Figure 21G).

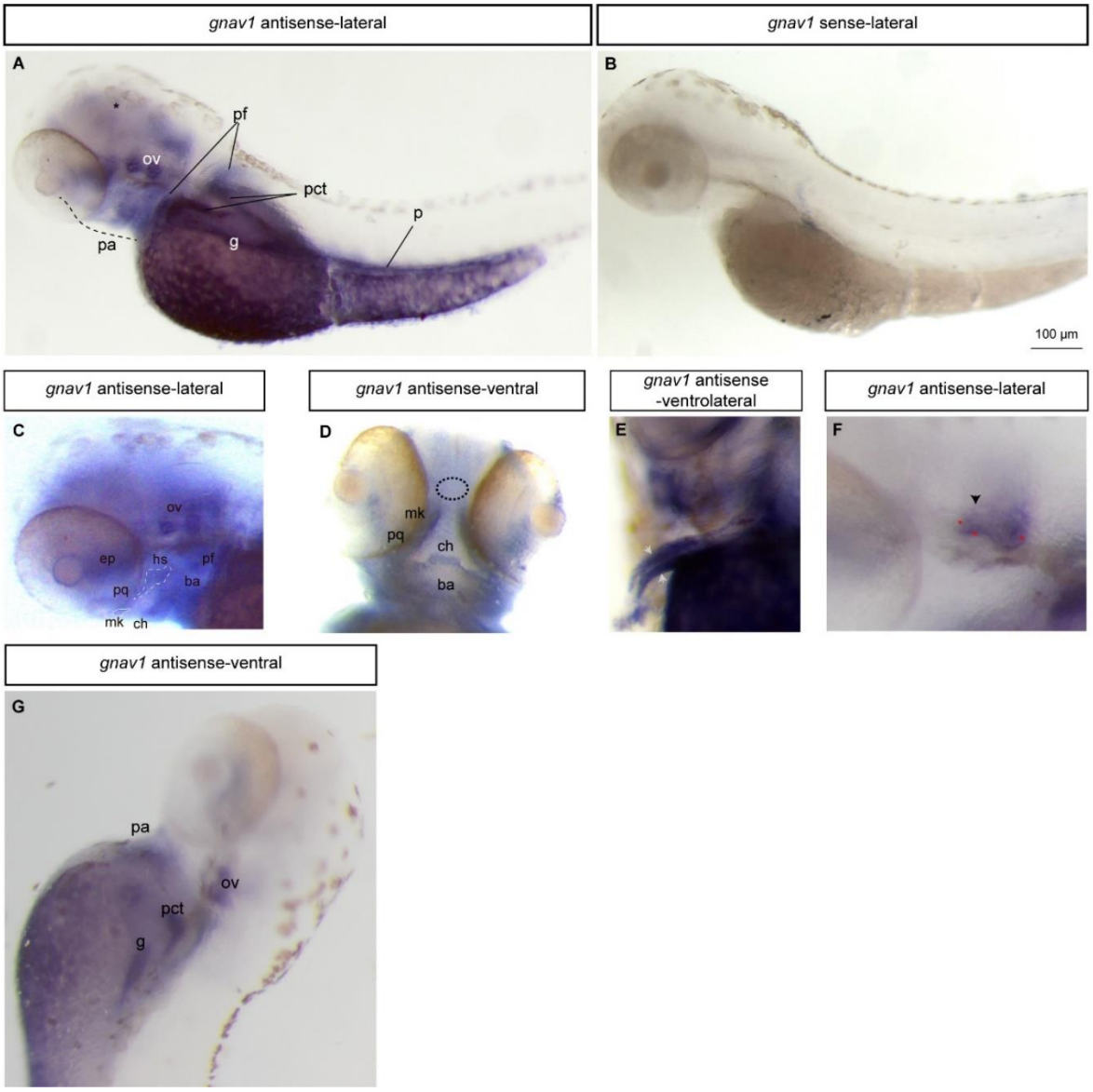


Figure 21: Expression pattern of zebrafish *gnav1* in 3 dpf wildtype larvae. WISH analysis of *gnav1* expression in 3 dpf larvae using antisense (A, C-G) and sense (B). (A) dorsolateral view anterior to the left, *gnav1* expression is distinct in the otic vesicle, pronephros, gut, and yolk, pectoral fin, and pharyngeal arch; whereas it is weak in MHB. (B) lateral view, no background signal is observed in the negative control (sense probe). (C-D) lateral and ventral views, respectively, anterior is top, prominent staining is visible in the ethmoid plate, the ceratobranchial arches. Also, the signal is still evident in the Meckel's cartilage, palatoquadrate, ceratohyal cartilages, and the hyosymplectic. (E) ventrolateral view anterior

is top, staining appears in the pectoral fin, consisting of two parallel stripes at the border of the fin. (F) lateral view anterior to the left, staining is present in presumably the cristae and possibly the anterior macula; also, a weak expression is present in the semicircular canal. (G) dorsolateral view anterior is top, very strong expression for *gnav1* in the pronephros, particularly in PCT; while faint staining is present in the remaining parts of the pronephros. In addition, there is a distinct staining on the left side corresponds to the gut. ch: ceratohyal; et: ethmoid plate; ba: branchial arch region; g: gut; hy: hyosymplectic; pa: pharyngeal arches region; pq: palatoquadrate; p: pronephros; pct: proximal convoluted tubule; pf: pectoral fin; mk: Meckel's cartilage; ov: otic vesicles. Asterisk: midbrain-hindbrain boundaries (MBH); dashed circle: mouth; black arrowhead: semicircular canal; white arrow heads: possibly the actinotrichia forming cells; red asterisks: cristae.

A picturesque staining appears in the **pectoral fin**, consisting of two parallel stripes at the border of the fin. These most likely correspond to actinotrichia forming cells (AFCs) (Figure 21E). The collagenous actinotrichia fin rays are the first exoskeletal elements to form in the developing fin. A particular collagen, *col2a1b*, only found in fish, is expressed in these fin rays; which would be consistent with a possible role for Gv in extracellular matrix homeostasis (Lalonde & Akimenko, 2018)**.**

Moreover, Gv expression develops in other cartilaginous tissues such as the **ethmoid plate**, particularly in the trabeculae. The expression continues in the **branchial arches** and is more prominent in the ceratobranchial arches 1-3. Also, the signal is still evident in the Meckel's cartilage, palatoquadrate, and ceratohyal cartilages, as well as the hyosymplectic, which is well defined at this stage (Figure 21C, D).

Our expression results for 3 dpf larvae are mostly in agreement with the published data for this stage (Oka et al., 2009)**,** however we obtained additional signals because we optimized the reaction to be performed at higher temperature with longer exposure time. This allows more time for the probe penetration into the deep tissues; and enhances the staining reaction in the tissues that harbor weakly expressed Gv mRNA. The additional signals are observed in the yolk, gut, and pronephros; the expression in the palatoquadrate we found here could correspond to the 'bilateral cell clusters' described by Oka (Oka et al., 2009).

In the next stage analyzed, 5 dpf, overall *gnav1* expression continues to become more and more restricted, and is mostly confined to three organs (Figure 22). Specific expression is evident in the swim bladder and the gut (note that the gut resides more ventrally in this stage) (Figure 22). We also observe a strong signal in the PCT

segment of the pronephros (Figure 22B). The PCT already harbors multiple loops and coils compared to 3 dpf larvae, and the Gv expression faithfully mirrors these morphological specializations. The pronephric tubule posterior to the swim bladder appears to be negative for Gv expression. For the more anterior segments it is difficult to determine, because the pronephros extends dorsally directly on top of the swim bladder; thus, the strong expression observed in the swim bladder might hinder detection of signal in this region of the pronephric tubule.

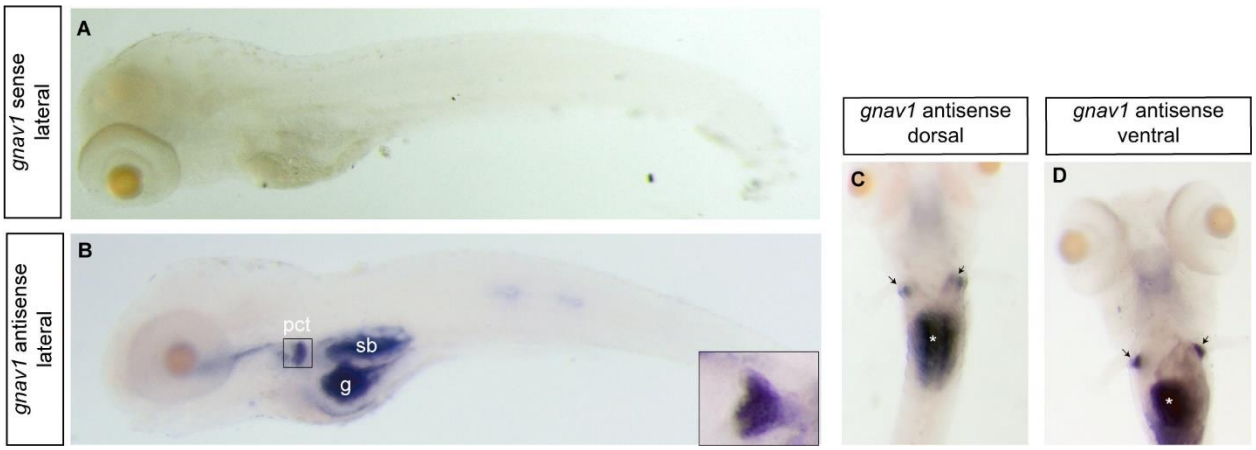


Figure 22: Expression pattern of zebrafish *gnav1* in 5 dpf wildtype larvae. WISH analysis of *gnav1* expression in 5 dpf larvae using sense (A) and antisense (B-D) probes. (A) lateral view, no background signal is observed in the negative control (sense probe). (B) lateral view anterior to the left, *gnav1* expression is distinct in the PCT, gut, and swim bladder. G: gut; pct: proximal convoluted tubule; sb: swim bladder; black arrows: pct; white arrowhead: the swim bladder and gut region.

To summarize, *gnav1* expression begins broadly in the early development (6 - 24 hpf), subsequently it becomes more and more tissue-specific (2 - 5 dpf) (Table 1). Gv mRNA is expressed *transiently* in: sensory tissues (12 - 72 hpf); some of the osmoregulatory tissues (12 hpf - 5dpf); cartilaginous tissues (24 - 72 hpf); and other tissues (somites, and notochord). On the other hand, some transient tissues express Gv mRNA during their entire existence (Kupffer's vesicle, and hatching gland). We find only one tissue with very late onset of expression (swim bladder). Overall, very few tissues still express Gv mRNA at the latest stage investigated, two of those during all stages analyzed (pronephros and gut). These patterns of Gv mRNA expression in the developing zebrafish implicate multiple roles for Gv in tissue formation and organogenesis. Intriguingly, expression in kidney was prominent in all

stages, suggesting Gv might have a role in both development and function of kidney. Therefore, we further examined Gv mRNA localization in the adult zebrafish kidney.

Table 1: *gnav1* expression pattern in 6-120 hpf zebrafish embryos/larvae.

hpf	**6**	**12**	**24**	**48**	**72**	**120**
General pattern in the whole embryo and larva	broad	broad	broad	restricted	restricted	restricted
Sensory tissues						
eye	/					
brain	/					
MHB	/					
otic vesicle	/					
Osmoregulatory tissues						
pronephros	/					
gut	/					
YSL	/					
Cartilaginous tissues						
Anterior pharyngeal arches	/					
posterior pharyngeal arches	/					
pectoral fin	/	/				
ethmoid plate	/	/	/			
Other tissues						
somites	/					
hatching gland	/				/	/
Kupffer's vesicle surrounds	/		/	/	/	/
notochord	/					
swim bladder	/	/	/	/		

/: not applicable; gray: unclear; white: absent; light purple: weakly expressed; dark purple: distinctly expressed.

3.1.4 Gv mRNA is expressed in the adult kidney of zebrafish

The adult kidney has undergone several differentiation steps beyond the larval stage, which just possesses a basic pronephros unit that comprises a single nephron on each body side. In the juvenile zebrafish kidney this pronephros has developed into a structure called mesonephros, which is composed of hundreds of nephrons (G. F. Gerlach & Wingert, 2013). In contrast to mammalian kidney development no further maturation into metanephros occurs in zebrafish, the mesonephros just increases in size and ramifications. Each nephron connects to a glomerulus, and its adjacent segment is called proximal tubule, whereas the following segment, which drains the urine into collecting ducts, is called distal tubule. Proximal and distal tubules can be distinguished according to the following morphological criteria: the proximal tubules encompass brush borders, elongated ‘more packed’ cells, dilated ‘occluded’ lumen, and the nucleus is closer to the lumen; on the contrary the distal tubules harbor narrow ‘clean’ lumen, and the brush borders are absent (McCampbell et al., 2015). Besides the glomerular units the mesonephros contains intervening stroma, which harbors the mesenchymal renal progenitors (from which new nephrons are continually added to the kidney), in addition to the blood and hematopoietic stem cells (McCampbell & Wingert, 2014).

Hence, we performed *in situ* hybridization (ISH) for cryosections for zebrafish kidney, particularly the trunk (Figure 23). As negative control, ISH was performed with sense probes. The background signal visible in the sense controls can be distinguished by its more brownish color from the purplish reaction product of the detection reaction (Figure 23D). The antisense signal is enriched in tubules albeit the intensity of the staining is heterogenous: the signal appears intense in some tubules and very light in others. Additionally, weak purplish signal is also observed in nontubular tissue (Figure 23A). The lighter staining might be an unspecific background signal due to the long chromogenic reaction time used in this experiment (overnight at room temperature). However, the more intense signal in some tubules (Figure 23A, C) is very likely to represent specific staining and thus expression of *gnav1*. Intense staining is observed

both in proximal and distal tubules (Figure 23A and B, black arrow and arrowhead respectively).

In zebrafish both the proximal and distal tubules are responsible for electrolytes reabsorption and secretion. The proximal tubule is the main part of zebrafish nephron for acid-base regulation; while the distal tubule is mainly responsible for the osmodilution of the zebrafish prourine (Kersten & Arjona, 2017). Therefore, we hypothesize that Gv is possibly involved in both the ionic and osmotic homeostasis in the adult zebrafish kidney. By inference it is also conceivable that Gv serves a similar functional role in the embryonic pronephros.

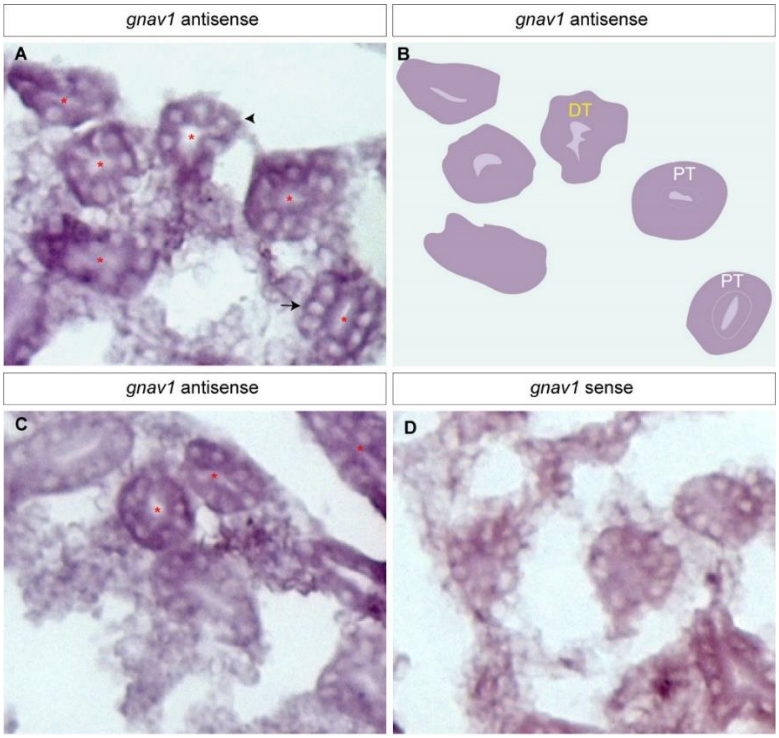


Figure 23: *gnav1* expression in the adult kidney of wildtype zebrafish. ISH analysis of *gnav1* expression in cryosections from the adult kidney of wildtype zebrafish using antisense (A, C) and sense (D) probes; the coronal sections for the kidney corresponds to the trunk region. (A) purplish (intense) staining signal is observed in the proximal and distal tubules. (B) scheme depict the stained tubules in A for better visualization of the marked distal and proximal tubules. The proximal tubules encompass brush borders, elongated 'more packed' cells, dilated 'occluded' lumen; while the distal tubules harbor narrow 'clean' lumen, and the brush borders are absent. (C) another cryosection that shows heterogenous staining. (D) brownish color background signal is visible in the negative control (sense probe) staining and can be distinguished from the purplish antisense staining in A and C. DT: distal tubule; PT: proximal

tubule; red asterisks: the positively stained tubules; black arrow: proximal tubule; arrowhead: distal tubule.

3.1.5 Quantitative evaluation of the time course of *gnav1* transcription in wildtype during development

We have shown the ontogenetic expression of *gnav1* during embryonic development by using WISH which only shows the qualitative changes in Gv mRNA localization. Therefore, we also evaluated the quantitative changes of Gv mRNA expression levels during the same period of development by qPCR. We carried out qPCR for 6 hpf, 12 hpf, 1, 2, 3, 4, and 5 dpf wildtype embryos as described (Materials and Methods 2.3.4).

We used geNorm algorithm to identify the most stable reference genes, since reference (housekeeping) genes are known to vary with experimental conditions. In the literature, variable sets of reference genes were used in qPCR during zebrafish development such as; *elf1* and *b2m* (McCurley & Callard, 2008); *elf1 and bactin* (Tang et al., 2007); *bactin* and *rpl8* (Rodrigues et al., 2022). Hence, we examined the stability ranking of six candidates (*gapdh*, *elf1*, *b2m*, *bactin*, *rpl8*, and *rpl37*) in samples of wildtype and mutant zebrafish embryos/larvae. Consequently, both *rpl37* and *rpl8* reference genes showed more stability than the others in our experimental conditions. We used both *rpl37* and *rpl8* to normalize *gnav1* to the mRNA levels between different samples; then we normalized *gnav1* values in 12-120 hpf stages to 6 hpf stage to evaluate the changes in Gv mRNA levels during development in reference to 6 hpf stage.

Interestingly, in all time points Gv mRNA expression is roughly comparable to its 6 hpf levels. Nevertheless, we observe slight fluctuation in *gnav1* levels in reference to 6 hpf; at 12 hpf its levels increase slightly; then the levels slightly decrease in 24 hpf (Figure 24A). The slight reduction in 24 hpf is followed by an increase in 48 hpf; then again decrease and increase in 72 hpf and 96-120 hpf (4-5 dpf), respectively (Figure 24A).

It is intriguing to compare our results with the published transcriptome data (Papatheodorou et al., 2018). Even though they score transcript expression levels as TPM (transcript per million), the expression values are roughly similar but alternate in a similar way to our data when it is compared with the 6 hpf stage. In reference to 6 hpf the TPM value increases in 12 hpf followed by a slight decrease in 24 hpf; albeit TPM does not increase in 48 hpf (compared to our data) it is still followed by slight fall and rise in 72 hpf and 96 hpf, respectively.

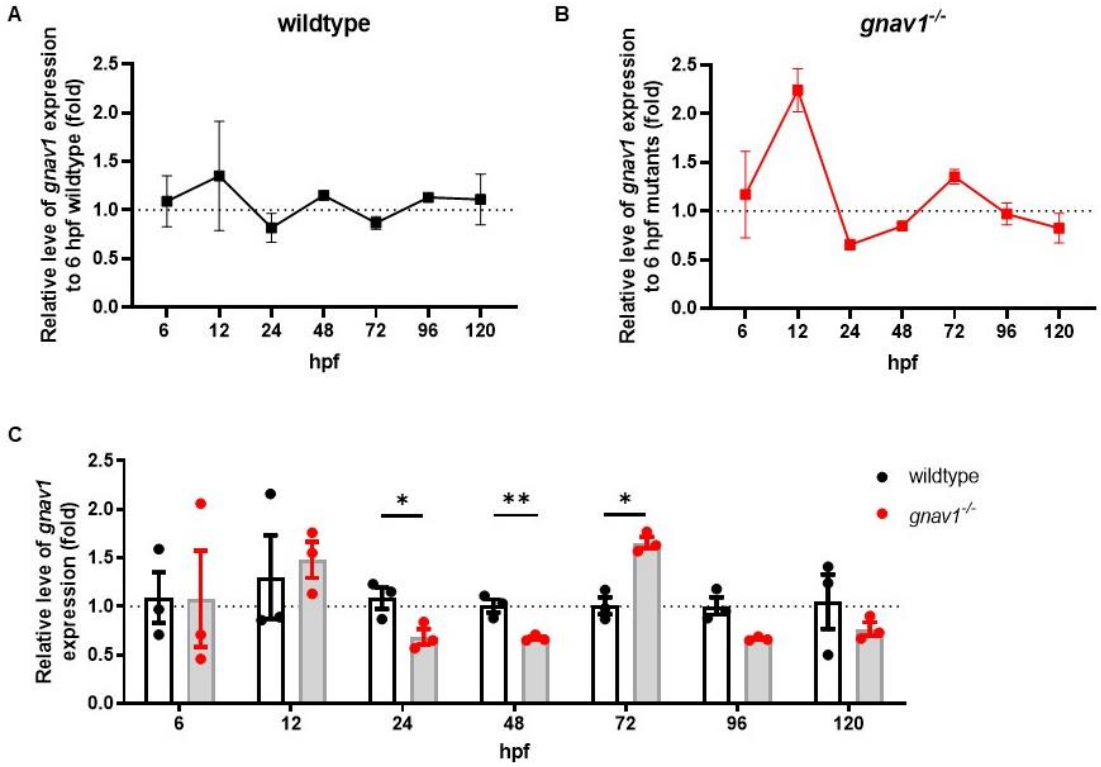


Figure 24: Changes in Gv mRNA levels during zebrafish embryonic and larval quantified by qPCR. (A) relative *gnav1* expression was at 6 hpf - 5 dpf in the wildtype; the relative ratio was calculated relative to 6 hpf wildtype embryos. B) relative *gnav1* at 6 hpf - 5dpf stages in the mutant; the relative ratio was calculated relative to 6 hpf mutant embryos. (C) relative *gnav1* expression at 6 hpf – 5 dpf stages in the mutant versus 6 hpf - 5 dpf wildtype, respectively (n = 3, each sample has a pool of 20 embryo2/larvae per stage). qPCR reactions were run in triplicates and RNA levels were normalized with the reference genes rpl37 and rpl8 expression. The significance was determined with two tailed t-test. *p = 0.01 to 0.05, **p = 0.001 to 0.01, and ***p < 0.001. Error bars denote SEM.

Overall, the qPCR data shows that Gv RNA levels do not drastically change in 12-120 hpf compared to 6 hpf stage, even though WISH data shows that *gnav1* expression begins broadly distributed and eventually is confined to specific organs (see 3.1.1-3).

There could be a maternal contribution in Gv expression at 6 hpf since it appears homogeneous by WISH; if that is the case, therefore, it appears that the embryo rather restores Gv mRNA levels in the stages following 6 hpf to similar levels to compensate for maternal RNA degradation. Also, taken together WISH and qPCR data sets we could infer that the expression turns out restricted, however, concomitant with an increase in the RNA transcription levels in particular organs. These results reinforce the conclusion from WISH data that Gv has an important role during embryogenesis.

In summary, the data obtained from both qualitative and quantitative analysis for Gv mRNA expression patterns suggest that Gv might have manifold functions. To gain a better understanding of these functions we employed the available *gnav1*$^{-/-}$ homozygous mutants.

3.2 Genotyping and examining the presence of Gv mRNA in the homozygous mutant line

A mutant line with preliminary characterization was available in the lab (Ivandic, 2015). First it was necessary to outcross and incross the mutants to further minimize the background/off-target effects. Secondly, it is important to analyze whether Gv mRNA is present in the mutants or is subjected to NMD, since it is not investigated so far.

3.2.1 Confirming the presence of 13 bp deletion in the genomic DNA of *gnav1*$^{-/-}$ mutant line

Ivandic 2015 produced *gnav1*$^{-/-}$ homozygous mutants (Ivandic, 2015). Briefly, one-cell stage zebrafish embryos (KS strain) were injected with *in vitro* transcribed Cas9 mRNA (capped) along with specific guide RNA (gRNA) against exon1. Subsequently, *gnav1*$^{-/-}$ homozygous mutants were outcrossed three consecutive times with wildtype KS strain. Since this strain in the meantime was no longer available, we further performed several outcrossings with the Ab/Tü strain in order to obtain a uniform Ab/Tü genetic background and to further reduce potential off-target effects. Furthermore, back-crossing of heterozygous fish generates mutant and wildtype siblings, which were

used for more stringent control of mutant phenotypes. However, for time reasons we also began to analyze the mutant phenotype using non-siblings (Ab/Tü strain) as control.

Foremost, it is necessary to validate the available mutant line obtained from the previous study, and the new progeny that we produced from outcrossing the mutants. The DNA samples were collected by swab method, followed by DNA extraction. The amplicons for the region surrounding the deletion site were sequenced. A representative example for the sequencing results derived from wildtype and mutant adult fish is depicted in Figure 25. Sequence alignment shows that the mutant harbors the 13 bp deletion at the expected position in exon 1, and the chromatogram illustrates that both genotype sequences are homogenous (Figure 25B, C).

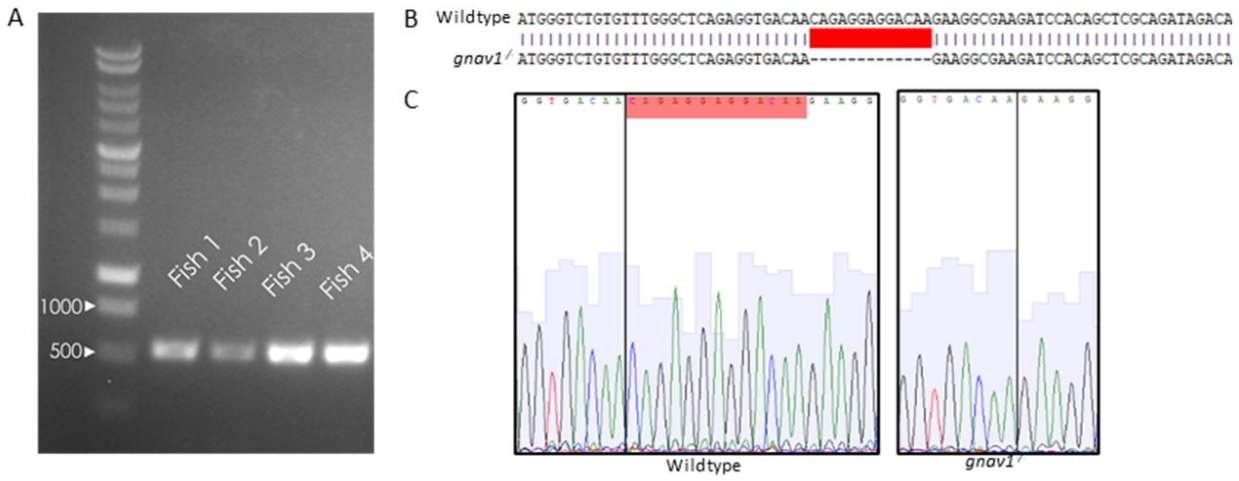


Figure 25: Validating *gnav1*$^{-/-}$ homozygous mutants. (A) PCR amplicons encompassing the deleted region of *gnav1* for progeny of heterozygous zebrafish; DNA is processed by swab method. (B) sequence alignment for genomic DNA of wildtype and mutants; the red bar indicates 13 bp deletion in the mutant sequence. (C) representative chromatogram for the sequences in B. Black line indicates the start of the deletion site. Red overlay indicates the 13 bp deletion site.

Here, we confirm the presence of the deletion (eventuating in frameshift mutation) in the genomic DNA of *gnav1*$^{-/-}$ line used in this study. Interestingly, the starting population of mutant fish contained a mixture of two closely related 13bp deletions differing in position by 3 bases, but resulting in the same premature stop codon. Therefore, no attempt was made to further differentiate these mutations. Next, we want

to examine whether the resulting nonsense mutation would dictate nonsense-mediated mRNA decay (NMD) and thus impact Gv mRNA levels.

3.2.2 The aberrant Gv mRNA bypasses NMD

The 13bp deletion in *gnav1* confirmed above results in a frameshift that introduces a premature stop codon in the first exon, which is expected to release the ribosome, i.e., before it gets to the normal stop codon. Consequently, the downstream exon junction complexes (EJCs) or other RNA binding proteins (RBPs) will not be evicted, which serve as signals to recruit NMD machinery in vertebrates including zebrafish (Wittkopp et al., 2009). Thus, the aberrant mRNA will be degraded, translation and accumulation of truncated proteins will be avoided (Dyle et al., 2020; Fatscher et al., 2015). However, it is not known whether NMD occurs for all truncated proteins.

Here we analyzed the Gv mRNA content of the mutant to check if NMD occurs. The RNA was extracted from 5 dpf wildtype and homozygous mutant larvae and reverse transcribed. Then, *gnav1* was amplified from the cDNA and sequenced using the same primers at the 5' and 3' ends for both the wild type and the mutant. We observe that *gnav1* is still transcribed in the 5 dpf mutants at very similar levels to wildtype (Figure 26). This means that the RNA of *gnav1* is not subjected to NMD. This is unexpected since the mutation fulfills the "50–55 bp rule" requiring the premature stop codons to be at least that distance from the last downstream exon-exon junction (Nagy & Maquat, 1998). However, NMD is often escaped from, when translation can be reinitiated at a downstream start codon (Buisson et al., 2006) either already in the wildtype or as a consequence of the premature stop codon being in close proximity to the initiation codon AUG (Neu-Yilik et al., 2011; Pereira et al., 2015). Both are the case for the *gnav1* mutant. For detailed evaluation of the result see Discussion.

To obtain a more quantitative result for the levels of Gv mRNA in mutant *vs.* wildtype we performed qPCR as previously described for the wildtype (see 3.1.5) also for the mutant for the same seven time points. We normalized Gv mRNA values to the total mRNA using two reference genes (*rpl8* and *rpl37*), then we determined the ratio of the mutant Gv mRNA change in all the stages either to 6 hpf or to the wildtype.

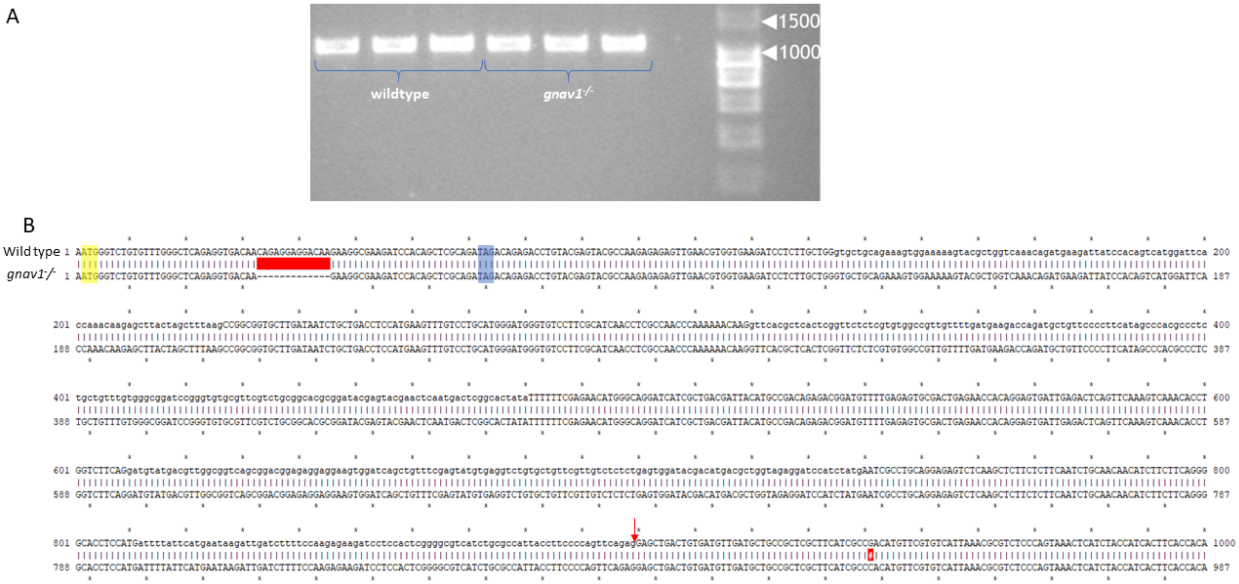


Figure 26: Gv mRNA is transcribed in the mutants and bypass NMD. (A) RT-PCR for 5 dpf wild type and mutants, pool of 20 larvae; Gv mRNA is transcribed in the mutants. (B) sequence alignment for the coding sequence (cDNA) of wildtype and mutant larvae, the premature stop codon (blue overlay) is > 55 bp upstream the final exon-exon borders (red arrow). * marks every tenth bp. The amplicon length is 1177 bp but only 1000 bp are presented due to sequencing limitation, and the quality of sequencing decreases at the end of the sequence so the mismatch is expected. Red bar marks the 13 bp deletion. Capital and small letters indicate alternating exons. Red arrow marks the final exon-exon borders.

When we normalized the expression values of all stages to 6 hpf for wildtype and mutants separately, we observe that the time course of *gnav1* expression in the mutant is different from the wildtype pattern, *cf*. (Figure 24A, B). At 12 hpf, the mutant Gv mRNA is over two folds of the 6 hpf levels; also, the mutant transcript levels increase in 3 dpf stage 1.35-fold, while the wildtype levels decrease slightly (0.8-fold of the 6 hpf levels) at this stage, *cf*. (Figure 24A, B). In the direct comparison of mutant and wildtype *gnav1* levels for each stage (i.e., normalizing the mutant to the wildtype levels), we observe a significant decrease in 1 and 2 dpf *gnav1*$^{-/-}$ levels; the mutants show only 69 and 75 % of the wildtype *gnav1* levels, respectively. However, we observe a significant increase in 3 dpf mutant levels to 150 % of the wildtype (Figure 24C). Taken together this would be consistent with a compensatory mechanism increasing mutant Gv mRNA levels at some stages (12 and 72hpf). The inverse effect

at other stages (24 and 48 hpf), i.e., a decrease in mutant Gv mRNA levels compared to wildtype suggests a complex pattern of regulation for Gv.

In practical terms, the (near) absence of NMD for the mutated *gnav1* means that we cannot confirm the knockout or distinguish the mutants by ISH, as the RNA will be detected in either case. More problematic, we cannot exclude the presence of potential truncated Gv protein. 5'-incomplete protein could be produced if an alternative start codon was used in the mutant Gv RNA. Therefore, it became necessary to examine whether Gv protein is still present in the mutants. To this end we first employed a polyclonal antiserum against a Gv peptide produced by a previous co-worker (Ivandic, 2015).

3.3 The polyclonal anti-Gv antibody does not allow reliable detection of Gv expression in tissues

Initial characterization of the custom-made antiserum appeared promising (Ivandic, 2015). We repeated and extended this analysis by comparing Western blot and whole mount immunofluorescence (IF) staining results with the Gv mRNA levels detected by RT-PCR and WISH, respectively. So far, expression of Gv has only been published at the mRNA level (Oka et al., 2009). There, it was demonstrated that Gv mRNA is expressed broadly, but not ubiquitously, in several adult tissues as well as 3 dpf larvae by RT-PCR and WISH respectively. Here, we confirmed by WISH the *gnav1* expression in 3 dpf larvae and showed for the first time expression in the swim bladder, gut, and kidney of 5 dpf larvae. We therefore evaluate the anti-Gv serum and affinity-purified antibody using these larval stages and tissues; both in wildtype (to determine specificity) and mutant (to examine possible retention of (truncated) Gv protein). We employed Western blot (for sensitive detection of denatured Gv in tissue extract) and IF staining (to analyze protein expression in the native tissue), as the epitope recognized by the primary antibody may not be identically available in Western blot and IF assays.

3.3.1 Western blot analysis

Polyclonal antibody against Gv protein was generated by immunizing rabbits with a synthetic peptide (Ivandic, 2015). Portion of the serum was purified on an affinity column by Ivandic, 2015. Here we used both antiserum and affinity-purified antibody for the Western blot analysis.

The expected molecular weight of Gv is 41.2 kDa (*ExPASy - Translate Tool*, 2019). In addition, two transcriptomes have annotated shorter versions of *gnav1* transcripts starting from the third exon (EST sequences EE310328, and CT736382, respectively). The predicted molecular weights for these short forms are 25.8 and 28 kDa, respectively. The polyclonal antibody recognizes an amino acid sequence that corresponds to exon 4, thus it would bind to all possible Gv forms.

To obtain a direct estimate for the expected band size of Gv protein in SDS-PAGE gel analysis, we have generated native and at the C-terminus histidine-tagged (His-tagged) Gv recombinant protein in collaboration with Dr. Thomas Hermanns and Ms. Israa Peker (see 3.4.2 for more details). SDS-PAGE for both purified native and His-tagged Gv was performed with pre-stained protein ladders at each side and gels were analyzed by Image Lab software 6.0.1 (Bio-Rad). The expected Gv molecular weight was estimated using the software. In repeat experiments maximal variability of molecular weight determinations for the same band was 2 kDa (data not shown). The native recombinant protein shows a band size of roughly 45 kDa (Figure 27), which is in good accordance with the predicted value of 41 kDa. Both the shape of the protein and technical aspects of the gel analysis can slightly shift the observed molecular weight value. The His-tagged Gv has 6 extra residues and accordingly a slightly larger band at estimated 48 kDa (Figure 27). In contrast to the native recombinant protein, the His-tagged protein also shows an additional band at 26 kDa (Figure 27). This 26 kDa band might correspond to an alternative start of translation as described above, but could also be a degradation product.

We expect that the endogenous Gv protein will differ somewhat from these values determined for the recombinant protein, because protein modification will shift

electrophoretic mobility. In particular, the endogenous Gv protein has putative sites for N-linked myristoylation and thio-palmitoylation. For other G alpha proteins myristoylation and/or thio-palmitoylation are known to occur and do result in apparent lower molecular weight (Linder et al., 1991; Smrcka, 2004). Later experiments with PRM (see 3.3.3) show the range of 40-48 kDa to encompass the Gv band. Taken together, we expect a variable band size (40-48 kDa) for Gv in the immunoblot.

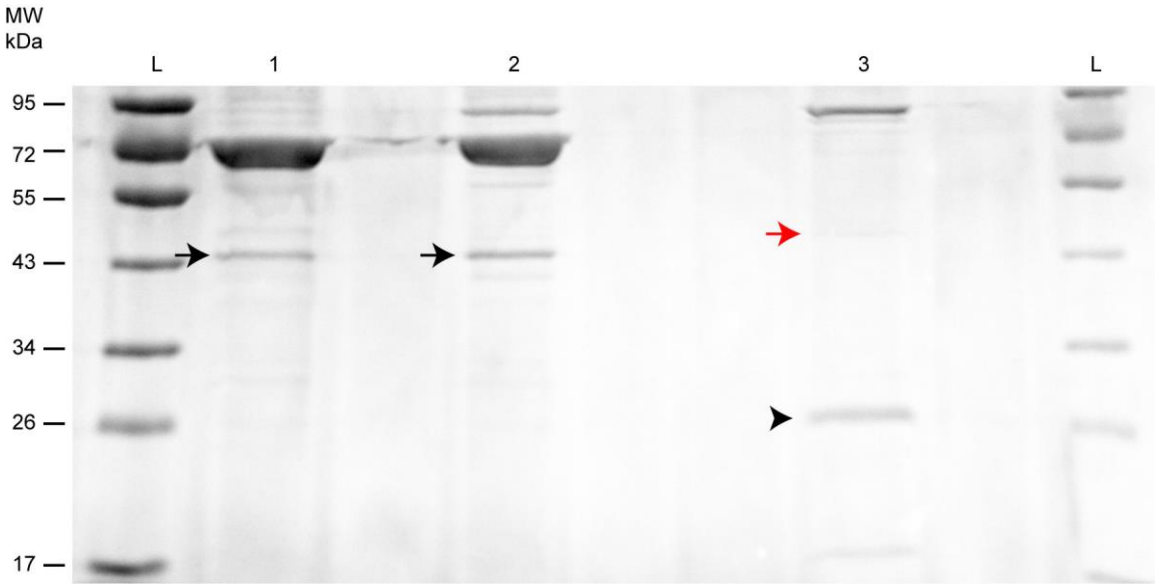


Figure 27: Evaluating the purified Gv recombinant proteins. The native Gv protein produced from pOPIN-K cloning (lanes 1 and 2) and the His-tagged Gv protein produced from pET24a (+) vector cloning (lane 3) were resolved by SDS-PAGE and subjected to Coomassie brilliant blue staining. The estimated molecular weight for the native recombinant protein (black arrows) and His-tagged Gv (red arrow) are roughly 45 and 48 kDa, respectively. The additional band at 26 kDa in lane 3 (black arrowhead) might correspond to degradation product or an alternative start of translation. MW; molecular weight, L; ladder.

We performed immunoblots with different larval stages and tissue types as described (Materials and Methods 2.7.2) to examine the specificity of the antiserum and purified antibody. Total protein was extracted from 3, 5 and 8 dpf larvae, as well as adult kidney and testis of wildtype and mutant. To increase the reliability, Western blot was performed repeatedly using either anti-Gv serum or purified anti-Gv antibody. The results differ somewhat for larvae and adult organs and will therefore be discussed separately.

Three larval stages were examined by Western blot (Figure 28). The affinity-purified antibody labels fewer bands than the antiserum, consistent with a reduction of unspecific binding in the purified antibody (Figure 28B). Unexpectedly, the staining with the purified antibody is distinctly fainter than that obtained with the antiserum, even though the purified antibody is enriched relative to the antiserum (Ivandic, 2015). Since we cannot exclude that the harsh elution conditions during purification have impaired the activity of the purified antibody we have continued the evaluation with both antiserum and purified antibody.

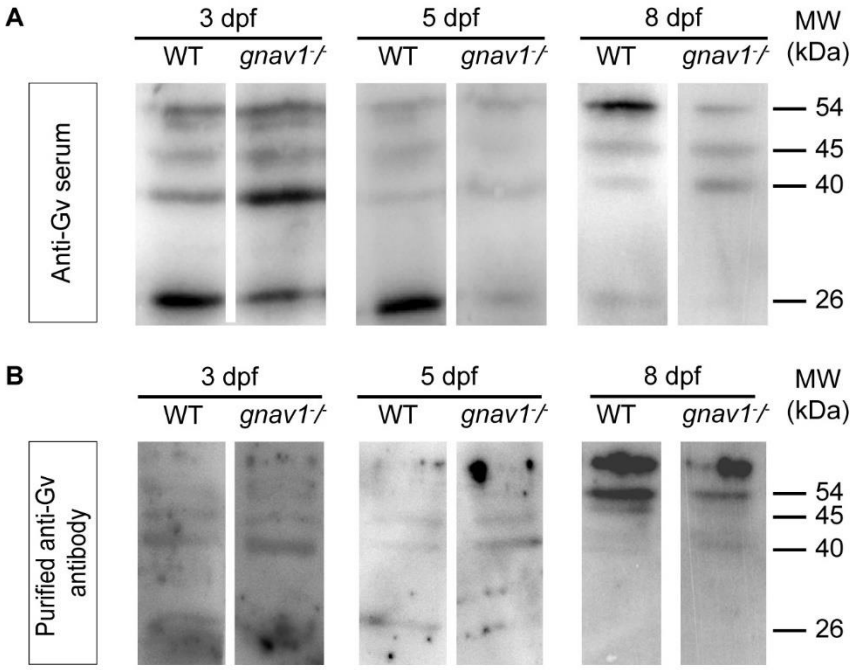


Figure 28: Western blot analysis to examine the specificity of the anti-Gv serum and affinity-purified anti-Gv antibody with different larval stages. Immunoblots were performed using total protein from 3, 5 and 8 dpf wildtype and mutant larvae with (A) anti-Gv serum and (B) anti-Gv antibody. The same bands are present in all wildtype and mutant larval stages with variable intensity; 40, 45, and 54 kDa fit to the expected molecular weight range of Gv; 26 kDa (smaller than expected) might correspond to Gv fragments either from alternative splicing or degradation, but also could result from cross-reactivity. 40 kDa band has higher intensity in 3, 5, and 8 dpf mutants compared to wild type (A); also, the 26 kDa band is decreased in the 3, 5, and 8 dpf mutant compared to wildtype. MW; molecular weight

Intriguingly, the same bands are present in all wildtype and mutant larval stages with variable intensities (Figure 28A, B). Three of the bands, 40, 45, and 54 kDa fit to the expected molecular weight range of Gv protein as they could correspond to lipid modification, unmodified Gv or glycosylation, respectively. All three bands are stained

with both antiserum and purified antibody, albeit the intensity is less with the purified antibody. They are present in all three larval stages investigated. Unfortunately, it is not possible to decide at this stage, whether these bands detect Gv protein (and a 5' truncated Gv protein in the mutant, see 3.1) or alternatively, represent cross-reactivity of antiserum and purified antibody. A band of 26 kDa (smaller than expected) molecular weight is also present, which could represent Gv fragments either from alternative splicing, secondary start of translation or degradation, but also could result from cross-reactivity.

Next, in an attempt to obtain less ambiguous results, adult tissues with stronger Gv expression (Oka et al., 2009) were analyzed by immunoblot (Figure 29). The immunoblots of kidney and testes showed overall more bands compared to the immunoblots with larval tissue (Figure 29A, B), possibly because the higher protein concentration used in the extracts from adult tissue (65 µg/lane compared to 40 µg/lane for the larval extracts) exacerbates the problem of cross-reactivity. Again, less bands are observed with the affinity-purified antibody compared to the antiserum, and again, the staining intensity is lower with the purified antibody (Figure 29B). Consistently, bands expected for Gv and a potential truncated form (26, 40, 45, and 54 kDa) are stained with antiserum (Figure 29A), but only the 26 kDa is stained consistently with the purified antibody (Figure 29B). Thus, it is still difficult to determine which of these bands, if any, are specific for Gv protein.

Therefore, we looked for further characteristics within the immunoblots to help with assignment of bands to molecular forms. The intensity of several bands appeared to be different between wildtype and mutant. Thus, we decided to quantitate the band intensity using Image Lab software 6.0.1 (data not shown) (Bio-Rad). All available immunoblots were analyzed, amounting to 2 to 4 repeat experiments per age and tissue. The intensity of the 26 and 40 kDa bands are considerably changed in mutants compared to wildtype, yet differently between larvae and tissue (Figure 28 and Figure 29). For instance, 40 kDa band has consistently over twofold higher intensity in 3, 5, and 8 dpf mutants compared to wild type, but not in kidney and testis. In contrast, the intensity of the 26 kDa band is increased in the mutant kidney and testis compared to wildtype, but not in the 3, 5, 8 dpf larval tissues of mutants, which actually show a

marked and consistent decrease in intensity (Figure 28 and Figure 29). Again, it is difficult to form a hypothesis to explain all results. One could hypothesize a compensatory increase of a truncated Gv in the mutant, but we do not see a shift in molecular weight in the mutant 40 kDa band. An increase of the 26 kDa band in the mutant could be explained by increased use of a secondary initiation of translation for Gv, but this would leave the decrease of the 26 kDa band in larval mutants unexplained. Furthermore, the intensity of the 26 kDa band relative to other bands is decreased with purified antibody compared to antiserum. Taken together these results suggest that the 26 kDa band (and possibly the 40 kDa band) are caused by cross-reactivity to unknown proteins.

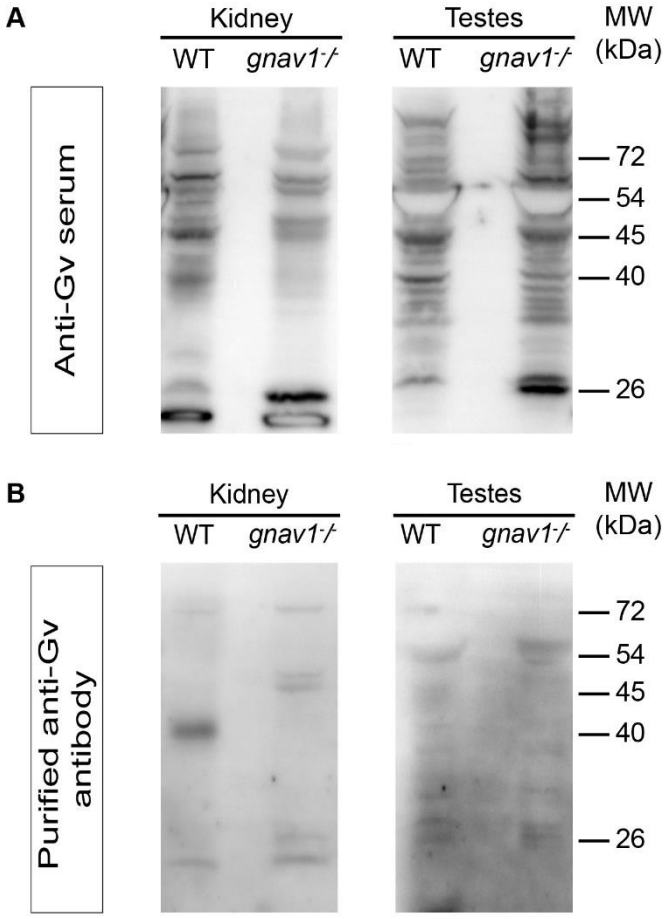


Figure 29: Western blot analysis to examine the specificity of the anti-Gv serum and affinity-purified anti-Gv antibody using adult organs. Immunoblots were performed using adult kidney (left) and testes (right) for wildtype and mutant with (A) anti-Gv serum and (B) anti-Gv antibody. Expected Gv bands (26, 40, 45, and 54 kDa) are stained with antiserum in immunoblots of both tissues (A). Only the 26 kDa band is stained consistently with the purified antibody (B) and its intensity increased (A) in the mutant kidney and testis compared to wildtype. Note: the

broad band at 54 kDa in (A) testes appears white due to high saturation. MW; molecular weight.

Overall, the immunoblot results with larval and adult tissues are more easily explained by assuming cross-reactivity than by detection of Gv. It is possible that Gv is under the detection limit under our experimental conditions, but it is also possible that the antibody is unable to detect Gv. To distinguish between these two possibilities, we performed Western blots with the recombinant Gv protein (see below).

However, we wish to note that one clear conclusion can be drawn from the manifold comparisons between mutant and wildtype: the consistent changes in intensity of the 26 and 40 kDa bands in mutant compared to wildtype suggest that levels of the underlying unknown proteins are altered in mutant animals, i.e., constitute a consequence of the absence of functional Gv.

We performed Western blots with the recombinant Gv protein to evaluate whether anti-Gv serum and the affinity-purified anti-Gv antibody are able to recognize Gv protein. Staining with pre-serum is included to examine whether unspecific binding originates from rabbit antibodies. We observe that both the anti-Gv serum and the purified antibody bind to Gv recombinant protein at the predicted molecular weight (45 kDa) (Figure 30). Both the anti-Gv serum and the pre-serum stain additionally a larger band at 72 kDa, which presumably corresponds to a bacterial contaminant not completely removed during the purification process (see 3.4.2 for more details). This band is absent when using purified antibody showing that the affinity purification of the antibody did strongly increase specificity. Moreover, the intensity of the 45 kDa band is strongly increased with the purified antibody compared to antiserum, which is expected due to the higher concentration of antibody in the purified preparation (see (Ivandic, 2015), p.39-40). Importantly this is in clear contrast to the situation with larval and adult tissues, where the affinity-purified antibody always had lower intensity bands compared to the antiserum, and as such makes a strong point that all bands appearing weaker with affinity-purified antibody (i.e., all bands) are caused by cross-reactivity and do not represent Gv protein.

Taken together, the antibody could be validated by Western blot with recombinant protein to detect Gv. However, Western blots with larval and adult tissues show several cross-reacting bands and no unambiguous Gv band, suggesting that Gv is under the detection limit in these tissues using our experimental protocols. Although Western blot as a method is considered to be more sensitive compared to IF, an expression in a sparse cell population might be easier to detect with IF. Therefore, we next turned to IF staining.

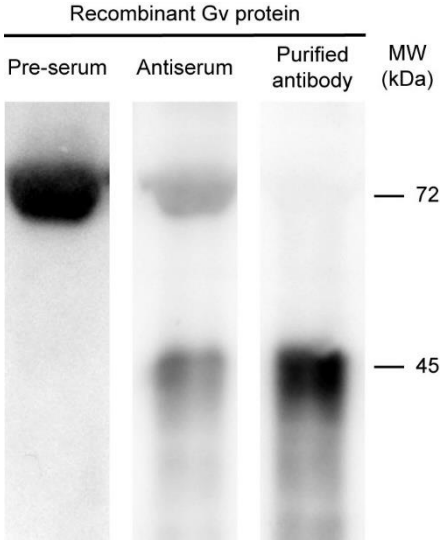


Figure 30: Evaluating whether anti-Gv serum antibody and the purified antibody recognize the recombinant protein (positive control) by Western blot assay. The anti-Gv serum and the pre-serum stain large band at 72 kDa (bacterial contaminant not completely removed during the purification process), yet absent with the purified antibody blot. However, both the anti-Gv serum and the purified antibody recognize Gv recombinant protein at the predicted molecular weight (45 kDa). MW: molecular weight.

3.3.2 Immunofluorescence staining

IF staining coupled with confocal imaging is an invaluable approach, it permits for excellent sensitivity and amplification of signal. We attempted to visualize the three-dimensional distributions for Gv protein and determine its cellular localization in the intact tissues through the entire organism using IF staining. However, multiple factors have to be considered and optimization steps have to be taken to ensure both sensitive and specific staining.

Wholemount IF staining was carried out for 3 and 5 dpf larvae, and adult kidney, using both anti-Gv serum and purified anti-Gv antibody. Three negative and one positive control were used. Anti-acetylated α-tubulin antibody was used for labeling as a **positive control** for the technique. One **negative control** staining was performed with only a secondary antibody. In addition, control IF staining with **pre-immune serum** and **rabbit IgG** were performed to test for the specificity of the primary anti-Gv serum. The staining was visualized under a confocal microscope after treatment with a fluorescent secondary antibody against immunoglobulins of the species of origin of the primary antibody.

The negative control staining with only a **secondary antibody** showed no fluorescence in all of the performed experiments (Figure 31 and Figure 32). Therefore, we exclude the possibility of non-specific signals that could be produced from the secondary antibody. The **positive control** (anti-acetylated α-tubulin antibody) expectedly labels axonal tracts in 3 and 5 dpf larvae (Figure 31 and Figure 32). This indicates that the IF staining protocol is reliable and sufficiently sensitive. The **pre-immune serum** staining for 3 dpf wildtype in Figure 31D as well as **rabbit IgG antibody** staining for 3 and 5 dpf wildtype larvae in Figure 31C and Figure 32C show general background labeling in the brain region (presumably the optic tectum), around the eye, and skin, therefore any of these signals emerging with anti-Gv serum and purified anti-Gv have to be considered as unspecific. For completeness sake, we stained also the 5 dpf mutant larvae with rabbit IgG. Interestingly, the skin at the head area looks somewhat different in mutant compared to wildtype (Figure 32C, D). The keratinocytes cover a smaller head area in the mutant, and they appear less structured. This could mean that the background staining of the rabbit IgG picks up a difference in (non Gv) protein expression between wildtype and mutant, similar to the different band intensities for clearly unspecific bands seen in the immunoblots between wildtype and mutant.

The anti-Gv serum staining for 3 dpf larvae reproduces the unspecific signal observed with the pre-serum staining (Figure 31E, F), but also labels several additional structures in the wildtype: a faint staining in the pronephros (Figure 31, white arrows), intensely labeled dot-like structures on the skin and branchial arches (possibly blood

cells or ionocytes) (Figure 31, red arrowheads), and intensely labeled pectoral fin (Figure 31, white arrowheads). The WISH results (Figure 21) show signals in the pronephros, branchial arch and pectoral fin, thus the IF signals in these regions might correspond to Gv protein. However, the blood cell or ionocyte IF signals on the skin do not have a counterpart in the WISH signals, suggesting that at least these signals are caused by cross-reacting proteins. Moreover, there are two regions clearly labeled in WISH, the palatoquadrate cartilage and the inner ear, which do not show IF signals. Therefore, we hypothesize that the protein expression is not high enough to be detected by IF staining at this stage. The mutant Gv genotype generally shows the same pattern of labeling, albeit less intense (Figure 31).

Using WISH, we showed that *gnav1* is mainly expressed in the gut, swim bladder, and kidney (pronephric tubule) of 5 dpf larvae (Figure 22). In contrast, the IF staining with anti-Gv serum for 5 dpf wildtype shows no staining in the gut and swim bladder (Figure 32), although a very faint staining in the pronephric tubules was observed (only visible if projection was done for a limited z stack, not with maximum intensity projection of the full stack) (Figure 32E, F white arrow). Moreover, we observe prominent IF labeling on structures, which are not stained in WISH, such as ionocytes on the branchial arch (Figure 32, red arrowheads) and neuromasts (Figure 32, white arrowheads). Mutant genotype shows very similar staining to wildtype. It is not clear, whether the antibody is not suitable to detect native, non-denatured Gv protein or whether it may be not sensitive enough to detect possibly low Gv protein expression in 3 and 5 dpf larvae. In the latter case, higher concentrations of antibody could help, but also might increase the unspecific staining leading to higher background intensity.

Therefore, we further stained wildtype 3 and 5 dpf larvae with the purified antibody. We considered staining to be specific for the purified antibody if it was not present in the rabbit IgG negative control, *cf*. (Figure 31 and Figure 32). No staining was observed at concentrations used for Western Blot. At 2 - 4 fold higher concentration we did observe faint staining in the kidney and pectoral fin as well as some labeling of the quadrate cartilage and branchial arches (Figure 33), which are regions expected to express Gv from WISH results (Figure 21 and Figure 22).

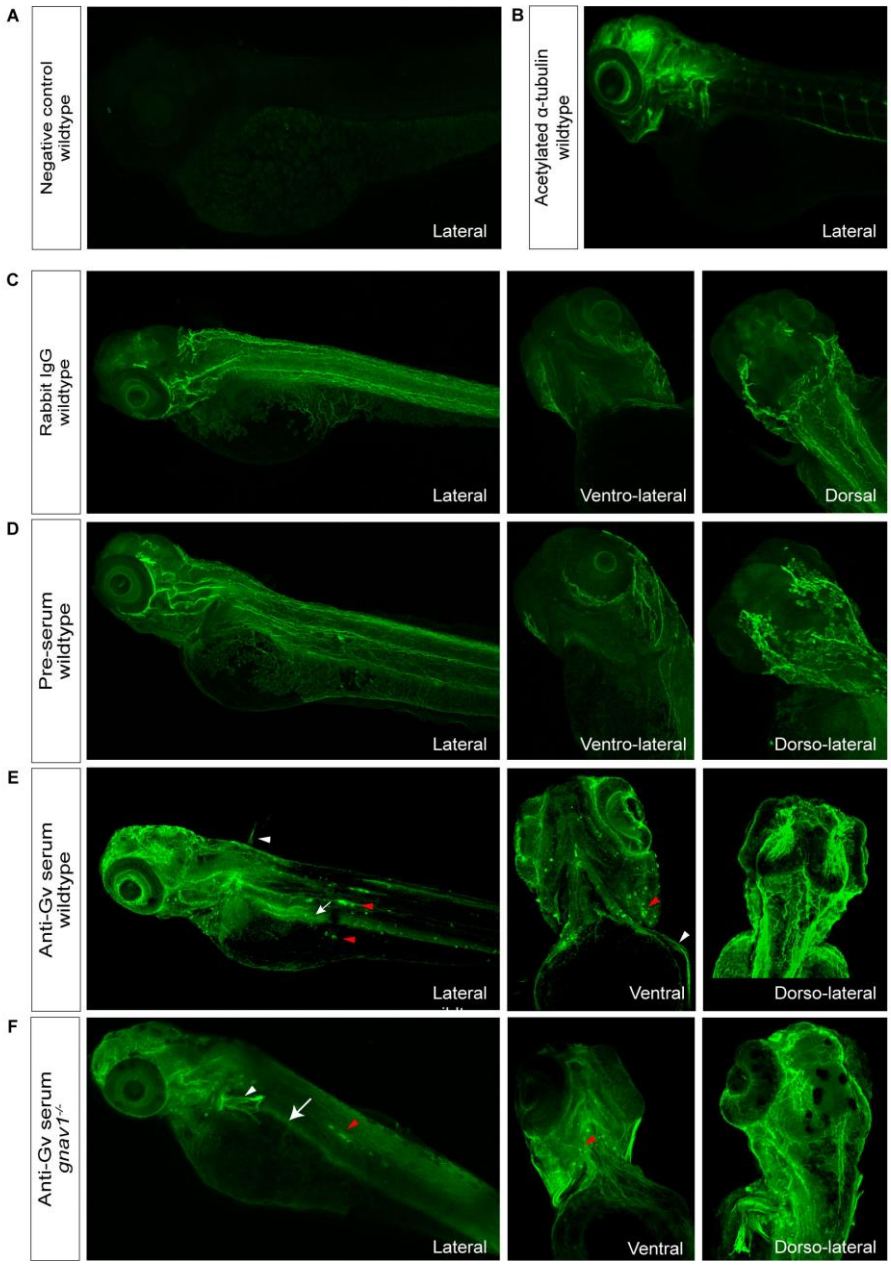


Figure 31: Wholemount IF staining for 3 dpf wildtype zebrafish larvae to examine the anti-Gv serum. (A and B) are lateral views; from (C) to (F) the left column is lateral view, the middle is ventral or ventrolateral, and the right column is dorsal or dorsolateral view of a maximum intensity projection confocal image. (A) negative control staining with only a secondary antibody shows no signal. (B) anti-acetylated α-tubulin antibody staining is used as a positive control and the antibody stains the axonal tracts. (C) rabbit IgG antibody and (D) pre-immune serum IF assays show general background labeling in the brain region, around the eye, and skin. (E) and (F) are wildtype and mutant larvae, respectively, stained with anti-Gv serum and labels the brain region, around the eye, skin, pronephros, and pectoral fin similarly in both

genotypes. White arrows: pronephros; white arrowheads: pectoral fin; red arrowheads: possibly blood cells or ionocytes.

However, in both developmental stages, unexpected labeling for the muscular filaments, notochord and optic nerve is prominent, as well as quadrate cartilage in 3 dpf (Figure 33, white arrows). Moreover, from the WISH data we expect Gv to be present in inner ear, lower lips, and yolk surface, but no IF staining was observed in these tissues with the purified anti Gv antibody (Figure 33). Hence, the pronephros, pectoral fin, branchial arch and quadrate cartilage signals could correspond to Gv expression, but this is less probable due to the absence of signals in ear, lower lip, and yolk cells, and the presence of IF signals in regions where WISH signals are absent. Moreover, we did not obtain more intense IF signals in kidney and gut than with antiserum, although the concentration of anti Gv antibodies is expected to be much higher in the purified antibody. Thus, signals overall might be caused by cross-reactivity and the purified antibody still might not be sensitive enough to detect endogenous Gv in the larvae.

As a final attempt we implemented IF staining for whole mounts of adult zebrafish kidney, because it should be more enriched there (Oka et al., 2009) and (Figure 23). The whole mount labeling appears to be sufficiently specific, since the negative control (only secondary antibody) shows no signal. The basic kidney structure of zebrafish is comparable to the mammalian metanephros. It is composed of several hundred nephrons that are mainly epithelial tubules (Figure 11). A prominent staining is observed in the tubular cells of the wildtype kidney that colocalize with DAPI-stained nuclei (Figure 34). While it cannot unambiguously be concluded that this staining represents presence of Gv, it is interesting to note that the mutant kidney still exhibits staining with anti-Gv antibody, albeit with considerably less intensity (Figure 34). A possible explanation could be the (less efficient) translation of a truncated Gv protein in the mutant.

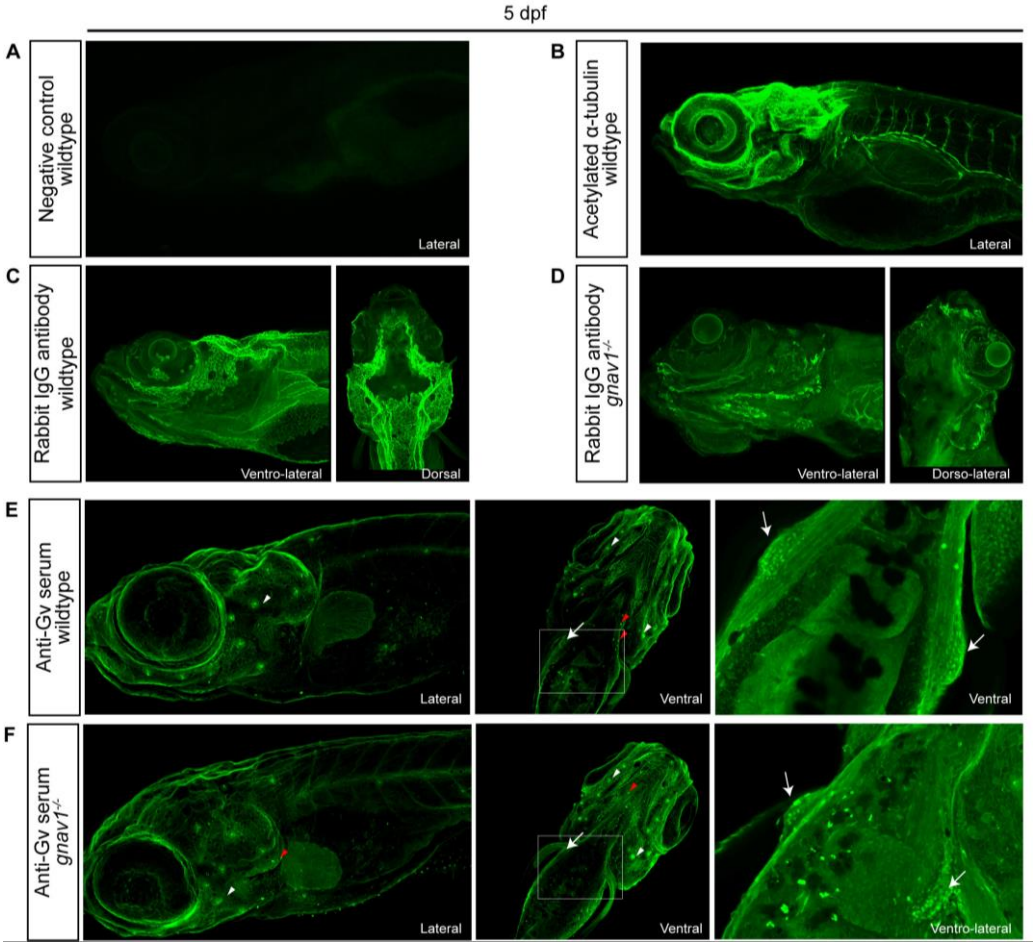


Figure 32: Wholemount IF staining for 5 dpf wildtype zebrafish larvae to examine the anti-Gv serum. (A and B) are lateral views; (C and D) left is lateral, right is dorsal or dorsolateral views; (E and F) left column is lateral view, the middle column in is ventral, and the right column is magnified dorsal or dorsolateral views of a maximum intensity projection confocal image. (A) negative control staining with only a secondary antibody shows no signal. (B) anti-acetylated α-tubulin antibody staining is used as a positive control and the antibody stains the axonal tracts. (C and D) rabbit IgG antibody staining for wildtype (C) and mutant (D) larvae show general background labeling in the brain region, around the eye, and skin. Skin keratinocytes cover a smaller head area in the mutant, and they appear less structured in (D) compared to the wildtype in (C). (E) and (F) are wildtype and mutant larvae, respectively, stained with anti-Gv serum and labels structures around the eye, skin, and the branchial arches possibly ionocytes or neuromasts, pectoral fin, and pronephros (right column depicts selected z stack not with the max intensity projection of full stack) similarly in both genotypes. White arrows: pronephros; white arrowheads: pectoral fin; red arrowheads: possibly blood cells or ionocytes.

Taken together, the results of both Western Blot analysis and IF suggest that the custom-made antibody against Gv is neither sufficiently specific nor sufficiently sensitive to detect Gv, with the possible exception of tissues with relatively high Gv content. Therefore, we turned to an alternative method for detection of Gv protein, i.e.,

MS and PRM, with the final goal of determining whether and how expression of Gv is altered in the mutant Gv genotype.

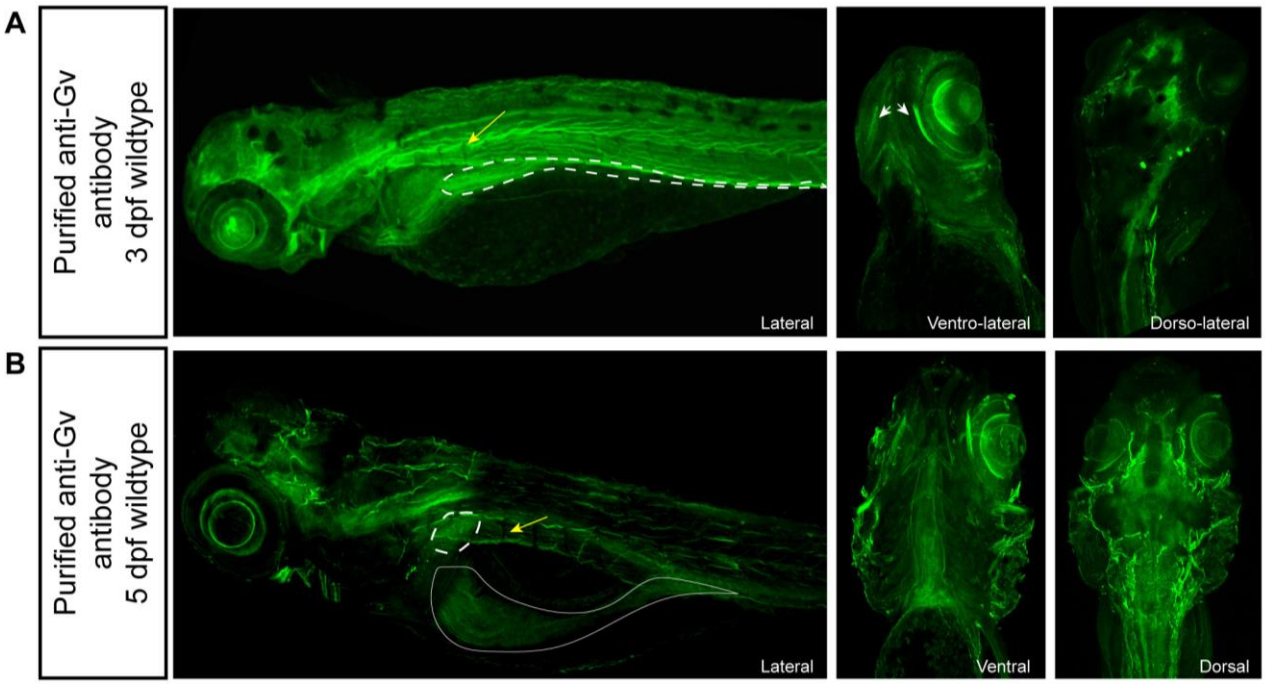


Figure 33: Wholemount IF staining for 3 and 5 dpf wildtype zebrafish larvae to examine the purified anti-Gv antibody. In (A and B) left column is lateral view, the middle column is ventral or ventrolateral view, right column is dorsal or dorsolateral view of a maximum intensity projection confocal image. (A and B) rabbit IgG antibody staining for 3 dpf (A) and 5 dpf (B) larvae. In 3 dpf larvae (A) labeling for the muscular filaments, notochord, palatoquadrate, and optic nerve is prominent; in 5 dpf larvae (B) labeling for the muscular filaments, notochord, and optic nerve is observed. Faint staining in the kidney, pectoral fin, and branchial arches is observed in both stages (A and B). White arrows: palatoquadrate; yellow arrows: muscular filaments; dashed line: pronephros; white continuous line: gut.

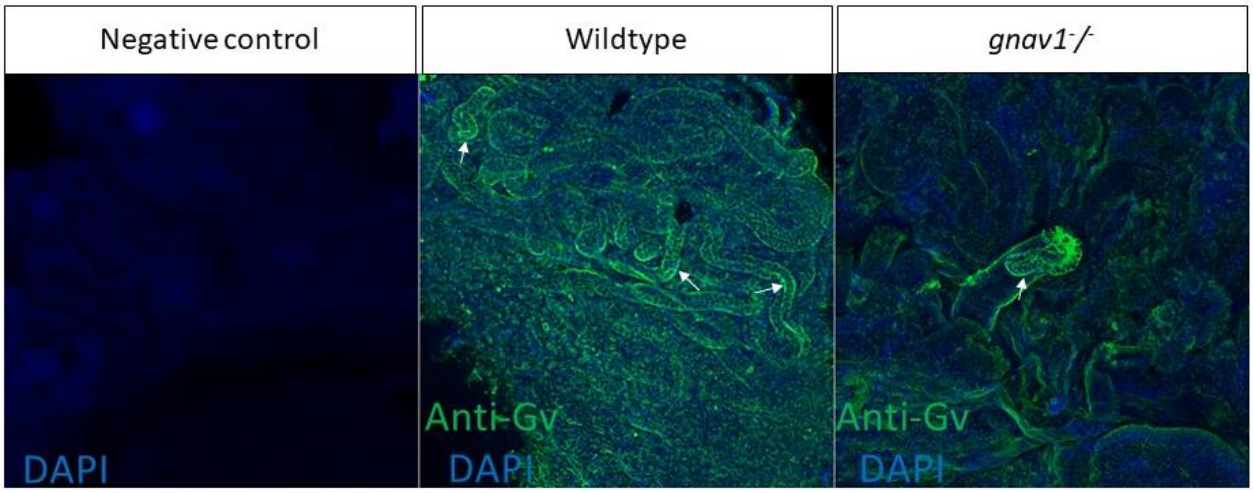


Figure 34: Wholemount IF staining for adult kidneys of wildtype and mutants to examine the anti-Gv serum antibody. The figures present a maximum intensity projection confocal image for the trunk part of adult zebrafish kidney stained with anti-Gv serum. Left column is negative control staining with only a secondary antibody and it shows no signal. Middle: wildtype kidney is stained with anti-Gv serum; it shows staining in the tubular cells that colocalize with DAPI-stained nuclei. Right: *gnav1*$^{-/-}$ adult zebrafish kidney stained with anti-Gv serum; it shows less intensity staining in the tubular cells colocalizing with DAPI-stained nuclei. White arrows marks examples of stained epithelial tubular cells.

3.4 Gv protein detection with mass spectrometry

3.4.1 Standard mass spectrometry is not sensitive enough to detect Gv protein in tissues

The ambiguous data obtained from Western blot and IF staining suggests that the custom-made antibody is not robust enough to localize Gv protein expression, and cannot be used to investigate the presence or absence of Gv in the knockout mutants. As an alternative approach to detect the Gv protein we performed MS. This approach should enable us to determine whether retention of Gv occurs in the mutants. In addition, it provides quantitative analysis that is useful to gain insights into changes in the protein profile in the mutants compared to the wildtype.

Initially, we began by performing MS for total protein extracted from four samples of 5 dpf wildtype larvae (known to express Gv, see 3.1.3). Two different lysis buffers were

used, Urea and RIPA buffers, in order to allow detection of as many different proteins as possible, which should increase the probability of Gv being included. As shown in Table 2, 2388-3738 different proteins were found per sample. In total 5342 proteins appear at least once. More proteins were present in the samples that were lysed with RIPA buffer. Many unique proteins were detected including several G alpha proteins from the other four families as following: G15 and G11 (Gq family), G13 (G12 family), Go, Gi and Gt (Gi family), and Golf (Gs family). Their molecular weights are 40-46 kDa. However, Gv was not found, although it is obvious from WISH and qPCR data that Gv is present in 5 dpf larvae. The most likely reason is a low abundance of Gv in the sample, below the detection limit. Therefore, we repeated MS with specific tissues where Gv is expected to be more enriched such as the adult kidney and testes (cf. Oka et al., 2009).

Table 2: Protein quantification results measured by MS for wildtype samples including 5 dpf larvae, adult kidney, and adult testes.

Sample type	Lysis buffer	# of peptide sequences identified	# of proteins identified
5dpf	Urea	17051	2388
5dpf	Urea	20424	2794
5dpf	RIPA	27773	3227
5dpf	RIPA	30300	3738
kidney	RIPA	29145	4064
kidney	RIPA	28669	4533
testis	RIPA	36252	4877
testis	RIPA	38170	5508

As shown in Table 2, more proteins were detected using kidney and testes samples, for instance 7094 proteins in total. Similar to 5dpf larvae, G alpha proteins from the other four families were detected, nevertheless, Gv again was not found. The MS is probably not sensitive enough for the detection of Gv protein in wildtype. To overcome this limitation, we decided to implement targeted PRM, which selects only fragments

from the protein of interest for detection (targeted-MS). Consequently, the background is much reduced in PRM, leading to much higher sensitivity, reproducibility and quantitative accuracy. A prerequisite for this method is the availability of pure protein of interest, which is required for calibration. Hence, recombinant Gv protein has to be generated.

3.4.2 Production of recombinant Gv protein

Besides using recombinant Gv to calibrate PRM, it can also be used for producing new antibody, which could be more specific than those produced against a single short peptide sequence as an antigen. Furthermore, the recombinant Gv protein can be used as a positive control in Western blots with Gv antibody. Unfortunately, purification of G***α*** proteins can be challenging, in part because they are localized at the inner leaflet of the membrane and require to be in the soluble form for successful purification (Smrcka, 2004).

Initially, we cloned the zebrafish *gnav1* sequence obtained from cDNA of adult kidney in pET24a (+) vector, which contains His-tag at the C-terminus for affinity purification. We performed the expression under six different conditions for optimization. We incubated the cultures at three different temperatures 16, 25 and 30 °C; and induced the expression with either 0.1nM or 1nM IPTG. Thereafter, Western blot was performed using anti His-tag antibody to examine Gv expression in the bacterial cell lysates from the cultures before and after inducing the expression. We observe that two bands at 48 and 40 kDa appear in the immunoblot before and after inducing the expression with IPTG (Figure 35, the rows marked with arrow and arrowhead, respectively). An additional band 25 kDa appears only after induction (Figure 35, the row marked with red arrowhead). The 48 kDa band corresponds to His-tagged Gv protein, and the other two bands could correspond to degraded Gv protein. It is not clear why the 25 kDa band appears only after induction. Gv levels do not notably go up after induction, but this is a known consequence of leaky expression (Gv protein is already expressed before the induction). This results in overall low levels of Gv protein. Several attempts were conducted to purify the protein on an anti-His-tag affinity column in collaboration with Ms. Israa Peker (AG Riemer Lab, Institute of Genetics).

The purification was performed at three different temperatures, 4, 25, and -20 °C. Coomassie staining for the preparation at 4 °C shows very faint band at 48 kDa which corresponds to Gv, also a fragment with higher intensity at 26 kDa, possibly a degradation product, and a prominent band at 92 kDa, presumably a contaminant bacterial protein eluted during Gv purification (Figure 27, lane 3). Western blot was carried out with anti-His-tag antibody for the purified protein at three different temperatures 4, 25, and -20 °C. For all purification temperatures the immunoblot shows a weak smeared band at 40 kDa in addition to 26 kDa, probably attributed to degradation of the protein (Figure 36). Signal was best for the 4°C purification, but still very faint and the majority of protein was not intact. Therefore, we repeated the generation of recombinant protein with a different vector and tagging method in collaboration with Dr. Thomas Hermman (AG Hofmann's Lab, Institute of Genetics).

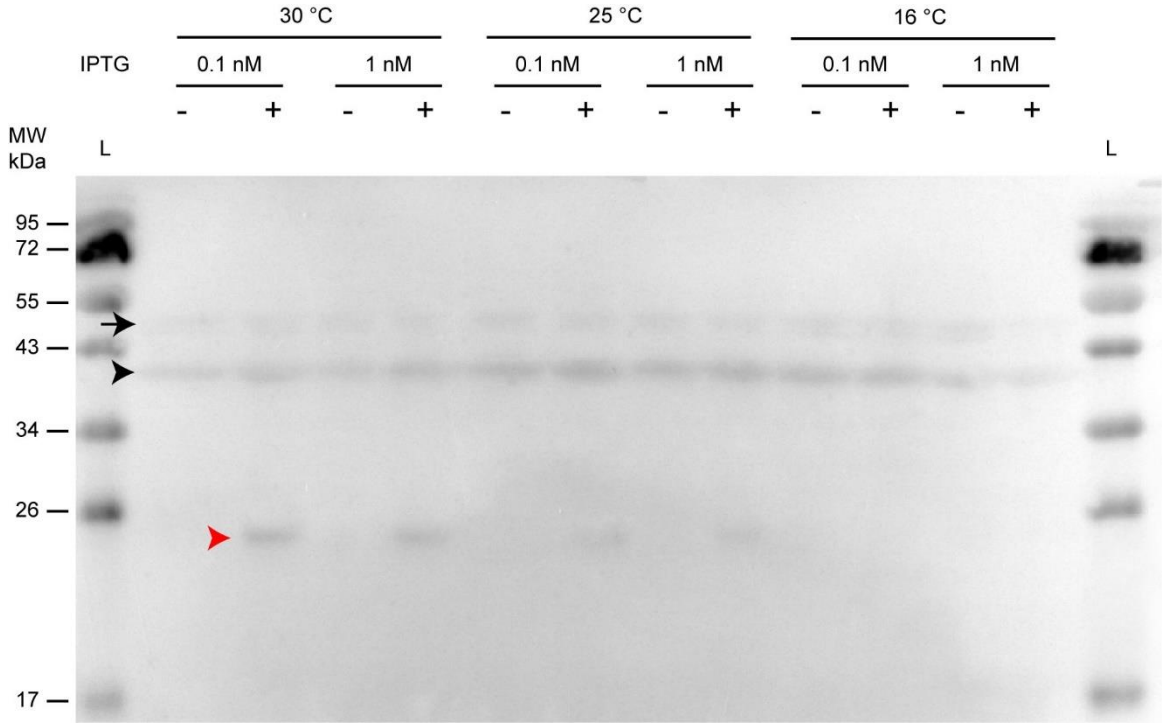


Figure 35: Optimization for *gnav1* expression cloned in pET24a (+) vector. Western blot stained with anti His-tag antibody for total protein extracted from bacterial lysate before and after induction of *gnav1* expression. *gnav1* expression was induced with either 0.1 or 1nM IPTG and bacterial cultures were incubated at different temperatures (16, 25, or 30 °C). The staining shows three bands; at 48 kDa (arrow) which possibly corresponds to Gv protein; 40 and 25 kDa (black and red arrowheads respectively) could correspond to Gv fragments resulted from degradation. Gv levels are not increased after induction in all tested conditions

suggesting leaky expression. MW; molecular weight, L; ladder, IPTG; isopropyl β-d-1-thiogalactopyranosides.

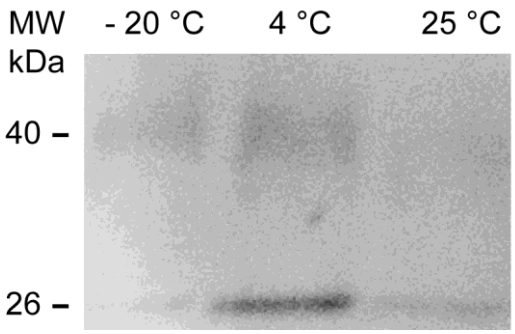


Figure 36: Evaluating the His-tagged recombinant Gv protein with Western blot stained with anti-His-tag antibody. Gv protein is precipitated at different temperatures (RT, 4 °c or -20 °c). The purified protein at the three different temperatures shows a weak smeared band at 40 kDa (with the best signal at 4 °c) in addition to 26 kDa which is probably a degradation product or an alternative start of translation. MW; molecular weight.

Gnav1 was cloned into pOPIN-K vector with an N-terminal His-glutathione S-transferase (GST)- tag. The GST-tagged fusion protein folds rapidly into a stable and highly soluble protein upon translation, thus it would promote greater expression and solubility of Gv recombinant protein (Smith & Johnson, 1988). The tag sequence also includes 3C protease cleavage site for the easy removal of the His-GST-tag from the recombinant protein following affinity purification. The protein expression was induced by 0.2 mM IPTG at 18°C. In total 12 L of culture were used for the production of Gv protein. The supernatant was affinity-purified on HisTrap FF columns. Subsequently, the His-GST tag was removed by incubation with 3C protease and the liberated tag sequences were then removed by a second round of affinity purification with HisTrap FF columns (Figure 37, lanes 1 and 2). Thereafter Gv protein was purified with a final size exclusion chromatography (Figure 37, lanes 3 and 4). We could successfully produce a prominent amount of Gv recombinant protein with the expected molecular weight at 45 kDa (Figure 37, lanes 6 and 7), however the purified mixture also contains contaminant bacterial proteins at even much higher levels (e.g., an intense band at 72 kDa), presumably because the overall efficiency of Gv synthesis in this bacterial system is rather low. To overcome the bacterial protein contamination, we performed in-gel digestion for the specific band that corresponds to Gv protein. This turned out to be sufficient for the PRM assay.

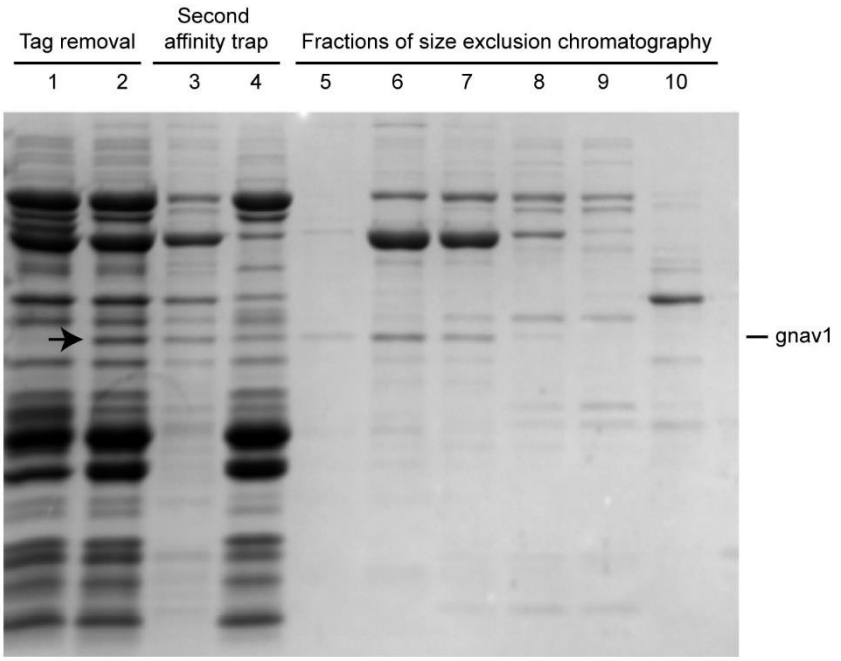


Figure 37: Coomassie blue staining shows the purification steps of the expressed Gv recombinant protein. g*nav1* expression was induced by 0.2 mM IPTG at 18°C. Supernatant produced from 12 L of culture was affinity-purified on HisTrap FF columns (lane 1); next it was incubated with 3C protease to remove the His-GST tag (lane 1 (before) and lane 2 (after)). Successful His-GST tag removal is indicated by the presence of new band after tag removal (black arrow in lane 2). Next, a second round of affinity purification was performed to remove the liberated tag, lane 3 (bound) and 4 (unbound). Next, a final size exclusion chromatography was performed to purify the Gv protein (lanes 5-10). The second fraction was used for further experiments since it has the highest amount of the recombinant protein.

3.4.3 Alteration of Gv protein levels in *gnav1*$^{-/-}$ adult kidney and 2 dpf embryos

Since MS was not sensitive enough for Gv detection we implemented PRM assay (see 3.4.1) using the generated Gv recombinant protein (see 3.4.2). This approach allows the detection of specific Gv peptides in a more sensitive and accurate manner. Thus, we will be able to evaluate whether Gv protein is absent in the mutant or whether a 5' truncated variant might be present, and if so, its relative abundance.

As a first step, the PRM assay was performed for the recombinant Gv protein in order to identify the Gv protein peptides obtained after trypsin digestion. This step serves to select proteotypic peptides that represent unique sequence regions in Gv, which then will be used to detect Gv in the heterogeneous sample. In order to enrich the

recombinant protein in the sample, SDS-PAGE was performed and the Gv band at around 45 kDa was excised from the gel (Figure 27, black arrow). Then, the sample was subjected to in-gel digestion prior to PRM assay. As a result, twenty-nine peptide sequences were detected by PRM, out of which eleven peptides correspond to Gv protein sequence (Table 3). The other peptides have lower intensities and presumably result from a contaminant bacterial protein. We performed MSA for Gv peptides with all other zebrafish G alpha protein sequences; we then selected five peptide sequences unique for Gv to be used for the calibration of the subsequent assays. A scheme for *gnav1* transcript in Figure 38A illustrates the position of each distinctive Gv peptide used in the assay design, also they are highlighted in Table 3 and Figure 38B. Overall, they are Gv specific and dispersed throughout the translated regions.

Table 3: Gv peptide sequences identified by first PRM assay performed for calibration.

#	UniProt ID	Gene names	Sequence	Start position	End position	Intensity (log_2)
1	F1QXX8	*gnav1*	ILLLGAAESGK	42	52	29.13
2	F1QXX8	*gnav1*	INLANPK	99	105	28.93
3	F1QXX8	*gnav1*	IIADDYMPTETDVLR	169	183	27.71
4	F1QXX8	*gnav1*	IIADDYMPTETDVLRVR	169	185	26.57
5	F1QXX8	*gnav1*	LRTTGVIETQFK	186	197	23.79
6	F1QXX8	*gnav1*	TTGVIETQFK	188	197	28.15
7	F1QXX8	*gnav1*	MYDVGGQR	205	212	26.67
8	F1QXX8	*gnav1*	GTSMILFMNK	268	277	23.15
9	F1QXX8	*gnav1*	IDLFQEK	278	284	25.82
10	F1QXX8	*gnav1*	HYLPQFR	294	300	26.73
11	F1QXX8	*gnav1*	ENLEAVSLL	355	363	21.21

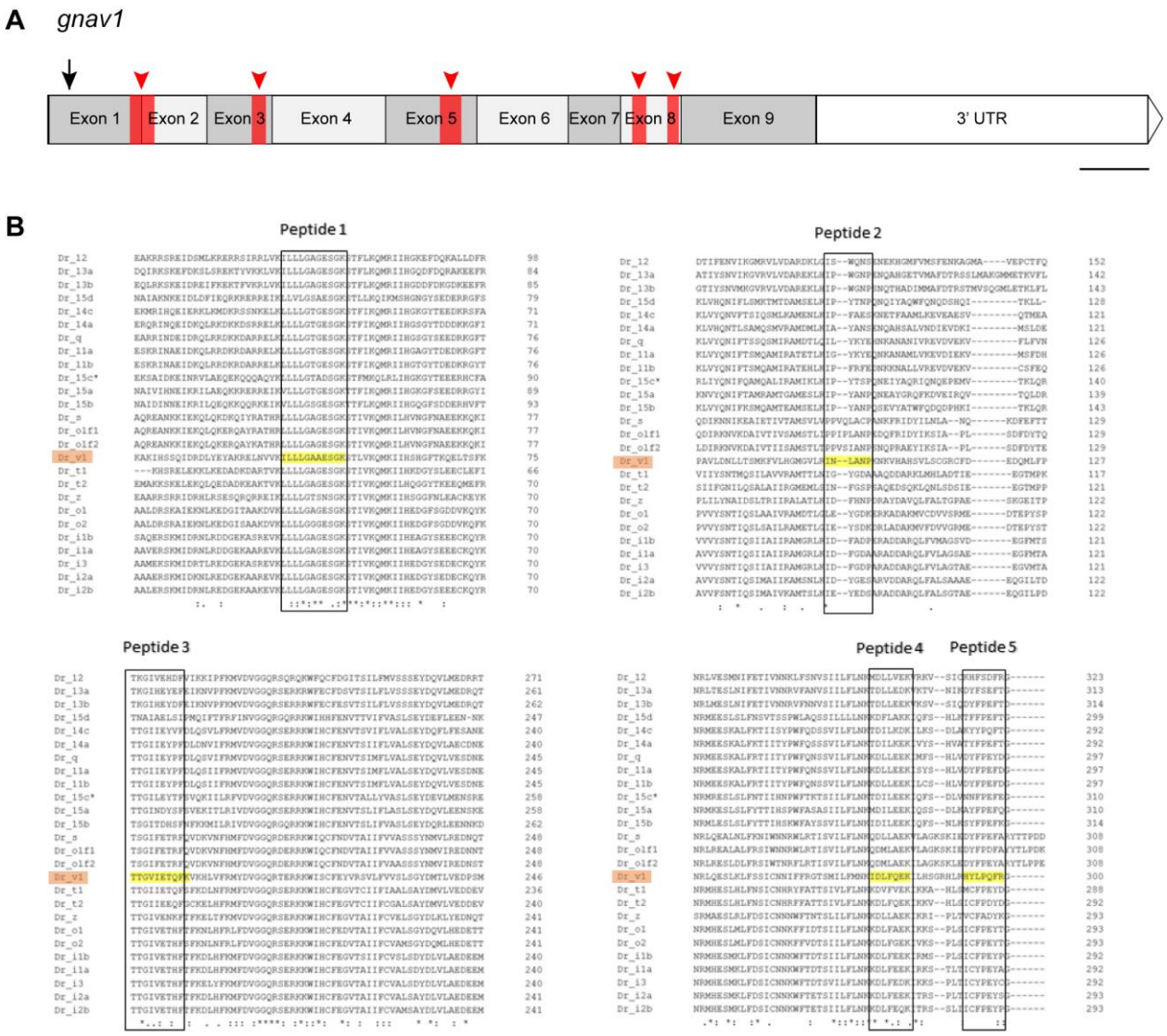


Figure 38: The distinctive proteotypic peptides of Gv protein selected for PRM assay to detect Gv protein in larvae and tissues lysate. (A) scheme illustrate *gnav1* mRNA composed of 9 exons, the red overlay and arrowheads indicate the position of the peptide fragments relative to their mRNA sequence. Note that peptide 1 spans exon1 and 2. Scale bar 100 bp. Black arrow indicates the mutation site. (B) MSA for all other zebrafish G alpha protein sequences. The Gv proteotypic peptides are overlayed with yellow color and the peptide name above the rectangle corresponds to Gv peptides; the black rectangles encompass Gv peptides used in PRM aligned to all other G alpha proteins.

In the second step, the PRM assay was applied to evaluate the levels of the selected peptides in adult kidney pools for wildtype (n=4) and mutant (n=3), and subsequently determine the relative Gv protein levels in these samples. We loaded an equal amount

of total protein lysate (50 ug) and performed SDS-PAGE. We excised the approximate molecular weight range from 40 to 50 kDa, followed by in-gel digestion. This molecular weight range most probably encompasses Gv - recombinant Gv runs at 45 kDa (Figure 27), and both potential 5' truncation and palmitoylation could lead to (apparent) molecular weight decrease. It should be noted that if occurring, a major truncation would not be picked up in the 40-50 kDa range.

The PRM data shows the presence of all five Gv peptides in the wildtype kidney (Figure 39). Unexpectedly, all five peptides are also present in the mutant sample. This suggests that the mutant can somehow circumvent the premature stop codon presumably by employing a secondary initiation site. The most N-terminal peptide fragment (peptide 1) corresponds to mRNA sequence that is located downstream the mutation site (Figure 38A), and it spans exon 1 and 2 borders. Since peptide 1 is similarly frequent as the other peptides in the mutant, the secondary initiation site should be located in exon 1, upstream of peptide 1 and downstream of the mutation site. Transcripts of peptide 2 and 3 correspond to the third and fifth exons respectively; peptides 4 and 5 represent exon 8 (Figure 38A). Thus, the truncated Gv sequence in *gnav1*$^{-/-}$ mutant kidneys appears to encompass a partial exon1 and could even be C-terminally complete (Gv has 9 exons, but no suitable peptide for exon 9 was found). However, we observe that the intensity of Gv peptides is significantly reduced in the mutant kidneys compared to wildtype, suggesting decreased levels of truncated Gv protein in the mutants compared to the levels of intact Gv in wildtype siblings (Figure 39). This could be caused by a lower efficacy of the alternative initiation site, which is sometimes observed, *cf.* (Neu-Yilik et al., 2011).

We then wished to examine whether the same truncated Gv protein was also present in mutant embryos, in particular at the 2 dpf stage, where the Gv mRNA is significantly decreased compared to wildtype. So, we repeated the PRM assay and collected an equivalent number of 2 dpf embryos from wildtype and mutant siblings raised in artificial water. For each genotype, proteins were extracted from four biological replicates, each representing a pool of 50 embryos. The extraction was followed by SDS-PAGE and in-gel digestion for the respective bands. The samples were treated equally in all steps. Similar to the previous findings, all the selected peptides used in

the PRM assay are detected in both wildtype and mutant embryos (Figure 40). However, in contrast to adult kidney, we observe a significant increase in the intensity of the Gv peptides in 2 dpf mutants compared to wildtype siblings (Figure 40). This finding is unexpected because the Gv mRNA levels are significantly decreased in 2 dpf mutant embryos. Reasons for this discrepancy are unknown. While an increase in mutant Gv protein could suggest a compensatory mechanism for the absence of a functional Gv (see below), this could not explain the decrease in the mRNA levels.

Taken together, alternative initiation - made possible by the near escape of Gv mRNA from NMD - results in a truncated Gv protein in the zebrafish mutants, both in embryonic and adult stages. The expression of a large Gv segment encompassing part of exon 1 and extending at least up to exon 8 raises the question, whether this truncated Gv might be functional. However, part of the alpha N motif and the predicted sites for thio-palmitoylation and N-myristoylation are missing in the truncated protein, as they occur upstream of the mutation site (see 4.2). Therefore, the expressed Gv protein is most probably not functional in the mutants (see Discussion for detailed evaluation, 4.2). The considerable increase of Gv protein in the mutant embryo and a contrary decrease in the mutant adult kidney both reflect a change resulting from the mutation supporting the notion that the truncated Gv cannot functionally mimic the intact Gv.

Thus, it appeared feasible to analyze the phenotype of *gnav1*$^{-/-}$ mutants, despite the presence of the truncated Gv protein.

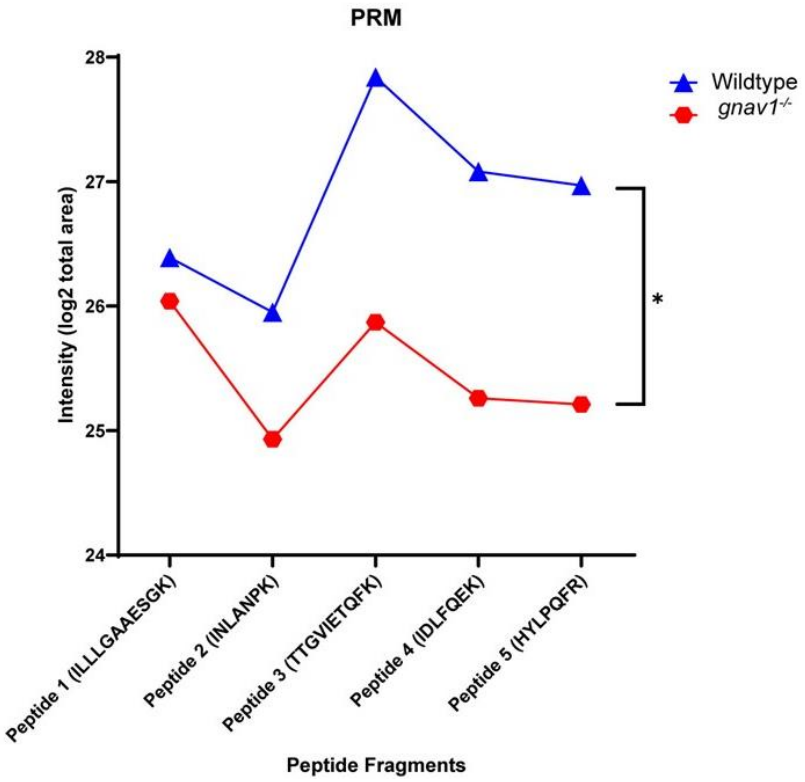


Figure 39: Gv protein is produced in the kidneys of the wildtype and mutant siblings. The figure depicts the intensity level of the five peptides measured by targeted-PRM MS assay, and subsequently represents the relative Gv protein levels. The intensity of Gv peptides is significantly reduced in the mutant kidneys compared to wildtype. Blue triangle and red circle represent kidney samples and each one represents a pool of (n=4) for the wildtype and (n=3) for the mutants, respectively. Equal amounts of total protein lysates (50 µg) from both genotypes were used for Gv measurements. The significance was determined with two-way ANOVA. *p = 0.01.

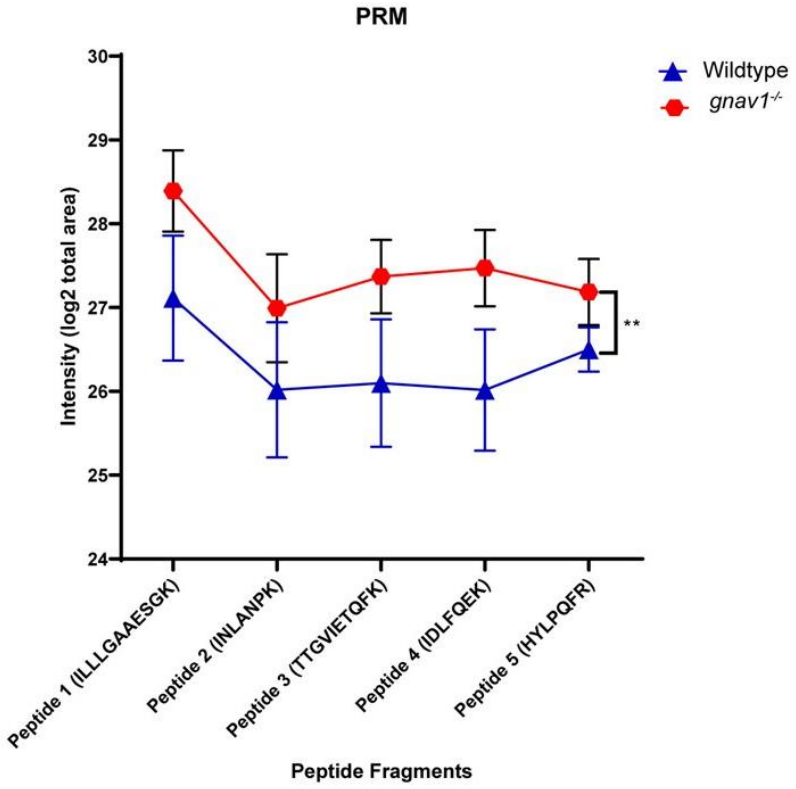


Figure 40: Gv protein expression in 2 dpf embryos from the progeny of the wildtype and mutant siblings. The figure depicts the mean intensity of the five peptides measured by targeted-PRM MS assay. Red circles and blue triangles represent the mean of four biological replicates, and each sample is a pool of 50 embryos from mutant and wildtype embryos, respectively. The intensity of Gv peptides is significantly increased in the mutant embryos compared to wildtype. Equal amounts of total protein lysates (50 µg) from both genotypes were used for Gv measurements. The significance was determined with two-way ANOVA. **p =0.0055. Error bars denote SEM.

3.5 Characterization of the mutant phenotype in larvae and adult zebrafish

In reference to our expression analysis data, we have shown that *gnav1* is expressed broadly in the gastrulation at 6 hpf and early segmentation at 12 and 24 hpf, and eventually becomes more restricted in 2-5 dpf embryos/larvae. Oka et al 2009 showed its expression in several adult tissues such as kidney, stomach and testes. Therefore, Gv might have a role in both embryonic development and adult kidney; hence we evaluated *gnav1*$^{-/-}$ phenotype in embryos/larvae and also in adults. During development we assessed several parameters including survival rate, hatching rate,

and craniofacial morphology. For the adult fish we performed body measurements and assessed oviposition.

3.5.1 Gv protein might not be crucial for embryonic survival

Here, we evaluated whether the alteration of Gv would affect the survival of the mutant zebrafish embryos. Thus, the wildtype and mutant were raised in the same conditions in artificial freshwater, which mimics the natural environment wherein the fish lives. The survival of the wildtype and mutants initially was monitored from 1 to 5 dpf. The majority of the mortality was observed in the first 24 hpf, both for wildtype and mutant, so in later experiments we evaluated the survival rate only at 1 dpf embryonic stage.

Due to time constraints, we began to evaluate embryos, which were obtained from non-sibling wildtype and mutant parents. We observed a pronounced and significant decrease in the viability of the mutant embryos compared to the wildtype (Figure 41A). Unexpectedly, this difference vanished, when the progeny from sibling wildtype and knockout was examined (Figure 41B). Since the latter experimental paradigm should constitute the most stringent control (mutants most similar to wildtype in all but *gnav1*), we conclude that Gv mutation does not influence the eggs viability. The deviant result obtained with non-sibling parents might result from strain differences, as the wildtype strain of the non-siblings was Ab/Tüb, whereas the mutants were generated from KS line and afterwards crossed out with Ab/Tüb. In contrast, the genotype of progeny from siblings should be more similar. Furthermore, the progeny from siblings was obtained after two additional crosses (outcross of homozygous, followed by incross of heterozygous fish), which could have further diluted any potential deleterious off-target effect of the CRISPR/Cas9 treatment. However, the viability of the mutant embryos was very similar between the first and the second set of experiments (Figure 41), which argues against the second possibility. Thus, the more likely explanation for the different wildtype viabilities are strain differences.

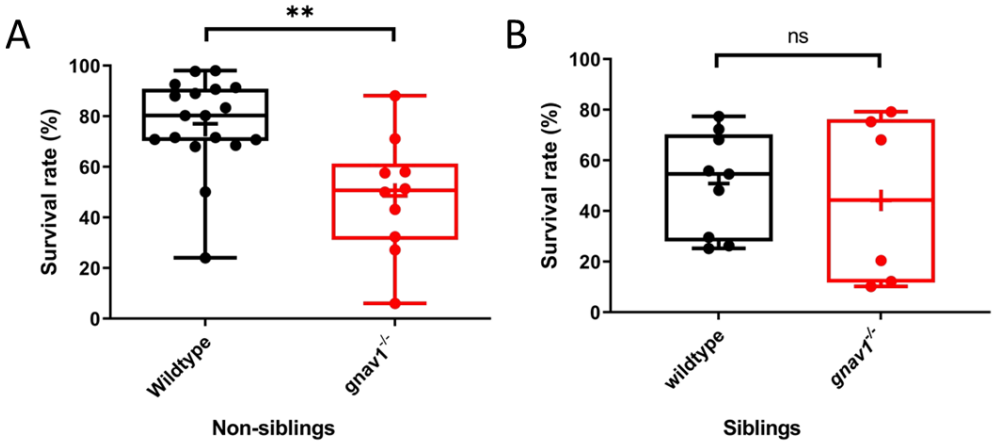


Figure 41: The survival rates of progenies from non-sibling and sibling parents of wildtype and mutants. Survival rate was measured at 1 dpf. (A) the data are obtained from progeny of non-sibling wildtype and mutant parents, and there is significant decrease in the survival rate of the mutant embryos compared to the wildtype. (B) data are obtained from the progeny of sibling wildtype and mutant parents, and there is no significant difference between wildtype and mutants' viability. The data are presented as box-and-whisker plot. Black circle represents the survival rate of embryos from one clutch of a wildtype couple, and red circle represents the survival rate of embryos from one clutch of a mutant couple. The significance was determined with two tailed t-test. **p = 0.001 to 0.01, and ns= not significant.

3.5.2 Mutant embryos hatch up to a day earlier than wildtype

When collecting samples at 2 dpf stage we observed an extremely high number of hatched mutant embryos, whereas the wildtype embryos mostly hatched on the third day, consistent with published observations (Nüsslein-Volhard & Dahm, 2002). Therefore, we quantified the hatching rate (the percentage of hatched eggs) at 2 dpf for the wildtype and mutant raised in E3 medium or artificial freshwater. E3 medium is fortified with the essential ions that are required during early development, including Na^{+}, Mg^{+2}, and Ca^{+2}, whereas the artificial freshwater has lower salt concentrations and mimics the natural environment wherein the fish lives.

The quantification showed that the *gnav1*$^{-/-}$ mutants hatch significantly faster compared to the progeny from wildtype siblings (Figure 42) and non-siblings as well (data not shown). This effect was observed both when embryos were raised in E3 medium and in artificial freshwater (Figure 42A, B). About 44% of mutants have already hatched at 2 dpf, compared to only 1.5% of the wildtype, when raised in E3 medium ($p<0.0001$) (Figure 42A). In freshwater the effect is somewhat smaller (24%

for mutants vs. 10% for wildtype, but still significant ($p<0.05$) (Figure 42B). Taken together these data open the possibility that Gv exerts an inhibitory effect on the hatching process. Noteworthy, Gv mRNA is expressed in the hatching gland (24-48 hpf, see section 3.1.2 and 3.1.3) which produces the hatching enzyme. Thus, Gv might have an inhibitory role in differentiation of the hatching gland and production of the hatching enzymes. It will be interesting to delineate the role of Gv in hatching in further experiments.

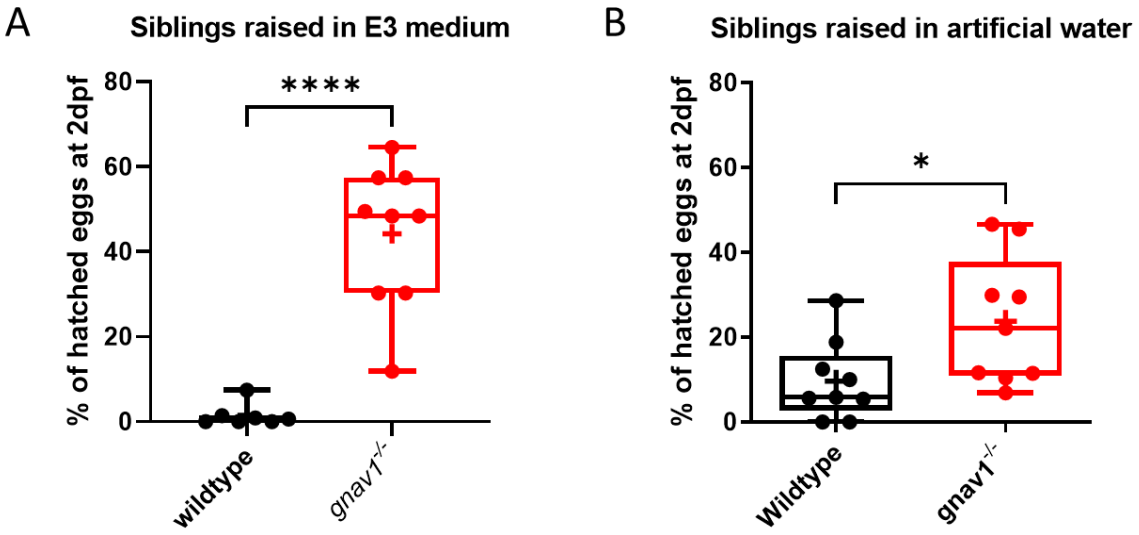


Figure 42: *gnav1*$^{-/-}$ mutants shows significantly higher hatching rate compared to the progeny from wildtype siblings. (A) the mutant and wildtype embryos were raised in E3 medium; the mutants show highly significant increase in hatching rate compared to the progeny from wildtype siblings in E3 medium. The hatching rate was measured for 56-57 hpf embryos. (B) the mutant and wildtype embryos were raised in artificial freshwater; the mutants also show significantly faster hatching compared to the progeny from wildtype siblings in artificial freshwater. The hatching rate was measured for 54-55 hpf embryos. The data are presented as box-and-whisker plot. Black circles represent the percentage of hatched embryos from one wildtype clutch, and red circles represent the percentage of hatched progeny from one mutant clutch. The significance was determined with two tailed t-test. *p = 0.01 to 0.05, and ****p<0.0001.

Interestingly, both mutant and wildtype hatching rates are different between E3 medium and fresh water, but in opposing directions. The mutants hatch earlier in E3 compared to freshwater, whereas wildtype tends to hatch earlier in freshwater. E3 medium has around 10 folds of Na^+ contents compared to freshwater. It has been reported that hatching is delayed with increased water salinity in zebrafish (Farhana et al., 2019; Sawant et al., 2001). This would fit with my observations for the wildtype.

The observation that the mutant behaves differently to changes in salinity further supports the notion that Gv is involved in the regulation of the hatching process.

Next, we examined the phenotype of subsequent stages in which Gv mRNA is expressed (3 and 5 dpf).

3.5.3 Abnormal pharyngeal morphology and overall fragility of *gnav1*$^{-/-}$ mutants

Since Gv mRNA is expressed in the craniofacial cartilage during development, it would be intriguing to investigate the effect of *gnav1* mutation on the craniofacial cartilage development. To stain the cartilage, we employed alcian blue, which binds to proteoglycans of epithelial and connective tissue (Adolphe et al., 1997) and thereby stains cartilage (and some other tissue). We evaluated the alcian blue staining beginning in 3 dpf larvae, the first stage in which all cartilages are clearly visible. Gv mRNA is expressed in the craniofacial cartilages at this stage. We imaged progeny from both wildtype and mutant siblings in the same field of view, to eliminate any distortion emerging from different exposure to light.

In 3 dpf larvae, the alcian blue staining shows that all craniofacial cartilages (ethmoid plate, Meckel's, palatoquadrates, ceratohyals, and ceratobranchial cartilages) are clearly visible in progeny from both wildtype and mutant siblings (Figure 43). The intensity of staining is a little less in the mutant, this difference will become more pronounced in 5 dpf larvae (see below). Moreover, we observed differences in the shape of the head cartilage between wildtype and mutants at 3 dpf. Both in the left-to-right and the dorsal-ventral axes the shape of the head cartilage is wider in the mutant, compared to wildtype.

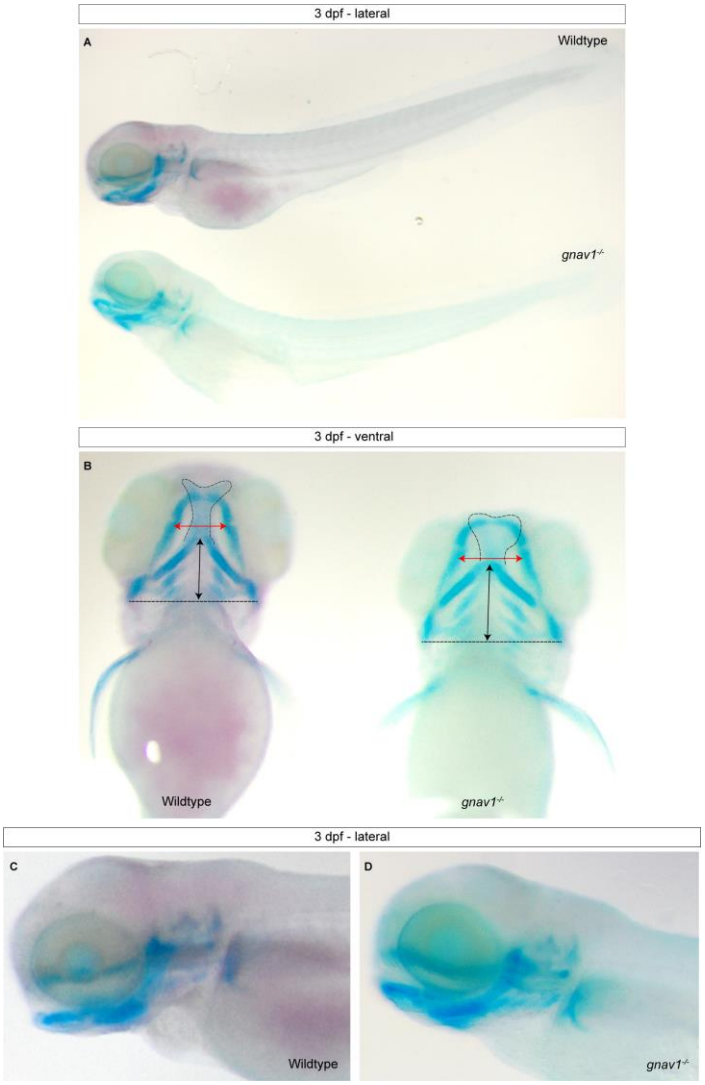


Figure 43: Alcian blue/ARS staining of 3 dpf wildtype and mutant larvae shows defect in the mutant craniofacial morphology. (A) observation from lateral perspective for the wildtype (top) and mutants (bottom) shows less intense alcian blue and ARS staining in the mutant head and yolk sac. (B) observation from ventral perspective for the wildtype (left) and mutant (right) shows differences in the facial cartilage morphology. (C and D) magnified lateral view for the wildtype and mutant embryos, respectively. The dashed black (straight) line is a reference line between the posterior edges of the hyosymplectic cartilage to the anterior edge of the ceratohyal cartilage; and the dashed irregular line marks the ethmoid plate. The double-headed arrow marks the ICD; and the black double-headed arrow marks the CCL (see Materials and Methods 2.5.5).

We then examined 5 dpf larvae, the stage where the larvae begin to feed independently. At this stage, the craniofacial cartilages and bones are well developed, albeit size and ossification increase in subsequent stages. Similar to 3 dpf larvae, we observe that the 5 dpf mutant head is widened in the dorsal-ventral and left-to-right axes (the distance between the eyes is clearly increased) (Figure 44). To quantify the

phenotype differences, we compared the ventral head structure and measured the craniofacial cartilage distance parameters for 5 dpf progeny from siblings; LJL, ICD, and CCL (according to the scheme in Figure 45A).

We found that the CCL distance - which represents the distance between the anterior tip of the ceratohyal and the reference line between the posterior edges of the hyosymplectic cartilages - is significantly shorter in the mutant larvae (Figure 45B). Also, the angle between the bilateral ceratohyals cartilages is increased, and they are bent towards ventral (Figure 44C). Thus, the angle between the lower jaw and the ceratohyal is increased resulting in a gap, hence the mutant head got widened in the dorsal-ventral axis. At the same time we observe from the lateral view that the basihyal cartilage (median position and also the anterior-most cartilage of the ventral hyoid arch) is protruding in the mutant whereas it is hindered by ceratohyal cartilage in the wildtype siblings (Figure 44B, asterisk).

We also observe that the most anterior part of the mutant head ICD is significantly wider than those of the wildtype siblings which means that the cranium is widened; the eyes appear further apart which is consistent with this notion (Figure 45C and Figure 44C). In addition, we observe that the ethmoid plate (upper jaw) appears shorter while the Meckel's cartilage (lower jaw) is protruding, as well as an increased gap between jaws in 5 dpf mutant compared to wildtype; both progenies are from siblings (Figure 44B, arrow). The overall defects eventuate in the gaping jaw phenotype described by (Neuhauss et al., 1996) for another mutant.

Furthermore, we also noticed that the processed mutant larvae (fixed in 4% PFA) for alcian blue staining were fragile and prone to damage easily, particularly the eye and the head area, in contrast to the wildtype (Figure 46). Rigorous care was required while handling 5 dpf mutant samples. In addition, *gnav1*$^{-/-}$ mutants show overall reduction in alcian blue staining of cartilage matrix in all cartilaginous elements of 5 dpf and 3 dpf larvae as well *vs* wildtype (Figure 43 and Figure 44). Taken together this might reflect decreases in the abundance of the proteoglycans in the body and cartilage ECM which would negatively impact mechanical stability.

It has been reported (Mobasheri et al., 1998) that synthesis and degradation of cartilage components is influenced by ionic homeostasis, which indeed appears to be altered in the mutants (see 3.6). Furthermore, Gv mRNA is expressed early in the development (1-3 dpf) in all cartilages that do show defects in the mutants (ethmoid plate, palatoquadrates, Meckel's, ceratohyals, and hyosemplectic cartilages, see 3.1). Taking together all these observations, we suggest that Gv is involved in the morphogenesis and ionic homeostasis of the craniofacial cartilages.

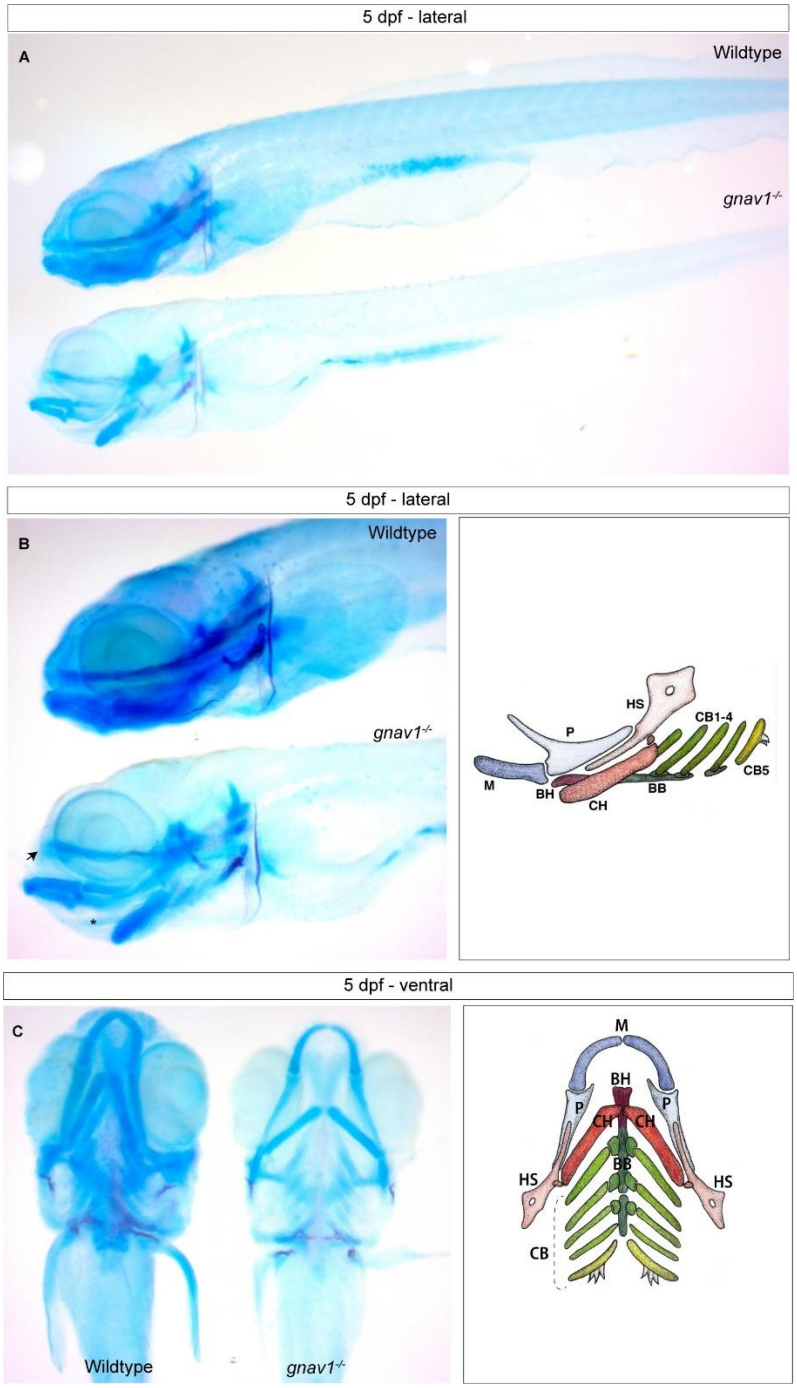


Figure 44: Alcian blue/ARS staining of 5 dpf wildtype and mutant larvae shows defect in the craniofacial cartilage parameters in mutant head. (A) observation from lateral perspective for the wildtype (top) and mutants (bottom) shows less intense alcian blue staining; and morphological defect in the mutant head. (B) magnified lateral view for the wildtype (top) and mutant (bottom) embryos; note the shorter ethmoid plate (black arrowhead) and the wide angle between the ceratohyal and the basihyal (black asterisk); also, the mutant head is widened in the dorsoventral axis. (C) observation from ventral perspective for the wildtype (left) and mutant (right) shows differences in the facial cartilage morphology and intensity of the cartilage and whole body alcian blue staining; note that mutant head is widened left-to-

right axes. The right column in (B and C) illustrate schematic drawings for the craniofacial cartilage in zebrafish modified from (Kimmel et al., 1998). black asterisk denoted basihyal; black arrow denotes ethmoid plate; BB: basibranchial; BH: basihyal; CB1–5: ceratobranchials 1–5; CH: ceratohyal; HS: hyosymplectic; M: Meckel's; P: palatoquadrate.

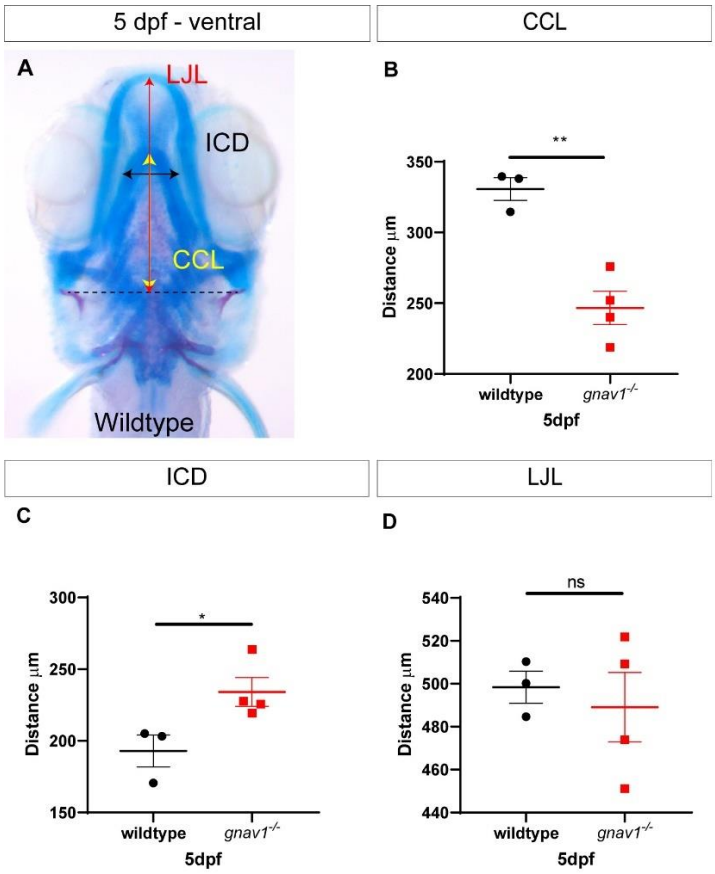


Figure 45: The craniofacial cartilage distance parameters (CCL and ICD) are significantly changed in the 5 dpf *gnav1*$^{-/-}$ larvae *vs* wildtype sibling larvae. (A) wildtype larvae stained with alcian blue/ARS; the figure depicts the distance parameters measured, i.e., LJL, ICD, and CCL as described in (Materials and Methods 2.5.5). (B) CCL is significantly shorter in the mutant larvae. (C) ICD is significantly wider in the mutant larvae. (D) the LJL is not changed in the mutant larvae *vs* wildtype. The dashed line is a reference line between the posterior edges of the hyosymplectic cartilage to the anterior edge of the ceratohyal cartilage; the red line with double arrowheads indicates LJL; and the yellow double-headed arrow hindered by the red line indicates CCL; and the black double-headed arrow indicates ICD (see Materials and Methods 2.5.5 for the measurements). The significance was determined with two tailed

unpaired t-test. *p = 0.01 to 0.05, **p= 0.001 to 0.01, and ns = not significant. Error bars denote SEM.

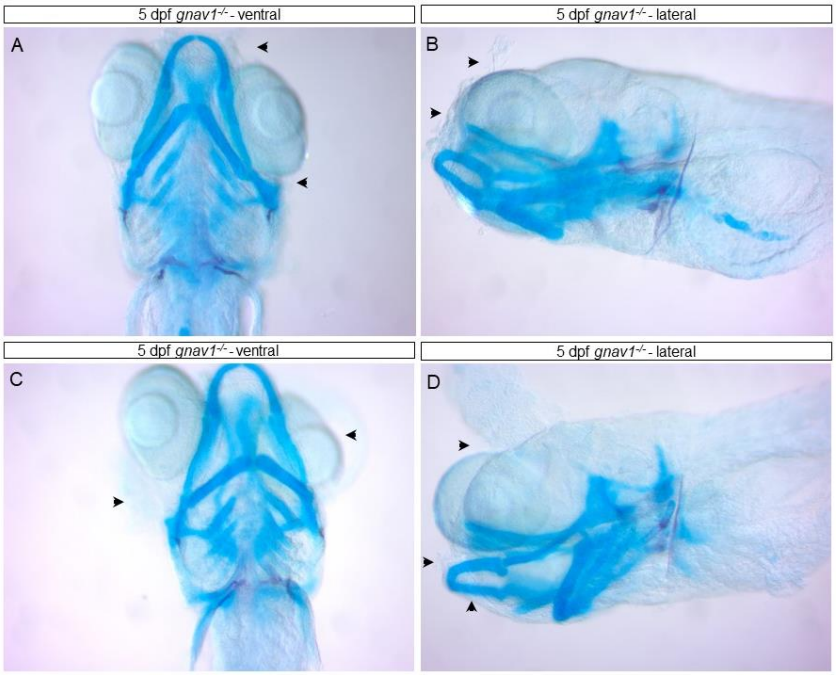


Figure 46: *gnav1*$^{-/-}$ 5 dpf larvae are fragile and prone to damage easily following alcian blue/ARS staining. (A and B) shows an example for the same mutant larvae that are fragile and were prone to damage; particularly the eye and the head. ventral (A) and lateral (B) views for the same larvae. (C and D) another example for fragile mutant; ventral (A) and lateral (B) views for the same larvae. The arrowheads denote damaged areas.

3.5.4 Oviposition is remarkably diminished by *gnav1*$^{-/-}$ mutation

The mutant embryos were able to survive until adulthood, thus it was possible to study the adult mutants' phenotype. Overall mutant fish appeared similar to wildtype in growth and behavior that were raised under identical conditions. To quantify the growth, we performed body measurements for wildtype and mutant male siblings; their ages were 15 months. We found no significant differences in the length and weight of the mutant and wildtype siblings (Figure 47A, B).

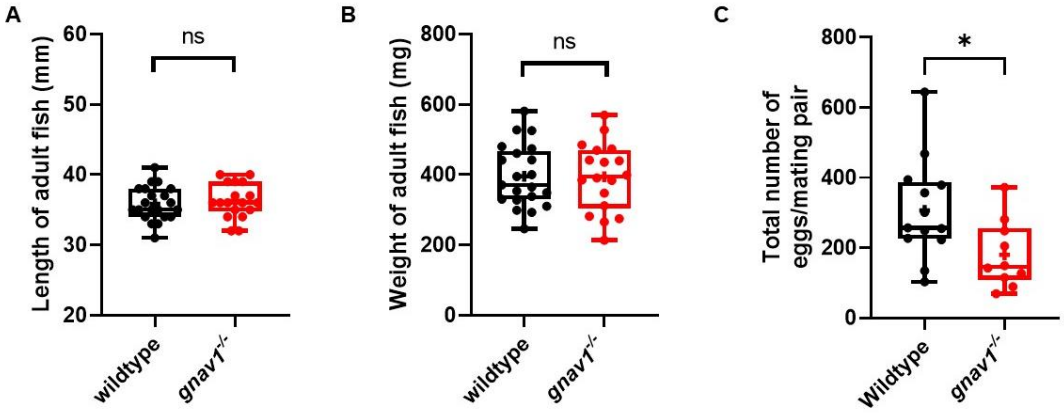


Figure 47: Oviposition is remarkably diminished in the mutants. (A) no significant difference is observed in the length between 15-month wildtype (black) and mutant (red) siblings. The length was measured from the tip of the mouth to the tip of the caudal fin. (B) no significant difference is observed in the weight between 15-month wildtype (black) and mutant (red) siblings. (C) The clutch size produced by the mutant siblings is significantly decreased. The data are presented as box-and-whisker plot, + is the mean. The significance was determined with two tailed t-test. *p = 0.01 to 0.05, and n= not significant.

However, while collecting specimens we observed an effect on the reproductive process. *gnav1*$^{-/-}$ mutants seemed to have a curtailment in egg production compared to wildtype. A single wildtype zebrafish female may produce clutches of several hundred eggs that could reach around 700 eggs in a single mating (Spence & Smith, 2006). To quantify the fecundity, we analyzed 1 year old siblings. In the early evening, single-pairs were separated by dividers in the crossing tank. Next morning, the dividers were removed and after 90 minutes the spawned eggs were counted, regardless of viability.

Intriguingly, we found that there is a significant difference in the oviposition of the mutants *vs*. wildtype siblings (Figure 47C). The clutch size produced by *gnav1*$^{-/-}$ mutants is markedly decreased, by 41 ± 17% (mean ± SEM) compared to their siblings. Egg maturation and ovulation are well studied and known to involve G protein signaling (Nagahama & Yamashita, 2008), thus it is conceivable that the reduced egg production results from a defect in the egg maturation process in *gnav1*$^{-/-}$ mutants.

3.6 *gnav1*$^{-/-}$ mutants show ionic homeostasis defects

At 5 dpf, Gv is mainly expressed in the osmoregulatory organs (kidney and gut), therefore we hypothesize that Gv may be involved in osmoregulation in this stage. This could be a cause of the observed cartilage defects, tissue fragility, and overall reduced alcian blue staining in the mutant (see 3.5.3), because cartilage metabolism is known to depend on ionic homeostasis (Mobasheri et al., 1998). Hence, we next evaluated whether there is an ionic homeostasis defect in *gnav1*$^{-/-}$ mutant.

In aquatic vertebrates such as zebrafish, the adult gills and ionocytes (skin of embryonic yolk sac) are the major sites involved in ionic homeostasis in addition to the kidney and intestine (Whittamore, 2012; Ying Jey et al., 2015). We previously showed that Gv mRNA is expressed in all these osmoregulatory organs (kidney, gut, skin (YSL)) in the wildtype embryos/larvae and in the adult kidney as well (see 3.1.1-3.1.4). Also, we confirmed the presence of Gv protein in wildtype embryos and adult kidneys (see 3.4.3). Moreover, we revealed that *gnav1*$^{-/-}$ mutants exhibit some defects that might be attributed to disturbance in ionic homeostasis (e.g., hatching, craniofacial defects). Therefore, we examined whether the mutant would show defects in ionic homeostasis. Firstly, we evaluated the total body cations of 5 dpf larvae by atomic absorption spectroscopy and Alizarin Red S (ARS) staining. Secondly, we explored the expression of iono-regulatory genes in the mutants by qPCR.

3.6.1 Total body cations are altered in *gnav1*$^{-/-}$ mutant larvae

A qualitative evaluation for the calcium contents was feasible by performing ARS staining concomitant with the previously described alcian blue staining. Thus, we performed ARS staining which detects the mineralized (mainly calcium) regions in larvae (Puchtler et al., 1969). The previously shown brightfield images of the ARS/alcian blue staining for 3 dpf larvae (Figure 43C) show mineralized areas in the skull, yolk sac, pronephros, and gut of the wildtype; whereas these regions show extremely reduced staining intensity in the mutant siblings (Figure 43D). We then obtained fluorescent images for 5 and 11 dpf progeny from mutant and wildtype siblings; 5 dpf siblings were adjacent in the same field of view for robust evaluation.

Consistently, ventral and lateral perspectives of 5 dpf mutant larvae show prominently reduced mineralization in the craniofacial bones (Figure 48). We also observe in 11 dpf mutant severe ossification malformation in craniofacial bones (Figure 49), and noteworthy, the absence of mineralization in the mutant vertebrae compared to wildtype (data not shown). This indicates that the abnormalities persist in 11 dpf mutant siblings.

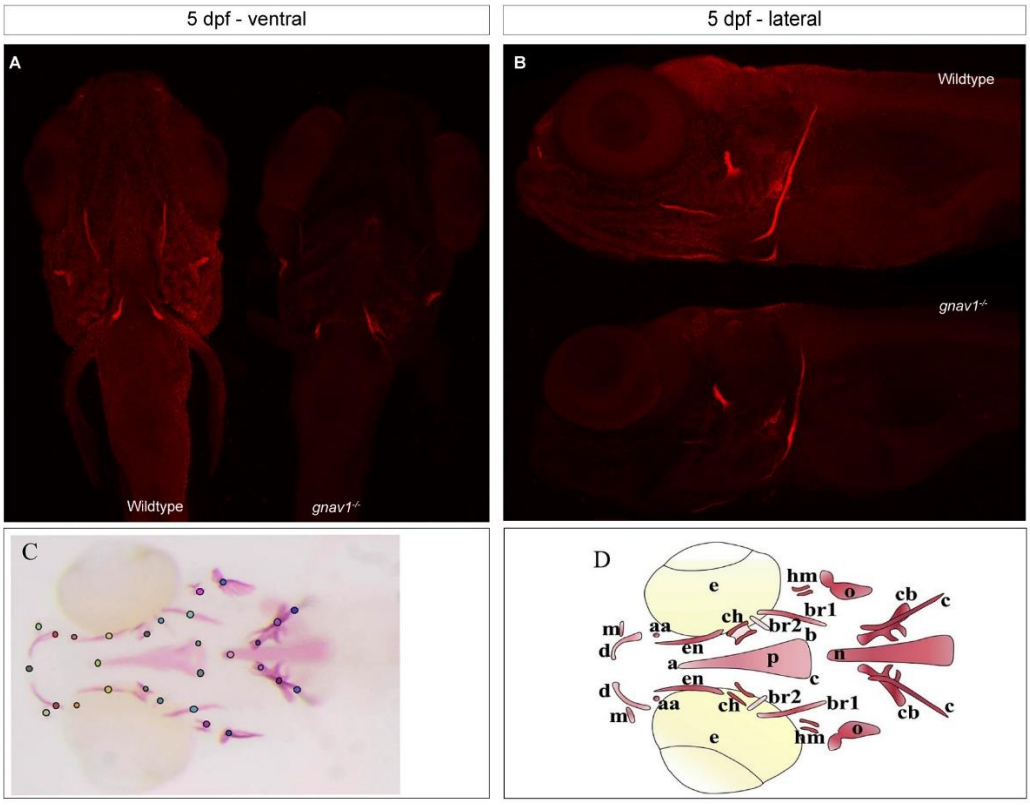


Figure 48: *gnav1*$^{-/-}$ 5 dpf larvae exhibit severe ossification malformation in craniofacial bones. (A) ventral view of maximum intensity projection confocal image stacks shows mineralized bone in red in wildtype (left) and mutant (right) sibling larvae. (B) lateral view for the wildtype (top) and mutant (bottom) shows bone ossification by ARS. (C and D) schematic representation of cranial bone elements in 10 dpf zebrafish larvae; the cranial bones in (D) are represented with the landmarks in (C) which is wildtype larva stained with ARS. Abbreviations; aa: anguloarticular, br: branchiostegal ray; en: entopterygoid; m: maxilla; n: notochord; o: opercle; p: parasphenoid; c: cleithrum; cb: ceratobranchial; ch: ceratohyal; d: dentary, hm: hyomandibular; a, b, and c: the parasphenoid is a triangular bone defined by its anterior summit (a) and two posterior summits (b,c); (C and D) Figures are modified from (Aceto et al., 2015).

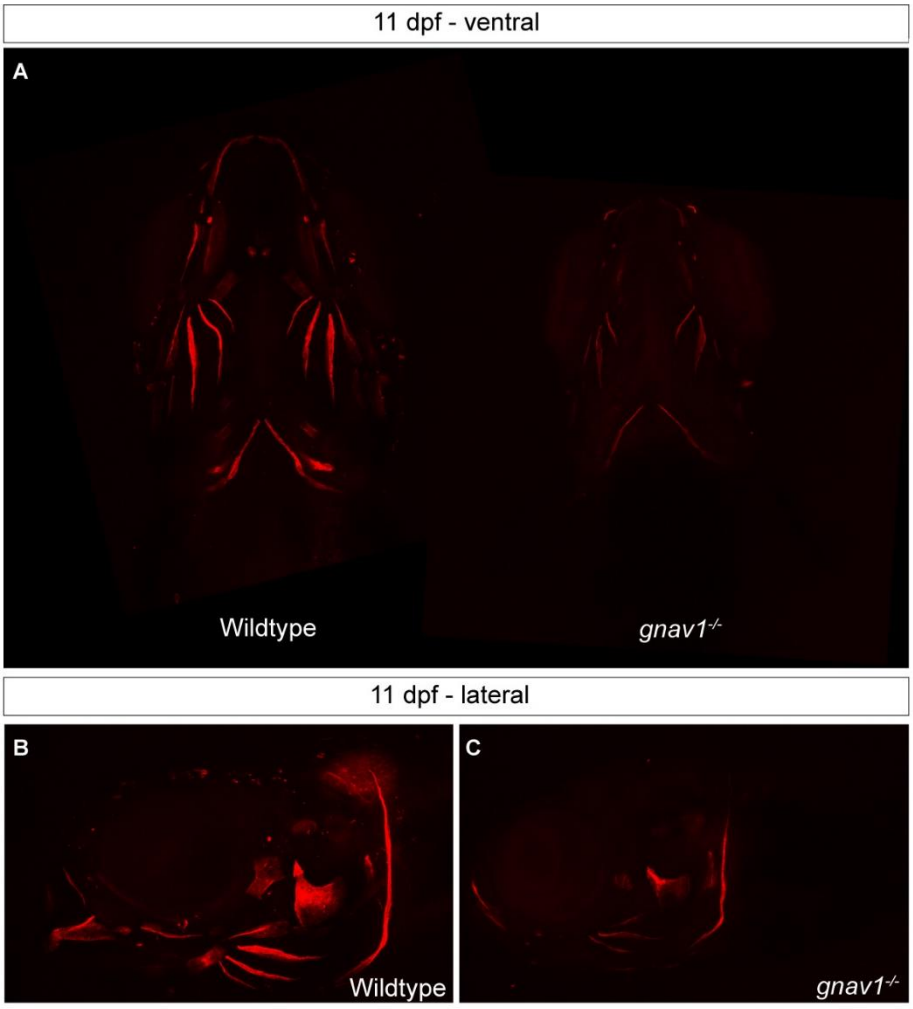


Figure 49: *gnav1*$^{-/-}$ 11 dpf larvae exhibit severe ossification malformation in craniofacial bones. (A) ventral view (anterior is top) of maximum intensity projection confocal image stacks shows mineralized bone in red in wildtype (left) and mutant (right) sibling larvae. (B and C) lateral view, anterior to the left, for the wildtype (B) and mutant (C) shows bone ossification by ARS. Note that bone elements in the mutant larvae are present and only the ossification is decreased. Concerning the overall body size, we observed individual variances in the size of both mutant and wildtype larvae which appears that it depends on feeding behavior; therefore, the size of the larvae in this Figure is not representative for all larvae, but the decreased ossification is reproducible in mutant larvae.

This finding is consistent with previous thesis findings that *gnav1*$^{-/-}$ mutant larvae have altered total body calcium levels (Ivandic, 2015). Therefore, we want to further examine whether the other body cations are also influenced. Atomic absorption spectroscopy was implemented to measure the total body cations (calcium, magnesium, sodium, and potassium); for the same number (n=25) of 5 dpf progeny from wildtype and mutant siblings acclimatized in artificial water. We also measured their wet and dry weights in order to assess their development rates. Both wet and dry

weights are *not* significantly different between both siblings, suggesting that the total volume and body weight are probably not altered in mutant larvae (Figure 50A, B).

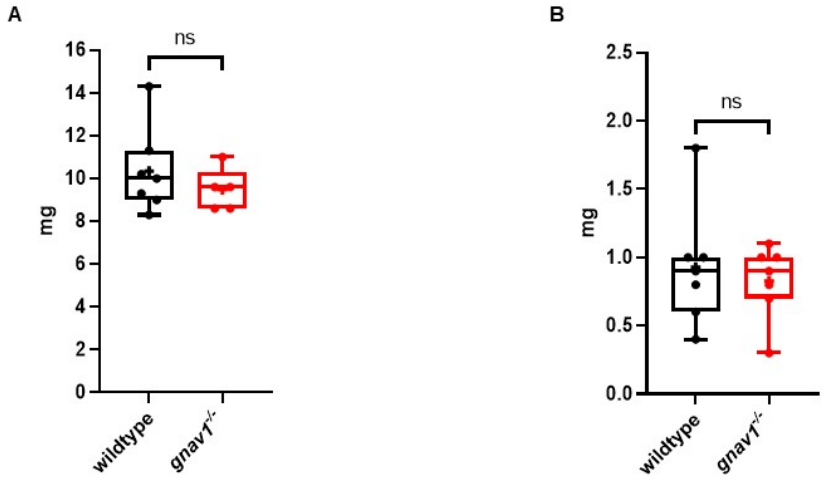


Figure 50: The total volume and body weight are not altered in mutant larvae. (A) the wet weight (before drying the larvae) was measured for the wildtype (black) and the mutant (red) larvae produced from siblings; there is no significant difference observed. (B) the dry weight was measured after drying the larvae overnight at 50 °C; there is no significant difference in the dry weight between the wildtype (black) and mutant (red) larvae produced from siblings. The data are presented as box-and-whisker plot, + is the mean. The significance was determined with two tailed t-test. n= not significant.

Interestingly, the mutant larvae show significantly reduced levels ($p<0.0001$) of calcium compared to wildtype (Figure 51A). This is in agreement with earlier results from a former graduate student in the lab (Ivandic 2015). The mutant larvae also show strikingly reduced levels of magnesium and potassium ($p=0.0001$ and $p=0.0021$, respectively). Sodium levels also appear reduced in the mutants; however, this effect did not reach significance (Figure 51B, C, D). This drastic depletion in cation levels in the mutant larvae confirms the Gv involvement in ionic homeostasis, which was suggested above based on indirect observations (3.5.2-3.5.3).

The disturbed cation levels might result from enhanced excretion or reduced reabsorption in kidney and/or intestinal epithelia. Also, changes in transepithelial transport and tight junction efficiency might play a role, possibly generated by altered cell-cell contact due to the observed deficiencies in ECM generation; noteworthy Gv mRNA is expressed in the skin of adult zebrafish (Oka et al., 2009). In the freshwater teleost, as zebrafish, most divalent ions such as calcium and magnesium are absorbed

in the adult gills/embryonic skin ionocytes and reabsorbed in the proximal tubule, while minor secretion into prourine takes place (Kersten & Arjona, 2017; Takvam et al., 2021). Therefore, the decline of cation levels could indicate a defect in the cation uptake or reabsorption. In order to explore possible deficiencies in ion pumps and transporters, we next analyzed the differential expression of many genes involved in ionic regulation and transport in mutant and wildtype adult kidneys.

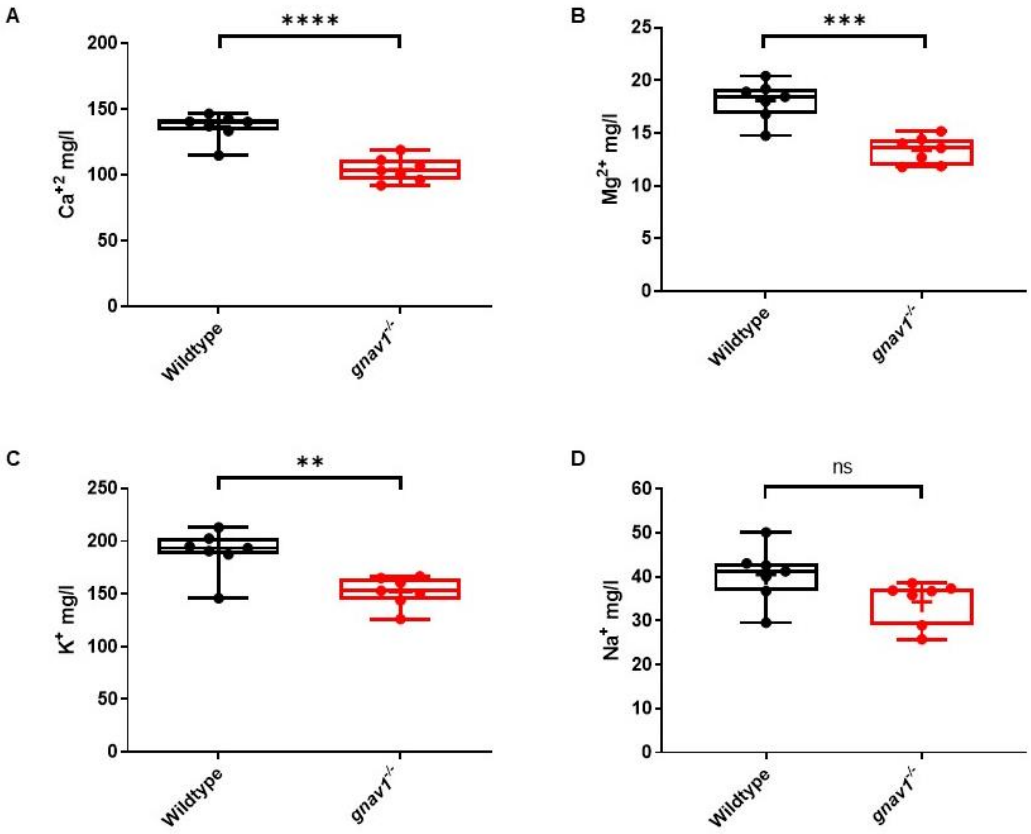


Figure 51: Total body cations are altered in *gnav1*$^{-/-}$ mutant larvae. The total body (A) calcium, (B) magnesium, (C) potassium, and (D) sodium levels were measured by atomic absorption spectrometry in seven biological replicates (each one is a pool of 25 larvae) of 5 dpf progeny from wildtype (black) and mutant (red) siblings acclimatized in artificial water. There is a significant decrease in calcium, magnesium, and potassium levels in the mutant larvae. The

significance was determined with two tailed unpaired t-test. ns = not significant, **p = 0.001 to 0.01, ***p 0.0001-0.001, ****<0.0001. Error bars denote SEM.

3.6.2 Alteration in iono-regulatory gene expression in the mutant adult kidney and larvae evaluated by qPCR

The ionic transportation in gills/skin ionocytes is similar to renal tubular cells, and many ion channels and transporters are coexpressed in both tissues (P.-P. Hwang & Chou, 2013). Since of those tissues the adult kidney is the major source of Gv in the adult zebrafish, we herein investigated the differential expression of iono-regulatory genes in the mutant kidney and of *gnav1* itself by qPCR (Oka et al., 2009). GeNorm algorithm was used to identify the most stable reference genes and used four reference genes (*rpl8*, *elf1*, *b2m* and *bactin)* for most stringent normalization of the siblings' data, unless otherwise mentioned.

Initially, we examined the ratio of *gnav1* in the mutant versus wildtype kidney (non-siblings) using two sets of primers, which covers exons 1-4 and exons 6-9, respectively. The transcription level of *gnav1* seems moderately reduced in the mutant kidney (≈ 85% of the wildtype) (Figure 52A). This is consistent with the reduced Gv protein levels in the adult mutant kidney as determined by MS (3.4.3) and the near absence of NMD for the *gnav1* mRNA (3.2.2).

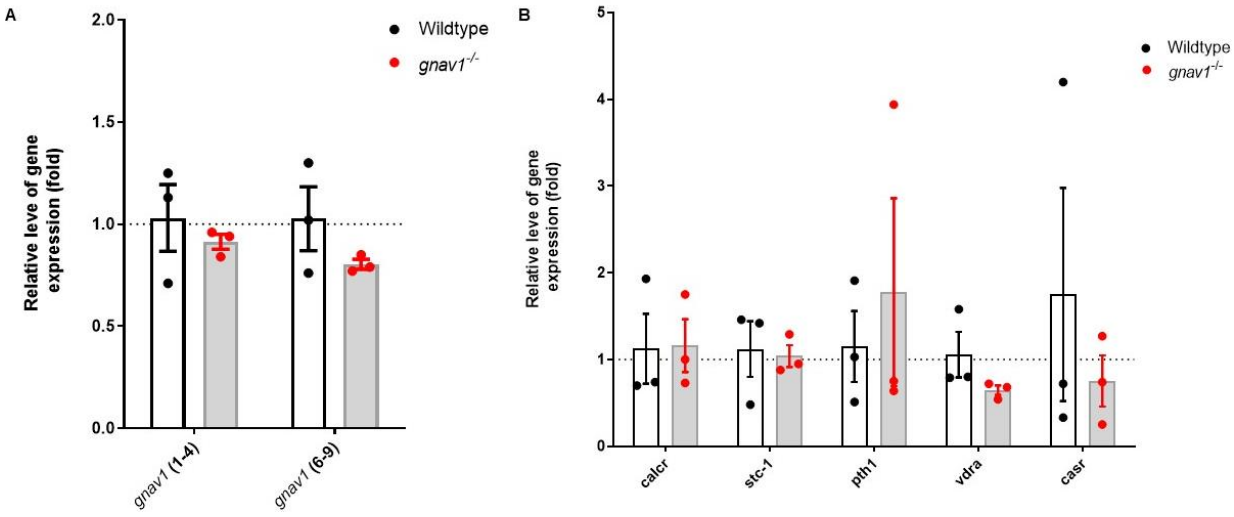


Figure 52: qPCR measurements for the relative expression levels of *gnav1* and genes involved in calcium homeostasis in kidneys show no significant difference between wildtype and mutant

RNA levels. (A) relative *gnav1* expression in mutant kidney was measured by qPCR using two set of primers which covers exons 1-4, and exons 6-9 respectively; and calculated relative to wildtype. No significant difference in *gnav1* expression levels is observed between wildtype and mutant kidneys. (B) the relative expression levels of genes involved in calcium homeostasis *calcr* (calcitonin receptor), *casr* (calcium sensing receptor), *pth1*(parathyroid hormone 1), *stc-1* (stanniocalcin1), and *vdra* (vitamin D receptor a) are not changed in the mutants. The relative ratio was normalized with *bactin* and *b2m* expression. Significance estimated by two tailed unpaired t-test. Error bars denote SEM.

Thereafter, calcium regulatory genes including GPCRs were assayed in the adult kidneys (non-siblings). No significant variation was observed in *stc-1*, *pth1*, *calcr*, *casr*, and *vdra* genes*,* albeit the latter appeared to show a small decrease (Figure 52B). The calcitonin receptor-like genes *calcrla* and *calcrlb*, which are involved in calcium and additionally in chloride homeostasis, do not show any change (Figure 53A) (Wang et al., 2016).

Another G protein gene, *gnai11a* and two phospholipase C effectors, *plcg1* and *plcb3* showed moderate albeit nonsignificant decrease suggesting that other G protein signaling pathways might be influenced in the mutant kidney (Figure 53B).

We then analyzed the expression of other iono-regulatory and iono-transporter genes in the siblings' adult kidneys that are also expressed in the ionocytes. In zebrafish, at least five types of ionocytes expressing different sets of ion transporters have been identified: H^+-ATPase-rich (HR), Na^+/K^+-ATPase-rich (NaR), Na^+/Cl^- cotransporter (NCC), K^+-secreting (KS), and solute carrier family 26 (SLC26)-expressing cells (Ying Jey et al., 2015). These ionocytes perform trans-epithelial H^+ secretion/Na^+ uptake/NH_4^+ excretion, Ca_2^+ uptake, Na^+/Cl^- uptake, K^+ secretion, and Cl^- uptake/HCO_3^- secretion, respectively.

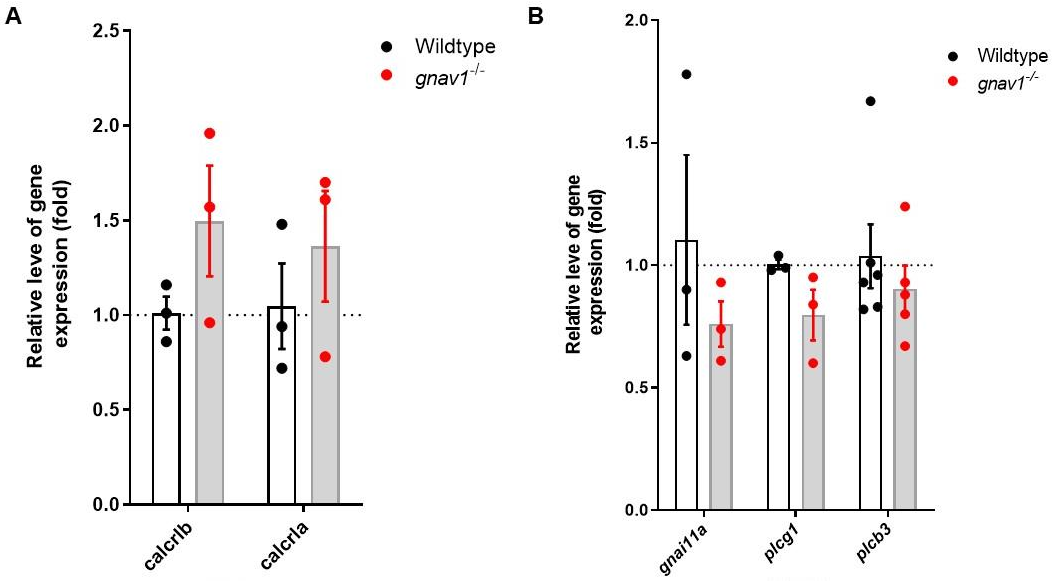


Figure 53: qPCR measurements for the relative expression level of calcitonin like receptors and genes involved in G protein signaling in kidneys show no significant difference between wildtype and mutant levels. (A) the relative expression level of *calcrlb* (calcitonin like receptor b), *calcrla* (calcitonin like receptor a). (B) the relative expression levels of *gnai11a* (G-alpha q), *plcg1* (phospholipase C-gamma 1), and *plcb1* (phospholipase C-betta 1). The relative ratio was normalized with *bactin* and *b2m* expression. Significance estimated by two tailed unpaired t-test. Error bars denote SEM.

As a result, on one hand, some genes do not show altered expression; such as *ca2* and *slc9a3.2* (present in HR); *ecac*, *ncx1b*, and *atp1a1a.1* (present in NaR); *atp1a1a.2* (present in NCC); and *atp1a1a.4* (present in KS) (Figure 54A). On the other hand, *slc12a3* (NCC) and *atp1a1a.5* (present in HR) are significantly increased whereas *scl26a4* (SLC26-expressing cells marker) is significantly decreased in the mutant siblings (Figure 54B and Figure 55A, B).

The NCC (*slc12a3*) is a sodium-chloride cotransporter involved in calcium and magnesium homeostasis (Arjona et al., 2019; Dong et al., 2020). Since we showed decreased magnesium and calcium levels in the mutant larvae (3.6.1), a corresponding decrease in adults could result in compensatory increased expression of *slc12a3* in *gnav1*$^{-/-}$ mutants. Since both NKA.5 (*atp1a1a.5*) and NCC (*slc12a3*) mechanism of action involve sodium reabsorption and potassium secretion (only NKA.5), their enhanced expression leads to an increase in sodium reabsorption and a (weaker) increase in potassium secretion. Since sodium levels are close to normal in the mutants, (3.6.1), we hypothesize that the absence of Gv leads to loss of sodium

by so far unknown mechanisms, which is compensated by the increase in sodium reabsorption due to the observed upregulation of NKA.5 and NCC. On the other hand, potassium levels *are* decreased in the mutant (3.6.1), which could be a direct effect of the increased potassium secretion caused by NKA.5 upregulation.

NCC function also involves chloride reabsorption thereby resulting in increased chloride contents. We did not measure chloride content directly, but the significant decrease in *Scl26a4* ($Cl^-/HCO3^-$ exchanger) expression in mutant larvae could be a compensatory response to decrease the chloride overload.

The largest changes in mRNA levels in adult kidney were seen for NKA.5 ($p<0.01$, (Figure 55A) and NKA.2, albeit the latter was not significant (Figure 54A). We therefore wished to see whether similar changes could be visible in larvae, i.e., in the cumulative effect of all iono-regulatory tissues. We examined NKA.5 and NKA.*2* levels in 3 dpf progeny from mutant and wildtype siblings. Again, NKA.2 was not significantly different in the mutant (Figure 56B). In contrast, NKA.5 is significantly *decreased* in 3 dpf mutants (Figure 56A). This is the opposite effect compared to the adult kidney (increase in NKA.5). Possibly the effect in the larval kidney (pronephros) is overruled by the opposite regulation in the other osmoregulatory organs (gills, skin), or alternatively, the NKA.5 might be regulated differently in the larvae.

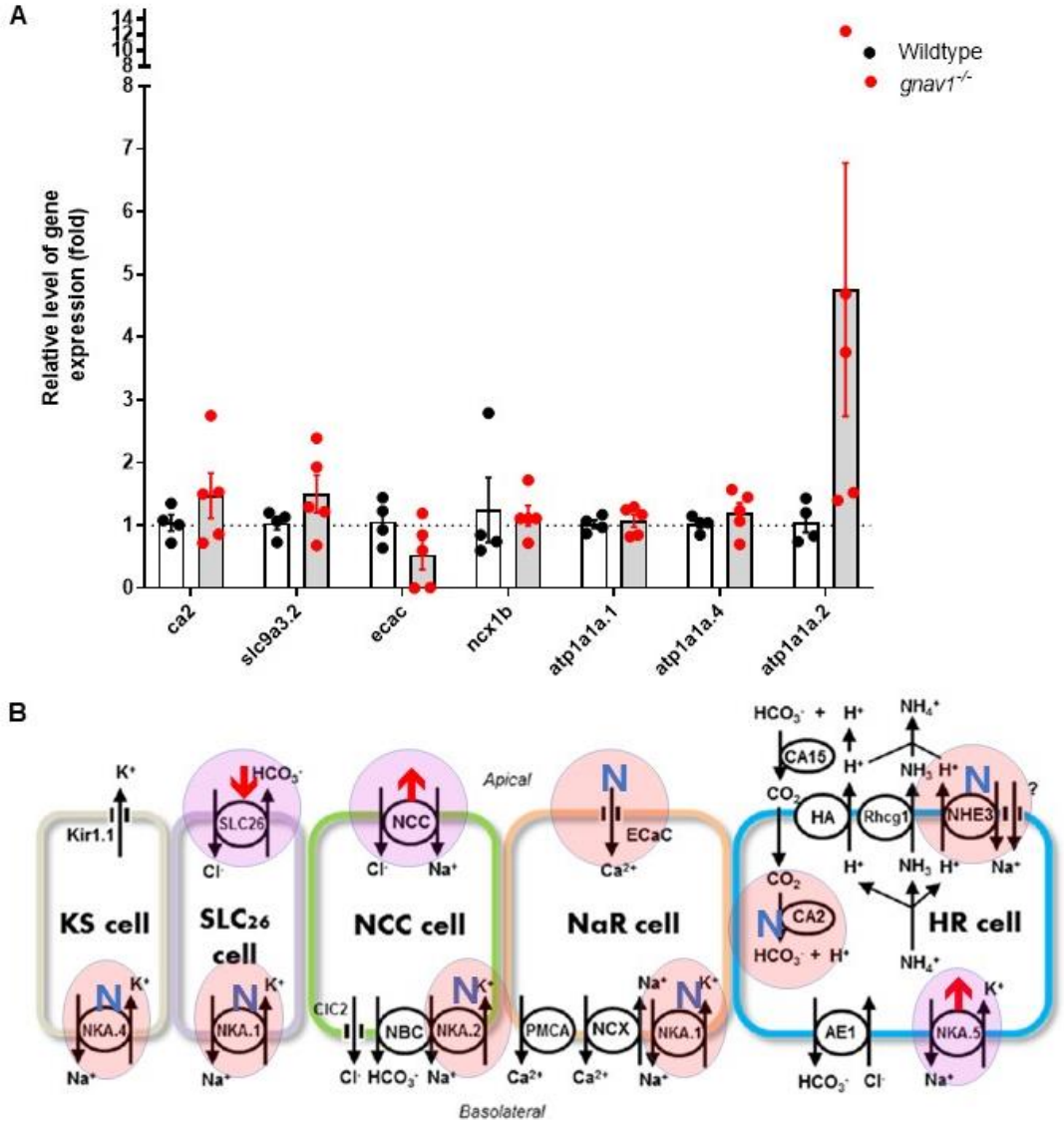


Figure 54: qPCR measurements for the relative expression level of iono-regulatory and iono-transporter genes in the sibling's adult kidneys that are also expressed in the ionocytes. (A) the genes *ca2* (carbonic anhydrase 2), *slc9a3.2* (solute carrier family 9, subfamily A, member 3, tandem duplicate 2), *ecac* (epithelial calcium channel), *ncx1b2* (solute carrier family 8-member 1b), *atp1a1a.1* (ATPase Na^+/K^+ transporting subunit alpha 1a, tandem duplicate 1), *atp1a1a.4*, and *atp1a1a.2* do not show altered expression in the mutant kidneys. (B) schematic representation of ionocytes and ion transport pathways in zebrafish modified from (Ying Jey et al., 2015). The genes corresponding to iono-transporters or pumps (in the scheme) examined in this study are marked by the purple and orange overlays. The orange overlays with N represent genes whose expression is not altered in *gnav1*$^{-/-}$ mutant while the purple overlays denote the genes whose expression was altered in *gnav1*$^{-/-}$ mutant kidneys; the red arrows indicate the decrease (down) or increase (up). The relative ratio was normalized with *rpl8*, *elf1*, *bactin* and *b2m* expression. Significance estimated by two tailed unpaired t-test.

Error bars denote SEM.N: not changed. For abbreviations the reader is referred to the list of abbreviations.

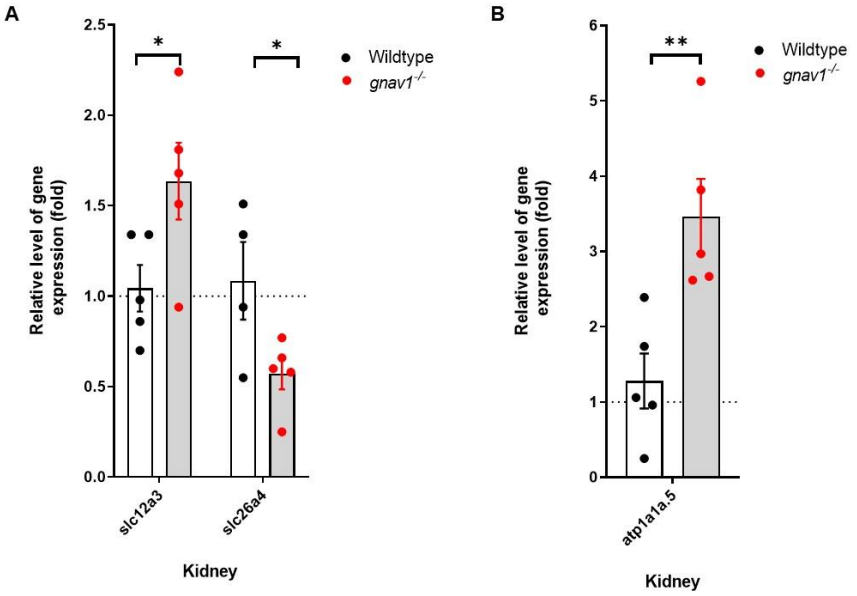


Figure 55: The relative expression level of *slc26a4* (encodes SLC26), *slc12a3* (encodes NCC), and *atp1a1a.5* (encodes NKA.5) are significantly different in mutant *vs* wildtype sibling kidneys. (A) the relative expression level of *slc12a3* (solute carrier family 12 (Sodium/Chloride Transporter), subfamily A, Member 3) and *slc26a4* (solute carrier family 26, subfamily A, member 4). (B) the relative expression level of *atp1a1a.5* (ATPase Na^+/K^+ transporting subunit alpha 1a, tandem duplicate 5). The relative ratio was normalized with *rpl8*, *elf1*, *bactin* and *b2m* expression. Significance estimated by two tailed unpaired t-test. *p = 0.01 to 0.05, **p = 0.001 to 0.01. Error bars denote SEM.

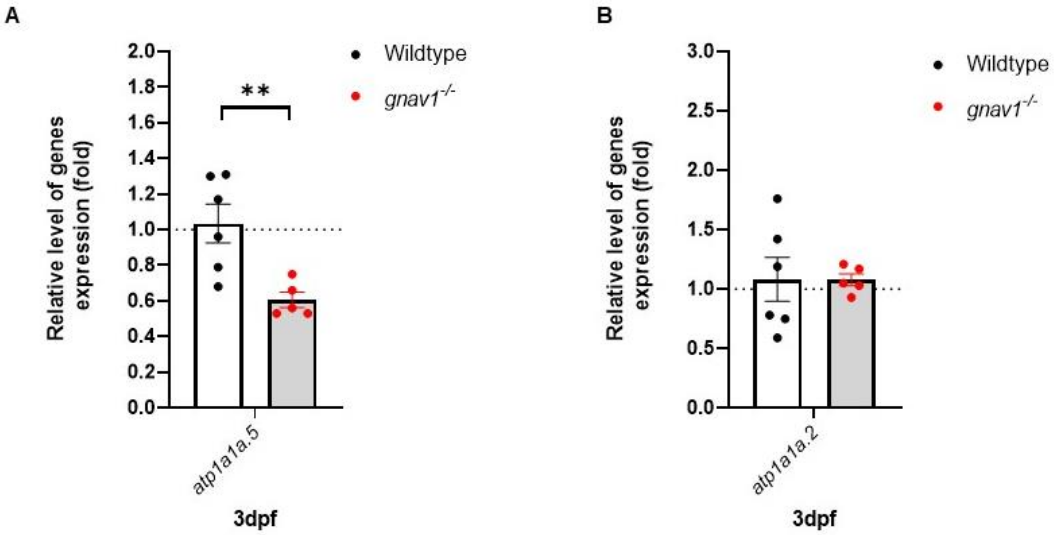


Figure 56: The relative expression level of *atp1a1a.5* (encodes NKA.5) is significantly different in 3 dpf mutant *vs* wildtype larvae from sibling couples, yet not *atp1a1a.2* (encodes NKA.2). The relative expression level of (A) *atp1a1a.5* and (B) *atp1a1a.2* in 3 dpf mutant and wildtype

larvae from sibling couples. The relative ratio was normalized with *rpl8*, *elf1*, *bactin* and *b2m* expression. Significance estimated by two tailed unpaired t-test. **p = 0.001 to 0.01. Error bars denote SEM.

To summarize the *gnav1*$^{-/-}$ mutant phenotype described in 3.5 to 3.6, the adult mutants have curtailment in oviposition; the mutant embryos increasingly hatch one day earlier than the wildtype; the larvae exhibit craniofacial cartilage defects and also ossification malformation in bones, and these defects could be direct or indirect effects of Gv mutation resulted from disruption in ionic homeostasis. Indeed, the mutant larvae have deficiency in cations, and altered expression of cation/anion transporter genes which corroborate our speculations that the impact of disturbed homeostasis resulted in the mutant phenotypes. Taken together all the results, we suggest that iono-regulation could be the predominant role for Gv in zebrafish adults and embryos/larvae.

3.7 Preliminary results for a novel knockout strategy

To eliminate the potential complication by the presence of truncated Gv protein in the current knockout (see 3.4.3) we plan to generate a new Gv knockout line besides the previously established line in the lab. Our new strategy is adapted and modified from (Hisano et al., 2015) to achieve a whole knockout for *gnav1* and simultaneous eGFP knockin as illustrated in (Figure 57). Hisano et al. could successfully achieve a precise knockin for the eGFP into keratin type 1 c19e (*krtt1c19e*) expressed in the keratinocytes. Additionally, in our new strategy we attempt to achieve whole gene knockout concurrent with the eGFP knockin. This new approach will be an important validation of results achieved with the first mutant line, and at the same time we will be able to trace the live eGFP expression in Gv-specific tissues thus circumventing the need for a specific Gv antibody (see 3.3). Although Gv protein levels were under the detection limits by conventional MS, a fluorescent signal in a sparse cell population might be easier to detect visually. Also, amplification of this signal by anti-eGFP antibody staining would be possible (Toth et al., 2007). If successful, our novel strategy would be generalizable and could be employed for other genes as well.

A donor vector was provided by Hisano et al., which contains an eGFP sequence followed by a polyadenylation (pA) signal to stabilize the RNA. The eGFP and pA

sequences were introduced between two eGFP-gRNA target sequences, in order to prevent the integration of the vector backbone into the gene locus. The eGFP sequence will be surrounded by short *gnav1* sequences (40bp) homologous to regions at both cut sites (homology arms) (Figure 57). We expect that the homology arms would be precisely integrated into the target locus by homologous recombination. When the donor vector, gRNAs and Cas9 protein will be co-injected into one cell-stage embryos, eGFP will be integrated between the homology sequences in *gnav1* resulting in in-frame integration of eGFP.

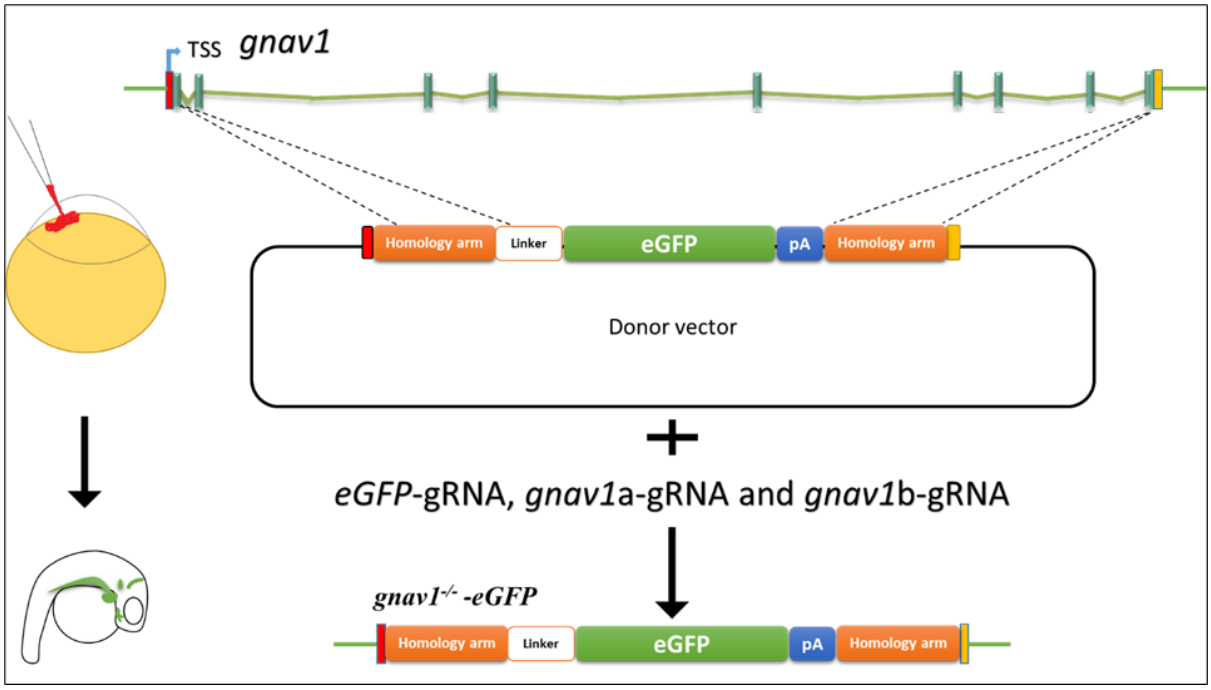


Figure 57: Strategy of *gnav1* deletion and precise integration of eGFP into the *gnav1* locus. A schematic representation of the *gnav1* locus and the donor vector consisting of eGFP-gRNA target sequences (red and yellow blocks), homology arms, eGFP and polyA (pA) signal. The *gnav1*a- and *gnav1*b-gRNAs will target the region after the transcription start site (TSS) and last exon (red and yellow block, respectively). Sequence from the first and last exons of *gnav1* will be inserted as homology arms (brown) flanking the eGFP+pA sequence in the donor vector. When the donor vector, gRNAs and Cas9 mRNA will be co-injected into 1 cell-stage embryos, eGFP will be integrated into the targeted genomic locus between the homology sequences in *gnav1*.

We have successfully generated the plasmid that harbors *gnav1* homologous arms and eGFP sequence (Figure 58) (see Materials and Methods 2.4). The first gRNA (*gnav1*a-gRNA), which targets the first exon, is already evaluated and has high efficiency (the knockout line available in the lab is generated with this gRNA). The

gRNA for exon 9 (*gnav1*b-gRNA) is designed using CHOPCHOP website tool predicting no off-target. The latter gRNA cuts just before the stop codon. The sequence surrounding the cut site was sequenced and validated. To test the efficiency of *gnav1*b-gRNA, one cell stage wildtype embryos were co-injected with gRNA and Cas9 protein. The survival rate for the embryos injected with the gRNA was high and nearly as good as control, indicating that the injection technique did not impair the survival of injected embryos (Table 4). Mutants were detected by genotyping; the amplified fragment of 650bp containing the target site was sequenced for randomly chosen embryos (Figure 59). Microinjection with *gnav1*b-gRNA successfully produced indel mutation in 3/4 of the embryos (Table 5), demonstrating high efficiency of this gRNA. Unfortunately, it was not possible to perform the subsequent steps (plasmid Co-injection with both gRNAs and Cas9 protein, genotyping, and raising the larva) within the time frame of this study.

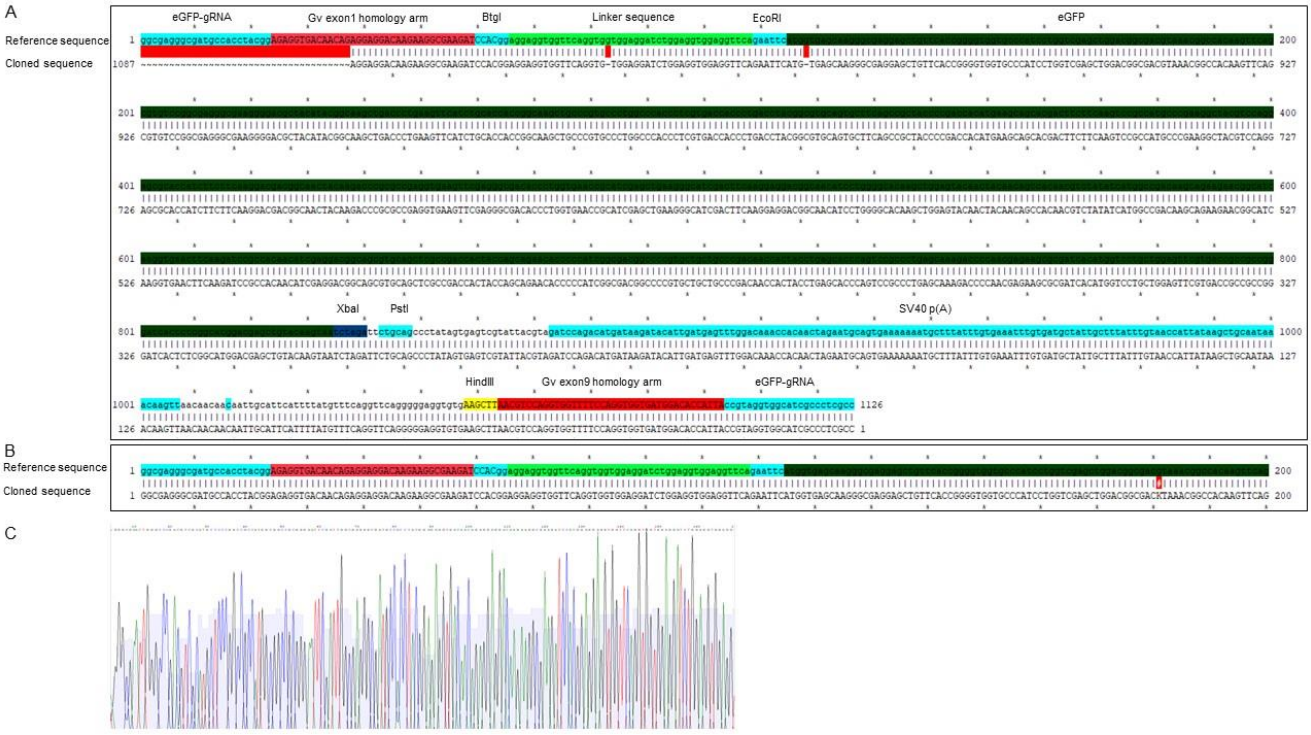


Figure 58: Integration cassette generated for CRISPR/Cas9 gene editing. (A) alignment for the generated plasmid with the reference sequence; sequencing was performed using the reverse primer. Note that the mismatches at the beginning of the sequence are not present in (B) in which the plasmid was sequenced with the forward primer and (C) shows the good

sequence quality of the chromatogram for the first 160 bp of the plasmid sequence shown in (B). color code is indicated. BtgI, EcoRO, XbaI, PstI, and HindIII indicate the cut site for the respective enzyme.

Table 4: Survival and death rates of embryos subjected to injections for CRISPR/Cas9 assay.

Type of injection	**% Dead**	**% alive**
Not injected (CTR)	4.0	96.0
Injected with PR (CTR)	8.4	91.6
Injected with *gnav1*b-gRNA	7.3	92.7

CTR, control; PR, phenol red

Table 5: Mutation frequency after injection with *gnav1*b-gRNA analyzed by forward and reverse sequencing (20 embryos).

% of wildtype	**% of homozygous mutant**	**% of heterozygous mutant**
21.4	21.4	57.1

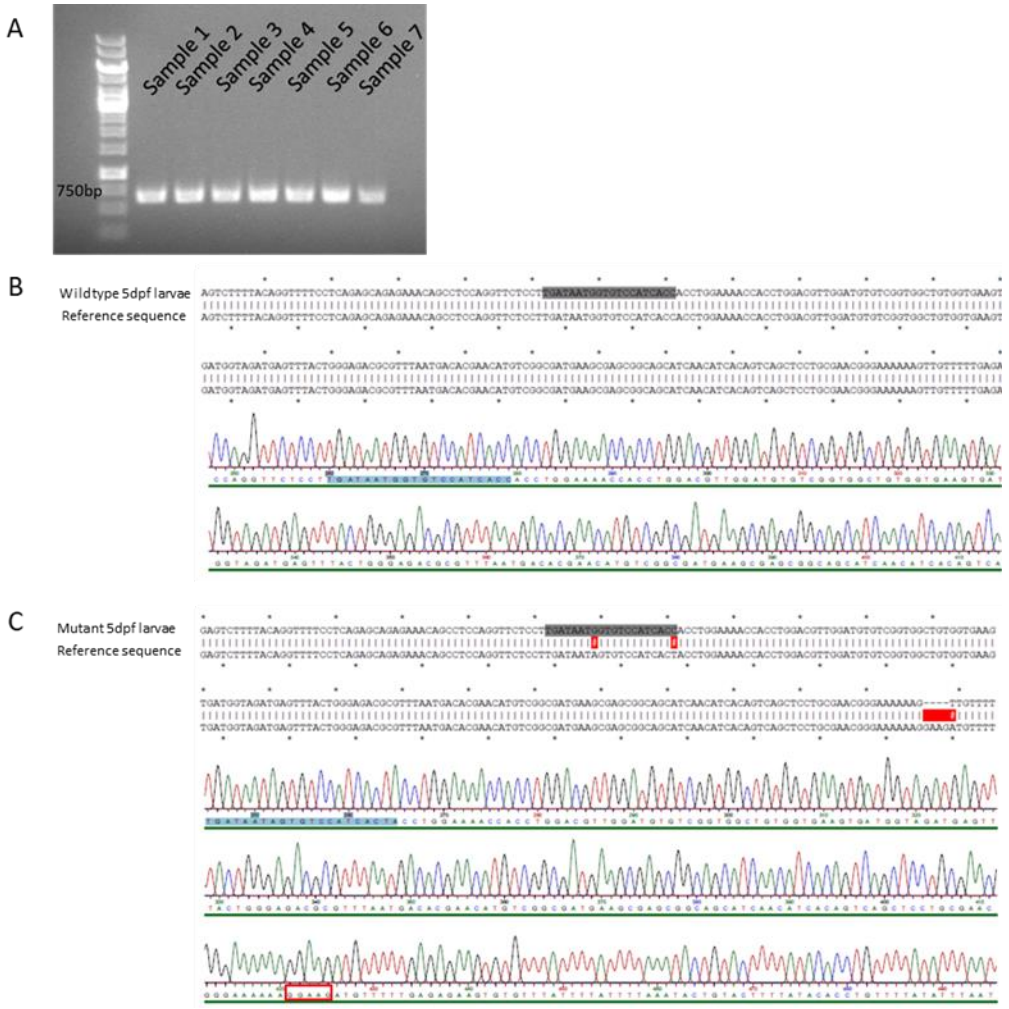


Figure 59: Detection of mutations in zebrafish embryos subjected to CRISPR/Cas9 using *gnav1*b-gRNA. (A) PCR amplicons for DNA extracted from injected embryos encompassing target region of *gnav1*b-gRNA. (B and C) representative example for non-mutant (B) and mutant (C) sequences aligned to *gnav1* reference sequence (wildtype) surrounding the gRNA target site, mutations were generated by CRISPR/Cas9. The shaded region is the gRNA sequence, and the red rectangles indicates the indel region).

3.8 A thorough phylogenetic analysis shows fractal pattern of gene loss for metazoan Gv as well as for the other four G alpha families

The fifth G alpha, Gv, is the most recently discovered member of G alpha proteins by a former lab worker (Oka et al., 2009). It was reported that vertebrates (including neoteleosts), arthropods, mollusks, and annelids possess Gv but that it has been lost in many lineages such as nematodes, fruit fly, jawless fish, and tetrapods. In 2009 the

number of available genomes was very limited and with that limited data set Gv appeared to have an unusually high number of gene loss events compared to the other four classes, Gs, Gi, Gq and G12. Interestingly, all the studied species possessing *gnav* genes are aquatic animals, except one. In particular, in insects, which are predominantly non-aquatic, only a single species was found to possess Gv, *Tribolium castaneum* (Oka et al., 2009).

We took the advantage of the plethora of new genomes to examine in detail whether Gv orthologs presence is correlated with aquatic living. For this search we focused not only on insects but extended the search to arthropods. We also re-investigated the evolutionary origin of Gv making use of the unicellular holozoan genomes (eg. Ichthyosporea, Corallothytrea, and Filasterea, etc.) that are currently available and allow to extend the evolutionary analysis beyond Metazoa. Moreover, we extended this systematic study to all G alpha families to delineate whether evolutionary dynamics of the Gv class are deviant from the other four G alpha classes in both Holozoa and more closely in Arthropoda. Hence, we searched the five G alpha families in the NCBI genomic databases rigorously in 873 species which led to the identification of 4479 *gna* genes (for the use of terms family and class see Materials and Methods 2.9.2). Herein, we only consider the major partial loss events that constitute over 60% of the species in the respected phyla.

The Holozoa clade comprises animals (Metazoa) and several unicellular lineages (Teretosporea, Filasterea, and Choanoflagellatea) (Figure 60). For the construction of the phylogenetic trees, we searched the genomic databases of selected species from animal phyla (Porifera, Placozoa, Chordata, Ctenophora, Cnidaria), and also animal-related eucaryotes (Ichthyosporea, Corallothytrea, Filasterea, and Choanoflagellatea) (Figure 60). This represents many major phylogenetic subdivisions in Holozoa, in particular early-diverging ones. The search in genomic databases also included all the subphyla of Arthropoda currently available, covering a wide evolutionary range from Chelicerata (spiders, scorpions, ticks, etc.) over Myriapoda (centipedes) and Crustacea (crabs, lobsters, crayfish, shrimp, etc.) to Hexapoda (dragonflies, damselflies, bugs, ants, beetles, butterflies, moth, flies, etc.). To compare the situation in arthropods to other protostomes, we used Ecdysozoa genomes (Nematoda and

Tardigrada) and one phylum of Lophotrochozoa (Mollusca), the sister clade to Ecdysozoa, as reference genomes. And from the deuterostomes we searched in the Ambulacraria which encompass two phyla, Echinodermata and Hemichordata.

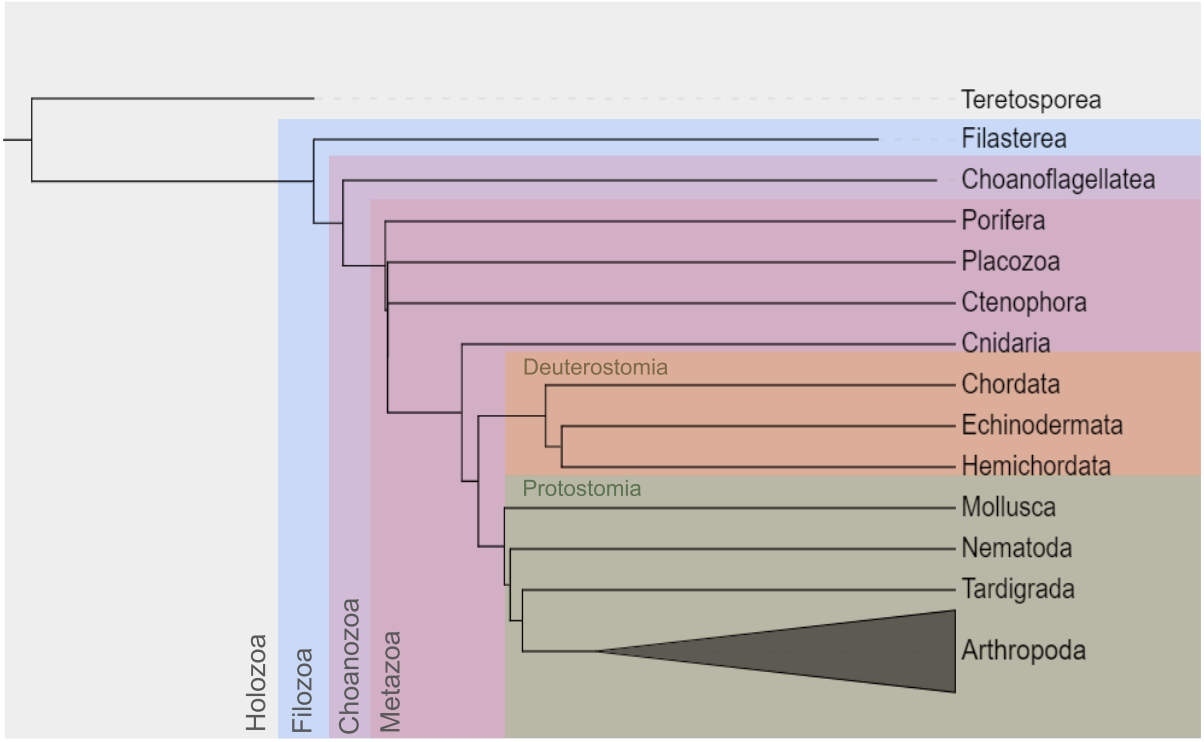


Figure 60: Species tree illustrates all phylogenetic branches investigated in this study. Arthropoda encompasses fifteen groups separately analyzed; Echinodermata and Hemichordata were analyzed together, they constitute the group of Ambulacraria. The tree was drawn using the web tool TimeTree by (Kumar et al., 2017).

3.8.1 The evolutionary origin of Gv is in the most recent common ancestor of holozoa, earlier than previously reported

The first analysis for the fifth class of G alpha families, Gv, demonstrated that it is present in sponges suggesting that it must be present in the most recent common ancestor (MRCA) of all animals (Oka et al., 2009). A follow-up study for the evolution of GPCRs signaling system also looked at the evolution of G alpha proteins (de Mendoza et al., 2014); they reported *gnav* genes in Filasterea (*Capsaspora owczarzaki*), choanoflagellates, Porifera, Ctenophora, but not in Ichthyosporea. This

seemed to suggest that *gnav* genes originated in the MRCA of the Filozoa, which encompass the Filasterea and all Choanozoa (de Mendoza et al., 2014). This interpretation was supported by two further phylogenetic studies with smaller species numbers, which looked at 27 and 65 species, respectively (Krishnan et al., 2015; Lokits et al., 2018). Both groups found *gnav* orthologs in Filasterea, which they named *GNAV*-like (Krishnan et al., 2015) and preGNAV (Lokits et al., 2018). However, both groups did not investigate earlier-diverging (non-filozoan) Holoza such as Corallochytrea and Ichthyosporea (both are called Teretosporea) (Figure 60). Nevertheless, all three studies (de Mendoza et al., 2014; Krishnan et al., 2015; Lokits et al., 2018) concluded that Gv originated in the common ancestor of Holozoa. But in fact de Mendoza and colleagues did not find *gnav* transcripts in *Ichthyosporea* (Teretosporea), analyzing RNAseq data. The two latter studies did not analyze Teretosporea which comprise earlier-diverging holozoa groups than Filasterea. However, all three studies included fungi, which constitute a sister group to Holozoa and found that fungi possess four distinct G alpha proteins (GPA1-4) which are not orthologous to any of the five G alpha classes of Holozoa (de Mendoza et al., 2014; Krishnan et al., 2015; Lokits et al., 2018). Together these data exclude an evolutionary origin of Gv in Opisthokonts (fungi and Holozoa), but the question was still unanswered, how early in Holozoa evolution Gv emerged.

Our study included genomes of Teretosporea, four species for *Ichthyosporea* and one for *Corallochytrea*, together with a Filasterea genome (*Capsaspora owczarzaki)* and two choanoflagellate genomes. We confirmed the presence of Gv in the filasterean genome (as a fragment) and choanoflagellate species (full length) (Figure 61). Furthermore, we report that *gnav* genes are present in all four Ichthyosporea species, but not in Corallochytrea (Figure 61). There is one *gnav* fragment in *Ichthyophonus hoferi*, *Sphaeroforma arctica*, *Sphaeroforma sirkka* and two *gnav* fragments, which may belong to a single *gnav* gene in *Pirum gemmata*. We consider it likely that the fragments are due to technical reasons, i.e., incompleteness of the corresponding genome assemblies. The coverage for the Filasterea *Capsaspora owczarzaki* genome is very low, only about tenfold (Suga et al., 2013), this could explain the presence of incomplete *gnav* gene sequence. Despite being fragments, the *gnav* genes of the four

holozoan and one filozoan species cluster basal within the Gv clade with maximal branch support in repeated phylogenetic tree analyses (Figure 61). Also, we identified *gnav* genes in each of the choanoflagellate *Monosiga brevicollis* and *Salpingoeca rosetta* with maximal branch support (Figure 61). Therefore, we can draw the novel conclusion that the *gnav* gene is born earlier than previously reported in the unicellular species and that its birth precedes the origin of Filozoa.

For comparison we investigated the presence of the other four G alpha families in these unicellular organisms. We found Gi orthologs in all the holozoan species including Teretosporea and Filasterea that are basal to both metazoan Gi and Go which together constitute the Gi family (Figure 61). Thus, the Gi family emerged in the common ancestor of holozoans, but the split into Gi and Go classes appears to have happened inside metazoan (animal) evolution.

We also examined the Gs family in the unicellular holozoan species and found Gs orthologs on the node basal to Gs and Gf classes (Figure 61) (for details on Gf see 3.8.4). We identified the presence of such Gs orthologs in Corallochytrea and two Ichthyosporea species (*Ichthyophonus hoferi* and *Sphaeroforma arctica*) and also in the Filasterean species (*Capsaspora owczarzaki*) (Figure 61). Hence, it appears that the Gs family originates from the MRCA of Holozoa.

When we analyzed the Gq/G12 origins in the earliest-diverging Holozoa we found a well-defined Gq clade (maximal branch support) to contain choanoflagellate and filasterean genes (*Monosiga brevicollis* and *Capsaspora owczarzaki*, respectively) (Figure 61). Similarly, the metazoan G12 clade also has sister genes in *Monosiga brevicollis, Salpingoeca rosetta* and *Capsaspora owczarzaki*, with maximal branch support in some cases (Figure 61). Thus, Gq and G12 likely can be recognized as separate genes at least as early as the MRCA of Filozoa. Four of the five teretosporean species (*Sphaeroforma arctica*, *Sphaeroforma sirkka*, and *Ichthyophonus hoferi*) possess Gq/G12 orthologs (Figure 61 and Figure 62), but it is unclear, whether these are ancestral to both Gq and G12 (Figure 62) or whether they constitute G12 orthologs (Figure 61). The second possibility fits to the tree topology in Figure 61, and would imply that the corresponding Gq genes have been lost in these

species. However, the low branch support for the placement of the teretosporean Gq/G12 genes does not allow to draw a firm conclusion. Moreover, a single gene in *Ichthyophonus hoferi* was found to lay clearly basal to both the Gq and the G12 clade in Figure 62, with maximal branch support, which would argue for the split into Gq and G12 to have occurred only inside the Filozoa. Similarly, de Mendoza and colleagues reported an Ichthyosporea sequence basal to both Gq and G12 clades (de Mendoza et al., 2014).

The above conclusions about the birth of the 5 families and their subfamilies were drawn from the presence or absence of a gene in the taxonomic groups Teretosporea, Filasterea, and Choanoflagellatea. However, we note that the *gna* genes (G alpha) sometimes are only found in a subset of species in the respective species group. For example, Gs, Gq and G12 are absent from two ichthyosporean species (*Pirum gemmata* and *Sphaeroforma sirkka*), and Gv is absent from one of the teretosporean Holozoa (*Corallochytrium limacisporum*) (Figure 61, Table 6). All three species have good to very good genomic coverage (870x, 180x, and 800x, respectively). Thus, it appears that some gene losses have occurred already in these early-diverging Holozoa. This pattern of gene loss is even more extreme in choanoflagellates, which have lost Gs and Gi in all species, and retain Gv, G12 and Gq (Figure 61). However, in this case, low genomic coverage (33x and 8x) could play a role. On the other hand, in the filasterean species *Capsaspora owczarzaki* (very low coverage, 10.6x) we find all five G alpha families, each represented by a single gene (Figure 61).

The detailed analysis of all teretosporean species also shows several more localized gene gains. Three of five Teretosporea harbor more than one Gi ortholog (*Ichthyophonus hoferi* four Gi genes, *Sphaeroforma arctica* four genes, and *Sphaeroforma sirkka* 3 Gi genes), but two species have a single Gi gene (the Ichthyosporea *Pirum gemmata* and the Corallochytrea species) (Figure 61). One Ichthyosporea species (*Ichthyophonus hoferi*) possesses four G12 genes, whereas all other Ichthyosporea only possess a single G12 (Figure 61). In contrast, no such duplications were seen for Filasterea (Figure 61).

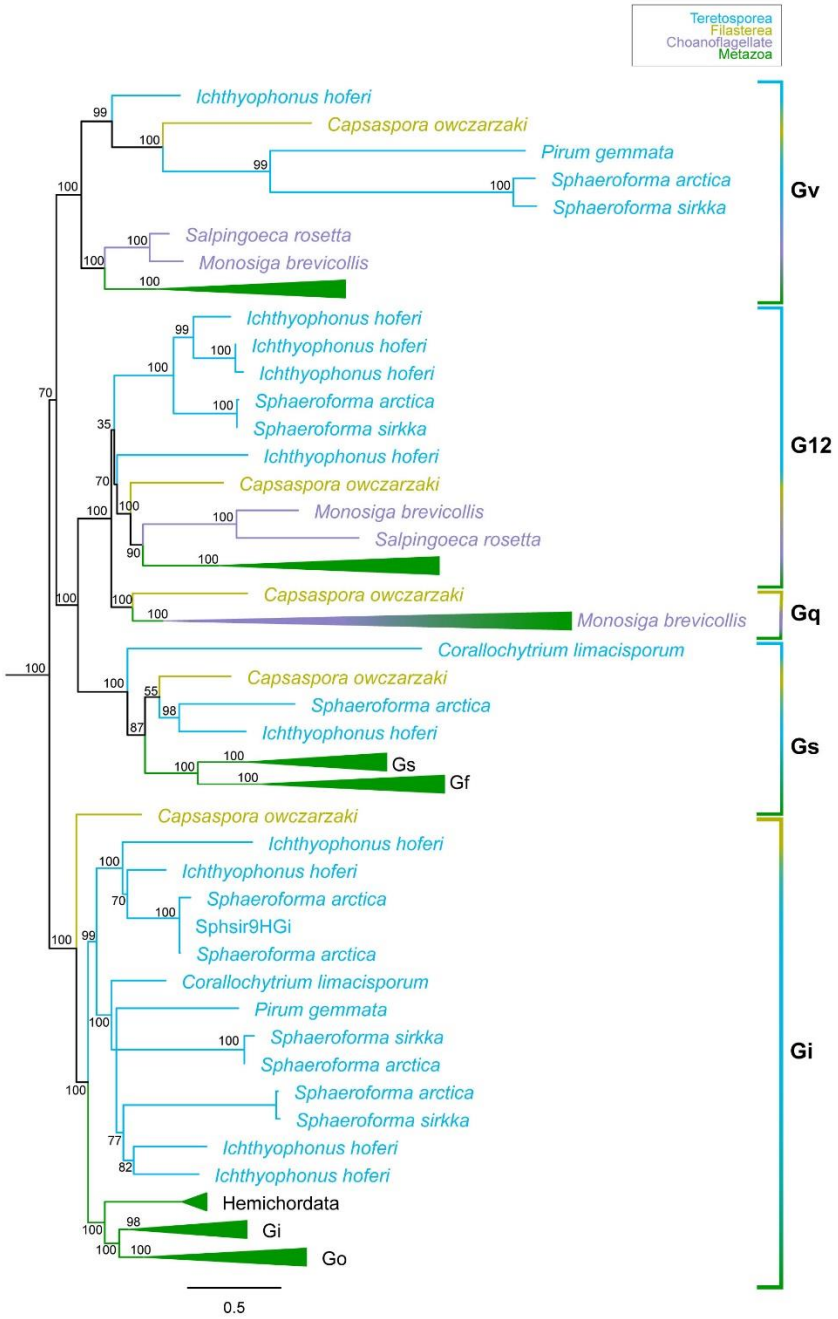


Figure 61: Phylogenetic tree shows the evolutionary origin of the five G alpha families in Holozoa. The metazoan clades are collapsed. The color code is indicated. The phylogenetic tree of the G alpha families was constructed using MAFFT for alignment and maximum likelihood algorithm PhyML-aLRT (for details see Materials and Methods 2.9.2). The tree was

drawn by FigTree, numbers on each branch indicate percent branch support for basal nodes. Scale bar indicates number of amino acid substitutions per site. See Appendix 6.2.

Taken together, three G alpha families Gi, Gs, and Gv emerge already in the common ancestor of all Holozoa, which also exhibits a G protein ancestral to both Gq and G12. The Gq and G12 families seem to be born later, in the common ancestor of Filozoa, but further phylogenetic studies are needed to unequivocally establish this point. Noteworthy, the Gv family is the only one present in all three early-diverging unicellular Holozoa groups (Teretosporea, Filasterea, and Choanoflagellatea).

Table 6: The presence of G alpha proteins in different taxonomic groups, all groups belong to Holozoa. The percentage of the species that possess the respective G alpha gene in the species group, is given. Color code follows conditional formatting with three color scale; red color denotes absence of the gene, white was set at 20 percentile, and blue color denotes maximal presence (100%).

	Classification	Group	# of species	# of genes	Gi		Gs	G12	Gq	Gv
					Go	Gi				
Metazoa	Clade	**Teretosporea**	5	28	0	100	60	60	0	80
	Class	**Filasterea**	1	5	0	100	100	100	100	100
	Class	**Choanoflagellatea**	2	5	0	0	0	100	50	100
	Phylum	**Porifera**	2	16	100	0	100	100	100	100
	Phylum	**Placozoa**	1	7	100	100	0	100	100	0
	Phylum	**Ctenophora**	3	28	0	100	0	100	100	100
	Phylum	**Cnidaria**	6	35	67	100	33	100	67	0
	2 phyla	**Ambulacraria**	6	39	100	100	17	100	100	83
	Phylum	**Chordata**	1	26	100	100	100	100	100	100
	Phylum	**Mollusca**	9	69	89	78	89	100	89	78
	Phylum	**Nematoda**	101	524	52	71	26	17	52	0
	Phylum	**Tardigrada**	2	25	100	100	100	50	100	0
	Phylum	**Arthropoda**	735	3698	36	74	83	89	67	16

3.8.2 Evolutionary dynamics of Gv across metazoan evolution and comparison with the other four classes

Herein, we evaluated Gv evolution in the Metazoa i.e., after the separation from the choanoflagellates (Figure 60). We included many earlier-diverging animal phyla (Porifera, Placozoa, Ctenophora, and Cnidaria), and among the bilateria the phyla Chordata, (Echinodermata and Hemichordata - together superphylum Ambulacraria), (Nematoda, Tardigrada, and Arthropoda - together Ecdysozoa) and Mollusca (a phylum in Lophotrochozoa, the sister clade to Ecdysozoa).

Our data analyses revealed that Gv orthologs are found in Porifera (sponges) and Ctenophora (sea gooseberry, sea walnut, and cigar comb jelly), i.e., must have been present at the very origin of metazoan evolution (sponges) and at the next major divergence (Ctenophora, Placozoa, Cnidaria, Bilateria) (Figure 63). On the other hand, Gv appears to be lost in several of the non-bilaterian phyla: Placozoa "flat animals" and Cnidaria (hydra, jellyfish, sea anemone, stony coral) (Figure 63). In Bilateria we identified Gv orthologs in both phyla of Ambulacraria (acorn worm, starfish, sea urchin); also, in the phylum Mollusca (giant squid, nautilus, octopus, cuttlefish), and in Arthropoda (Figure 63). On the other hand, despite our recursive search we couldn't find Gv orthologs in any of the currently available 96 nematodes and two tardigrade genomes.

Taken together, we observed four gene loss events for Gv at the phylum level in Animalia (Figure 63). For Porifera, Ctenophora, and Echinodermata no Gv losses were observed inside the respective phylum, whereas partial losses were observed in Hemichordata, Mollusca, and Arthropoda (over 80% of arthropod species have lost it, see also 3.8.3 for detailed analysis). In our analysis we did not focus on chordates, but previous studies have reported several loss events also in Chordata (Krishnan et al., 2015; Lokits et al., 2018; Oka et al., 2009). Hence, Gv appears to have undergone several independent loss events at different levels of metazoan evolution. This feature presents an interesting contrast to the observed presence of Gv orthologs in all the non-Metazoan Holozoa species examined (3.8.1) - the absence of Gv in the fourth class, Corallochytrea, may have technical reasons, i.e., incompleteness of the genome

assemblies. Although Corallochytrea species has good genomic coverage (800x), we could identify only two G alpha sequences (Gi and Gs), on the other hand, *Sphaeroforma arctica* has only 180x genomic coverage and we could find seven G alpha genes which encompass four G alpha families. However, to draw a firm conclusion, more species of non-Metazoan Holozoa would need to be identified (and sequenced).

On the other hand, we observed several Gv gene gains at late levels of evolution, some at the species level; three *gnav* genes in a sponge (*Ephydatia muelleri*), two in a mollusc (*Octopus sinensis*), also a gene gain in ctenophores (*Mnemiopsis leidyi* and *Beroe forskalii*) and several independent *gnav* gene gains in arthropods.

To evaluate whether the evolutionary dynamics of Gv as investigated here is similar to that of the other four classes, we analyzed the evolution of Gi, Gq, G12 and Gs classes in Metazoa, using the same set of species as analyzed for Gv.

The Gi family which encompasses Gi and Go classes is present in all phyla of the Metazoa (Figure 61), we only observe loss of Gi class in Porifera and Go class in Ctenophora (Table 6). The Gq family is also present in all phyla of Metazoa, albeit about 33-50% of the species in Cnidaria, nematodes, and arthropods appear to have lost Gq genes (Table 6). We also observe that the G12 family is present in all metazoan phyla, but only 17% of nematodes and only one of the two tardigrade species possess G12 sequences. In contrast, the Gs family shows gene losses already at the level of phyla (absent in Ctenophora, Placozoa, and Hemichordata) and partial losses in three other phyla (Cnidaria, Echinodermata, and Nematoda) (Table 6).

Overall, The Gi family is represented through all metazoan phyla, while the Gq, G12 families appear to have undergone partial losses during metazoan evolution, but not at the phylum level. Gv and Gs families show complete losses in some metazoan phyla and partial losses in others. Therefore, we conclude that Gv evolutionary dynamics are within the range of those observed for the other four G alpha classes.

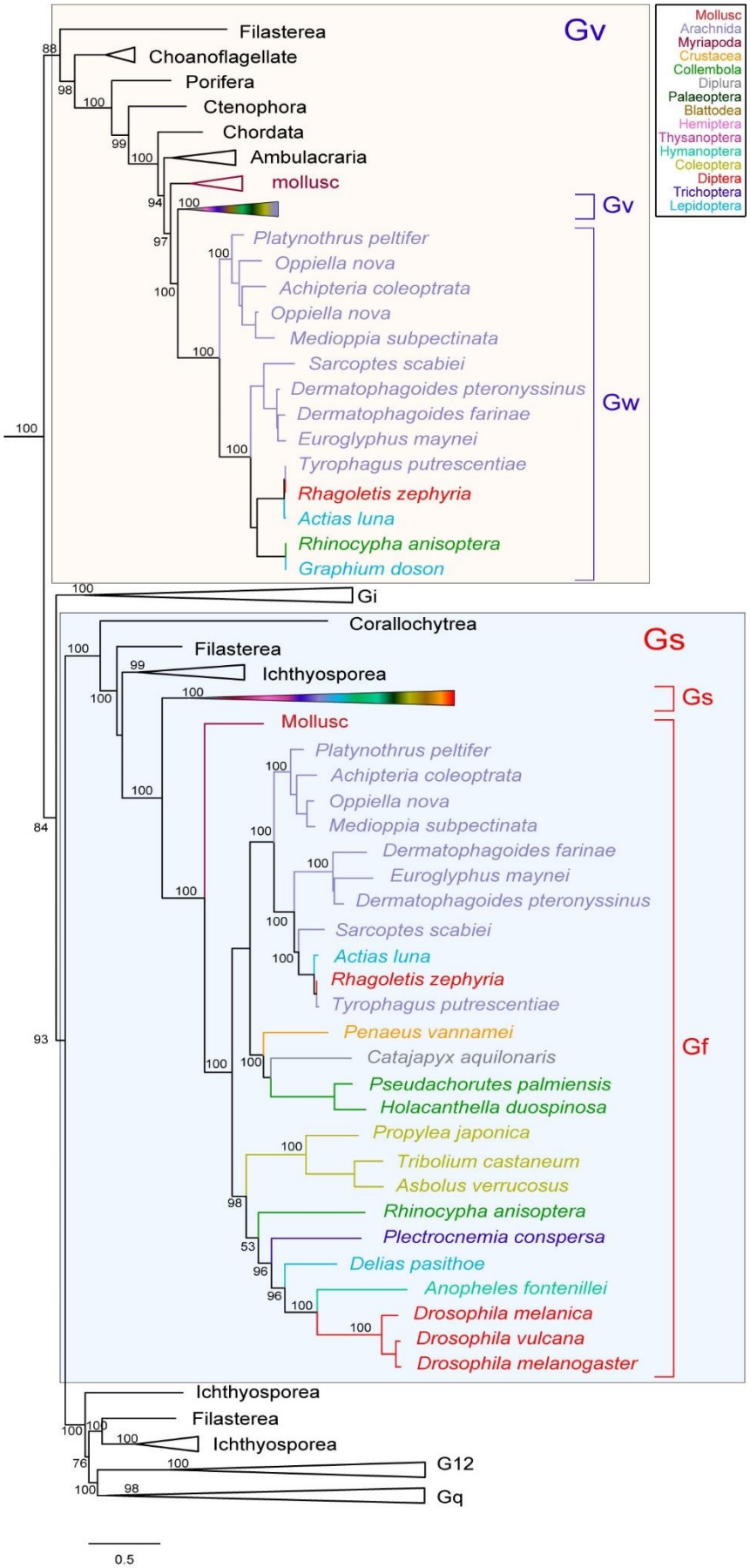

Gv
Filasterea
Choanoflagellate
Porifera
Ctenophora
Chordata
Ambulacraria
mollusc
Gv
Platynothrus peltifer
Oppiella nova
Achipteria coleoptrata
Oppiella nova
Medioppia subpectinata
Sarcoptes scabiei
Dermatophagoides pteronyssinus
Dermatophagoides farinae
Euroglyphus maynei
Tyrophagus putrescentiae
Rhagoletis zephyria
Actias luna
Rhinocypha anisoptera
Graphium doson
Gw
Gi
Gs
Corallochytrea
Filasterea
Ichthyosporea
Gs
Mollusc
Platynothrus peltifer
Achipteria coleoptrata
Oppiella nova
Medioppia subpectinata
Dermatophagoides farinae
Euroglyphus maynei
Dermatophagoides pteronyssinus
Sarcoptes scabiei
Actias luna
Rhagoletis zephyria
Tyrophagus putrescentiae
Penaeus vannamei
Catajapyx aquilonaris
Pseudachorutes palmiensis
Holacanthella duospinosa
Propylea japonica
Tribolium castaneum
Asbolus verrucosus
Rhinocypha anisoptera
Plectrocnemia conspersa
Delias pasithoe
Anopheles fontenillei
Drosophila melanica
Drosophila vulcana
Drosophila melanogaster
Gf
Ichthyosporea
Filasterea
Ichthyosporea
G12
Gq
0.5
Mollusc
Arachnida
Myriapoda
Crustacea
Collembola
Diplura
Palaeoptera
Blattodea
Hemiptera
Thysanoptera
Hymanoptera
Coleoptera
Diptera
Trichoptera
Lepidoptera

Figure 62: Phylogenetic tree shows the evolutionary origin the novel Gf and Gw clades and their sister-related clades Gs and Gv, respectively. All the G alpha clades are collapsed except Gs and Gv. The color code is indicated. Metazoan Gs clade is collapsed and it encompasses species from the following groups: Porifera, Cnidaria, Chordata, Echinodermata, Mollusca, Nematoda, Tardigrada, Arachnida, Crustacea, Collembola, Diplura, Palaeoptera, Blattodea, Hemiptera, Thysanoptera, Hymenoptera Coleoptera, Diptera, Trichoptera, and Lepidoptera. The collapsed Gv clade encompasses Gv of arthropods including (Arachnida, Myriapoda, Diplura, Palaeoptera, Blattodea, Hemiptera, Coleoptera, Diptera, and Trichoptera). The tree was constructed the same as in Figure 61. Species groups are visualized by color code as indicated. Numbers on each branch indicate percent branch support for basal node; scale bar indicates number of amino acid substitutions per site. Generally, the gene tree topology follows the topology expected from species relationships. See Appendix 6.3.

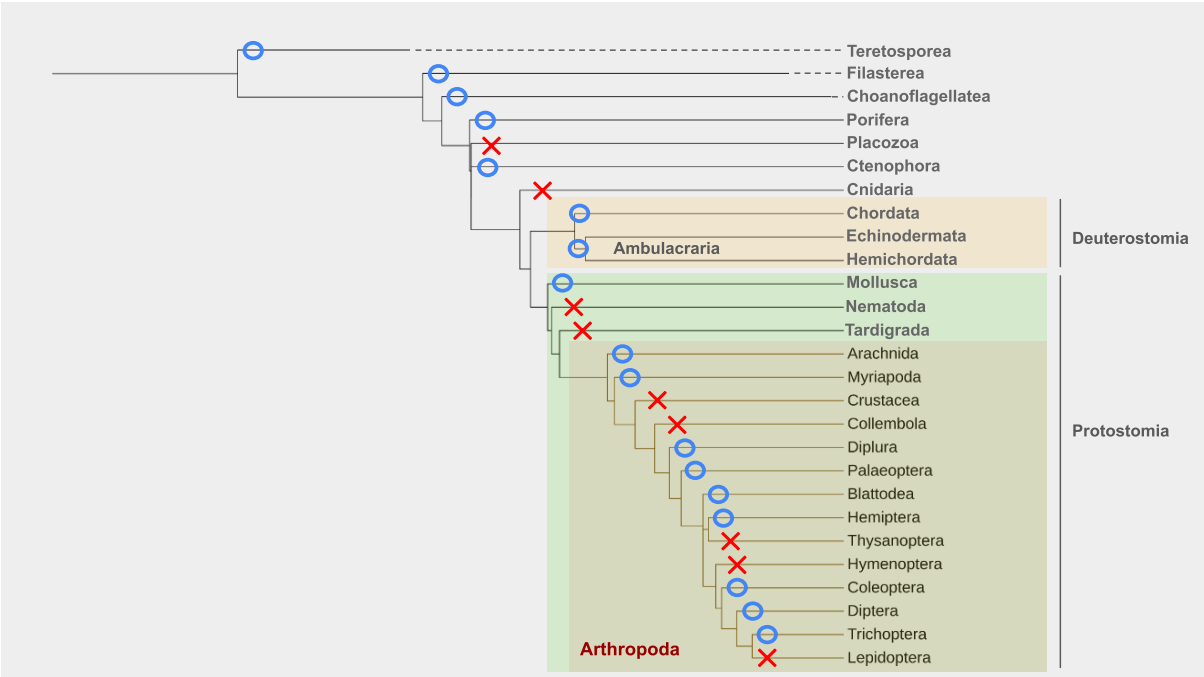


Figure 63: Gv gene losses visualized in species tree. The clades shown in the tree include all species analyzed in this study. The tree was drawn using the web tool TimeTree by (Kumar et al., 2017). Blue circles denote presence of Gv in the respective clade and red crosses depict independent gene loss events.

3.8.3 Gene loss events are a recurrent feature in the evolution of Gv in arthropods

Gv appeared to be lost in many branches of the tree-of-life including tetrapods, Urochordata, and Cnidaria (see 3.8.1) (Oka et al., 2009). Only a single insect species, the red flour beetle was found to possess Gv, whereas fruit fly, mosquito, bee, or moth

had lost it. It was hypothesized that there was a tendency towards gene loss in non-aquatic living species (Oka et al., 2009), since the vast majority of insects have a terrestrial lifestyle. To evaluate this hypothesis, we investigated the presence of Gv in other insect species. Initial analysis with small subsets of species showed that Gv indeed is present in insect groups other than beetles such as bugs, dragonflies, and damselflies. Insects are an extremely species-rich animal group and in the meantime many insect genomes have been assembled. Here we have performed a comprehensive analysis of the entire G alpha protein repertoire of all 622 insect genomes currently available for a thorough test of this hypothesis. Moreover, we have extended this analysis to include almost all non-insect arthropods such as Chelicerata (arachnids), Crustacea (crustaceans), other Hexapoda (springtails), and Myriapoda (millipedes and centipedes), some of which show an aquatic lifestyle (e.g., crustaceans). We used a representative set of query genes (see Materials and Methods 2.9.2) and tblastn (Altschul et al., 1990) to search WGS databases maintained at NCBI at the subphylum, (sub) class, or order levels.

We identified Gv orthologs in many non-insect arthropod groups such as Arachnida, Myriapoda, and Diplura, and also in many insect groups such as Palaeoptera, Blattodea, Hemiptera, Coleoptera, Diptera and Trichoptera (Figure 63). However, we observed major partial losses in Hemiptera and Diptera since only 10% and 11%, respectively, of the species in these groups possess *gnav* genes (Table 7). Since the other G alpha classes are present in a very high percentage of the species in Hemiptera (85 - 100%) and Diptera (80 - 98%) (Table 7), we expect the absence of Gv in these orders to reflect the absence of the gene itself, not inadequacies of the respective databases. We conclude that the Gv family is lost in about 90% of the species in Hemiptera and Diptera (Table 7). The Lepidoptera group appears to have lost the Gv class genes completely (but see Figure 62 and 3.8.4) and also exhibits major partial losses in Gi (99% of the species), and Gq (97% of the species), while most of the species in this group harbor Gs and G12 orthologs (Table 7). Again, the presence of Gs and G12 argues against low quality databases having a major effect on these findings.

Furthermore, we couldn't find any Gv orthologs in several non-insect arthropods, i.e., Collembola (springtails), and Crustacea (small and large crustaceans, shrimps, prawns, water fleas, etc.), and in the insect groups Thysanoptera (thrips) and Hymenoptera (sawflies, wasps, bees, and ants) (Figure 63). Hymenoptera is the largest insect group (Forbes et al., 2018) and we searched in all of the available 225 species genome databases. Overall, 738 G alpha genes were identified (Table 7), but there are no Gv orthologs among them. Likewise, we could identify 281 genes from searches in the 43 species crustacean group (Table 7), yet no Gv orthologs are found. In spite of our recursive searches, we could not also find *gnav* genes in Collembola (16 species, 109 genes) and Thysanoptera (3 species, 14 genes) (Table 7). On the other hand, it appears that there is a good representation for the other four G alpha families in Crustacea, Collembola, Thysanoptera, and Hymenoptera. Thus, the absence of Gv in these clades likely reflects the true absence of *gnav*.

Table 7: The presence of G alpha proteins in different taxonomic groups, all groups belong to Arthropoda. The percentage of the species that possess the respective G alpha gene in the species group, is given. Color code follows conditional formatting with three color scale; red color denotes absence of the gene, white was set at 8 percentile, and blue color denotes maximal presence (100%).

	Classification	Group	# of species	# of genes	Gi		Gs	G12	Gq	Gv
					Go	Gi				
	Class	**Arachnida**	49	433	94	94	92	84	96	57
	Subphylum	**Myriapoda**	3	19	100	100	33	100	100	100
	Class	**Crustacea**	43	281	79	95	58	93	65	0
	Subphylum	**Collembola**	16	109	94	88	94	94	100	0
	Order	**Diplura**	2	18	100	100	100	100	100	100
	Infraclass	**Palaeoptera**	6	56	100	100	83	100	100	100
	Order	**Blattodea**	4	28	100	100	100	100	100	100
	Order	**Hemiptera**	39	207	85	100	87	97	97	10
	Order	**Thysanoptera**	3	14	100	100	67	100	100	0
Insecta	Order	**Hymenoptera**	225	738	1.3	74	78	74	72	0
	Order	**Coleoptera**	50	377	66	100	84	94	72	90
	Order	**Diptera**	172	1065	43	98	96	95	80	11
	Order	**Trichoptera**	6	31	83	0	67	100	83	100
	Order	**Lepidoptera**	117	322	3.4	0.9	78	99	2.6	0

Therefore, it appears that at least five independent Gv losses occurred at three different taxonomic levels within arthropod evolution: the most basal Gv loss has taken place in the subphylum Crustacea after the divergence of Crustacea from Hexapoda, then another loss occurred in Collembola after the emergence of this subclass in Hexapoda, and then another 3 losses within Insecta for the orders Thysanoptera (a Hemimetabola), Hymenoptera (a Holometabola) and Lepidoptera (another Holometabola). Moreover, major partial losses appear to have occurred within the orders Hemiptera (Holometabola) and Diptera (Hemimetabola). The three orders Hymenoptera, Lepidoptera and Diptera together comprise the large majority of insect

species, and the total or partial losses in these orders therefore results in a low overall retention of Gv (16% of arthropod species).

On the other hand, we observed recurrent gene gains for Gv genes in Arthropoda: three *gnav* genes were observed in one Arachnida species and two *gnav* genes/species were observed in two Arachnida, one Myriapoda, one Palaeoptera, ten Coleoptera, one Trichoptera, and one Diptera species.

This pattern of recurrent gene gains and losses suggests that the Gv class shows birth-and-death mode of evolution. This parallels earlier results obtained by Oka and colleagues for mostly non-arthropod species. When comparing the presence/absence of Gv to aquatic/terrestrial lifestyle no correlation is apparent: the overwhelming majority of Crustacea lives aquatically (woodlice are the only exception in the species studied here), but none of them possess Gv. On the other hand, the majority of Arachnida and Coleoptera possess Gv notwithstanding their terrestrial lifestyle.

In order to explore whether the evolutionary dynamics of Gv as investigated here is similar to that of the other four classes, we analyzed their evolution in Arthropoda, using the same set of species as analyzed for Gv.

We observe that the G12 family is present in all arthropod groups analyzed and in the large majority of species in each group (74 - 100%). Similarly, the Gi family (Gi and Go) is present in all arthropod groups analyzed and all but one possesses both Gi and Go genes (Trichoptera lack Gi) (Table 7). However, in Lepidoptera both genes are only represented in very few species, 1% for Gi and 3% for Go. The Gq family is also present in all arthropod groups, however we observe more frequent gene losses: a partial loss in 28 - 35 % of the species in Crustacea, Hymenoptera and Coleoptera, and major loss (97% of the species) in Lepidoptera (Table 7). Similarly, Gs is present in all arthropod groups analyzed, however, partial gene losses are more frequent: two of the three myriapods species lost Gs; 42% of the crustaceans lost Gs; one of the three of the Thysanoptera and two of the four Trichoptera species appear to have lost Gs class (Table 7). Comparing these gene loss patterns with those described above

for Gv, we see a much broader pattern of gene loss for Gv which is lost completely in 5 arthropod groups and partially in two more.

3.8.4 The comprehensive search for G alpha genes across a wide swath of species unearthed two novel gene clades, Gf and Gw

Besides the well-known five families of G alpha proteins there have been isolated reports of a G protein similar to Gs, but not within the Gs clade. This protein has been first determined in *Drosophila melanogaster* by analyzing cDNA coding for fly G alpha subunits by degenerative PCR (Quan et al., 1993) and was named Gf for fly G protein. Also, a homolog for Gf was found in mosquito *Anopheles gambiae* G alpha, which was named gm (Rützler et al., 2006). However no systematic phylogenetic study for this gene had been done so far.

Therefore, we included the Gf sequence from *Drosophila melanogaster* into the different representative sets of query genes during our searches in the various databases. Interestingly, we found 258 Gf homologs in ten arthropod groups including Diptera (where it was originally determined). The earliest-diverging arthropod group (Arachnida) already possesses Gf, and some other non-insect arthropods (Crustacea and Collembola) also exhibit Gf, but myriapods do not. Several insect groups (Diplura, Palaeoptera, Coleoptera, Hymenoptera, Diptera, Trichoptera, and Lepidoptera) possess Gf, while Gf was absent in Blattodea, Thysanoptera, and Hemiptera. This pattern could be explained by independent losses in Myriapoda, Blattodea, and the common ancestor of Thysanoptera, and Hemiptera (Figure 64).

Since the earliest-diverging arthropod group (Arachnida) already possesses Gf, we also investigated the presence of Gf in three Protostomia groups that diverged earlier than arthropods (Figure 62 and Figure 64). We find Gf orthologs in molluscs (2 of 9 species), but not in nematodes (101 species) and tardigrades. The mollusc genes are fragments and might be pseudogenes. Nevertheless, the entire Gf clade as described here, presents with maximal branch support as sister to the Gs clade (Figure 62).

Concluding from the species composition in the Gf clade, the Gf gene has to be at least as old as the divergence between Mollusca and Ecdysozoa. In fact, the tree topology suggests an even earlier origin in the common ancestor of Metazoa, since the sister clade contains Gs genes from the earliest diverging metazoa (Porifera) and also from Cnidaria. However, such an early origin would require independent gene losses for Gf in Porifera, Cnidaria, Ctenophora and deuterostomes in addition to those necessitated by assuming a birth of Gf in the common ancestor of Ecdysozoa and Mollusca (several protostome phyla). Currently it is not possible to decide between an origin in the common ancestor of all Metazoa coupled with many independent losses of Gf and a later origin in the common ancestor of Mollusca and Ecdysozoa which requires fewer independent gene losses. While one could argue that a hypothesis with less assumption of gene losses is more parsimonious, we know from investigation of the five major G protein families that complete and partial gene losses are a common and recurrent feature.

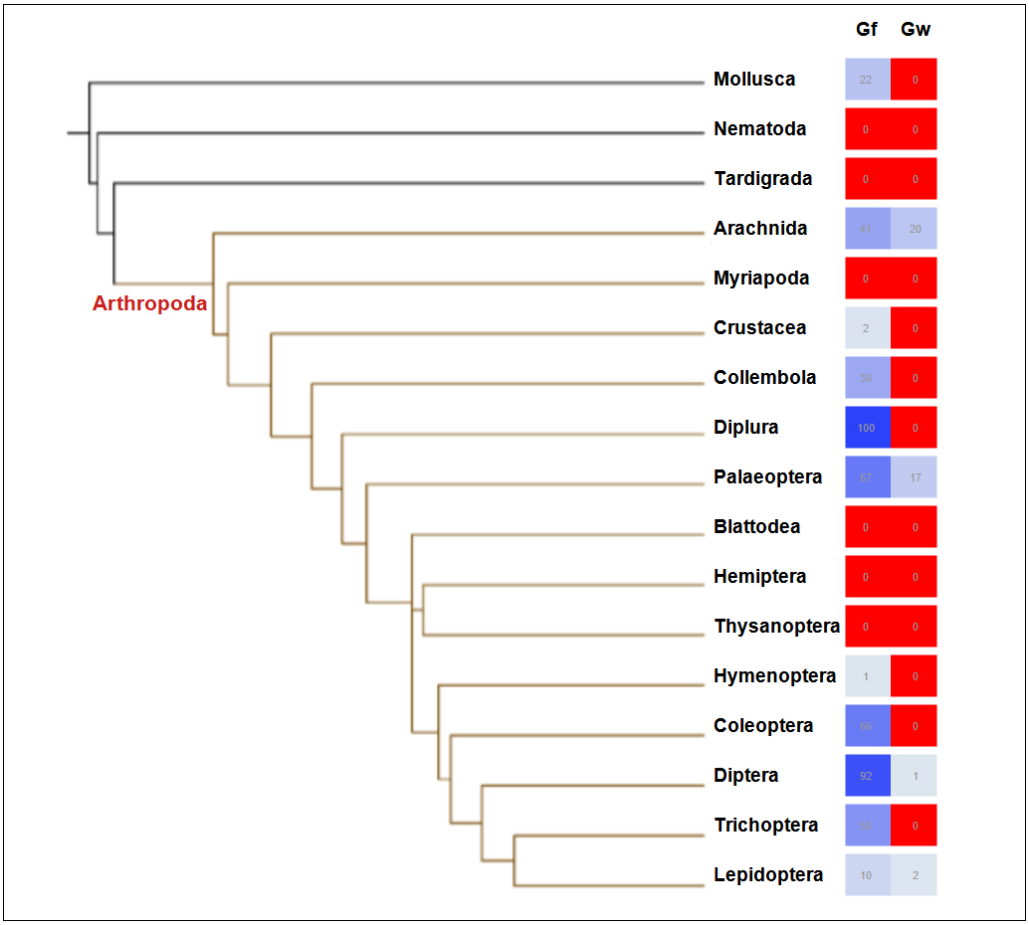


Figure 64: Species tree is annotated with percentage of the species that possess Gf and Gw in each group. The tree was drawn using the web tool TimeTree by (Kumar et al., 2017). The percentage of the species that possess the respective G alpha gene in the species group, is given. Color code follows conditional formatting with three color scale; red color denotes absence of the gene, white was set at 20 percentile, and blue color denotes maximal presence (100%).

The situation is different for the unicellular Holozoa (*Corallochytrium limacisporum*, *Ichthyophonus hoferi*, and *Sphaeroforma arctica*). All their Gs/f orthologs cluster basal to both Gs and Gf clades (Figure 62). Here the tree topology and the species composition of the clades lead to the same conclusion, i.e., the duplication of an ancestral Gs/f gene into Gs and Gf happened only inside metazoa.

Moreover, our searches led to the identification of a group of genes that had not been described in the literature so far to the best of our knowledge. This group is limited to arthropods and clusters as a sister clade to the arthropod Gv with maximal branch support (Figure 62). We obtained this result repeatedly in several phylogenetic tree

analyses using different species groups. We named this group of genes the Gw class due to the proximity to Gv class. Accordingly, the Gv family encompasses two classes, Gv and Gw.

Gw orthologs were identified in ten Arachnida, and also in very few insect species: two Lepidoptera (out of 117 species), one Diptera (out of 172 species), and one out of six Paleoptera species. This finding firstly shows that the origin of the Gv/Gw duplication must have been in the common ancestor of all arthropods, since Arachnida are the earliest-diverging group and they do possess both Gv and Gw (Figure 62). Secondly, the complete absence of Gw in Myriapoda, Crustacea, Collembola, Diplura, and six insect groups (Blattodea, Hemiptera, Thysanoptera, Hymenoptera, Coleoptera, and Trichoptera) and the near complete loss in Lepidoptera and Diptera suggests many independent gene losses for Gw, much more extensive than observed for other classes (Figure 64).

Moreover, we observed few gene gain for the novel Gf class, i.e., in only ten species in Diptera, one in each of Coleoptera, Palaeoptera, and Lepidoptera. On the other hand, gene gains seem very limited for the novel class Gw, a single gene gain was observed in an Arachnida, *Oppiella nova*.

4 Discussion

Gv proteins constitute one of the five canonical G alpha families in animals. Gv was discovered 13 years ago and is so far functionally uncharacterized. The study at hand deciphers for the first-time biological roles of this novel Gv protein using zebrafish as an experimental system. First, we investigated the ontogenetic and tissue-specific expression of Gv mRNA during embryogenesis and larval development by the histochemical technique WISH. We thereafter evaluated phenotypic and molecular changes in *gnav1*$^{-/-}$ mutants with an early 13bp deletion in the mutant genome. These approaches constituted the starting blocks to elucidate the potential role of Gv, focusing on developing larvae and adult kidney.

4.1 Developmental course and tissue specificity of expression suggest potential functions for Gv

We have shown that *gnav1* exhibits broad expression in early embryonic stages from gastrulation through the pharyngula period with enhanced levels in specific tissues. Thereafter, Gv transcripts become confined to few organs in hatching to larvae periods, 2 to 5 dpf respectively.

At 6 hpf (gastrulation period), Gv mRNA is expressed broadly through the blastoderm including the embryonic shield, where the cells form two layers of the epiblast and hypoblast. Fate maps show that epiblast and hypoblast give rise to many different organs and tissues; thus, Gv expression in both layers suggests that Gv is involved in various organs and tissues development. Subsequently in the early segmentation stage at 12 hpf, the signal is sustained broadly yet most intense in the head (axial mesoderm, eye and brain primordia), ventral to the head where the hatching gland forms, lateral plate mesoderm (pronephros, and possibly the gut), notochord primordium and tail bud. The expression of *gnav1* in these organ primordia (eye, brain, hatching gland, pronephros, notochord, and tail bud) appears more distinct in the next 24 hpf.

At the prime-5 stage (24 hpf), Gv is also expressed broadly through the AP axis and extends to the YSL and somites. The expression is distinct in the sensory organs (eye, brain, floor plate, and otic vesicles); the cartilaginous tissues (pharyngeal arches regions and pectoral fin bud); osmoregulatory organs (pronephros and YSL); and hatching gland.

In the next day of development (long-pec stage at 48 hpf), overall Gv mRNA signal is not broad anymore and becomes more restricted than 1 dpf stage. Gv mRNA expression disappears from the notochord, eyes, and somites; decreases in the brain, YSL of the yolk sac, and hatching gland; does not change in pronephros and YSL of yolk extension; whereas increases in some other tissues (otic vesicle, pharyngeal arches, pectoral fin) and newly appears in the gut. After the embryo hatches at 3 dpf stage, *gnav1* expression remains and become more distinct (compared to 48 hpf) in otic vesicles, craniofacial cartilages such as hyoid arches and ethmoid plate, pectoral fin, pronephros, gut, and YSL. And in 5 dpf larvae it is confined to only three organs where it is distinct (pronephros - in particular PCT, gut, and swim bladder).

It appears that Gv expression is transient in some tissues such as sensory organs, cartilaginous tissues, somites, and notochord during the time period studied (6 hpf - 5 dpf). In contrast, after an early onset the expression was persistent in the osmoregulatory organs (pronephros and gut) through all the stages examined. For some transient tissues such as Kupffer's vesicle and hatching gland after an early onset no decrease was seen during the existence of these structures. A special case concerns the swim bladder, which starts to develop before 3 dpf, but only shows Gv expression at 5 dpf.

These expression patterns suggest manifold functions for Gv in the developing zebrafish embryo/larvae. It appears that it is involved in the development of sensory, osmoregulatory, cartilaginous organs and muscles. Also, its expression in the hatching gland suggests its involvement in the hatching process. From these data, one could hypothesize the cellular processes in which Gv could be involved.

Firstly, Gv might be involved in cellular movements since its expression is evident in several tissues at particular stages where their development involves cell migration. These tissues and migration processes include: the blastoderm during gastrulation (Warga & Kimmel, 1994); the brain, in part forebrain, during eversion (Folgueira et al., 2012); the pronephros and PCT during cell migration and convolution respectively (Vasilyev et al., 2009); YSL through yolk elongation (Virta & Cooper, 2011); hatching gland through its development which also involves cell migration (Hatzold et al., 2021); pharyngeal arches region that harbors cranial neural crest cells (CNCCs) which migrate and give rise to the craniofacial skeleton (Schilling & Kimmel, 1994); ethmoid plate during extension (Dougherty et al., 2013)**.**

Secondly, we speculate that Gv might also have a role in cell proliferation. The organogenesis of some of the above organs (where Gv is expressed) involve cell migration coupled to cell proliferation. For instance, Dougherty et al., 2013 showed that there is coordination between many processes during ethmoid plate formation which include extension and proliferation. Also, Vasilyev et al., 2009 reported that the cell movement is coupled to proliferation in the distal tubule of the pronephros. Moreover, we showed that Gv mRNA is expressed in the MBH, brain ventricular zone, and pharyngeal arches wherein proliferation is known to take place (Birkholz et al., 2009; Folgueira et al., 2012; März et al., 2010; Sasaki et al., 2013; Tallafuß & Bally-Cuif, 2003).

Thirdly, we showed that Gv is expressed in the embryonic osmoregulatory organs (gut, pronephros (kidney), and YSL) and the adult kidney, therefore we suggest that Gv could have a role in ionic homeostasis. A recent study showed that the pronephric cell migration and tubule convolution depends on lumenal fluid flow, thus connecting pronephric morphogenesis to its function in ionic homeostasis (Vasilyev et al., 2009). Also, it has been shown in the literature review that ion transport activity is required for gut morphogenesis and kidney tubulogenesis in zebrafish and other organisms (Krupinski & Beitel, 2009)**.**

Last but not least, Gv might also have a role in other cellular processes such as cellular differentiation in the somites and anterior pharyngeal arches as it is expressed in these

tissues during maturation. Additionally, we cannot also exclude the possibility that Gv might have a role in signal patterning of adjacent, Gv-negative, tissues inferred from its expression in the MBH, Kupffer's vesicle, and notochord; as these tissues are known to control the cell fate of the surrounding cells (Ellis, Bagwell, et al., 2013; Essner et al., 2005; Kimmel et al., 1995; Rhinn & Brand, 2001; Warga & Kane, 2018)

To the best of our knowledge, Wnt signaling (secreted glycoproteins) orchestrates cellular morphogenesis during development e.g., cell migration, proliferation, polarity, and cell fate determination, and Gv might play a role in these cellular processes (as discussed above) (Clevers, 2006; Eisenmann, 2005; Gruber et al., 2016). Therefore, there could be a correlation between Gv function and Wnt signaling. Indeed, we observe several examples of Gv being expressed (bold) in parallel to known Wnt roles in organogenesis. *wnt8a* has a role in posterior mesoderm patterning that includes the **pronephros** (Naylor et al., 2017); *wnt1*, *wnt3a*, *wnt10b*, and *wnt3a* are required for **MHB** patterning; the latter is also involved in **somitogenesis** (Aulehla et al., 2003; Buckles et al., 2004; Lun & Brand, 1998); *wnt4a*, *wnt11 (slb)*, and *wnt11r* have a role in midline convergence that includes **the gut** (Matsui et al., 2005)**;** *wnt9a* is required early during **pharyngeal arches** (palate and lower jaw), **pronephric ducts**, **sensory** organs and **pectoral fin buds** development (Cox et al., 2010; Curtin et al., 2011; Dougherty et al., 2013); *wnt5b is* involved in patterning of swim bladder (Yin et al., 2011); Also there is a requirement for wnt5 (*ppt*) and *wnt11* (*slb*) genes in posterior convergence and extension that involves **tail** formation (Marlow et al., 2004).

Among Wnt interacting partners the leucine-rich repeat-containing G protein-coupled receptors (LGRs, class A GPCRs) have essential roles in development. In particular, Lgr4 has a crucial role in development (Barker et al., 2013; Ordaz-Ramos et al., 2021). Recent studies showed that Lgr4 regulates R-Spondin-mediated Wnt signaling (Carmon et al., 2011; Glinka et al., 2011; Ordaz-Ramos et al., 2021; Ruffner et al., 2012). Interestingly, so far only signal transduction pathways involving lgr4 have been found that do not include the activation of a G protein. However, lgr4 is a GPCR, thus one might expect that it also signals *via*, so far unknown, G protein pathways.

The Lgr4 mRNA is expressed dynamically in the developing zebrafish embryo, in a similar fashion to Gv mRNA from about 12 - 72 hpf. Like *gnav1* at around 12 hpf, *lgr4* is expressed in the head, axial mesoderm, Kupffer's vesicle, and notochord. At 24 hpf, *lgr4* is expressed in the brain, ventricular zone, floor plate, otic vesicles and eyes. The expression of *lgr4* becomes more restricted, in 48 hpf to MHB, brain ventricular zone, otic vesicles, pharyngeal arches, pectoral fin buds, intestine, reminiscent of *gnav1* expression pattern, and also in pancreas. In 72 hpf larvae, *lgr4* signal is present in the otic vesicle, pharyngeal arches (Meckel's cartilage, palatoquadrate, ceratohyal, and the branchial arches), pectoral fin, and gut similar to Gv expression; with additional signal in the liver (Hirose et al., 2011). It appears that Gv mRNA expression shows extended overlap to Lgr4 mRNA expression pattern during embryogenesis; with additional signals for *gnav1* in hatching gland, pronephros, YSL, and somites. Even though Lgr4 is not expressed in the embryonic zebrafish kidney, it has been shown that Lgr4 is required for embryonic nephrogenesis in mice (Hirose et al., 2011; Mohri et al., 2012).

It has been shown in mice that the orphan Lgr4 has a role in cell proliferation and migration which is required for tissue elongation and convolution (Jin et al., 2008; Mendive et al., 2006). Lgr4 mutants show defects in the development or function of multiple organs including the kidneys, intestine, male and female reproductive tracts and skin (Barker et al., 2013; Mohri et al., 2012). Interestingly, loss of function mutation of Lgr4 result in bone defects and reduced fertility in female mice, and we observe both these defects in *gnav1*$^{-/-}$ mutants as well (see 3.5.4 and 3.6.1 for details) (Luo et al., 2009; Mohri et al., 2010). Therefore, we speculate that *lgr4* might be upstream of *gnav1* in the signaling cascade and that Gv could be an effector G protein for Lgr4 in teleosts. To test this hypothesis, further experiments are required. For example, constitutively active Lgr4 mutants (Gao et al., 2006) could be employed to assess a possible activation of Gv. It might also be instructive to generate a double knockout for Lgr4 and Gv and examine whether effects are additive. The downstream effectors of Gv are not known, but it might be possible to compare the metabolome of cells transfected with the constitutively active Lgr4 in Gv mutant and wildtype background.

4.2 A truncated Gv protein is present in the mutant but most likely non-functional

The detailed characterization of Gv expression in the developing zebrafish (see 4.1) showed a range of Gv-expressing organs and developmental stages, suggesting that Gv has multifaceted roles during organogenesis. In order to further explore the function of Gv we employed the available mutant line obtained in a previous thesis work (Ivandic, 2015). Foremost, we validated and confirmed the presence of the 13 bp deletion described in (Ivandic, 2015) in the genomic DNA of the homozygous line used in this study. Then we evaluated the mutant Gv mRNA and protein contents.

The mutant Gv mRNA harbors a frameshift mutation (resulting from the 13 pb deletion in *gnav1*) which in turn introduces a premature stop codon (nonsense mutation), thus the aberrant mRNA is expected to undergo degradation by NMD machinery. However, we showed that the NMD is nearly bypassed in *gnav1*$^{-/-}$ zebrafish embryos, larvae, and adult kidney. The qPCR data showed that Gv mRNA in 1 and 2 dpf mutants is only reduced to 69 and 75 % of the wildtype *gnav1* levels, respectively. This is not anticipated since the NMD is triggered efficiently in zebrafish, as in mammals, when the premature stop codon lies at least 50 to 55 bp away from the downstream exon-exon boundaries (Dyle et al., 2020; Wittkopp et al., 2009). This condition is fulfilled by the *gnav1*$^{-/-}$ mutation. However, several studies have postulated exceptions to this role; such as the ‘AUG-proximity effect’ when the premature stop codon occurs very early in the coding sequence. Furthermore, translation may be reinitiated at a downstream initiating codon either with or without the presence of the proximity effect (Buisson et al., 2006; Pereira et al., 2015). Both are the case for the *gnav1* mutant.

The AUG-proximity can override the 50–55 bp boundary rule in initiating NMD-resistant mRNA for a particular mutation. When the premature stop codon is in close proximity to the initiation codon AUG the assembly of termination machinery at the stop codon is impaired, which represses NMD, subsequently translation can be reinitiated downstream at a secondary start codon (Fatscher et al., 2015). A well-studied example for ‘AUG-proximity effect’ is the nonsense mutations of human β- and

α-globin genes; the boundary of the AUG-proximity effect for NMD resistance in β-globin mRNA is at codons < 23, and at codons < 30 in the α-globin mRNA (Inácio et al., 2004; Neu-Yilik et al., 2011; Pereira et al., 2015; Romão et al., 2000). In this study, *gnav1* nonsense mutation is at codon 20, therefore we suggest that the juxtaposition of the premature stop codon to the initiating codon in mutants (AUG-proximity effect) could inhibit NMD. The boundaries of the proximity effect for NMD resistance is determined by the ORF structure and whether translation reinitiation is possible downstream the premature stop codon (Neu-Yilik et al., 2011; Pereira et al., 2015). In particular, the predicted secondary structure of the 5' Gv mRNA region by LinearFold-V (L. Huang et al., 2019) appears almost relaxed (at least at the 5' of ORF sequence where the deletion site is present); the more relaxed structure allows a faster elongation rate which in turn extends the boundary of AUG-proximity effect for NMD resistance (Figure 65) (Pereira et al., 2015).

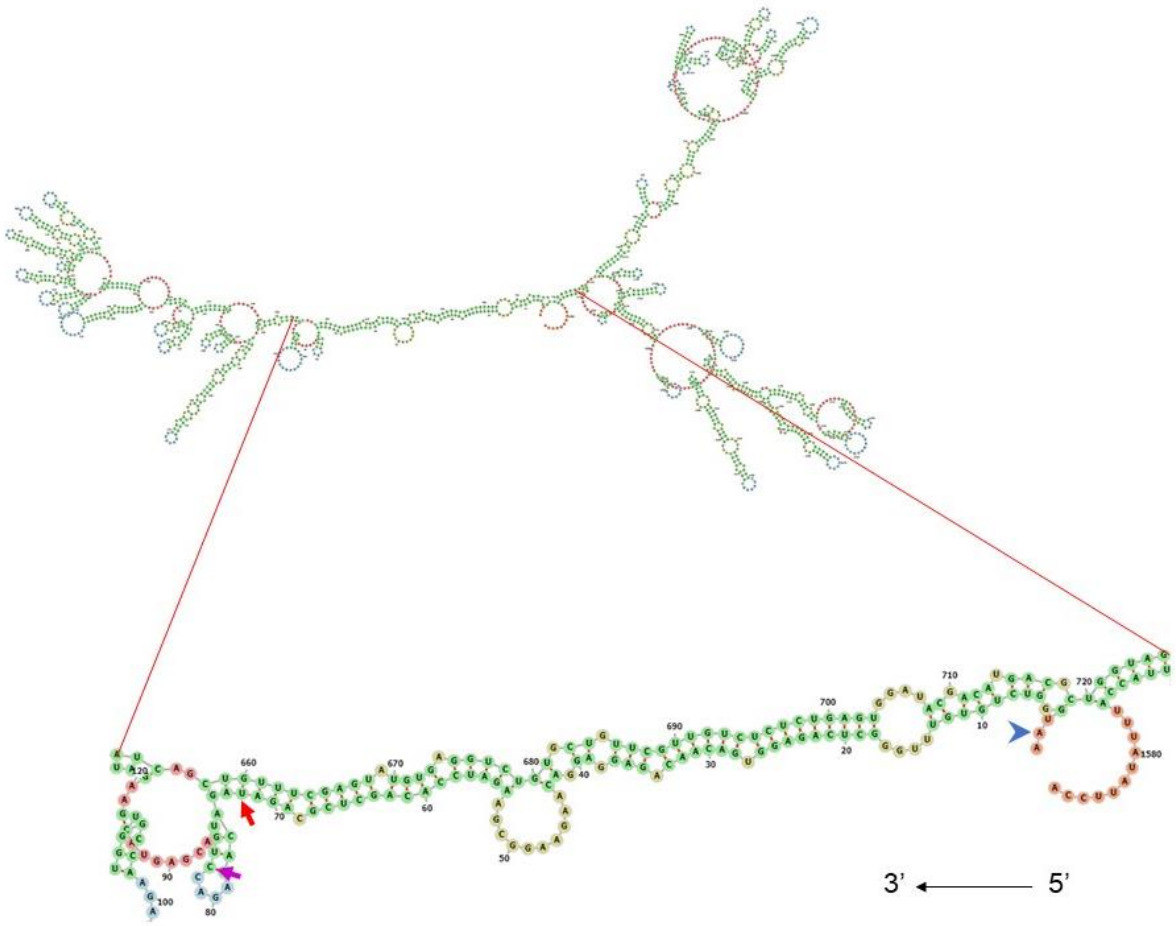


Figure 65: The secondary structure of *gnav1* mRNA prediction. The figure represents a prediction for the folding of *gnav1* mRNA sequence by LinearFold-V. The magnified region at the bottom encompasses the first 160 bp in the coding sequence. The direction of tranlation is indicated by the black arrow from 5' to 3'; and the start codon is indicated by the blue arrowhead. The red arrow indicates the premature stop codon position resulting from the frame shift mutation. The purple arrow indicates potential alternative start codon (non-AUG) which encodes leucine at codon 28 (CUG). The secondary structure of this sequence appears relaxed; no hairpins or pseudoknots are present. The numbers represent the nucleotide position in the sequence. Color coding of the sequence: Green: Stems (canonical helices); Red: Multiloops (junctions); Yellow: Interior Loops; Blue: Hairpin loops; Orange: 5' and 3' unpaired region.

Even without the presence of the proximity effect, translation may be reinitiated at a downstream initiating codon (Buisson et al., 2006; Pereira et al., 2015). The ability to reinitiate translation after termination at a downstream AUG or non-AUG start codon has a pivotal role in evading NMD (Buisson et al., 2006; Neu-Yilik et al., 2011). The Gv ORF harbors several methionine (AUG) codons downstream the mutation site; the first alternative AUG is present at codon 58 in exon 2. In addition, two non-AUG

initiation codons at 28 (see Figure 65) and 36 are present downstream the premature stop codon in exon 1 (leucines coded by CUG and AAG, respectively). Several studies have shown that translation can be initiated despite the absence of conventional AUG codons in different types of cells (Damme et al., 2014; X. Gao et al., 2015; Malarkannan et al., 1999; Ronsin et al., 1999; Schwab et al., 2003, 2004). Particularly, the triplet CUG can be used by leucine- or methionine-tRNAs and AAG by methionine-tRNAs to initiate the translation (Kearse & Wilusz, 2017; Starck et al., 2012)**.** Taken together, the potential AUG-proximity effect and translation reinitiation using one of the previously mentioned downstream start codons after the premature termination presumably resulted in evading the NMD in *gnav1*$^{-/-}$ mutant. Consequently, it cannot be excluded that the mutation can still generate an, albeit 5'-incomplete, protein, whose sequence begins from an alternative start codon subsequent to the nonsense mutation. Therefore, we evaluated whether *gnav1*$^{-/-}$ mutants still produce Gv protein.

Extensive work was conducted to validate the custom-made anti-Gv antibody. The final outcome was that the antibody is not specific and possibly also not sensitive enough to detect Gv protein (see 3.3 for details). As an alternative method for detection of Gv protein we next used a conventional MS approach, which however was not able to detect Gv protein in wildtype either, presumably due to low abundance in the sample below the sensitivity of the method (see 3.4.1). Finally, we turned to an advanced MS approach, the PRM assay, which uses the recombinant Gv protein for calibration, and thus achieves a much higher sensitivity.

The PRM data showed that Gv protein was expressed in the wildtype kidney and 2 dpf embryos which is inferred by the presence of the five Gv peptides that cover most of the Gv protein sequence. This is consistent with the detection of Gv mRNA by WISH and qPCR in these tissues and stages and shows that Gv mRNA indeed is translated into Gv protein.

Unexpectedly, all five Gv peptides were also found in the mutant kidneys and 2 dpf embryos, despite the repeatedly confirmed frameshift mutation which results in several early stop codons. Theoretically it could also happen that the stop codon(s) in the mutant frame are read through (Palma & Lejeune, 2021), but since this would not shift

the frame, a completely different protein sequence would result, which would contain none of the five peptides (Figure 16C). Since the mutation is upstream of the first peptide this suggests that the mutants could mostly evade the NMD machinery, and subsequently secondary initiation of translation could take place. This would result in a 5' truncated Gv protein.

Since the first peptide (which spans the first exon-exon border) is at similar levels as the other peptides (which correspond to exons 3-8), the truncated Gv protein encompasses N-terminally at least the first peptide, i.e., starts somewhere in exon1, downstream of the nonsense mutation and upstream of peptide 1 (located in the middle of the β1 sheet). The exact starting point is unclear. If a secondary methionine was used as an alternative start (first possibility of an in-frame AUG in exon 2 at codon 58), the protein sequence would be starting from within exon 2, which can be excluded due to the presence of peptide 1. Alternatively, the leucine at codon 28 (in-frame CUG in exon 1) may have been used for secondary initiation (Figure 65). Codon 28 is downstream of the mutation site and upstream of peptide 1, which is consistent with the presence of peptide 1 in the mutant Gv. In this case, the mutant Gv protein would miss the entire αN helix.

The last peptide is located in exon 8, thus strictly speaking we cannot draw a conclusion concerning the presence of the C-terminus (in exon 9) in the mutant protein. However, there is no reason to assume translation would stop prematurely, thus we expect the mutant Gv protein to be C-terminally complete. Taken together, a 5' truncated Gv protein with likely complete C-terminus is produced in the embryonic and adult mutants.

This raises the question, whether the mutant Gv protein may be expected to be functional. We consider this highly unlikely, for two reasons: firstly, a portion of the N-terminus of the αN motif should be absent from the truncated Gv protein, since the 13 bp deletion of the mutant corresponds to the first part of the αN helix (Figure 66). The αN helix is one of the major contact sites with the receptor which contributes in transmitting the receptor-induced conformational changes to the nucleotide-binding pocket of G alpha subunit (Chung, 2013; Oldham & Hamm, 2008). Therefore, deletion

or truncation of this vital motif in Gv protein would affect the Gv-GPCR binding and subsequently the activation of Gv, thereby interrupting signal transduction in all Gv-mediated intracellular pathways.

Secondly, since the potential alternative initiation should occur in exon 1 downstream of the nonsense mutation, all posttranslational modifications occurring upstream of the mutation site cannot be present in the mutant Gv protein, in particular the crucial lipid modification motifs (Thio-Palmitoylation and N-Myristoylation), which occur at the N-terminus (Figure 66).

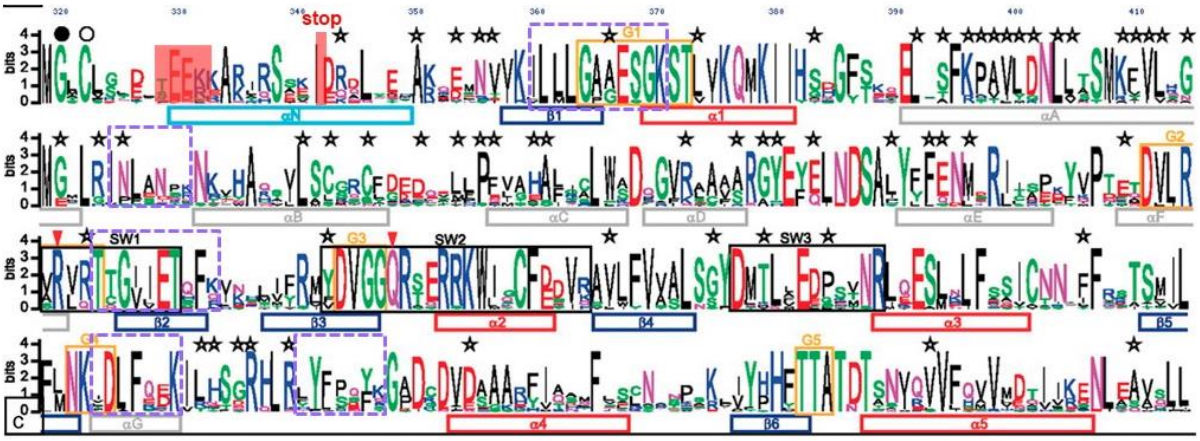


Figure 66: Web logo for Gv proteins illustrates the five peptide sequences used in the PRM assay and the mutant nonsense mutation. The premature stop codon is indicated by the red stop above the thin bar, the corresponding 13 pb deletion is indicated by the red transparent box. The dashed boxes (purple) indicate the position of the peptide sequences used in the PRM assay for Gv protein detection in the wildtype and mutant samples. Secondary structures are indicated under the sequence inside the bars. The corresponding color codes and symbols are same as in Figure 7. Black and white circles above the logo indicates the putative sites for N-linked myristoylation and thio-palmitoylation, respectively, which are possibly absent in Gv mutants. Modified from (Oka et al., 2009).

Both Thio-Palmitoylation and N-Myristoylation motifs have a significant role in promoting membrane tethering and specifying location to the membrane (Rocks et al., 2005; Webb et al., 2000; Wedegaertner & Bourne, 1994). Membrane localization of the protein is essential for mediating its function within the cellular context and mislocalization can offset cellular balance (Alamer et al., 2018). Also, they play a role in regulating proteins stability (Xia et al., 2021), modulating protein-protein interactions (Linder et al., 1991) and enzyme activity (Michaelson et al., 2002; Smotrys & Linder,

2004; Wang et al., 2021). Thus, the mutant's Gv protein would presumably lack the attachment to the plasma membrane, resulting in a drastically lowered effective concentration at the plasma membrane. Consequently, the signal transduction from receptor to effector would be much impaired or even abrogated. Intriguingly, the mutant Gv protein levels were significantly changed in both the mutant adult kidneys and embryos, albeit in opposing ways, suggesting differences in regulation of expression between these organs/stages. In adult kidneys the Gv mRNA level is slightly, but insignificantly decreased in the mutant, whereas the GV protein level is significantly decreased, suggesting a reduced efficiency of the alternative initiation in the mutant kidneys. In contrast, in embryos the Gv mRNA level is significantly decreased, but the Gv protein level is significantly *increased*, possibly as compensation for a nonfunctional mutant Gv protein. Insofar this interpretation is correct, it serves to substantiate our conclusion that the truncated Gv protein produced in the mutants does not function properly compared to the wildtype version.

4.3 Gv may have a role in egg maturation, hatching, and craniofacial development

As discussed above, the aberrant Gv protein is presumably 5' truncated and lacks essential motifs, consequently it is most likely not functional in the mutant. Therefore, we further evaluated the mutant phenotypes and investigated whether there are possible defects in the embryos/larvae and adults, since Gv mRNA is expressed in both adult and embryonic tissues. Indeed, we observed that the *gnav1*$^{-/-}$ mutants exhibit some alterations at various stages including embryos, larvae, and adults. The mutant phenotypes include: the number of eggs produced per clutch was significantly reduced in the adult *gnav1*$^{-/-}$ mutants; the mutant embryos hatched one whole day earlier than the wildtype siblings (highly significant); 3 and 5 dpf mutants showed craniofacial defects and overall decrease in the alcian blue staining (decreased glycoproteins); 5 and 11 dpf mutants also showed bone ossification defects.

Complex interacting factors, including hormonal and environmental factors, contribute to oogenesis, the development and maturation of oocytes (Hoo et al., 2016;

Nagahama & Yamashita, 2008). The ovulation in female zebrafish is induced by the male gonadal pheromones (G. Gerlach, 2006; G. Gerlach & Lysiak, 2006; van den Hurk et al., 1987), and the egg maturation (the progression of the meiotic cell cycle from prophase I to metaphase II) is a prerequisite to the ovulation process (Nagahama & Yamashita, 2008). The male pheromones are perceived by means of olfaction in the female (van den Hurk et al., 1987). Then the luteinizing hormone (LH) is produced by the pituitary gland in the brain which induces the production of maturation-inducing hormone (MIH), a ligand for a GPCR (mPRalpha) signaling through Gi in some fish species, which then activates maturation-promoting factor (MPF) that is the final inducer of oocyte maturation (Nagahama & Yamashita, 2008; Zhu et al., 2003). Thus, defects in G protein signaling could impair egg maturation and subsequently oviposition. Indeed, we observe as prominent phenotype a reduction in clutch size for the Gv mutant compared to the wildtype.

Gv mRNA is present in the adult zebrafish ovaries transcriptome and is expressed in the unfertilized egg (which is still in metaphase II) (Fu et al., 2017; Pasquier et al., 2016) zebrafish and flounder, respectively. Furthermore, Gv is expressed strongly in the testis and weekly in the brain of adult zebrafish (Oka et al., 2009). All three places of expression could have a connection to egg maturation, either directly (ovaries and oocyte) or indirectly (brain for the olfactory processing of male pheromones and testis for the synthesis of male pheromones). Therefore, Gv could have a role in egg maturation (involving meiotic cell divisions, see 4.1), albeit the exact stage(s) where Gv might be involved cannot be concluded from the present data set.

The reduction in mutant clutch size notwithstanding, the fertilized eggs could survive until adulthood which allowed us to study mutant phenotypes in further developmental stages.

The second major phenotype observed in the mutant embryos was a drastically faster hatching rate compared to the wildtype. It occurred in two different media, E3 and artificial freshwater. Hatching is known to be a critical period in zebrafish embryogenesis. The chorion mainly protects the embryo from the surrounding environment, but at some point, the embryo needs to secure nutrients from the

environment before the yolk is depleted. The hatching process involves biochemical softening of the chorion caused by enzymatic degradation together with physical force exerted by the embryo to pierce the softened chorion (Yamagami, 1996). Several factors could influence the hatching rate of zebrafish embryos: (1) either related to chorion hardening/softening by the enzymes - metalloproteases - and membrane proteins; (2) altered punching potency of the embryo; (3) or environmental stressors - oxidative or ionic- which affect the metabolic status of the embryo, leading to less or more expeditious consumption of the yolk, and thus impact hatching (Sano et al., 2008; Small et al., 2020; Trikić et al., 2011; Villamizar et al., 2012; Yamagami, 1996). Under extraordinary circumstances, the earliest developmental time, when the fish is capable of hatching is determined at roughly 36 hpf (Robertson et al., 2014)**.**

We have shown that Gv mRNA is expressed in the hatching gland of 1 and 2 dpf embryos. Also, there was an increase in the mutant Gv protein levels at 2 dpf stage indicating a compensation for the truncated and presumably not functional protein. Therefore, we propose two hypotheses: (1) Gv could be involved in inhibiting the hatching process (hatching gland development and production of hatching enzymes), thus the presence of nonfunctional Gv in the mutants would result in premature release for the hatching enzymes, accordingly triggering earlier hatching. (2) alternatively, since Gv is expressed in the embryonic kidney which begins filtration at 2 dpf, Gv mutation might lead to defect in the internal acid-base balance and ionic homeostasis, consequently ‘emergency’ pathways are triggered, which accelerate the hatching process.

The literature contains several examples of the effect of alteration in acid-base balance, ionic homeostasis, and other stressors on hatching. The changing pH has been shown to influence the hatching rate of the fish eggs (Andrade et al., 2017; Y. Gao et al., 2011), one reason could be the hatching enzyme, whose activity was shown to be pH-dependent in whitefish and carp (EL-Fiky, 2002; Luczynski et al., 1987). Otherwise, the pH might alter the hardness of the chorion and thereby influence hatching. Additionally, ionic stress was also shown to promote earlier hatching in zebrafish embryos (Farhana et al., 2019; Ord, 2019; Sawant et al., 2001) and other fish (Albert et al., 2004; Nadirah et al., 2014; Z. Yang & Chen, 2006). Oxidative stress

also could lead to significant earlier hatching (Robertson et al., 2014; Small et al., 2020; Uren Webster et al., 2014). Overall, it is tempting to propose that Gv influences hatching by both regulating ionic homeostasis, hatching gland differentiation, and hatching enzyme production/release.

The third major phenotype observed in the Gv mutants concerns the cartilage and bone composition at later stages. Alcian blue staining for 3 dpf mutant larvae showed craniofacial defects. Furthermore, the mutant 5 dpf larvae showed craniofacial defects such as gaping jaw, increased angle between ceratohyals, and short ethmoid plate. The mutant larvae also exhibited tissue fragility, and overall decrease in alcian blue staining. The alcian blue binds to proteoglycans, which are essential components of the ECM in all tissues.

Proteoglycans are involved in a variety of cellular processes; in zebrafish they contribute to development and tissue homeostasis (Jessen, 2015). Importantly, they have a role in bone and cartilage formation, also they act as hydrated gel that resists compressive forces, stabilize, and organize collagen fibers in ECM (Couchman & Pataki, 2012; Eames et al., 2011; Frank Eames et al., 2010; Lamoureux et al., 2007). In fact, some of the defects manifesting in proteoglycan mutations (see below) may be related to disorganized collagen fibers, since in zebrafish mutations in collagen showed cartilage and bone defects (Lawrence et al., 2018), and the fish appeared more fragile and easily damaged (Gistelinck et al., 2018). It has been shown that mutations affecting proteoglycan biosynthesis in ECM result in pleiotropic symptoms, including craniofacial dysmorphism, skeletal dysplasia, and skin fragility (Bammens et al., 2015; Delbaere et al., 2020; Malfait et al., 2013; Nakajima et al., 2013). For instance, mutation of TMEM 165, a protein involved in ECM glycosylation and proteoglycans biosynthesis, results in craniofacial defects in zebrafish larvae including altered chondrocyte differentiation and increased angle between ceratohyals (Bammens et al., 2015). Overall, these effects observed in several species bear considerable similarity to the craniofacial defects we observe in the Gv mutants. Moreover, we observed Gv mRNA expression in the mandibular and hyoid cartilages, particularly the chondrocytes in early embryonic development up to and including 3 dpf (see 3.1). Therefore, we may infer that Gv plays a role in the morphogenesis of

the craniofacial cartilages (and more generally in biosynthesis/deposition/stability of ECM components) during early development.

It is well known that there is a correlation between the proper cartilage development and ionic homeostasis and it has been reported that the metabolism of cartilage ECM components is influenced by ionic homeostasis (Mobasheri et al., 1998). Also, the proteins involved in ECM biosynthesis also regulates ionic and pH homeostasis as TMEM 165 (Bammens et al., 2015; Demaegd et al., 2013; Kornak et al., 2008; Rosnoblet et al., 2013). This opens a second potential mechanism for Gv to influence ECM composition, via its role in the kidney in ionic homeostasis (see 4.4). Since the craniofacial defects observed in 3 and 5 dpf mutant larvae coincide with Gv expression in kidney and gut, we suggest that the mutant cartilage defects, tissue fragility, and overall reduced alcian blue staining resulted from ionic homeostasis defects that impact the ECM protein synthesis. For 3 dpf this mechanism may act in addition to a direct effect in cartilage itself, whereas for 5 dpf it seems to be the major mechanism, since Gv expression is no longer detectable in cartilage at this stage.

4.4 Potential role for Gv in cation homeostasis and bone formation

Gv mRNA is prominently present in osmoregulatory organs (gills, kidney, and intestine) of the adult zebrafish (3.1.4) (Oka et al., 2009), and in the kidney, intestine, and YSL (embryonic skin) of zebrafish larvae (3.1.3), thereby Gv could be involved in osmoregulation. These organs are crucial for maintaining the internal acid-base and ionic homeostasis; since the aquatic environment is diverse in the context of ionic contents and pH levels, fish should have the capacity to survive under stressful and harsh conditions. Interestingly, *gnav1*$^{-/-}$ mutants showed some defects that could result from ionic imbalance e.g., hatching and cartilage defects. To investigate this possibility, we examined whether Gv plays a role in ionic/acid-base regulation by evaluating the total body calcium deposition and total body cations for the progeny of wildtype and mutant siblings by ARS staining and atomic absorption spectroscopy,

respectively. Moreover, we analyzed the differential expression of some iono-regulatory genes.

Indeed, 5 dpf and 11 dpf *gnav1*$^{-/-}$ mutants did show defects in bone mineralization, in particular highly reduced calcium deposits in craniofacial bones. In contrast to wildtype, the 11 dpf mutant vertebrae were non-ossified. Consistent with these results, the 5 dpf mutants showed severe deficiency in total body calcium levels in the cation levels measurements, which parallels previous thesis results (Ivandic, 2015). Moreover, the mutants also showed significantly reduced magnesium and potassium levels, and a moderate reduction (did not reach significance) in sodium levels compared to wildtype. The drastic depletion of cations in the mutant larvae corroborates our hypothesis that Gv plays a role in ionic homeostasis.

Calcium and magnesium are the main cations stored in zebrafish bone (Bijvelds et al., 1998), they are also shown to stimulate bone formation (Díaz-Tocados et al., 2017; Qiao et al., 2022). Decrease of these cations is known to be associated with skeletal defects (Elizondo et al., 2010), thereby the impaired bone ossification in the Gv mutants is expected to occur as a result of deficiency in these minerals. This indicates that Gv mutation resulted in cation deficiency and subsequent bone defects. Since cation levels are regulated by a meshwork of ion pumps, transporters and channels, we anticipated altered expression of some of these components in the Gv mutant.

Indeed, the adult mutant kidneys showed (compared to wildtype siblings) increased mRNA levels of NCC and NKA.5. Both proteins play an (indirect) role in calcium and magnesium homeostasis (Arjona et al., 2019; Dong et al., 2020); in particular NCC mutation in humans leads to many defects including hypokalemia, hypomagnesemia, hypocalciuria and metabolic alkalosis (Dong et al., 2020). Thus, the increase in mRNA levels of NCC and NKA.5 appears to be a compensatory mechanism for the decreased divalent cation levels in larvae and supposedly also in adults. The link between NCC and NKA.5 levels reported here is reminiscent of similar links in various other conditions; for instance it has been shown that both NCC and NKA.5 are downregulated in prolactin deficient zebrafish larvae (Shu et al., 2016), and NKA.5 is

upregulated in response to low sodium environments and when zebrafish are deficient in NCC (Esbaugh et al., 2019; Liao et al., 2009).

It should be noted that the direct effect of NCC and NKA.5 is on sodium and potassium levels: NCC reabsorbs sodium and chloride ions and is expressed in the distal tubule of zebrafish (Shu et al., 2016; Wingert et al., 2007); NKA.5 (specifically found in zebrafish) has sodium/potassium-exchanging ATPase activity, it provides a driving force in a variety of osmoregulatory epithelia, and is expressed in the gills, the proximal and distal tubules of zebrafish (Liao et al., 2009; Shu et al., 2016). Hence, increased levels of NCC and NKA.5 coincide with enhanced sodium reabsorption (by both) and potassium secretion (only by NKA.5). Therefore, the decreased potassium levels in the mutant could be a secondary effect caused by the compensatory upregulation of NKA.5 in the mutants. By the same reasoning we would expect enhanced sodium levels in the mutants, which are, however, close to normal in mutant larvae. It is conceivable that the apparently unchanged sodium levels result from a combination of the NCC and NKA.5 increase and another, so far unknown mechanism decreasing sodium levels in the mutants.

Next, we examined larval NKA.5 expression levels in 3 dpf progeny of wildtype and mutant siblings. The larval expression pattern of NKA.5 - including eye, brain, otic vesicle, pectoral fin, gills, and pronephros, see (Shu et al., 2016; Thisse et al*., 2001*), is remarkably similar to that of *gnav1*. The NKA.5 mRNA levels were significantly reduced in mutant 3 dpf larvae which is opposite to the effect seen for NKA.5 in the adult kidney (see above). However, it needs to be considered that body fluid osmoregulation is different between the adult and larvae (Charmantier et al., 2022). Furthermore, the decreased mRNA levels in larvae reflects the cumulative effect of all iono-regulatory tissues: even if NKA.5 mRNA levels might be increased in larval ‘kidney’ (pronephros), a more pronounced reduction in e.g., gills could override that effect.

Such opposing regulation was indeed seen in prolactin-deficient mutants (Shu et al., 2016). NKA.5 levels measured by qPCR were reduced in samples from the whole larvae and also in samples from larval body from which the head was removed thus

indicating that the decrease of NKA.5 is in the pronephros and excluding the gills. Whereas the expression of NKA.5 was shown to be enhanced in the gills of the same mutants when evaluated by WISH (Shu et al., 2016).

Interestingly, NKA pump subunits are highly expressed in human chondrocytes, and it was shown that NKA pump regulates chondrocyte differentiation and bone ossification in mice (Marchini et al., 2021; Mobasheri et al., 2012). It was also implied that NKA regulates ECM homeostasis by modulating the expression of collagens and proteins associated with ECM degradation (Marchini et al., 2021). This indicates that the NKA.5 deficiency in *gnav1*$^{-/-}$ mutant larvae could be the reason for craniofacial defects that we observed.

Another important finding was that the mutant kidney showed decreased mRNA levels of SLC26 which is an anion exchanger that involves chloride uptake in the distal tubules (Bayaa et al., 2009). We propose that SLC26 decrease could be a compensatory response to avoid the chloride overload for the anticipated increase in chloride absorption by NCC upregulation. Another possible explanation could be that Gv is involved in acid-base regulation, thereby the nonfunctional Gv might cause metabolic acidosis, thus SLC26 expression might be decreased in *gnav1*$^{-/-}$ mutants as a response to metabolic acidosis in order to reduce the base excretion (HCO_3^-), see (Bayaa et al., 2009; Lewis & Kwong, 2018; Petrovic et al., 2003). The cations deficiency in mutants could be caused by the disturbed acid-base homeostasis, since increased body fluid acidity is known to affect the renal function and exert wasting of ions such as calcium, magnesium, potassium and sodium (Lewis & Kwong, 2018; Nijenhuis et al., 2006). In zebrafish, acid exposure inhibits calcium uptake and affects the integrity of the epithelial tight junctions resulting in paracellular loss of sodium (Lewis & Kwong, 2018). Therefore, there is a complex interaction between ionic and acid-base homeostasis and it is not clear in which direction Gv exerts its effect.

As mentioned above, organs other than kidney are also involved in cation homeostasis, such as gills which is the main site for calcium uptake in addition to embryonic skin of the yolk sac (Pan et al., 2005); and body fluid osmoregulation is different between the adult and larvae (Charmantier et al., 2022). It would be

informative to examine in future studies the influence of Gv on the expression of cation regulatory genes in the adult gills, and also in larval head (includes gills) *vs.* trunk (includes pronephros) preparations. To thoroughly delineate Gv role in homeostasis, it will also be necessary to measure anionic levels such as chloride, and to investigate additional ion-regulatory genes such as other cation (e.g., magnesium) transporters and acid-base regulatory genes.

Taken together all the findings, Gv mRNA is expressed in several tissues during embryogenesis and in adult zebrafish including osmoregulatory organs. Gv mutation resulted in several defects in the mutants such as diminished oviposition, earlier egg hatching, craniofacial defects, reduced ossification, cation deficiency, and altered expression of kidney ions transporters and pumps in adult fish. These findings suggest that Gv alters ionic/acid-base homeostasis resulting in manifold effects in different organs including extracellular matrix structures. Moreover, Gv mRNA is expressed in the brain of embryos (see 3.1.2) and adult zebrafish (Oka et al., 2009)**,** opening the possibility that Gv might be involved in modulation of hormonal signaling originating in the brain. The manifold effects of Gv described here would be in line with multifaceted functions of other G alpha proteins (Neves et al., 2002; Weinstein, 2001).

Furthermore, some other G proteins and downstream signaling genes seemed to be influenced by Gv mutation in the kidney, hence further experiments are required to examine all other G proteins differential expression in order to fully understand the extent Gv removal has on G protein signaling in general. It has been reported that mutations in other G alpha proteins result in complex phenotypes that include craniofacial and bone defects. For instance, mutations in Gs cause multiple clinical manifestations which include growth, endocrine tissue, and bone defects in humans (Weinstein, 2001). Also, the presence of aberrant Go protein causes multiple neurodevelopmental phenotypes in mice and humans (Kehrl et al., 2014; Talvik et al., 2015); while gain-of-function mutation of mice Go also results in complex phenotypes that involve many organs as bone and brain, and comprise e.g. reduced viability, low birth weight, growth retardation, and behavioral hyperactivity (X. Huang et al., 2006). Mutant phenotypes described for human Gi (Romanelli Tavares et al., 2015) appear closest to the Gv mutation effects in zebrafish. This is interesting, because Gi is the G

protein class most closely related to Gv, and could suggest that Gv might have an inhibitory role in the signaling cascade similar to Gi. A defect in human Gi gives rise to auriculocondylar syndrome, which is characterized by craniofacial defects mainly in the mandibular and hyoid pharyngeal arches (Romanelli Tavares et al., 2015). This is very reminiscent of the distortion seen in zebrafish mandibular and hyoid pharyngeal arches for the Gv mutation.

Taken together we have shown several significant phenotypes for the Gv mutant, many of which may be linked to primary defects in ionic homeostasis. To ascertain that indeed the absence of (functional) Gv (and not some off-target effects or an effect of truncated Gv with a novel function) is causing the observed phenotypes, and as a follow up for this dissertation, we intend to perform a rescue experiment. An analysis of the rescued phenotypes will show, whether the *gnav1*$^{-/-}$ mutant phenotype is Gv-specific. Furthermore, to eliminate the potential complication by the presence of truncated Gv protein in the current knockout we have developed a novel strategy for a whole gene knockout for *gnav1* and simultaneous eGFP knockin (Figure 57), based on (Hisano et al., 2015), who used CRISPR/Cas9 with homologous recombination. So far, we have selected and validated a highly efficient gRNA and prepared the required integration cassette containing *gnav1* homologous arms and the eGFP sequence. Unfortunately, it was not possible to perform the subsequent steps (plasmid co-injection with both gRNAs and Cas9 protein and generating the mutant line) within the time frame of this study. This new approach will be an important validation of results achieved with the first mutant line, and at the same time we will be able to trace the live eGFP expression in Gv-specific tissues thus circumventing the need for a specific Gv antibody (see 3.3). If successful, our novel strategy would be generalizable and could be employed for other genes as well.

4.5 A meticulous characterization of metazoan, particularly protostomean G alpha proteins in nearly one thousand species shows a fractal pattern of gene losses and novel phylum-specific G alpha (sub)classes

This study grew out of the hypothesis, based on the phylogenetic analysis, that the presence of Gv proteins is correlated with the aquatic lifestyle (a single beetle species was the only exception) and also the observation that Gv class seemed to have atypical high number of gene loss events compared to the other four major classes, Gs, Gi, Gq and G12 (Oka et al., 2009). Since then, a plethora of new genomes has become available, thereby, we have re-investigated the hypothesis rigorously whether *gnav* genes have a tendency towards loss in non-aquatic living species. The present study was also designed to examine Gv evolutionary dynamics in relation to the other four G alpha families. Moreover, the literature is scarce on the question of Gv emergence in the earliest-diverging Holozoa, the Teretosporea. In recent years many genomes from this group have become available, enabling us to study the evolutionary origin of the Gv family.

To address all three questions, we constructed the phylogenetic trees by rigorously searching the five G alpha classes in NCBI genomic databases of 873 species from many major phylogenetic subdivisions in Holozoa. We used a representative set of query genes from all five G alpha families, and additionally the Gf sequence of *Drosophila melanogaster*, which is related to Gs, but whose phylogenetic origin was not clear. The tblastn search in WGS led to the identification of thousands of sequences which are filtered according to specific criteria (Materials and Methods 2.9.2) and eventually 4478 *gna* genes were validated.

Origin of the Gv family in the MRCA of all Holozoa

The current study showed that Gv family birth is earlier than previously reported, as we identified Gv orthologs in the earliest-diverging Teretosporea species (Ichthyosporea) (Figure 67). Also, we report the presence of Gv orthologs in the other

unicellular Holozoa, Filasterea and Choanoflagellatea species (Figure 67). The first phylogenetic analysis for Gv class was restricted to Animalia and showed that Gv is present in the MRCA of all animals (Oka et al., 2009). The following study identified Gv orthologs in the unicellular Filasterea and choanoflagellates (de Mendoza et al., 2014), this finding was supported by another two studies (Krishnan et al., 2015; Lokits et al., 2018). Thus, their studies pushed up the origin of the Gv class to Filozoa, nevertheless they concluded that Gv originated in the common ancestor of Holozoa. However, de Mendoza and colleagues did not find Gv orthologs in *Ichthyosporea* using RNAseq data, and the other two studies did not analyze the earliest-diverging Teretosporea (Krishnan et al., 2015; Lokits et al., 2018). This could be attributed to the lack of complete genomic assembly back then (NCBI Resource Coordinators, 2013). On the other hand, the three studies included fungi which corresponds to Holomycota, a basal Opisthokonta clade as sister of the Holozoa (Figure 68). The three studies showed four distinct G alpha proteins (GPA1-4) in fungi which are not orthologous to any of the five G alpha classes of Holozoa (de Mendoza et al., 2014; Krishnan et al., 2015; Lokits et al., 2018). Therefore, taking together the presence of Gv in the earliest-diverging Holoza (Ichthyosporea) [this study] and excluding a birth of Gv in holomycota (de Mendoza et al., 2014; Lokits et al., 2018) suggests that Gv evolutionary origin is in the common ancestor of all Holozoa, prior to the evolution of Filasterea.

Noteworthy, compared to the other families, Gv class appears to be the only one which is present in all of the unicellular groups (Teretosporea, Filasterea, and choanoflagellates) (Figure 67). Interestingly, all of these unicellular species exhibit an aquatic lifestyle; Ichthyosporea are all parasites of aquatic animals including fish and crustaceans (Ragan et al., 1996); the filasterean *Capsaspora* species are endosymbiont amoeba-like protists that lives in the haemolymph of the freshwater snail (a mollusc) (Hertel et al., 2002); the choanoflagellates, the closest unicellular relatives of the animals (Figure 68), lives in different aquatic environments including freshwater and brackish water (Hoffmeyer & Burkhardt, 2016). Therefore, it seems that Gv proteins in the unicellular organism might be correlated with their aquatic lifestyle and perhaps play a role in ionic homeostasis as we observed in zebrafish. In order to examine such a proposed functional role in unicellular Holozoa one might take

the advantage of the genome editing tool (CRISPR/Cas9) that has been recently established for a choanoflagellate species (Booth & King, 2020).

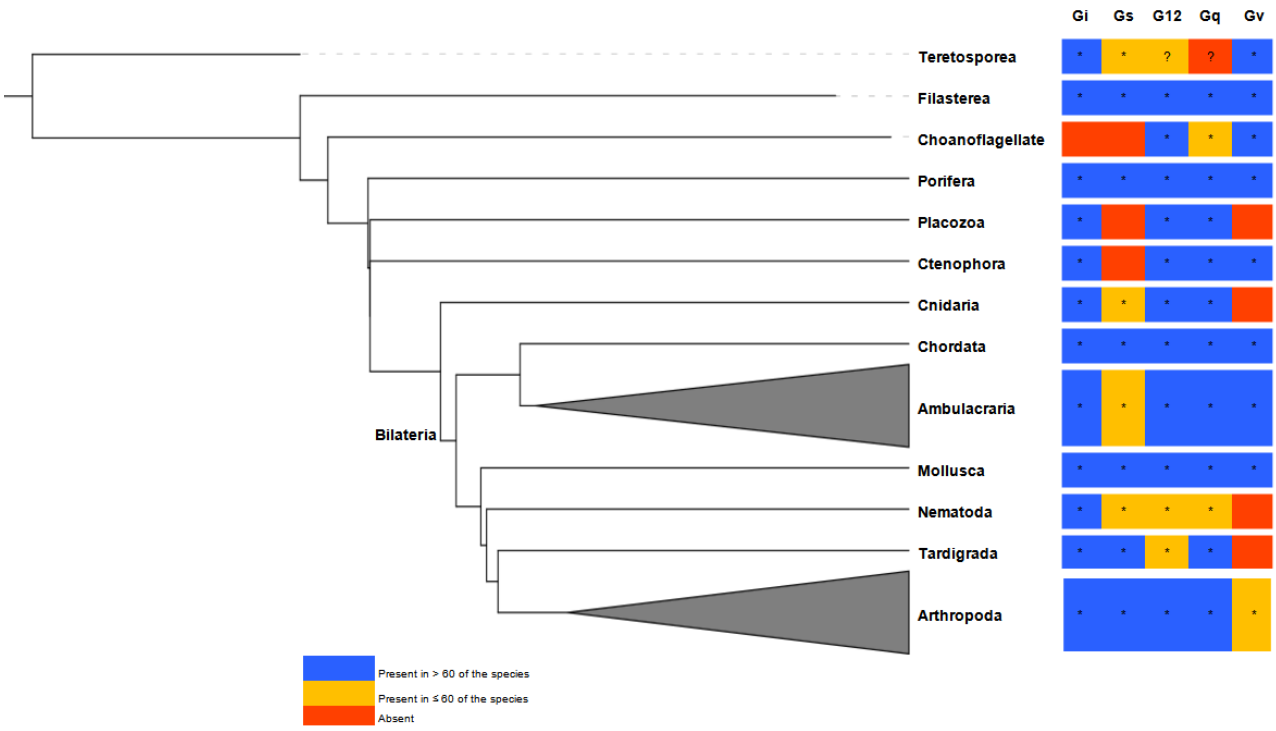


Figure 67: G alpha gene losses visualized in phylogenetic species tree. The clades shown in the tree include all species analyzed in this study. The tree shows the presence, complete loss, and partial loss of the respective G alpha gene in the corresponding clade. Arthropoda and Ambulacraria clades are collapsed. The tree was drawn using the web tool TimeTree by (Kumar et al., 2017). Color code is indicated. * denotes presence of the respective gene; ? denotes uncertainty.

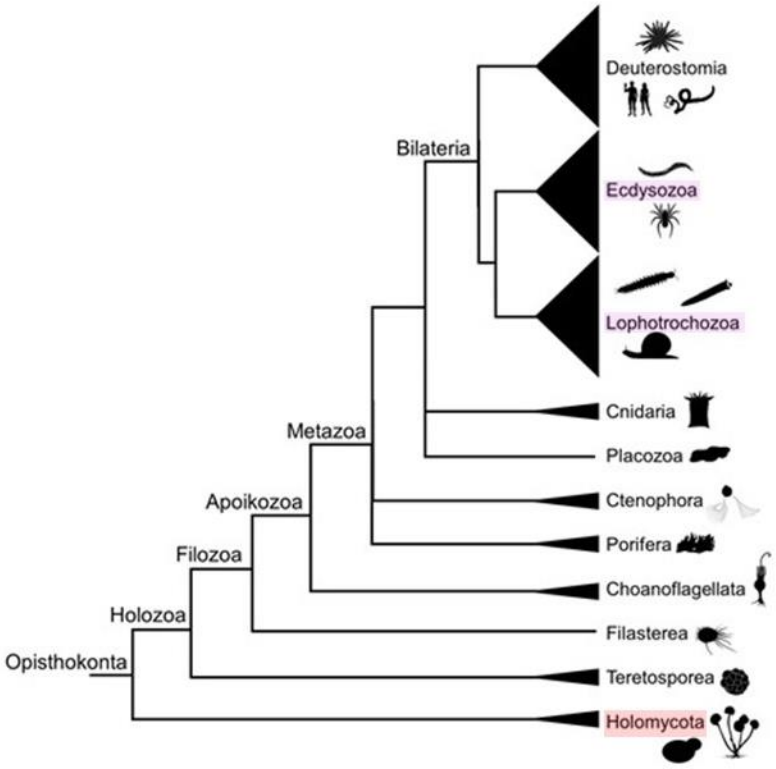


Figure 68: Species tree shows Opisthokonta, which include fungi (Holomycota), and Holozoa (Metazoa and unicellular Holozoa). The red overlay indicates a group analyzed in (de Mendoza et al., 2014; Krishnan et al., 2015; Lokits et al., 2018); the purple overlay emphasizes the basic subdivision of protostomes. Modified from (Paps, 2018).

In this study we also showed the birth of the G alpha families Gi, and Gs in the common ancestor of all Holozoa, whereas Gq and G12 may have diverged later in Filasterea, because the Teretosporea in some trees only contain a G protein ancestral to both Gq and G12 (see below). Moreover, when we analyzed the four G alpha dynamics in the unicellular Holozoa; the Gv class appears to exhibit no gene loss or gain events; the Gi class exhibit several localized gene gains in Teretosporea, and was lost in choanoflagellates (Figure 67). The Gs class showed partial and complete loss events in Ichthyosporea and choanoflagellates, respectively. We also showed that partial gene losses and gains for Gq and G12 have occurred already in early-diverging Holozoa (Ichthyosporea).

The finding we obtained about the Gi family evolutionary origin (emerged in the unicellular Holozoa - Ichthyosporea, Corallochytrea, and Filasterea), is consistent with the previous study (de Mendoza et al., 2014). Similarly, it appears that the Gs family

originates in the common ancestor of unicellular Holozoa, which is also in line with the previous report (de Mendoza et al., 2014). Our finding about the absence of Gi/s genes in choanoflagellates is consistent with both (de Mendoza et al., 2014; Lokits et al., 2018) while (Krishnan et al., 2015) reported uncertainty over the absence of the Gi and Gs genes in the choanoflagellates.

On the other hand, the situation was not conclusive about the evolutionary origin for the Gq/12 families. The Teretosporea (i.e., Ichthyosporea) possess Gq/G12 orthologs, but it is ambiguous whether these are ancestral to both Gq and G12 or whether they constitute G12 orthologs (see 3.8.1). Since (only) the phylogenetic tree placing a single Ichthyosporea gene distinctly basal to both G12 and Gq exhibits maximal branch support (Figure 62) we consider the first possibility more likely. Nevertheless, both Gq and G12 occur as separate clades at least as early as the MRCA of Filozoa, as we identified both Gq and G12 orthologs in Filasterea and in choanoflagellates.

The previous reports about the emergence of Gq/12 orthologs in Filasterea and choanoflagellates are controversial. Krishnan and colleagues reported the presence of one Gq and one G12 orthologs in Filasterea consistent with our findings, on the other hand they reported two Gq genes while indicating uncertainty about the absence of G12 in choanoflagellates (Krishnan et al., 2015) - we find one Gq and two G12 in choanoflagellates. Contrary to us and to (Krishnan et al., 2015), Lokits and colleagues showed the absence of G12 orthologs but the presence of two Gq orthologs in Filasterea, and choanoflagellates possessing two G12 but no Gq orthologs. Moreover, (de Mendoza et al., 2014) reported one gene for each Filasterea and Choanoflagellatea basal to both Gq and G12 clades, also a Gq ortholog in only Filasterea. Overall, these controversial findings might in part be due to different availability, genome coverage and quality of genome assemblies at the time of the respective study. Also, the conflicting results could arise from the different methods, research aims, and different type and number of species used to construct the phylogenetic trees, for instance the three studies cited (de Mendoza et al., 2014; Krishnan et al., 2015; Lokits et al., 2018) included 75, 27, 65 species in total, respectively, while we included nearly thousand species (see Appendix 6.1).

Our results concerning the tree topology and comparing different trees (cf. Figure 61 and Figure 62) could in principle be explained by two hypotheses; (1) the ancestral Gq/G12 gene emerges in the ancestor of Holozoa and a gene duplication within Filasterea resulted in separate Gq and G12 genes; (2) Teretosporea genes are G12 orthologs, which means that both Gq and G12 were already present in the MRCA of Holozoa and the Gq was lost in Teretosporea. We consider the first hypothesis more likely because the gene tree supporting this hypothesis (a single *gna* gene in Ichthyosporea basal to both Gq and G12) exhibits maximal branch support, in contrast to the gene tree supporting the second hypothesis (Ichthyosporea *gna* gene inside the G12 clade). Furthermore, Mendoza and colleagues also reported an Ichthyosporea sequence basal to both Gq and G12 clades (de Mendoza et al., 2014). To ascertain the exact evolutionary origin for the Gq/12 families unambiguously, further analysis is required, particularly when more genomic databases for unicellular species with good quality become available.

Presence of Gv is not correlated to aquatic life style

Next, we used a selection of phyla during animal evolution ranging from the earliest-diverging (Porifera) over non-bilaterian phyla (Placozoa, Ctenophora, and Cnidaria) to early-diverging deuterostomes (Echinoderms and Hemichordata) and several protostomes (Mollusca, Nematoda, Tardigrada, and Arthropoda) to further investigate the aquatic hypothesis. This selection encompasses species that live in different habitats, i.e., terrestrial, aquatic, and semi-aquatic environments. Arthropods are the most derived protostomes and also the most species-rich group, and will be treated separately from the other groups.

The results of this study showed that Gv is present in the aquatic animal sponges and ctenophores (sea gooseberry, sea walnut, and cigar comb jelly), the earliest and second-earliest diverging animals, respectively. This finding indicates that Gv must have been present in the common ancestor of Metazoa which is consistent with the previous report (Oka et al., 2009). In addition, we found Gv genes in the bilaterian Mollusca (giant squid, nautilus, octopus, cuttlefish) and Ambulacraria (acorn worm, starfish, sea urchin). All of these species are aquatic or at most semi-aquatic animals,

thus these findings seem to support the Gv-aquatic hypothesis that the presence of the Gv gene could be correlated with the aquatic lifestyle. However, the non-bilaterian Placozoa (trichoplax) and Cnidaria (hydra, jellyfish, sea anemone, stony coral), and the bilaterian tardigrades live in aquatic environments, but none of them appear to harbor Gv genes. Thus, the aquatic environment is not strictly linked to the retention of Gv.

Another test of the aquatic hypothesis was performed with nematodes. None of the available 101 species genomes (among them some terrestrial and many parasitic species) possess Gv. Again, it appears that there is no correlation between Gv presence/absence and an aquatic environment (which would be the case for internal parasites).

We next examined the species-rich phylum Arthropoda, which contains mostly terrestrial (e.g., hymenopterans) species, but also some aquatic species (e.g., crustaceans). We focused not only on insects but extended the search to non-insect Hexapoda and also to non-Hexapoda, thus representing many major phylogenetic subdivisions with aquatic or terrestrial living.

Our data analysis showed that Gv genes are found in several insects and non-insect groups, whereas they are absent in three insect groups and three non-insect arthropod groups (Figure 63). The **insect** groups which *possess* Gv are Palaeoptera (mayflies, damselflies and dragonflies), Blattodea (cockroaches and termites), Hemiptera (bugs), Coleoptera (beetles), Diptera (flies) and Trichoptera (caddisflies) which all are predominantly terrestrial. Also, the **Non-insect** groups that *harbor* Gv are Arachnida (spiders, scorpions, ticks, and mites) which are mostly terrestrial, Myriapoda (millipedes and centipedes), and Diplura (two-pronged bristletail) that both are terrestrial. From these results, it appears that Gv is present in many of the predominantly terrestrial groups of Arthropoda. On the other hand, the **insect** groups which are terrestrial and exhibit *absence* of Gv are Thysanoptera (thrips), Hymenoptera (sawflies, wasps, bees, and ants), and Lepidoptera (butterflies and moths). Also, the **non-insect** groups that *lost* Gv are Crustacea (crabs, lobsters, crayfish, shrimp, etc.) whose vast majority are aquatic species, and Collembola

(springtails) whose species are terrestrial. Therefore, no correlation is apparent when comparing the presence/absence of Gv to aquatic/terrestrial lifestyle; particularly, the largest groups (Hymenoptera and Lepidoptera) which constitute the majority of the examined insect species are predominantly terrestrial and lost Gv, and at the same time, the mostly aquatic group (Crustacea) also does not possess Gv. Once a database with lifestyle and other parameters for all the different arthropod species would become available, our huge dataset for presence/absence of all G alpha proteins could be usefully looked at and compared to the corresponding species parameters to find any kind of correlation.

It is worth noting that the absence of correlation with aquatic life style does not contradict the role of Gv in ionic homeostasis in the aquatic organism zebrafish, since regulating the ionic homeostasis is not only relevant to aquatic organisms but is extremely important in all biological systems.

The evolutionary dynamics of Gv are within the range observed for other G alpha classes

The third aim of this study was to examine the evolutionary dynamics of Gv in relation to the other G alpha classes. There has been no detailed investigation of the G alpha families in protostome evolution. The recent studies analyzing the evolutionary origin of G alpha proteins focused on deuterostomes and included very few protostome species (de Mendoza et al., 2014; Krishnan et al., 2015; Lokits et al., 2018). Hence, we focused on the Protostome phyla and added selected deuterostomes, non-Bilateria and all earlier-diverging animal species for reference. From the protostomes we included almost all the available genome databases from (Nematoda, Tardigrada, and Arthropoda - together Ecdysozoa) and cephalopods in Mollusca (a phylum in Lophotrochozoa, the sister clade to Ecdysozoa) (Figure 68). Herein, we only consider the major partial loss events that constitute over 60% of the species in the respective phyla.

The size of gene repertoires varies significantly among the G alpha families, both due to gene loss, and to gene birth events. Some gene expansions are phylum-specific,

for instance, the three ctenophores (phylum Ctenophora) possess on average four Gq genes whereas the trichoplax (phylum Placozoa) possesses a single Gq gene. When considering all analyzed species, G12 and Gq have the highest number of *gna* genes and Gv has the smallest repertoire size.

We report several independent gene loss events for Gv class at different levels of metazoan evolution, both early (at phylum level) and late (at order level). This is in contrast to three of four classes of non-Metazoan Holozoa, where Gv is found in all species analyzed (Figure 63 and Table 6) - the absence of Gv in the fourth class, Corallochytrea, may have technical reasons (see 3.8.1). However, to draw a firm conclusion, more species of non-Metazoan Holozoa would need to be identified (and sequenced).

We report three to four independent losses of Gv at phylum level in non-Bilateria (in Placozoa and Cnidaria) and Bilateria (in Nematoda and Tardigrada). Such losses, by their nature, may affect a large number of species (e.g., a million nematode species, *cf*. (Blaxter, 2016). Comparing these losses with the tree of life (Figure 67) shows that in the Ecdysozoa, an independent *gnav* gene loss event must have occurred after the divergence of Nematoda and another after the divergence of Tardigrada. Depending on the placement of Cnidaria in phylogenetic trees one or two gene losses are inferred in non-bilateria: sometimes Cnidaria appear as sister clade to Bilateria (Figure 60), which would necessitate an independent gene loss for Cnidaria, but other trees show Cnidaria more related to Placozoa, which would also be consistent with a single gene loss (Figure 68).

These Gv losses in Metazoa at phylum level are similar to those observed for the other four G alpha families. Gs class appears to have undergone two independent losses in metazoan phyla (Placozoa and Ctenophora). The Gi and Go classes both exhibit one gene loss at phylum level (Gi in sponges and Go in ctenophores), while the Gq and G12 families showed no losses at the phylum level.

Arthropods constitute the most diverged protostome phylum, and due to the large number of genomes available we could analyze the evolutionary dynamics of Gv in

depth in this phylum. Several independent loss events at several taxonomic levels, subphylum, class, and three orders (which constitute the majority of the arthropod species) have led to the loss of Gv in the large majority of arthropod species (Figure 63). In comparison the other four classes of G alpha proteins do not show losses at the subphylum and class levels, and only one loss (Gi in Trichoptera) at the order level. In contrast, gene losses within order (Table 7) occur in Gv well within the range observed for the other four G alpha classes.

Taken together, Gv genes are subjected to frequent losses somewhat more but within the range of the other classes.

Gene family repertoires are determined by the balance of gene loss and gene gain events. Gv gene gains do occur in some metazoan phyla, with an overall frequency similar to Gs, but smaller than observed for Gi, Gq and G12. The Gi class shows two independent gain events at phylum level in Placozoa and Ctenophora (Table 8). Similarly, Gq appears to have undergone two independent gene gains at phylum level (Porifera and Ctenophora) (Table 8). The other families show no gene gain events at phylum level, with the exception of a single gene gain in G12 (Placozoa).

Table 8: Gene gains of G alpha families in different taxonomic groups, all groups belong to Holozoa. The percentage of the species that possess more than one copy of the respective G alpha gene in the species group, is given. Color code follows conditional formatting with three color scale; yellow color denotes absence of gene gains, white was set at 40 percentile, and green color denotes maximal gene gains (100%).

	Classification	Group	# of species	# of genes	Gi		Gs	G12	Gq	Gv
					Go	Gi				
Metazoa	Clade	**Teretosporea**	5	28	0	60	0	20	0	20
	Class	**Filasterea**	1	5	0	0	0	0	0	0
	Class	**Choanoflagellatea**	2	5	0	0	0	0	0	0
	Phylum	**Porifera**	2	16	0	0	50	0	100	50
	Phylum	**Placozoa**	1	7	0	100	0	100	0	0
	Phylum	**Ctenophora**	3	28	0	100	0	0	100	67
	Phylum	**Cnidaria**	6	35	17	33	0	17	17	0
	2 phyla	**Ambulacraria**	6	39	0	50	0	33	0	0
	Phylum	**Chordata**	1	26	100	100	100	100	100	0
	Phylum	**Mollusca**	9	69	33	22	0	0	56	11
	Phylum	**Nematoda**	101	524	2	17	2	1	2	0
	Phylum	**Tardigrada**	2	25	50	50	50	0	50	0
	Phylum	**Arthropoda**	735	3698	3	9	5	18	20	2

At subphylum level all five families show some gene gains, with Gv together with Gs at the lower end of the range. The Gv class show gene gain in four phyla, Porifera, Ctenophora, Mollusca, Arthropoda, with frequency of species ranging between 2 and 67%, similar to Gs gene gains in four phyla, Porifera, Nematoda, Tardigrada and Arthropoda (in 2 to 50% of species in the respective phylum). The G12 family shows gene gains in four phyla, Cnidaria, Ambulacraria, Nematoda, and Arthropoda (1 to 33% of species) (Table 8). The Gi class shows gene duplication events at subphylum level in six phyla (i.e., Cnidaria, Ambulacraria, Mollusca, Nematoda, Tardigrada, and Arthropoda). Gq shows gene gains in five phyla similar to Gi except Ambulacraria (Table 8).

Taken together, it generally appears that the G alpha families evolved dynamically; several gene duplications, but also frequent gene loss were observed during the course of the metazoan evolution. Gi and Gq appear to have undergone fewer gene losses but more duplications, whereas Gv and Gs were subjected to frequent gene loss and fewer gene gains. For G12 the frequency of losses and gains is similar. Gene losses are inevitable, but in order for them to be fixed in the species, they would have to be compensated by corresponding neofunctionalization of paralog genes, or gene losses could also reflect adaptations to environmental changes: they may be either a consequence of environmental adaptations or could even contribute to such adaptations, e.g. a loss of genes with epidermis-related functions in cetaceans as an adaptation to the aquatic environment (Sharma et al., 2018). Similarly, gene gains may allow adaptation to new environments due to potential neofunctionalization of the duplicated genes.

Interestingly, our recursive searches in arthropods unearthed two groups of genes, which present with maximal branch support as sister to the Gs and Gv clades. One group clusters with a *Drosophila melanogaster* gene, which was named Gf (Quan et al., 1993), but has not been classified so far into a G alpha family (Krishnan et al., 2015). In keeping with the original name, we designate this group as class Gf. We identified Gf genes in ten arthropod groups including Diptera (Figure 64). The arthropod groups include **insects**: Diplura, Palaeoptera, Coleoptera, Hymenoptera, Diptera, Trichoptera, and Lepidoptera, and **non-insects**: Arachnida, Crustacea and Collembola (Figure 64). The other group of genes is completely novel, it has to the best of our knowledge not been described in the literature so far. We designate this group as Gw to refer to its close relationship to Gv. Gw genes were found both in insect groups (Lepidoptera, Diptera, and Palaeoptera) and the non-insect Arachnids.

In order to identify the evolutionary origin of the novel genes, we investigated the presence of these genes in other protostome phyla that are earlier-diverging than arthropods: Ecdysozoa (Tardigrada and Nematoda) and Lophotrochozoa (Mollusca) the sister clade to Ecdysozoa. No Gf genes were found in Tardigrada and Nematoda, but two fragments were found in two of six Mollusca species investigated, which appear to be pseudogenes, since the quality of other predicted G alpha genes is

generally very good in these species. Both fragments place consistently with the Gf clade, one of them with maximal branch support. Since molluscs are Lophotrochozoa, it appears that the Gf class is at least as old as the split between the Lophotrochozoa and Ecdysozoa. Since in the gene phylogenetic tree a separate Gs clade comprises the earliest-diverging metazoa (sponges) but no unicellular Holozoa, we suggest that the ancestral Gs/f gene duplicated in the metazoan common ancestor, and that Gs has been kept in many phyla whereas Gf was lost repeatedly in Porifera, Cnidaria, Ctenophora, deuterostomes, and several protostome phyla.

Drosophila melanogaster Gf mRNA is expressed during embryonic stages and in adults only present at low levels, thereby it was suggested to have a role in insect development (Quan et al., 1993). A recent study showed that Gf (here named G alpha 37B) is a downstream effector in Janus kinase/signal transducer and activator of transcription (JAK/STAT) signaling which is known to regulate cell proliferation and haematopoiesis (Bausek & Zeidler, 2014). Therefore, the new class of Gf genes might have a role in metazoan development, but its function may have become unimportant or taken over by other G proteins in most phyla except arthropods. The presence of presumed Gf pseudogenes in only some molluscan species would be consistent with a transitional stage en route to complete gene loss in molluscs.

In reference to the novel Gw class, it appears that this class is arthropod-specific, since we found Gw genes only in Arthropoda and as a sister group to arthropod Gv. No separate Gv and Gw genes were found in earlier-diverging protostomes, or any other metazoan phyla. Within Arthropoda, Gw was found in Arachnids (non-insect) and three of nine orders/infraclasses of insects (Lepidoptera, Diptera, and Palaeoptera) (Figure 64). The presence of both Gv and Gw in Arachnida suggests that duplication of the ancestral Gv/Gw gene must have occurred in the most recent common ancestor of all arthropods. Thereafter, Gw was subjected to frequent independent losses in Myriapoda, Crustacea, Diplura, Blattodea, Hymenoptera, Coleoptera, Trichoptera, and the common ancestor of Thysanoptera, and Hemiptera. In fact, in all three insect orders possessing Gw, only 1-2 species retained Gw, which amounts to 0.6-1.7 % of species in Diptera and Lepidoptera, respectively. Such a low percentage makes it very difficult to generate hypotheses as to possible functions for Gw. In Palaeoptera and

Arachnida about one fifth of species retained Gw. The arachnids that possess Gw all belong to the order Sarcoptiformes (10 out of 12 species), and are all parasitic mites, so a connection with host seeking or adaptation is conceivable.

It is interesting that all the species which possess Gv-related Gw genes also possess Gs-related Gf genes. This could be related to specialized needs of these species for more fine-tuned G protein signalling, which could result in a higher efficiency of retaining both these gene duplications.

4.6 Outlook

The study at hand unveiled the ontogenetic and tissue-specific expression patterns for the novel Gv protein during organogenesis of zebrafish. Also, this research suggested potential roles for Gv protein in the early developing and also in the adult zebrafish. These findings constitute a starting point in performing further investigations such as uncovering the signalling pathway of the fifth G alpha family (Gv): which GPCRs activate Gv, what are the downstream effectors, regulators (RGS), and corresponding beta/gamma subunits? Transcriptomic analysis for wildtype and mutants (using either specific organs such as kidney or whole embryo) might give the opportunity to identify differentially expressed genes, which would be potential candidates for any of these roles. These candidates could be validated by qPCR and fluorescent-double ISH. They could be further analysed by assessing their position in Gene Ontology networks. A promising candidate would be examined for direct interaction with Gv using an *in vitro* model, which is better suited to studies of protein interaction.

Moreover, the ontogenetic expression data could be a foundation for further studies exploring the mechanism of action for Gv during organogenesis, and examining its role in organs other than kidney. For example, the expression of Gv in the inner ear could suggest the role of Gv in hearing/balance.

Furthermore, we have shed light on the evolutionary origin and dynamics for the five G alpha families in a large number of species comprising different taxonomic levels. The huge data set resulting from our study can be employed in comparative genomics

to associate genomic differences with phenotypic differences between species; correlating the genomic dynamics in this way may allow extraction of information about function of the different G alpha genes in a wide evolutionary context.

Noteworthy, we used a half-automated approach to retrieve the amino acid sequences for the candidate genes. This approach allowed the analysis of a much larger number of species compared to previous evolutionary studies of G alpha proteins (de Mendoza et al., 2014; Krishnan et al., 2015; Lokits et al., 2018; Oka et al., 2009), and has already enabled the identification and delineation of novel G alpha classes and subclasses. Our approach can be easily extended to other gene families and species to identify novel protein clades, and thus could serve more generally to provide new insights paving the way for future experimental studies.

5 References

Aceto, J., Nourizadeh-Lillabadi, R., Marée, R., Dardenne, N., Jeanray, N., Wehenkel, L., Aleström, P., Loon, J. J. W. A., & Muller, M. (2015). Zebrafish Bone and General Physiology Are Differently Affected by Hormones or Changes in Gravity. *PLoS ONE*, *10*. https://doi.org/10.1371/journal.pone.0126928

Adolphe, M., Thenet-gauci, S., & Demignot, S. (1997). 8 - Chondrocyte Culture: A Target System to Evaluate: Pharmacotoxicological Effects of Drugs. In J. V. Castell & M. J. Gómez-Lechón (Eds.), *In Vitro Methods in Pharmaceutical Research* (pp. 181–207). Academic Press. https://doi.org/10.1016/B978-012163390-5.50009-0

Albert, A., Vetemaa, M., & Saat, T. (2004). Effects of salinity on the development of Peipsi whitefish Coregonus lavaretus maraenoides Poljakow embryos. *Annales Zoologici Fennici*, *41*(1), 85–88. https://www.jstor.org/stable/23736190

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Andrade, T. S., Henriques, J. F., Almeida, A. R., Soares, A. M. V. M., Scholz, S., & Domingues, I. (2017). Zebrafish embryo tolerance to environmental stress factors—Concentration–dose response analysis of oxygen limitation, pH, and UV-light irradiation. *Environmental Toxicology and Chemistry*, *36*(3), 682–690. https://doi.org/10.1002/etc.3579

Arjona, F. J., Latta, F., Mohammed, S. G., Thomassen, M., van Wijk, E., Bindels, R. J. M., Hoenderop, J. G. J., & de Baaij, J. H. F. (2019). SLC41A1 is essential for magnesium homeostasis in vivo. *Pflügers Archiv - European Journal of Physiology*, *471*(6), 845–860. https://doi.org/10.1007/s00424-018-2234-9

Aulehla, A., Wehrle, C., Brand-Saberi, B., Kemler, R., Gossler, A., Kanzler, B., & Herrmann, B. G. (2003). Wnt3a Plays a Major Role in the Segmentation Clock Controlling Somitogenesis. *Developmental Cell*, *4*(3), 395–406. https://doi.org/10.1016/S1534-5807(03)00055-8

Bammens, R., Mehta, N., Race, V., Foulquier, F., Jaeken, J., Tiemeyer, M., Steet, R., Matthijs, G., & Flanagan-Steet, H. (2015). Abnormal cartilage development and altered N-glycosylation in Tmem165-deficient zebrafish mirrors the phenotypes associated with TMEM165-CDG. *Glycobiology*, *25*(6), 669–682. https://doi.org/10.1093/glycob/cwv009

Barker, N., Tan, S., & Clevers, H. (2013). Lgr proteins in epithelial stem cell biology. *Development*, *140*(12), 2484–2494. https://doi.org/10.1242/dev.083113

Bassilana, F., Nash, M., & Ludwig, M.-G. (2019). Adhesion G protein-coupled receptors: Opportunities for drug discovery. *Nature Reviews Drug Discovery*, *18*(11), 869–884. https://doi.org/10.1038/s41573-019-0039-y

Bausek, N., & Zeidler, M. P. (2014). Gα73B is a downstream effector of JAK/STAT signalling and a regulator of Rho1 in Drosophila haematopoiesis. *Journal of Cell Science*, *127*(1), 101–110. https://doi.org/10.1242/jcs.132852

Bayaa, M., Vulesevic, B., Esbaugh, A., Braun, M., Ekker, M. E., Grosell, M., & Perry, S. F. (2009). The involvement of SLC26 anion transporters in chloride uptake in zebrafish (Danio rerio) larvae. *Journal of Experimental Biology*, *212*(20), 3283–3295. https://doi.org/10.1242/jeb.033910

Bedell, V. M., Wang, Y., Campbell, J. M., Poshusta, T. L., Starker, C. G., Krug II, R. G., Tan, W., Penheiter, S. G., Ma, A. C., Leung, A. Y. H., Fahrenkrug, S. C., Carlson, D. F., Voytas, D. F., Clark, K. J., Essner, J. J., & Ekker, S. C. (2012). In vivo genome editing using a high-efficiency TALEN system. *Nature*, *491*(7422), 114–118. https://doi.org/10.1038/nature11537

Bijvelds, null, Velden, null, Kolar, null, & Flik, null. (1998). Magnesium transport in freshwater teleosts. *The Journal of Experimental Biology*, *201 (Pt 13)*, 1981–1990. https://doi.org/10.1242/jeb.201.13.1981

Bird, L. (2012). *OPPF-UK Standard Protocols: Cloning and Expression Screening*. 26.

Birkholz, D. A., Olesnicky Killian, E. C., George, K. M., & Artinger, K. B. (2009). Prdm1a is necessary for posterior pharyngeal arch development in zebrafish. *Developmental Dynamics : An Official Publication of the American Association of Anatomists*, *238*(10), 2575–2587. https://doi.org/10.1002/dvdy.22090

Blaxter, M. (2016). Imagining Sisyphus happy: DNA barcoding and the unnamed majority. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1702), 20150329. https://doi.org/10.1098/rstb.2015.0329

Blumer, J. B., Oner, S. S., & Lanier, S. M. (2012). Group II activators of G-protein signalling and proteins containing a G-protein regulatory motif. *Acta Physiologica*, *204*(2), 202–218. https://doi.org/10.1111/j.1748-1716.2011.02327.x

Booth, D. S., & King, N. (2020). Genome editing enables reverse genetics of multicellular development in the choanoflagellate Salpingoeca rosetta. *ELife*, *9*, e56193. https://doi.org/10.7554/eLife.56193

Bourmaud, A., Gallien, S., & Domon, B. (2016). Parallel reaction monitoring using quadrupole-Orbitrap mass spectrometer: Principle and applications. *PROTEOMICS*, *16*(15–16), 2146–2159. https://doi.org/10.1002/pmic.201500543

Breacker, C., Barber, I., Norton, W. H. J., McDearmid, J. R., & Tilley, C. A. (2017). A Low-Cost Method of Skin Swabbing for the Collection of DNA Samples from Small Laboratory Fish. *Zebrafish*, *14*(1), 35–41. https://doi.org/10.1089/zeb.2016.1348

Brown, L. A., Rodaway, A. R. F., Schilling, T. F., Jowett, T., Ingham, P. W., Patient, R. K., & Sharrocks, A. D. (2000). Insights into early vasculogenesis revealed by expression of the ETS-domain transcription factor Fli-1 in wild-type and mutant zebrafish embryos. *Mechanisms of Development*, *90*(2), 237–252. https://doi.org/10.1016/S0925-4773(99)00256-7

Bruder, V., Ludwig, T., Opitz, S., Christoffels, R., Fischer, T., & Maleki, H. (2021). Hierarchical Assembly of Surface Modified Silk Fibroin Biomass into Micro-, and Milli-Metric Hybrid Aerogels with Core-Shell, Janus, and Composite Configurations for Rapid Removal of Water Pollutants. *Advanced Materials Interfaces*, *8*(5), 2001892. https://doi.org/10.1002/admi.202001892

Buckles, G. R., Thorpe, C. J., Ramel, M.-C., & Lekven, A. C. (2004). Combinatorial Wnt control of zebrafish midbrain–hindbrain boundary formation. *Mechanisms of Development*, *121*(5), 437–447. https://doi.org/10.1016/j.mod.2004.03.026

Buisson, M., Anczuków, O., Zetoune, A. B., Ware, M. D., & Mazoyer, S. (2006). The 185delAG mutation (c.68_69delAG) in the BRCA1 gene triggers translation reinitiation at a downstream AUG codon. *Human Mutation*, *27*(10), 1024–1029. https://doi.org/10.1002/humu.20384

Carmon, K. S., Gong, X., Lin, Q., Thomas, A., & Liu, Q. (2011). R-spondins function as ligands of the orphan receptors LGR4 and LGR5 to regulate Wnt/ -catenin signaling. *Proceedings of the National Academy of Sciences*, *108*(28), 11452–11457. https://doi.org/10.1073/pnas.1106083108

Carter, K. M., Woodley, C. M., & Brown, R. S. (2011). A review of tricaine methanesulfonate for anesthesia of fish. *Reviews in Fish Biology and Fisheries*, *21*(1), 51–59. https://doi.org/10.1007/s11160-010-9188-0

Charmantier, G., Nguyen-Chi, M., & Lutfalla, G. (2022). Ontogenetic Changes in Blood Osmolality During the Postembryonic Development of Zebrafish (Danio rerio). *Zebrafish*. https://doi.org/10.1089/zeb.2021.0075

Chung, K. Y. (2013). Structural Aspects of GPCR-G Protein Coupling. *Toxicological Research*, *29*(3), 149–155. https://doi.org/10.5487/TR.2013.29.3.149

Clevers, H. (2006). Wnt/β-Catenin Signaling in Development and Disease. *Cell*, *127*(3), 469–480. https://doi.org/10.1016/j.cell.2006.10.018

Couchman, J. R., & Pataki, C. A. (2012). An Introduction to Proteoglycans and Their Localization. *Journal of Histochemistry and Cytochemistry*, *60*(12), 885–897. https://doi.org/10.1369/0022155412464638

Cox, A. A., Jezewski, P. A., Fang, P.-K., & Payne-Ferreira, T. L. (2010). Zebrafish Wnt9a,9b paralog comparisons suggest ancestral roles for Wnt9 in neural, oral–pharyngeal ectoderm and mesendoderm. *Gene Expression Patterns*, *10*(6), 251–258. https://doi.org/10.1016/j.gep.2010.05.005

Curtin, E., Hickey, G., Kamel, G., Davidson, A. J., & Liao, E. C. (2011). Zebrafish wnt9a is expressed in pharyngeal ectoderm and is required for palate and lower jaw development. *Mechanisms of Development*, *128*(1), 104–115. https://doi.org/10.1016/j.mod.2010.11.003

Damme, P. V., Gawron, D., Criekinge, W. V., & Menschaert, G. (2014). N-terminal Proteomics and Ribosome Profiling Provide a Comprehensive View of the Alternative Translation Initiation Landscape in Mice and Men *. *Molecular & Cellular Proteomics*, *13*(5), 1245–1261. https://doi.org/10.1074/mcp.M113.036442

de Mendoza, A., Sebé-Pedrós, A., & Ruiz-Trillo, I. (2014). The Evolution of the GPCR Signaling System in Eukaryotes: Modularity,

Conservation, and the Transition to Metazoan Multicellularity. *Genome Biology and Evolution*, *6*(3), 606–619. https://doi.org/10.1093/gbe/evu038

Delbaere, S., De Clercq, A., Mizumoto, S., Noborn, F., Bek, J. W., Alluyn, L., Gistelinck, C., Syx, D., Salmon, P. L., Coucke, P. J., Larson, G., Yamada, S., Willaert, A., & Malfait, F. (2020). B3galt6 Knock-Out Zebrafish Recapitulate β3GalT6-Deficiency Disorders in Human and Reveal a Trisaccharide Proteoglycan Linkage Region. *Frontiers in Cell and Developmental Biology*, *8*, 1524. https://doi.org/10.3389/fcell.2020.597857

Demaegd, D., Foulquier, F., Colinet, A.-S., Gremillon, L., Legrand, D., Mariot, P., Peiter, E., Schaftingen, E. V., Matthijs, G., & Morsomme, P. (2013). Newly characterized Golgi-localized family of proteins is involved in calcium and pH homeostasis in yeast and human cells. *Proceedings of the National Academy of Sciences*, *110*(17), 6859–6864. https://doi.org/10.1073/pnas.1219871110

Díaz-Tocados, J. M., Herencia, C., Martínez-Moreno, J. M., Montes de Oca, A., Rodríguez-Ortiz, M. E., Vergara, N., Blanco, A., Steppan, S., Almadén, Y., Rodríguez, M., & Muñoz-Castañeda, J. R. (2017). Magnesium Chloride promotes Osteogenesis through Notch signaling activation and expansion of Mesenchymal Stem Cells. *Scientific Reports*, *7*, 7839. https://doi.org/10.1038/s41598-017-08379-y

Dong, B., Chen, Y., Liu, X., Wang, Y., Wang, F., Zhao, Y., Sun, X., & Zhao, W. (2020). Identification of compound mutations of SLC12A3 gene in a Chinese pedigree with Gitelman syndrome exhibiting Bartter syndrome-liked phenotypes. *BMC Nephrology*, *21*(1), 328. https://doi.org/10.1186/s12882-020-01996-2

Dorsam, R. T., & Gutkind, J. S. (2007). G-protein-coupled receptors and cancer. *Nature Reviews Cancer*, *7*(2), 79–94. https://doi.org/10.1038/nrc2069

Dougherty, M., Kamel, G., Grimaldi, M., Gfrerer, L., Shubinets, V., Ethier, R., Hickey, G., Cornell, R. A., & Liao, E. C. (2013). Distinct requirements for wnt9a and irf6 in extension and integration mechanisms during zebrafish palate morphogenesis. *Development*, *140*(1), 76–81. https://doi.org/10.1242/dev.080473

Drummond, I. A., Majumdar, A., Hentschel, H., Elger, M., Solnica-Krezel, L., Schier, A. F., Neuhauss, S. C., Stemple, D. L., Zwartkruis, F., Rangini, Z., Driever, W., & Fishman, M. C. (1998). Early development of the zebrafish pronephros and analysis of mutations affecting pronephric function. *Development*, *125*(23), 4655–4667. https://doi.org/10.1242/dev.125.23.4655

Dyle, M. C., Kolakada, D., Cortazar, M. A., & Jagannathan, S. (2020). How to get away with nonsense: Mechanisms and consequences of escape from nonsense-mediated RNA decay. *WIREs RNA*, *11*(1), e1560. https://doi.org/10.1002/wrna.1560

Eames, B. F., Yan, Y.-L., Swartz, M. E., Levic, D. S., Knapik, E. W., Postlethwait, J. H., & Kimmel, C. B. (2011). Mutations in fam20b and xylt1 Reveal That Cartilage Matrix Controls Timing of Endochondral Ossification by Inhibiting Chondrocyte Maturation. *PLoS Genetics*, *7*(8), e1002246. https://doi.org/10.1371/journal.pgen.1002246

Eisenmann, D. M. (2005). Wnt signaling. *WormBook*. https://doi.org/10.1895/wormbook.1.7.1

EL-Fiky, N. K. (2002). The influence of water phonthe embryonic development of grass carp, tenopharyngodonidell *4*. 28.

Elizondo, M. R., Budi, E. H., & Parichy, D. M. (2010). Trpm7 Regulation of in Vivo Cation Homeostasis and Kidney Function Involves Stanniocalcin 1 and fgf23. *Endocrinology*, *151*(12), 5700–5709. https://doi.org/10.1210/en.2010-0853

Ellis, K., Bagwell, J., & Bagnat, M. (2013). Notochord vacuoles are lysosome-related organelles that function in axis and spine morphogenesis. *The Journal of Cell Biology*, *200*(5), 667–679. https://doi.org/10.1083/jcb.201212095

Ellis, K., Hoffman, B. D., & Bagnat, M. (2013). The vacuole within: How cellular organization dictates notochord function. *Bioarchitecture*, *3*(3), 64–68. https://doi.org/10.4161/bioa.25503

Esbaugh, A. J., Brix, K. V., & Grosell, M. (2019). Na^+ K^+ ATPase isoform switching in zebrafish during transition to dilute freshwater habitats. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1903), 20190630. https://doi.org/10.1098/rspb.2019.0630

Essner, J. J., Amack, J. D., Nyholm, M. K., Harris, E. B., & Yost, H. J. (2005). Kupffer's vesicle is a ciliated organ of asymmetry in the zebrafish embryo that initiates left-right development of the brain, heart and gut. *Development*, *132*(6), 1247–1260. https://doi.org/10.1242/dev.01663

ExPASy—Translate tool. (2019). https://web.expasy.org/translate/

Farhana, T., Haque, F., Amin, F. B., Zahangir, Md. M., & Islam, M. S. (2019). Developmental pliability in zebrafish: An experimental enquiry of acute salinity stress on the early life of zebrafish.

Aquaculture Reports, *14*, 100189. https://doi.org/10.1016/j.aqrep.2019.100189

Fatscher, T., Boehm, V., & Gehring, N. H. (2015). Mechanism, factors, and physiological role of nonsense-mediated mRNA decay. *Cellular and Molecular Life Sciences*, *72*(23), 4523–4544. https://doi.org/10.1007/s00018-015-2017-9

Flik, G., Bevelander, G., & Klaren, P. (2009). *Regulation of calcium and magnesium handling in fishes* (pp. 151–180).

Folgueira, M., Bayley, P., Navratilova, P., Becker, T. S., Wilson, S. W., & Clarke, J. D. (2012). Morphogenesis underlying the development of the everted teleost telencephalon. *Neural Development*, *7*(1), 212. https://doi.org/10.1186/1749-8104-7-32

Forbes, A. A., Bagley, R. K., Beer, M. A., Hippee, A. C., & Widmayer, H. A. (2018). Quantifying the unquantifiable: Why Hymenoptera, not Coleoptera, is the most speciose animal order. *BMC Ecology*, *18*(1), 21. https://doi.org/10.1186/s12898-018-0176-x

Frank Eames, B., Singer, A., Smith, G. A., Wood, Z. A., Yan, Y.-L., He, X., Polizzi, S. J., Catchen, J. M., Rodriguez-Mari, A., Linbo, T., Raible, D. W., & Postlethwait, J. H. (2010). UDP xylose synthase 1 is required for morphogenesis and histogenesis of the craniofacial skeleton. *Developmental Biology*, *341*(2), 400–415. https://doi.org/10.1016/j.ydbio.2010.02.035

Fredriksson, R., & Schiöth, H. B. (2005). The Repertoire of G-Protein–Coupled Receptors in Fully Sequenced Genomes. *Molecular Pharmacology*, *67*(5), 1414–1425. https://doi.org/10.1124/mol.104.009001

Fu, Y., Jia, L., Shi, Z., Zhang, J., & Li, W. (2017). Gene expression patterns regulating embryogenesis based on the integrated de novo transcriptome assembly of the Japanese flounder. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, *22*, 58–66. https://doi.org/10.1016/j.cbd.2017.01.003

Gao, X., Wan, J., Liu, B., Ma, M., Shen, B., & Qian, S.-B. (2015). Quantitative profiling of initiating ribosomes in vivo. *Nature Methods*, *12*(2), 147–153. https://doi.org/10.1038/nmeth.3208

Gao, Y., Kim, S.-G., & Lee, J.-Y. (2011). Effects of pH on Fertilization and the Hatching Rates of Far Eastern Catfish Silurus asotus. *Fisheries and Aquatic Sciences*, *14*(4), 417–420. https://doi.org/10.5657/FAS.2011.0417

Garcia-Marcos, M. (2021). Complementary biosensors reveal different G-protein signaling modes triggered by GPCRs and non-receptor activators. *ELife*, *10*, e65620. https://doi.org/10.7554/eLife.65620

Gays, D., Hess, C., Camporeale, A., Ala, U., Provero, P., Mosimann, C., & Santoro, M. M. (2017). An exclusive cellular and molecular network governs intestinal smooth muscle cell differentiation in vertebrates. *Development*, *144*(3), 464–478. https://doi.org/10.1242/dev.133926

Gering, M., Yamada, Y., Rabbitts, T. H., & Patient, R. K. (2003). Lmo2 and Scl/Tal1 convert non-axial mesoderm into haemangioblasts which differentiate into endothelial cells in the absence of Gata1. *Development*, *130*(25), 6187–6199. https://doi.org/10.1242/dev.00875

Gerlach, G. (2006). Pheromonal regulation of reproductive success in female zebrafish: Female suppression and male enhancement. *Animal Behaviour*, *72*(5), 1119–1124. https://doi.org/10.1016/j.anbehav.2006.03.009

Gerlach, G. F., & Wingert, R. A. (2013). Kidney organogenesis in the zebrafish: Insights into vertebrate nephrogenesis and regeneration. *WIREs Developmental Biology*, *2*(5), 559–585. https://doi.org/10.1002/wdev.92

Gerlach, G., & Lysiak, N. (2006). Kin recognition and inbreeding avoidance in zebrafish, Danio rerio, is based on phenotype matching. *Animal Behaviour*, *71*(6), 1371–1377. https://doi.org/10.1016/j.anbehav.2005.10.010

Gistelinck, C., Kwon, R. Y., Malfait, F., Symoens, S., Harris, M. P., Henke, K., Hawkins, M. B., Fisher, S., Sips, P., Guillemyn, B., Bek, J. W., Vermassen, P., Saffel, H. D., Witten, P. E., Weis, M., Paepe, A. D., Eyre, D. R., Willaert, A., & Coucke, P. J. (2018). Zebrafish type I collagen mutants faithfully recapitulate human type I collagenopathies. *Proceedings of the National Academy of Sciences*, *115*(34), E8037–E8046. https://doi.org/10.1073/pnas.1722200115

Glinka, A., Dolde, C., Kirsch, N., Huang, Y.-L., Kazanskaya, O., Ingelfinger, D., Boutros, M., Cruciat, C.-M., & Niehrs, C. (2011). LGR4 and LGR5 are R-spondin receptors mediating Wnt/β-catenin and Wnt/PCP signalling. *EMBO Reports*, *12*(10), 1055–1061. https://doi.org/10.1038/embor.2011.175

Goldsmith, Z. G., & Dhanasekaran, D. N. (2007). G Protein regulation of MAPK networks. *Oncogene*, *26*(22), 3122–3142. https://doi.org/10.1038/sj.onc.1210407

Gruber, J., Yee, Z., & Tolwinski, N. S. (2016). Developmental Drift and the Role of Wnt

Signaling in Aging. *Cancers*, *8*(8), 73. https://doi.org/10.3390/cancers8080073

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology*, *59*(3), 307–321. https://doi.org/10.1093/sysbio/syq010

Hariharan, A., Weir, N., Robertson, C., He, L., Betsholtz, C., & Longden, T. A. (2020). The Ion Channel and GPCR Toolkit of Brain Capillary Pericytes. *Frontiers in Cellular Neuroscience*, *14*. https://www.frontiersin.org/article/10.3389/fncel.2020.601324

Hart, N. H., & Fluck, R. A. (1996). 11 Cytoskeleton in Teleost Eggs and Early Embryos: Contributions to Cytoarchitecture and Motile Events. In D. G. Capco (Ed.), *Current Topics in Developmental Biology* (Vol. 31, pp. 343–381). Academic Press. https://doi.org/10.1016/S0070-2153(08)60233-1

Hatzold, J., Wessendorf, H., Pogoda, H.-M., Bloch, W., & Hammerschmidt, M. (2021). The Kunitz-type serine protease inhibitor Spint2 is required for cellular cohesion, coordinated cell migration and cell survival during zebrafish hatching gland development. *Developmental Biology*, *476*, 148–170. https://doi.org/10.1016/j.ydbio.2021.03.017

Hauser, A. S., Attwood, M. M., Rask-Andersen, M., Schiöth, H. B., & Gloriam, D. E. (2017). Trends in GPCR drug discovery: New agents, targets and indications. *Nature Reviews Drug Discovery*, *16*(12), 829–842. https://doi.org/10.1038/nrd.2017.178

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F., & Vandesompele, J. (2007). QBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biology*, *8*(2), R19. https://doi.org/10.1186/gb-2007-8-2-r19

Hertel, L. A., Bayne, C. J., & Loker, E. S. (2002). The symbiont Capsaspora owczarzaki, nov. Gen. Nov. Sp., isolated from three strains of the pulmonate snail Biomphalaria glabrata is related to members of the Mesomycetozoea. *International Journal for Parasitology*, *32*(9), 1183–1191. https://doi.org/10.1016/S0020-7519(02)00066-8

Hilger, D., Masureel, M., & Kobilka, B. K. (2018). Structure and dynamics of GPCR signaling complexes. *Nature Structural & Molecular Biology*, *25*(1), 4–12. https://doi.org/10.1038/s41594-017-0011-7

Hillegass, J. M., Villano, C. M., Cooper, K. R., & White, L. A. (2008). Glucocorticoids Alter Craniofacial Development and Increase Expression and Activity of Matrix Metalloproteinases in Developing Zebrafish (Danio rerio). *Toxicological Sciences*, *102*(2), 413–424. https://doi.org/10.1093/toxsci/kfn010

Hirose, K., Shimoda, N., & Kikuchi, Y. (2011). Expression patterns of lgr4 and lgr6 during zebrafish development. *Gene Expression Patterns*, *11*(7), 378–383. https://doi.org/10.1016/j.gep.2011.04.002

Hisano, Y., Sakuma, T., Nakade, S., Ohga, R., Ota, S., Okamoto, H., Yamamoto, T., & Kawahara, A. (2015). Precise in-frame integration of exogenous DNA mediated by CRISPR/Cas9 system in zebrafish. *Scientific Reports*, *5*(1), 8841. https://doi.org/10.1038/srep08841

Hoffmeyer, T. T., & Burkhardt, P. (2016). Choanoflagellate models—Monosiga brevicollis and Salpingoeca rosetta. *Current Opinion in Genetics & Development*, *39*, 42–47. https://doi.org/10.1016/j.gde.2016.05.016

Holley, S. A., Geisler, R., & Nüsslein-Volhard, C. (2000). Control of her1 expression during zebrafish somitogenesis by a Delta-dependent oscillator and an independent wave-front activity. *Genes & Development*, *14*(13), 1678–1690. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC316735/

Hoo, J. Y., Kumari, Y., Shaikh, M. F., Hue, S. M., & Goh, B. H. (2016). Zebrafish: A Versatile Animal Model for Fertility Research. *BioMed Research International*, *2016*, 9732780. https://doi.org/10.1155/2016/9732780

Horsfield, J., Ramachandran, A., Reuter, K., LaVallie, E., Collins-Racie, L., Crosier, K., & Crosier, P. (2002). Cadherin-17 is required to maintain pronephric duct integrity during zebrafish development. *Mechanisms of Development*, *115*(1), 15–26. https://doi.org/10.1016/S0925-4773(02)00094-1

Huang, L., Zhang, H., Deng, D., Zhao, K., Liu, K., Hendrix, D. A., & Mathews, D. H. (2019). LinearFold: Linear-time approximate RNA folding by 5'-to-3' dynamic programming and beam search. *Bioinformatics*, *35*(14), i295–i304. https://doi.org/10.1093/bioinformatics/btz375

Huang, X., Fu, Y., Charbeneau, R. A., Saunders, T. L., Taylor, D. K., Hankenson, K. D., Russell, M. W., D'Alecy, L. G., & Neubig, R. R. (2006). Pleiotropic Phenotype of a Genomic Knock-In

of an RGS-Insensitive G184S Gnai2 Allele. *Molecular and Cellular Biology*. https://doi.org/10.1128/MCB.00314-06

Hwang, P.-P., & Chou, M.-Y. (2013). Zebrafish as an animal model to study ion homeostasis. *Pflugers Archiv*, *465*(9), 1233–1247. https://doi.org/10.1007/s00424-013-1269-1

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., Peterson, R. T., Yeh, J.-R. J., & Joung, J. K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature Biotechnology*, *31*(3), 227–229. https://doi.org/10.1038/nbt.2501

Inácio, Â., Silva, A. L., Pinto, J., Ji, X., Morgado, A., Almeida, F., Faustino, P., Lavinha, J., Liebhaber, S. A., & Romão, L. (2004). Nonsense Mutations in Close Proximity to the Initiation Codon Fail to Trigger Full Nonsense-mediated mRNA Decay*. *Journal of Biological Chemistry*, *279*(31), 32170–32180. https://doi.org/10.1074/jbc.M405024200

Ivandic, I. (2015). *CRISPR/Cas9 knockout and functional characterization of gnav1* [PhD thesis]. University of Cologne.

Jessen, J. R. (2015). Recent advances in the study of zebrafish extracellular matrix proteins. *Developmental Biology*, *401*(1), 110–121. https://doi.org/10.1016/j.ydbio.2014.12.022

Jin, C., Yin, F., Lin, M., Li, H., Wang, Z., Weng, J., Liu, M., Da Dong, X., Qu, J., & Tu, L. (2008). GPR48 Regulates Epithelial Cell Proliferation and Migration by Activating EGFR during Eyelid Development. *Investigative Ophthalmology & Visual Science*, *49*(10), 4245–4253. https://doi.org/10.1167/iovs.08-1860

Johnston, C. A., & Siderovski, D. P. (2007). Receptor-Mediated Activation of Heterotrimeric G-Proteins: Current Structural Insights. *Molecular Pharmacology*, *72*(2), 219–230. https://doi.org/10.1124/mol.107.034348

Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kearse, M. G., & Wilusz, J. E. (2017). Non-AUG translation: A new start for protein synthesis in eukaryotes. *Genes & Development*, *31*(17), 1717–1731. https://doi.org/10.1101/gad.305250.117

Kehrl, J. M., Sahaya, K., Dalton, H. M., Charbeneau, R. A., Kohut, K. T., Gilbert, K., Pelz, M. C., Parent, J., & Neubig, R. R. (2014). Gain-of-function mutation in Gnao1: A murine model of epileptiform encephalopathy (EIEE17)? *Mammalian Genome: Official Journal of the International Mammalian Genome Society*, *25*(5–6), 202–210. https://doi.org/10.1007/s00335-014-9509-z

Kersten, S., & Arjona, F. J. (2017). Ion transport in the zebrafish kidney from a human disease angle: Possibilities, considerations, and future perspectives. *American Journal of Physiology-Renal Physiology*, *312*(1), F172–F189. https://doi.org/10.1152/ajprenal.00425.2016

Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., & Schilling, T. F. (1995). Stages of embryonic development of the zebrafish. *Developmental Dynamics*, *203*(3), 253–310. https://doi.org/10.1002/aja.1002030302

Kimmel, C. B., Miller, C. T., Kruze, G., Ullmann, B., BreMiller, R. A., Larison, K. D., & Snyder, H. C. (1998). The shaping of pharyngeal cartilages during early development of the zebrafish. *Developmental Biology*, *203*(2), 245–263. https://doi.org/10.1006/dbio.1998.9016

Kornak, U., Reynders, E., Dimopoulou, A., van Reeuwijk, J., Fischer, B., Rajab, A., Budde, B., Nürnberg, P., Foulquier, F., Lefeber, D., Urban, Z., Gruenewald, S., Annaert, W., Brunner, H. G., van Bokhoven, H., Wevers, R., Morava, E., Matthijs, G., Van Maldergem, L., & Mundlos, S. (2008). Impaired glycosylation and cutis laxa caused by mutations in the vesicular H^+-ATPase subunit ATP6V0A2. *Nature Genetics*, *40*(1), 32–34. https://doi.org/10.1038/ng.2007.45

Krishnan, A., Mustafa, A., Almén, M. S., Fredriksson, R., Williams, M. J., & Schiöth, H. B. (2015). Evolutionary hierarchy of vertebrate-like heterotrimeric G protein families. *Molecular Phylogenetics and Evolution*, *91*, 27–40. https://doi.org/10.1016/j.ympev.2015.05.009

Krupinski, T., & Beitel, G. J. (2009). Unexpected Roles of the Na-K-ATPase and Other Ion Transporters in Cell Junctions and Tubulogenesis. *Physiology*, *24*(3), 192–201. https://doi.org/10.1152/physiol.00008.2009

Kumar, S., Stecher, G., Suleski, M., & Hedges, S. B. (2017). TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Molecular Biology and Evolution*, *34*(7), 1812–1819. https://doi.org/10.1093/molbev/msx116

Labun, K., Montague, T. G., Krause, M., Torres Cleuren, Y. N., Tjeldnes, H., & Valen, E. (2019). CHOPCHOP v3: Expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Research*, *47*(W1), W171–W174. https://doi.org/10.1093/nar/gkz365

Lagerström, M. C., & Schiöth, H. B. (2008). Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nature Reviews Drug Discovery*, *7*(4), 339–357. https://doi.org/10.1038/nrd2518

Lalonde, R. L., & Akimenko, M.-A. (2018). Contributions of 5'HoxA/D regulation to actinodin evolution and the fin-to-limb transition. *International Journal of Developmental Biology*, *62*(11–12), 705–716. https://doi.org/10.1387/ijdb.180248rl

Lamoureux, F., Baud'huin, M., Duplomb, L., Heymann, D., & Rédini, F. (2007). Proteoglycans: Key partners in bone cell biology. *BioEssays*, *29*(8), 758–771. https://doi.org/10.1002/bies.20612

Langenbacher, A., Huang, J., Chen, Y., & Chen, J.-N. (2011). Sodium pump activity in the yolk syncytial layer regulates zebrafish heart tube morphogenesis. *Developmental Biology*, *362*, 263–270. https://doi.org/10.1016/j.ydbio.2011.12.004

Lawrence, E. A., Kague, E., Aggleton, J. A., Harniman, R. L., Roddy, K. A., & Hammond, C. L. (2018). The mechanical impact of col11a2 loss on joints; col11a2 mutant zebrafish show changes to joint development and function, which leads to early-onset osteoarthritis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1759), 20170335. https://doi.org/10.1098/rstb.2017.0335

Lewis, L., & Kwong, R. W. M. (2018). Zebrafish as a Model System for Investigating the Compensatory Regulation of Ionic Balance during Metabolic Acidosis. *International Journal of Molecular Sciences*, *19*(4), 1087. https://doi.org/10.3390/ijms19041087

Liao, B.-K., Chen, R.-D., & Hwang, P.-P. (2009). Expression regulation of Na^+-K^+-ATPase α1-subunit subtypes in zebrafish gill ionocytes. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *296*(6), R1897–R1906. https://doi.org/10.1152/ajpregu.00029.2009

Lieschke, G. J., & Currie, P. D. (2007). Animal models of human disease: Zebrafish swim into view. *Nature Reviews Genetics*, *8*(5), 353–367. https://doi.org/10.1038/nrg2091

Linder, M., Pang, I.-H., Duronio, R., Gordon, J., Sternweis, P., & Gilman, A. (1991). Lipid modifications of G protein subunits: Myristoylation of Goα increases its affinity for βγ. *The Journal of Biological Chemistry*, *266*, 4654–4659.

Link, V., Shevchenko, A., & Heisenberg, C.-P. (2006). Proteomics of early zebrafish embryos. *BMC Developmental Biology*, *6*(1), 1. https://doi.org/10.1186/1471-213X-6-1

Lokits, A. D., Indrischek, H., Meiler, J., Hamm, H. E., & Stadler, P. F. (2018). Tracing the evolution of the heterotrimeric G protein α subunit in Metazoa. *BMC Evolutionary Biology*, *18*(1), 51. https://doi.org/10.1186/s12862-018-1147-8

Luczynski, M., Strzczek, J., & Brzuzan, P. (1987). Secretion of hatching enzyme and its proteolytic activity in coregoninae (Coregonus albula L andC. lavaretus L) embryos. *Fish Physiology and Biochemistry*, *4*(2), 57–62. https://doi.org/10.1007/BF02044314

Lun, K., & Brand, M. (1998). A series of no isthmus (noi) alleles of the zebrafish pax2.1 gene reveals multiple signaling events in development of the midbrain-hindbrain boundary. *Development*, *125*(16), 3049–3062. https://doi.org/10.1242/dev.125.16.3049

Luo, J., Zhou, W., Zhou, X., Li, D., Weng, J., Yi, Z., Cho, S. G., Li, C., Yi, T., Wu, X., Li, X.-Y., de Crombrugghe, B., Höök, M., & Liu, M. (2009). Regulation of bone formation and remodeling by G-protein-coupled receptor 48. *Development*, *136*(16), 2747–2756. https://doi.org/10.1242/dev.033571

Malarkannan, S., Horng, T., Shih, P. P., Schwab, S., & Shastri, N. (1999). Presentation of Out-of-Frame Peptide/MHC Class I Complexes by a Novel Translation Initiation Mechanism. *Immunity*, *10*(6), 681–690. https://doi.org/10.1016/S1074-7613(00)80067-9

Malbon, C. C. (2005). G proteins in development. *Nature Reviews Molecular Cell Biology*, *6*(9), 689–701. https://doi.org/10.1038/nrm1716

Malfait, F., Kariminejad, A., Van Damme, T., Gauche, C., Syx, D., Merhi-Soussi, F., Gulberti, S., Symoens, S., Vanhauwaert, S., Willaert, A., Bozorgmehr, B., Kariminejad, M. H., Ebrahimiadib, N., Hausser, I., Huysseune, A., Fournel-Gigleux, S., & De Paepe, A. (2013). Defective Initiation of Glycosaminoglycan Synthesis due to B3GALT6 Mutations Causes a Pleiotropic Ehlers-Danlos-Syndrome-like Connective Tissue Disorder. *The American Journal of Human Genetics*, *92*(6), 935–945. https://doi.org/10.1016/j.ajhg.2013.04.016

Marchini, M., Ashkin, M. R., Bellini, M., Sun, M. M.-G., Workentine, M. L., Okuyan, H. M., Krawetz, R., Beier, F., & Rolian, C. (2021). A Na^+/K^+ ATPase Pump Regulates Chondrocyte Differentiation and Bone Length Variation in Mice. *Frontiers in Cell and Developmental Biology*, *9*.

https://www.frontiersin.org/article/10.3389/fcell.2021.708384

Marlow, Florence, Gonzalez, , Encina M., Yin, C., Rojo, C., & Solnica-Krezel, Lilianna. (2004). No tail co-operates with non-canonical Wnt signaling to regulate posterior body morphogenesis in zebrafish. *Development*, *131*(1), 203–216. https://doi.org/10.1242/dev.00915

März, M., Chapouton, P., Diotel, N., Vaillant, C., Hesl, B., Takamiya, M., Lam, C. S., Kah, O., Bally-Cuif, L., & Strähle, U. (2010). Heterogeneity in progenitor cell subtypes in the ventricular zone of the zebrafish adult telencephalon. *Glia*, *58*(7), 870–888. https://doi.org/10.1002/glia.20971

Mathavan, S., Lee, S. G. P., Mak, A., Miller, L. D., Murthy, K. R. K., Govindarajan, K. R., Tong, Y., Wu, Y. L., Lam, S. H., Yang, H., Ruan, Y., Korzh, V., Gong, Z., Liu, E. T., & Lufkin, T. (2005). Transcriptome Analysis of Zebrafish Embryogenesis Using Microarrays. *PLOS Genetics*, *1*(2), e29. https://doi.org/10.1371/journal.pgen.0010029

Matsui, T., Raya, Á., Kawakami, Y., Callol-Massot, C., Capdevila, J., Rodríguez-Esteban, C., & Belmonte, J. C. I. (2005). Noncanonical Wnt signaling regulates midline convergence of organ primordia during zebrafish development. *Genes & Development*, *19*(1), 164–175. https://doi.org/10.1101/gad.1253605

McCampbell, K. K., Springer, K. N., & Wingert, R. A. (2015). Atlas of Cellular Dynamics during Zebrafish Adult Kidney Regeneration. *Stem Cells International*, *2015*, 547636. https://doi.org/10.1155/2015/547636

McCampbell, K. K., & Wingert, R. A. (2014). Using Zebrafish to Study Renal Regeneration. *Translational Research : The Journal of Laboratory and Clinical Medicine*, *163*(2), 109–122. https://doi.org/10.1016/j.trsl.2013.10.003

McCurley, A. T., & Callard, G. V. (2008). Characterization of housekeeping genes in zebrafish: Male-female differences and effects of tissue type, developmental stage and chemical treatment. *BMC Molecular Biology*, *9*(1), 102. https://doi.org/10.1186/1471-2199-9-102

Mendive, F., Laurent, P., Van Schoore, G., Skarnes, W., Pochet, R., & Vassart, G. (2006). Defective postnatal development of the male reproductive tract in LGR4 knockout mice. *Developmental Biology*, *290*(2), 421–434. https://doi.org/10.1016/j.ydbio.2005.11.043

Mobasheri, A., Mobasheri, R., Francis, M. J. O., Trujillo, E., Rosa, D. Á. de la, & Vasallo, P. M. (1998). Ion transport in chondrocytes: Membrane transporters involved in intracellular ion homeostasis and the regulation of cell volume, free [$Ca2^{+}$] and pH. *Histology and Histopathology*, *13*(3), 893–910. https://revistas.um.es/hh/article/view/130171

Mobasheri, A., Trujillo, E., Arteaga, M.-F., & Martín-Vasallo, P. (2012). Na^{+}, K^{+}-ATPase Subunit Composition in a Human Chondrocyte Cell Line; Evidence for the Presence of α1, α3, β1, β2 and β3 Isoforms. *International Journal of Molecular Sciences*, *13*(4), 5019–5034. https://doi.org/10.3390/ijms13045019

Mohri, Y., Oyama, K., Sone, M., Akamatsu, A., & Nishimori, K. (2012). LGR4 Is Required for the Cell Survival of the Peripheral Mesenchyme at the Embryonic Stages of Nephrogenesis. *Bioscience, Biotechnology, and Biochemistry*, *76*(5), 888–891. https://doi.org/10.1271/bbb.110834

Mohri, Y., Umezu, T., Hidema, S., Tomisawa, H., Akamatsu, A., Kato, S., Nawa, A., & Nishimori, K. (2010). Reduced fertility with impairment of early-stage embryos observed in mice lacking Lgr4 in epithelial tissues. *Fertility and Sterility*, *94*(7), 2878–2881. https://doi.org/10.1016/j.fertnstert.2010.05.050

Nadirah, M., Munafi, A. B. A., KhairulAnuar, K., Mohamad, R. Y. R., & Najiah, M. (2014). *Suitability of water salinity for hatching and survival of newly hatched larvae of climbing perch, Anabas testudineus*. 5.

Nagahama, Y., & Yamashita, M. (2008). Regulation of oocyte maturation in fish. *Development, Growth & Differentiation*, *50*(s1), S195–S219. https://doi.org/10.1111/j.1440-169X.2008.01019.x

Nagy, E., & Maquat, L. E. (1998). A rule for termination-codon position within intron-containing genes: When nonsense affects RNA abundance. *Trends in Biochemical Sciences*, *23*(6), 198–199. https://doi.org/10.1016/S0968-0004(98)01208-0

Nakajima, M., Mizumoto, S., Miyake, N., Kogawa, R., Iida, A., Ito, H., Kitoh, H., Hirayama, A., Mitsubuchi, H., Miyazaki, O., Kosaki, R., Horikawa, R., Lai, A., Mendoza-Londono, R., Dupuis, L., Chitayat, D., Howard, A., Leal, G. F., Cavalcanti, D., … Ikegawa, S. (2013). Mutations in B3GALT6, which Encodes a Glycosaminoglycan Linker Region Enzyme, Cause a Spectrum of Skeletal and Connective Tissue Disorders. *American Journal of Human Genetics*, *92*(6), 927–934. https://doi.org/10.1016/j.ajhg.2013.04.003

Naylor, R. W., Han, H. I., Hukriede, N. A., & Davidson, A. J. (2017). Wnt8a expands the pool of embryonic kidney progenitors in zebrafish. *Developmental Biology*, *425*(2), 130–141. https://doi.org/10.1016/j.ydbio.2017.03.027

NCBI Resource Coordinators. (2013). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, *41*(D1), D8–D20. https://doi.org/10.1093/nar/gks1189

Neuhauss, S. C., Solnica-Krezel, L., Schier, A. F., Zwartkruis, F., Stemple, D. L., Malicki, J., Abdelilah, S., Stainier, D. Y., & Driever, W. (1996). Mutations affecting craniofacial development in zebrafish. *Development (Cambridge, England)*, *123*, 357–367.

Neu-Yilik, G., Amthor, B., Gehring, N. H., Bahri, S., Paidassi, H., Hentze, M. W., & Kulozik, A. E. (2011). Mechanism of escape from nonsense-mediated mRNA decay of human β-globin transcripts with nonsense mutations in the first exon. *RNA*, *17*(5), 843–854. https://doi.org/10.1261/rna.2401811

Neves, S. R., Ram, P. T., & Iyengar, R. (2002). G Protein Pathways. *Science*, *296*(5573), 1636–1639. https://doi.org/10.1126/science.1071550

Ng, A., de Jong-Curtain, T. A., Mawdsley, D. J., White, S. J., Shin, J., Appel, B., Dong, P. D. S., Stainier, D. Y. R., & Heath, J. K. (2005). Formation of the digestive system in zebrafish: III. Intestinal epithelium morphogenesis. *Developmental Biology*, *286*(1), 114–135. https://doi.org/10.1016/j.ydbio.2005.07.013

Nijenhuis, T., Renkema, K. Y., Hoenderop, J. G. J., & Bindels, R. J. M. (2006). Acid-Base Status Determines the Renal Expression of $Ca2^+$ and $Mg2^+$ Transport Proteins. *Journal of the American Society of Nephrology*, *17*(3), 617–626. https://doi.org/10.1681/ASN.2005070732

Nüsslein-Volhard, C., & Dahm, R. (2002). *Zebrafish*. OUP Oxford.

O'Connor, C., & Adams, J. U. (2010). *Essentials of Cell Biology*. Cambridge, MA: NPG Education.

Oka, Y., & Korsching, S. I. (2011). Shared and Unique G Alpha Proteins in the Zebrafish Versus Mammalian Senses of Taste and Smell. *Chemical Senses*, *36*(4), 357–365. https://doi.org/10.1093/chemse/bjq138

Oka, Y., Saraiva, L. R., Kwan, Y. Y., & Korsching, S. I. (2009). The fifth class of Gα proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(5), 1484–1489. https://doi.org/10.1073/pnas.0809420106

Okashah, N., Wan, Q., Ghosh, S., Sandhu, M., Inoue, A., Vaidehi, N., & Lambert, N. A. (2019). Variable G protein determinants of GPCR coupling selectivity. *Proceedings of the National Academy of Sciences*, *116*(24), 12054–12059. https://doi.org/10.1073/pnas.1905993116

Oldham, W. M., & Hamm, H. E. (2008). Heterotrimeric G protein activation by G-protein-coupled receptors. *Nature Reviews Molecular Cell Biology*, *9*(1), 60–71. https://doi.org/10.1038/nrm2299

Ord, J. (2019). Ionic Stress Prompts Premature Hatching of Zebrafish (Danio rerio) Embryos. *Fishes*, *4*(1), 20. https://doi.org/10.3390/fishes4010020

Ordaz-Ramos, A., Rosales-Gallegos, V. H., Melendez-Zajgla, J., Maldonado, V., & Vazquez-Santillan, K. (2021). The Role of LGR4 (GPR48) in Normal and Cancer Processes. *International Journal of Molecular Sciences*, *22*(9), 4690. https://doi.org/10.3390/ijms22094690

Palma, M., & Lejeune, F. (2021). Deciphering the molecular mechanism of stop codon readthrough. *Biological Reviews*, *96*(1), 310–329. https://doi.org/10.1111/brv.12657

Pan, T.-C., Liao, B.-K., Huang, C.-J., Lin, L.-Y., & Hwang, P.-P. (2005). Epithelial $Ca2^+$ channel expression and $Ca2^+$ uptake in developing zebrafish. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *289*(4), R1202–R1211. https://doi.org/10.1152/ajpregu.00816.2004

Papatheodorou, I., Fonseca, N. A., Keays, M., Tang, Y. A., Barrera, E., Bazant, W., Burke, M., Füllgrabe, A., Fuentes, A. M.-P., George, N., Huerta, L., Koskinen, S., Mohammed, S., Geniza, M., Preece, J., Jaiswal, P., Jarnuczak, A. F., Huber, W., Stegle, O., … Petryszak, R. (2018). Expression Atlas: Gene and protein expression across multiple studies and organisms. *Nucleic Acids Research*, *46*(D1), D246–D251. https://doi.org/10.1093/nar/gkx1158

Paps, J. (2018). What Makes an Animal? The Molecular Quest for the Origin of the Animal Kingdom. *Integrative and Comparative Biology*, *58*. https://doi.org/10.1093/icb/icy036

Pasquier, J., Cabau, C., Nguyen, T., Jouanno, E., Severac, D., Braasch, I., Journot, L., Pontarotti, P., Klopp, C., Postlethwait, J. H., Guiguen, Y., & Bobe, J. (2016). Gene evolution and gene expression after whole genome duplication in fish: The PhyloFish database. *BMC Genomics*, *17*(1), 368. https://doi.org/10.1186/s12864-016-2709-z

Pathak, N. H., & Barresi, M. J. F. (2020). Zebrafish as a Model to Understand Vertebrate Development. In *The Zebrafish in Biomedical Research* (pp. 559–591). Elsevier. https://doi.org/10.1016/B978-0-12-812431-4.00045-2

Pereira, F. J. C., Teixeira, A., Kong, J., Barbosa, C., Silva, A. L., Marques-Ramos, A., Liebhaber, S. A., & Romão, L. (2015). Resistance of mRNAs with AUG-proximal nonsense mutations to nonsense-mediated decay reflects variables of mRNA structure and translational activity. *Nucleic Acids Research*, *43*(13), 6528–6544. https://doi.org/10.1093/nar/gkv588

Petrovic, S., Wang, Z., Ma, L., & Soleimani, M. (2003). Regulation of the apical Cl−/HCO 3 − exchanger pendrin in rat cortical collecting duct in metabolic acidosis. *American Journal of Physiology-Renal Physiology*, *284*(1), F103–F112. https://doi.org/10.1152/ajprenal.00205.2002

Prummel, K. D., Hess, C., Nieuwenhuize, S., Parker, H. J., Rogers, K. W., Kozmikova, I., Racioppi, C., Brombacher, E. C., Czarkwiani, A., Knapp, D., Burger, S., Chiavacci, E., Shah, G., Burger, A., Huisken, J., Yun, M. H., Christiaen, L., Kozmik, Z., Müller, P., … Mosimann, C. (2019). A conserved regulatory program initiates lateral plate mesoderm emergence across chordates. *Nature Communications*, *10*(1), 3857. https://doi.org/10.1038/s41467-019-11561-7

Puchtler, H., Meloan, S. N., & Terry, M. S. (1969). On the history and mechanism of Alizarin and Alizarin Red S stains for calcium. *Journal of Histochemistry & Cytochemistry*, *17*(2), 110–124. https://doi.org/10.1177/17.2.110

Qiao, W., Pan, D., Zheng, Y., Wu, S., Liu, X., Chen, Z., Wan, M., Feng, S., Cheung, K. M. C., Yeung, K. W. K., & Cao, X. (2022). Divalent metal cations stimulate skeleton interoception for new bone formation in mouse injury models. *Nature Communications*, *13*(1), 535. https://doi.org/10.1038/s41467-022-28203-0

Quan, F., Wolfgang, W. J., & Forte, M. (1993). A Drosophila G-protein alpha subunit, Gf alpha, expressed in a spatially and temporally restricted pattern during Drosophila development. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(9), 4236–4240. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC46481/

Ragan, M. A., Goggin, C. L., Cawthorn, R. J., Cerenius, L., Jamieson, A. V., Plourde, S. M., Rand, T. G., Söderhäll, K., & Gutell, R. R. (1996). A novel clade of protistan parasites near the animal-fungal divergence. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(21), 11907–11912. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC38157/

Rhinn, M., & Brand, M. (2001). The midbrain–hindbrain boundary organizer. *Current Opinion in Neurobiology*, *11*(1), 34–42. https://doi.org/10.1016/S0959-4388(00)00171-9

Robertson, C. E., Wright, P. A., Köblitz, L., & Bernier, N. J. (2014). Hypoxia-inducible factor-1 mediates adaptive developmental plasticity of hypoxia tolerance in zebrafish, Danio rerio. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1786), 20140637. https://doi.org/10.1098/rspb.2014.0637

Rodrigues, P., Cunha, V., Ferreira, M., Reis-Henriques, M. A., Oliva-Teles, L., Guimarães, L., & Carvalho, A. P. (2022). Differential Molecular Responses of Zebrafish Larvae to Fluoxetine and Norfluoxetine. *Water*, *14*(3), 417. https://doi.org/10.3390/w14030417

Rojas, J., Hinostroza, F., Vergara, S., Pinto-Borguero, I., Aguilera, F., Fuentes, R., & Carvacho, I. (2021). Knockin' on Egg's Door: Maternal Control of Egg Activation That Influences Cortical Granule Exocytosis in Animal Species. *Frontiers in Cell and Developmental Biology*, *9*. https://www.frontiersin.org/article/10.3389/fcell.2021.704867

Romanelli Tavares, V. L., Gordon, C. T., Zechi-Ceide, R. M., Kokitsu-Nakata, N. M., Voisin, N., Tan, T. Y., Heggie, A. A., Vendramini-Pittoli, S., Propst, E. J., Papsin, B. C., Torres, T. T., Buermans, H., Capelo, L. P., den Dunnen, J. T., Guion-Almeida, M. L., Lyonnet, S., Amiel, J., & Passos-Bueno, M. R. (2015). Novel variants in GNAI3 associated with auriculocondylar syndrome strengthen a common dominant negative effect. *European Journal of Human Genetics*, *23*(4), 481–485. https://doi.org/10.1038/ejhg.2014.132

Romão, L., Inácio, Â., Santos, S., Ávila, M., Faustino, P., Pacheco, P., & Lavinha, J. (2000). Nonsense mutations in the human β-globin gene lead to unexpected levels of cytoplasmic mRNA accumulation. *Blood*, *96*(8), 2895–2901. https://doi.org/10.1182/blood.V96.8.2895

Ronsin, C., Chung-Scott, V., Poullion, I., Aknouche, N., Gaudin, C., & Triebel, F. (1999). A Non-AUG-Defined Alternative Open Reading Frame of the Intestinal Carboxyl Esterase mRNA Generates an Epitope Recognized by Renal Cell Carcinoma-Reactive Tumor-Infiltrating

Lymphocytes In Situ. *The Journal of Immunology*, *163*(1), 483–490. https://www.jimmunol.org/content/163/1/483

Rosnoblet, C., Legrand, D., Demaegd, D., Hacine-Gherbi, H., de Bettignies, G., Bammens, R., Borrego, C., Duvet, S., Morsomme, P., Matthijs, G., & Foulquier, F. (2013). Impact of disease-causing mutations on TMEM165 subcellular localization, a recently identified protein involved in CDG-II. *Human Molecular Genetics*, *22*(14), 2914–2928. https://doi.org/10.1093/hmg/ddt146

Ruffner, H., Sprunger, J., Charlat, O., Leighton-Davies, J., Grosshans, B., Salathe, A., Zietzling, S., Beck, V., Therier, M., Isken, A., Xie, Y., Zhang, Y., Hao, H., Shi, X., Liu, D., Song, Q., Clay, I., Hintzen, G., Tchorz, J., … Cong, F. (2012). R-Spondin Potentiates Wnt/β-Catenin Signaling through Orphan Receptors LGR4 and LGR5. *PLoS ONE*, *7*(7), e40976. https://doi.org/10.1371/journal.pone.0040976

Rützler, M., Lu, T., & Zwiebel, L. J. (2006). Gα Encoding Gene Family of the Malaria Vector Mosquito Anopheles gambiae: Expression Analysis and Immunolocalization of AGαq and AGαo in Female Antennae. *The Journal of Comparative Neurology*, *499*(4), 533–545. https://doi.org/10.1002/cne.21083

Sánchez, R., Serra, F., Tárraga, J., Medina, I., Carbonell, J., Pulido, L., de María, A., Capella-Gutíerrez, S., Huerta-Cepas, J., Gabaldón, T., Dopazo, J., & Dopazo, H. (2011). Phylemon 2.0: A suite of web-tools for molecular evolution, phylogenetics, phylogenomics and hypotheses testing. *Nucleic Acids Research*, *39*(Web Server issue), W470-474. https://doi.org/10.1093/nar/gkr408

Sano, K., Inohaya, K., Kawaguchi, M., Yoshizaki, N., Iuchi, I., & Yasumasu, S. (2008). Purification and characterization of zebrafish hatching enzyme – an evolutionary aspect of the mechanism of egg envelope digestion. *The FEBS Journal*, *275*(23), 5934–5946. https://doi.org/10.1111/j.1742-4658.2008.06722.x

Sasaki, M. M., Nichols, J. T., & Kimmel, C. B. (2013). Edn1 and hand2 Interact in Early Regulation of Pharyngeal Arch Outgrowth during Zebrafish Development. *PLOS ONE*, *8*(6), e67522. https://doi.org/10.1371/journal.pone.0067522

Sawant, M. S., Zhang, S., & Li, L. (2001). Effect of salinity on development of zebrafish, Brachydanio rerio. *Current Science*, *81*(10), 1347–1350. https://www.jstor.org/stable/24105851

Schier, A. F., & Talbot, W. S. (2005). Molecular Genetics of Axis Formation in Zebrafish. *Annual Review of Genetics*, *39*(1), 561–613. https://doi.org/10.1146/annurev.genet.37.110801.143752

Schilling, T., & Kimmel, C. (1994). Segment and cell type lineage restrictions during pharyngeal arch development in the zebrafish embryo. *Development*.

Schwab, S. R., Li, K. C., Kang, C., & Shastri, N. (2003). Constitutive Display of Cryptic Translation Products by MHC Class I Molecules. *Science*, *301*(5638), 1367–1371. https://doi.org/10.1126/science.1085650

Schwab, S. R., Shugart, J. A., Horng, T., Malarkannan, S., & Shastri, N. (2004). Unanticipated Antigens: Translation Initiation at CUG with Leucine. *PLoS Biology*, *2*(11), e366. https://doi.org/10.1371/journal.pbio.0020366

Sharma, V., Hecker, N., Roscito, J. G., Foerster, L., Langer, B. E., & Hiller, M. (2018). A genomics approach reveals insights into the importance of gene losses for mammalian adaptations. *Nature Communications*, *9*(1), 1215. https://doi.org/10.1038/s41467-018-03667-1

Shu, Y., Lou, Q., Dai, Z., Dai, X., He, J., Hu, W., & Yin, Z. (2016). The basal function of teleost prolactin as a key regulator on ion uptake identified with zebrafish knockout models. *Scientific Reports*, *6*(1), 18597. https://doi.org/10.1038/srep18597

Sladky, K. K., Swanson, C. R., Stoskopf, M. K., Loomis, M. R., & Lewbart, G. A. (2001). Comparative efficacy of tricaine methanesulfonate and clove oil for use as anesthetics in red pacu (Piaractus brachypomus). *American Journal of Veterinary Research*, *62*(3), 337–342. https://doi.org/10.2460/ajvr.2001.62.337

Small, C. D., el-Khoury, M., Deslongchamps, G., Benfey, T. J., & Crawford, B. D. (2020). Matrix Metalloproteinase 13 Activity is Required for Normal and Hypoxia-Induced Precocious Hatching in Zebrafish Embryos. *Journal of Developmental Biology*, *8*(1), 3. https://doi.org/10.3390/jdb8010003

Smith, D. B., & Johnson, K. S. (1988). Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. *Gene*, *67*(1), 31–40. https://doi.org/10.1016/0378-1119(88)90005-4

Smrcka, A. V. (Ed.). (2004). *G Protein Signaling: Methods and Protocols*. Humana Press. https://doi.org/10.1385/1592594301

Solnica-Krezel, L. (2020). Chapter Thirteen—Maternal contributions to gastrulation in zebrafish. In F.

L. Marlow (Ed.), *Current Topics in Developmental Biology* (Vol. 140, pp. 391–427). Academic Press. https://doi.org/10.1016/bs.ctdb.2020.05.001

Spence, R., Gerlach, G., Lawrence, C., & Smith, C. (2008). The behaviour and ecology of the zebrafish, Danio rerio. *Biological Reviews*, *83*(1), 13–34. https://doi.org/10.1111/j.1469-185X.2007.00030.x

Spence, R., & Smith, C. (2006). Mating preference of female zebrafish, Danio rerio, in relation to male dominance. *Behavioral Ecology*, *17*(5), 779–783. https://doi.org/10.1093/beheco/arl016

Starck, S. R., Jiang, V., Pavon-Eternod, M., Prasad, S., McCarthy, B., Pan, T., & Shastri, N. (2012). Leucine-tRNA Initiates at CUG Start Codons for Protein Synthesis and Presentation by MHC Class I. *Science*, *336*(6089), 1719–1723. https://doi.org/10.1126/science.1220270

Stickney, H. L., Barresi, M. J. F., & Devoto, S. H. (2000). Somite development in zebrafish. *Developmental Dynamics*, *219*(3), 287–303. https://doi.org/10.1002/1097-0177(2000)9999:9999<::AID-DVDY1065>3.0.CO;2-A

Strähle, U., Scholz, S., Geisler, R., Greiner, P., Hollert, H., Rastegar, S., Schumacher, A., Selderslaghs, I., Weiss, C., Witters, H., & Braunbeck, T. (2012). Zebrafish embryos as an alternative to animal experiments—A commentary on the definition of the onset of protected life stages in animal welfare regulations. *Reproductive Toxicology*, *33*(2), 128–132. https://doi.org/10.1016/j.reprotox.2011.06.121

Streisinger, G., Walker, C., Dower, N., Knauber, D., & Singer, F. (1981). Production of clones of homozygous diploid zebra fish (Brachydanio rerio). *Nature*, *291*(5813), 293–296. https://doi.org/10.1038/291293a0

Suga, H., Chen, Z., de Mendoza, A., Sebé-Pedrós, A., Brown, M. W., Kramer, E., Carr, M., Kerner, P., Vervoort, M., Sánchez-Pons, N., Torruella, G., Derelle, R., Manning, G., Lang, B. F., Russ, C., Haas, B. J., Roger, A. J., Nusbaum, C., & Ruiz-Trillo, I. (2013). The Capsaspora genome reveals a complex unicellular prehistory of animals. *Nature Communications*, *4*(1), 2325. https://doi.org/10.1038/ncomms3325

Syrovatkina, V., Alegre, K. O., Dey, R., & Huang, X.-Y. (2016). Regulation, Signaling, and Physiological Functions of G-Proteins. *Journal of Molecular Biology*, *428*(19), 3850–3868. https://doi.org/10.1016/j.jmb.2016.08.002

Takvam, M., Wood, C. M., Kryvi, H., & Nilsen, T. O. (2021). Ion Transporters and Osmoregulation in the Kidney of Teleost Fishes as a Function of Salinity. *Frontiers in Physiology*, *12*. https://www.frontiersin.org/article/10.3389/fphys.2021.664588

Tallafuß, A., & Bally-Cuif, L. (2003). Tracing of her5 progeny in zebrafish transgenics reveals the dynamics of midbrain-hindbrain neurogenesis and maintenance. *Development*, *130*(18), 4307–4323. https://doi.org/10.1242/dev.00662

Talvik, I., Møller, R. S., Vaher, M., Vaher, U., Larsen, L. H., Dahl, H. A., Ilves, P., & Talvik, T. (2015). Clinical Phenotype of De Novo GNAO1 Mutation: Case Report and Review of Literature. *Child Neurology Open*, *2*(2), 2329048X15583717. https://doi.org/10.1177/2329048X15583717

Tang, R., Dodd, A., Lai, D., McNabb, W. C., & Love, D. R. (2007). Validation of Zebrafish (Danio rerio) Reference Genes for Quantitative Real-time RT-PCR Normalization. *Acta Biochimica et Biophysica Sinica*, *39*(5), 384–390. https://doi.org/10.1111/j.1745-7270.2007.00283.x

Thisse, C., & Thisse, B. (2008). High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nature Protocols*, *3*(1), 59–69. https://doi.org/10.1038/nprot.2007.514

Toth, Z. E., Shahar, T., Leker, R., Szalayova, I., Bratincsák, A., Key, S., Lonyai, A., Németh, K., & Mezey, É. (2007). Sensitive detection of GFP utilizing tyramide signal amplification to overcome gene silencing. *Experimental Cell Research*, *313*(9), 1943–1950. https://doi.org/10.1016/j.yexcr.2007.02.024

Trikić, M. Z., Monk, P., Roehl, H., & Partridge, L. J. (2011). Regulation of Zebrafish Hatching by Tetraspanin cd63. *PLOS ONE*, *6*(5), e19683. https://doi.org/10.1371/journal.pone.0019683

Trzaskowski, B., Latek, D., Yuan, S., Ghoshdastider, U., Debinski, A., & Filipek, S. (2012). Action of Molecular Switches in GPCRs—Theoretical and Experimental Studies. *Current Medicinal Chemistry*, *19*(8), 1090–1109. https://doi.org/10.2174/092986712799320556

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., & Leunissen, J. A. M. (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research*, *35*(Web Server issue), W71-74. https://doi.org/10.1093/nar/gkm306

Uren Webster, T. M., Laing, L. V., Florance, H., & Santos, E. M. (2014). Effects of Glyphosate and its Formulation, Roundup, on Reproduction

in Zebrafish (Danio rerio). *Environmental Science & Technology*, *48*(2), 1271–1279. https://doi.org/10.1021/es404258h

van den Hurk, R., Schoonen, W. G. E. J., van Zoelen, G. A., & Lambert, J. G. D. (1987). The biosynthesis of steroid glucuronides in the testis of the zebrafish, Brachydanio rerio, and their pheromonal function as ovulation inducers. *General and Comparative Endocrinology*, *68*(2), 179–188. https://doi.org/10.1016/0016-6480(87)90027-X

Vasilyev, A., Liu, Y., Mudumana, S., Mangos, S., Lam, P.-Y., Majumdar, A., Zhao, J., Poon, K.-L., Kondrychyn, I., Korzh, V., & Drummond, I. A. (2009). Collective Cell Migration Drives Morphogenesis of the Kidney Nephron. *PLOS Biology*, *7*(1), e1000009. https://doi.org/10.1371/journal.pbio.1000009

Villamizar, N., Ribas, L., Piferrer, F., Vera, L. M., & Sánchez-Vázquez, F. J. (2012). Impact of Daily Thermocycles on Hatching Rhythms, Larval Performance and Sex Differentiation of Zebrafish. *PLOS ONE*, *7*(12), e52153. https://doi.org/10.1371/journal.pone.0052153

Virta, V. C., & Cooper, M. S. (2011). Structural components and morphogenetic mechanics of the zebrafish yolk extension, a developmental module. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, *316B*(1), 76–92. https://doi.org/10.1002/jez.b.21381

Walker, M., & Kimmel, C. (2007). A two-color acid-free cartilage and bone stain for zebrafish larvae. *Biotechnic & Histochemistry*, *82*(1), 23–28. https://doi.org/10.1080/10520290701333558

Wang, Y.-F., Lafont, A.-G., Lee, Y.-C., & Hwang, P.-P. (2016). A novel function of calcitonin gene-related peptide in body fluid Cl− homeostasis. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1832), 20160684. https://doi.org/10.1098/rspb.2016.0684

Warga, R. M., & Kane, D. A. (2018). Wilson cell origin for kupffer's vesicle in the zebrafish. *Developmental Dynamics*, *247*(9), 1057–1069. https://doi.org/10.1002/dvdy.24657

Warga, R. M., & Kimmel, C. B. (1994). *Cell movements during epiboly and gastrulation in zebrafish*. 14.

Wedegaertner, P. B., Wilson, P. T., & Bourne, H. R. (1995). Lipid Modifications of Trimeric G Proteins (∗). *Journal of Biological Chemistry*, *270*(2), 503–506. https://doi.org/10.1074/jbc.270.2.503

Weinberg, E. S., Allende, M. L., Kelly, C. S., Abdelhamid, A., Murakami, T., Andermann, P., Doerre, O. G., Grunwald, D. J., & Riggleman, B. (1996). *Developmental regulation of zebrafish MyoD in wild-type, no tail and spadetail embryos*. 10.

Weinstein, L. S. (2001). The Stimulatory G Protein α-Subunit Gene: Mutations and Imprinting Lead to Complex Phenotypes. *The Journal of Clinical Endocrinology & Metabolism*, *86*(10), 4622–4626. https://doi.org/10.1210/jcem.86.10.8007

Weis, W. I., & Kobilka, B. K. (2018). The Molecular Basis of G Protein–Coupled Receptor Activation. *Annual Review of Biochemistry*, *87*, 897–919. https://doi.org/10.1146/annurev-biochem-060614-033910

Whittamore, J. M. (2012). Osmoregulation and epithelial water transport: Lessons from the intestine of marine teleost fish. *Journal of Comparative Physiology B*, *182*(1), 1–39. https://doi.org/10.1007/s00360-011-0601-3

Wingert, R. A., Selleck, R., Yu, J., Song, H.-D., Chen, Z., Song, A., Zhou, Y., Thisse, B., Thisse, C., McMahon, A. P., & Davidson, A. J. (2007). The cdx Genes and Retinoic Acid Control the Positioning and Segmentation of the Zebrafish Pronephros. *PLOS Genetics*, *3*(10), e189. https://doi.org/10.1371/journal.pgen.0030189

Wittkopp, N., Huntzinger, E., Weiler, C., Saulière, J., Schmidt, S., Sonawane, M., & Izaurralde, E. (2009). Nonsense-Mediated mRNA Decay Effectors Are Essential for Zebrafish Embryonic Development and Survival. *Molecular and Cellular Biology*, *29*(13), 3517–3528. https://doi.org/10.1128/MCB.00177-09

Yamagami, K. (1996). Studies on the Hatching Enzyme (Choriolysin) and Its Substrate, Egg Envelope, Constructed of the Precursors (Choriogenins) in Oryzias latipes: A Sequel to the Information in 1991/1992. *Zoological Science*, *13*(3), 331–340. https://doi.org/10.2108/zsj.13.331

Yang, D., Zhou, Q., Labroska, V., Qin, S., Darbalaei, S., Wu, Y., Yuliantie, E., Xie, L., Tao, H., Cheng, J., Liu, Q., Zhao, S., Shui, W., Jiang, Y., & Wang, M.-W. (2021). G protein-coupled receptors: Structure- and function-based drug discovery. *Signal Transduction and Targeted Therapy*, *6*(1), 1–27. https://doi.org/10.1038/s41392-020-00435-w

Yang, Z., & Chen, Y. (2006). Salinity tolerance of embryos of obscure puffer Takifugu obscurus. *Aquaculture*, *253*(1), 393–397. https://doi.org/10.1016/j.aquaculture.2005.08.014

Yin, A., Korzh, S., Winata, C. L., Korzh, V., & Gong, Z. (2011). Wnt Signaling Is Required for Early Development of Zebrafish Swimbladder. *PLoS*

ONE, *6*(3), e18431. https://doi.org/10.1371/journal.pone.0018431

Ying Jey, G., Lin, C.-H., & Hwang, P.-P. (2015). Osmoregulation in zebrafish: Ion transport mechanisms and functional regulation. *EXCLI Journal*, *14*, 627–659. https://doi.org/10.17179/excli2015-246

Thisse, B., Pflumio, S., Fürthauer, M., Loppin, B., Heyer, V., Degrave, A., Woehl, R., Lux, A., Steffan, T., Charbonnier, X.Q. and Thisse, C. (2001) Expression of the zebrafish genome during embryogenesis (NIH R01 RR15402). ZFIN Direct Data Submission (http://zfin.org).

Zhu, Y., Rice, C. D., Pang, Y., Pace, M., & Thomas, P. (2003). Cloning, expression, and characterization of a membrane progestin receptor and evidence it is an intermediary in meiotic maturation of fish oocytes. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(5), 2231–2236. https://doi.org/10.1073/pnas.0336132100

6 Appendix

6.1 Species list and abbreviations

Group	Species name	Abbrev.
Teretosporea	*Corallochytrium limacisporum*	Coli
Teretosporea	*Ichthyophonus hoferi*	Icho
Teretosporea	*Pirum gemmata*	Pige
Teretosporea	*Sphaeroforma arctica JP610*	Spar
Teretosporea	*Sphaeroforma sirkka*	Sphs
Filasterea	*Capsaspora owczarzaki*	Caow
Choanoflagellatea	*Monosiga brevicollis MX1*	Mobr
Choanoflagellatea	*Salpingoeca rosetta*	Saro
Porifera	*Amphimedon queenslandica*	Amqu
Porifera	*Ephydatia muelleri*	Epmu
Ctenophora	*Beroe forskalii*	Befo
Ctenophora	*Mnemiopsis leidyi*	Mnle
Ctenophora	*Pleurobrachia bachei*	Plba
Placozoa	*Trichoplax adhaerens*	Tric
Cnidaria	*Acropora millepora*	Acmi
Cnidaria	*Alatina alata*	Alal
Cnidaria	*Aurelia coerulea*	Auco
Cnidaria	*Exaiptasia diaphana*	Exai
Cnidaria	*Hydra vulgaris*	Hyvu
Cnidaria	*Stylophora pistillata*	Styl
Ambulacraria	*Acanthaster planci*	Acpl
Ambulacraria	*Asterias rubens*	Asru
Ambulacraria	*Lytechinus variegatus*	Lyva
Ambulacraria	*Ptychodera flava*	Ptfl
Ambulacraria	*Saccoglossus kowalevskii*	Sako
Ambulacraria	*Strongylocentrotus purpuratus*	Stpu
Mollusca	*Architeuthis dux*	Arcdu
Mollusca	*Euprymna scolopes*	Eupsc
Mollusca	*Hapalochlaena maculosa*	Hapma
Mollusca	*Nautilus pompilius*	Naupo
Mollusca	*Octopus bimaculoides*	Octbi
Mollusca	*Octopus sinensis*	Octsi
Mollusca	*Octopus vulgaris*	Octvu
Mollusca	*Sepia pharaonis*	Sepph
Mollusca	*Watasenia scintillans*	Watsc
Nematoda	*Acrobeloides nanus*	Acrnan
Nematoda	*Allodiplogaster sudhausi*	Allsud
Nematoda	*Ancylostoma caninum*	Anccan
Nematoda	*Ancylostoma ceylanicum*	Anccey
Nematoda	*Ancylostoma duodenale*	Ancduo
Nematoda	*Angiostrongylus cantonensis*	Angcan
Nematoda	*Angiostrongylus costaricensis*	Angcos
Nematoda	*Anisakis simplex*	Anisim
Nematoda	*Auanema rhodensis*	Auarho
Nematoda	*Bunonema sp. RGD898*	Bunsp.
Nematoda	*Bursaphelenchus okinawaensis*	Buroki
Nematoda	*Bursaphelenchus xylophilus*	Burxyl
Nematoda	*Caenorhabditis angaria*	Caeang
Nematoda	*Caenorhabditis becei*	Caebec

Nematoda	*Caenorhabditis bovis*	Caebov
Nematoda	*Caenorhabditis brenneri*	Caebre
Nematoda	*Caenorhabditis briggsae*	Caebri
Nematoda	*Caenorhabditis elegans*	Caeele
Nematoda	*Caenorhabditis japonica*	Caejap
Nematoda	*Caenorhabditis latens*	Caelat
Nematoda	*Caenorhabditis nigoni*	Caenig
Nematoda	*Caenorhabditis panamensis*	Caepan
Nematoda	*Caenorhabditis remanei*	Caerem
Nematoda	*Caenorhabditis sp.*	Caesp.
Nematoda	*Caenorhabditis tropicalis*	Caetro
Nematoda	*Deladenus siricidicola*	Delsir
Nematoda	*Dictyocaulus viviparus*	Dicviv
Nematoda	*Diploscapter coronatus*	Dipcor
Nematoda	*Diplogasteroides magnus*	Dipmag
Nematoda	*Diploscapter pachys*	Dippac
Nematoda	*Dirofilaria immitis*	Dirimm
Nematoda	*Ditylenchus destructor*	Ditdes
Nematoda	*Ditylenchus dipsaci*	Ditdip
Nematoda	*Dracunculus medinensis*	Dramed
Nematoda	*Enterobius vermicularis*	Entver
Nematoda	*Globodera ellingtonae*	Gloell
Nematoda	*Globodera pallida*	Glopal
Nematoda	*Globodera rostochiensis*	Gloros
Nematoda	*Haemonchus contortus*	Haecon
Nematoda	*Halicephalobus mephisto*	Halmep
Nematoda	*Halicephalobus sp. NKZ332*	Halsp.
Nematoda	*Heterorhabditis bacteriophora*	Hetbac
Nematoda	*Heterodera glycines*	Hetgly
Nematoda	*Hoplolaimus columbus*	Hopcol
Nematoda	*Levipalatum texanum*	Levtex
Nematoda	*Meloidogyne arenaria*	Melare
Nematoda	*Meloidogyne chitwoodi*	Melchi
Nematoda	*Meloidogyne enterolobii*	Melent
Nematoda	*Meloidogyne floridensis*	Melflo
Nematoda	*Meloidogyne graminicola*	Melgra
Nematoda	*Meloidogyne hapla*	Melhap
Nematoda	*Meloidogyne incognita*	Melinc
Nematoda	*Meloidogyne javanica*	Meljav
Nematoda	*Meloidogyne luci*	Melluc
Nematoda	*Mesorhabditis belari*	Mesbel
Nematoda	*Micoletzkya japonica*	Micjap
Nematoda	*Necator americanus*	Necame
Nematoda	*Nippostrongylus brasiliensis*	Nipbra
Nematoda	*Oesophagostomum dentatum*	Oesden
Nematoda	*Oscheius sp. TEL-2014*	Oscsp.
Nematoda	*Oscheius tipulae*	Osctip
Nematoda	*Panagrolaimus davidi*	Pandav
Nematoda	*Panagrellus redivivus*	Panred
Nematoda	*Panagrolaimus sp. JU765*	Pansp.
Nematoda	*Panagrolaimus superbus*	Pansup
Nematoda	*Parapristionchus giblindavisi*	Pargib
Nematoda	*Plectus murrayi*	Plemur
Nematoda	*Plectus sambesii*	Plesam
Nematoda	*Pristionchus exspectatus*	Priexs
Nematoda	*Pristionchus japonicus*	Prijap
Nematoda	*Pristionchus maxplancki*	Primax
Nematoda	*Pristionchus mayeri*	Primay
Nematoda	*Pristionchus pacificus*	Pripac
Nematoda	*Radopholus similis*	Radsim
Nematoda	*Rhabditida sp. S2_005_001R2*	Rhasp.
Nematoda	*Romanomermis culicivorax*	Romcul
Nematoda	*Rotylenchulus reniformis*	Rotren
Nematoda	*Soboliphyme baturini*	Sobbat

Nematoda	*Steinernema carpocapsae*	Stecar
Nematoda	*Steinernema diaprepesi*	Stedia
Nematoda	*Steinernema feltiae*	Stefel
Nematoda	*Steinernema glaseri*	Stegla
Nematoda	*Steinernema khuongi*	Stekhu
Nematoda	*Steinernema monticolum*	Stemon
Nematoda	*Steinernema scapterisci*	Stesca
Nematoda	*Strongyloides papillosus*	Strpap
Nematoda	*Subanguina moxae*	Submox
Nematoda	*Teladorsagia circumcincta*	Telcir
Nematoda	*Trichinella britovi*	Tribri
Nematoda	*Trichinella murrelli*	Trim
Nematoda	*Trichuris muris*	Trimur
Nematoda	*Trichinella nativa*	Trinat
Nematoda	*Trichinella nelsoni*	Trinel
Nematoda	*Trichinella papuae*	Tripap
Nematoda	*Trichinella patagoniensis*	Tripat
Nematoda	*Trichinella pseudospiralis*	Tripse
Nematoda	*Trichinella sp. T6 isolate*	Trisp.
Nematoda	*Trichinella spiralis*	Trispi
Nematoda	*Trichuris suis*	Trisui
Nematoda	*Trichuris trichiura*	Tritri
Nematoda	*Trichinella zimbabwensis*	Trizim
Tardigrada	*Hypsibius dujardini*	Hypduj
Tardigrada	*Ramazzottius varieornatus*	Ramvar
Arachnida	*Achipteria coleoptrata*	Achcol
Arachnida	*Aculops lycopersici*	Aculyc
Arachnida	*Brevipalpus yothersi*	Breyot
Arachnida	*Dermatophagoides farinae*	Derfar
Arachnida	*Dermatophagoides pteronyssinus*	Derpte
Arachnida	*Dinothrombium tinctorium*	Dintin
Arachnida	*Euroglyphus maynei*	Eurmay
Arachnida	*Hypochthonius rufulus*	Hypruf
Arachnida	*Leptotrombidium deliense*	Lepdel
Arachnida	*Leptotrombidium pallidum*	Leppal
Arachnida	*Medioppia subpectinata*	Medsub
Arachnida	*Oppiella nova*	Oppnov
Arachnida	*Panonychus citri*	Pancit
Arachnida	*Platynothrus peltifer*	Plapel
Arachnida	*Psoroptes ovis*	Psoovi
Arachnida	*Sarcoptes scabiei*	Sarsca
Arachnida	*Steganacarus magnus*	Stemag
Arachnida	*Tetranychus urticae*	Teturt
Arachnida	*Tyrophagus putrescentiae*	Tyrput
Arachnida	*Acanthoscurria geniculata*	Acagen
Arachnida	*Anelosimus studiosus*	Anestu
Arachnida	*Araneus ventricosus*	Araven
Arachnida	*Argiope bruennichi*	Argbru
Arachnida	*Dysdera silvatica*	Dyssil
Arachnida	*Latrodectus hesperus*	Lathes
Arachnida	*Loxosceles reclusa*	Loxrec
Arachnida	*Pardosa pseudoannulata*	Parpse
Arachnida	*Parasteatoda tepidariorum*	Partep
Arachnida	*Stegodyphus dumicola*	Stedum
Arachnida	*Stegodyphus mimosarum*	Stemim
Arachnida	*Trichonephila clavipes*	Tricla
Arachnida	*Dermanyssus gallinae*	Dergal
Arachnida	*Dermacentor silvarum*	Dersil
Arachnida	*Galendromus occidentalis*	Galocc
Arachnida	*Haemaphysalis longicornis*	Haelon
Arachnida	*Hyalomma asiaticum*	Hyaasi
Arachnida	*Ixodes persulcatus*	Ixoper
Arachnida	*Ixodes ricinus*	Ixoric
Arachnida	*Ixodes scapularis*	Ixosca
Arachnida	*Rhipicephalus annulatus*	Rhiann
Arachnida	*Rhipicephalus microplus*	Rhimic
Arachnida	*Rhipicephalus sanguineus*	Rhisan
Arachnida	*Tropilaelaps mercedesae*	Tromer

Arachnida	*Varroa destructor*	Vardes
Arachnida	*Varroa jacobsoni*	Varjac
Arachnida	*Androctonus mauritanicus*	Andmau
Arachnida	*Centruroides sculpturatus*	Censcu
Arachnida	*Cordylochernes scorpioides*	Corsco
Arachnida	*Mesobuthus martensii*	Mesmar
Arachnida	*Helicorthomorpha holstii*	Helhol
Myriapoda	*Strigamia maritima*	Strmar
Myriapoda	*Trigoniulus corallinus*	Tricor
Myriapoda	*Acartia tonsa*	Acaton
Crustacean	*Amphibalanus amphitrite*	Ampamp
Crustacean	*Apocyclops royi*	Aporoy
Crustacean	*Armadillidium nasatum*	Armnas
Crustacean	*Armadillidium vulgare*	Armvul
Crustacean	*Caligus rogercresseyi*	Calrog
Crustacean	*Caridina multidentata*	Carmul
Crustacean	*Cherax destructor*	Chedes
Crustacean	*Cherax quadricarinatus*	Chequa
Crustacean	*Daphnia carinata/Daphnia sinensis*	Dapcar
Crustacean	*Daphnia dubia*	Dapdub
Crustacean	*Daphnia magna*	Dapmag
Crustacean	*Daphnia obtusa*	Dapobt
Crustacean	*Daphnia pulex*	Dappul
Crustacean	*Eriocheir sinensis*	Erisin
Crustacean	*Eulimnadia texana*	Eultex
Crustacean	*Eurytemora affinis*	Euraff
Crustacean	*Gammarus roeselii*	Gamroe
Crustacean	*Hyalella azteca*	Hyaazt
Crustacean	*Lepidurus arcticus*	Leparc
Crustacean	*Lepidurus apus lubbocki*	Leplub
Crustacean	*Lepeophtheirus salmonis*	Lepsal
Crustacean	*Ligia exotica*	Ligexo
Crustacean	*Oithona nana*	Oitnan
Crustacean	*Orchestia grillus*	Orcgri
Crustacean	*Palaemon carinicauda*	Palcar
Crustacean	*Parhyale hawaiensis*	Parhaw
Crustacean	*Paralithodes platypus*	Parpla
Crustacean	*Penaeus japonicus*	Penjap
Crustacean	*Penaeus monodon*	Penmon
Crustacean	*Penaeus vannamei*	Penvan
Crustacean	*Platorchestia sp. MABIK*	Plasp.
Crustacean	*Pollicipes pollicipes*	Polpol
Crustacean	*Portunus trituberculatus*	Portri
Crustacean	*Procambarus virginalis*	Provir
Crustacean	*Semibalanus balanoides*	Sembal
Crustacean	*Tigriopus californicus*	Tigcal
Crustacean	*Tigriopus japonicus*	Tigjap
Crustacean	*Tigriopus kingsejongensis*	Tigkin
Crustacean	*Tisbe holothuriae*	Tishol
Crustacean	*Trachelipus rathkii*	Trarat
Crustacean	*Triops cancriformis*	Trican
Crustacean	*Trinorchestia longiramus*	Trilon
Crustacean	*Ceratophysella communis*	Cercom
Collembola	*Folsomia candida*	Folcan
Collembola	*Holacanthella duospinosa*	Holduo
Collembola	*Lipothrix lubbocki*	Liplub
Collembola	*Mesaphorura yosii*	Mesyos
Collembola	*Neelides sp. FZ-2019*	Neesp.
Collembola	*Oncopodura yosiiana*	Oncyos
Collembola	*Orchesella cincta*	Orccin
Collembola	*Pseudachorutes palmiensis*	Psepal
Collembola	*Pseudobourletiella spinata*	Psespi
Collembola	*Pygmarrhopalites habei*	Pyghab
Collembola	*Sinella curviseta*	Sincur
Collembola	*Sminthurides bifidus*	Smibif
Collembola	*Thalassaphorura encarpata*	Thaenc

Collembola	*Tomocerus minor*	Tommin	**Hemiptera**	*Nesidiocoris tenuis*	Nesten
Collembola	*Tomocerus qinae*	Tomqin	**Hemiptera**	*Nilaparvata lugens*	Nillug
Collembola	*Campodea augens*	Camaug	**Hemiptera**	*Oncopeltus fasciatus*	Oncfas
Diplura	*Catajapyx aquilonaris*	Cataqu	**Hemiptera**	*Orius insidiosus*	Oriins
Diplura	*Baetis rhodani*	Baerho	**Hemiptera**	*Pachypsylla venusta*	Pacven
Diplura	*Calopteryx splendens*	Calspl	**Hemiptera**	*Paracoccus marginatus*	Parmar
Palaeoptera	*Cloeon dipterum*	Clodip	**Hemiptera**	*Pentalonia nigronervosa*	Pennig
Palaeoptera	*Ephemera danica*	Ephdan	**Hemiptera**	*Phenacoccus solenopsis*	Phesol
Palaeoptera	*Ladona fulva*	Ladful	**Hemiptera**	*Pseudococcus longispinus*	Pselon
Palaeoptera	*Rhinocypha anisoptera*	Rhiani	**Hemiptera**	*Rhopalosiphum maidis*	Rhomai
Palaeoptera	*Blattella germanica*	Blager	**Hemiptera**	*Rhodnius prolixus*	Rhopro
Palaeoptera	*Coptotermes formosanus*	Copfor	**Hemiptera**	*Schizaphis graminum*	Schgra
Palaeoptera	*Cryptotermes secundus*	Crysec	**Hemiptera**	*Sipha flava*	Sipfla
Blattodea	*Periplaneta americana*	Perame	**Hemiptera**	*Sitobion miscanthi*	Sitmis
Blattodea	*Acyrthosiphon pisum*	Acypis	**Hemiptera**	*Stiretrus anchorago*	Stianc
Blattodea	*Aphis craccivora*	Aphcra	**Hemiptera**	*Trabutina mannipara*	Traman
Blattodea	*Aphis glycines*	Aphgly	**Hemiptera**	*Triatoma infestans*	Triinf
Hemiptera	*Aphis gossypii*	Aphgos	**Hemiptera**	*Trionymus perrisii*	Triper
Hemiptera	*Apolygus lucorum*	Apoluc	**Hemiptera**	*Trialeurodes vaporariorum*	Trivap
Hemiptera	*Aulacorthum solani*	Aulsol	**Hemiptera**	*Aptinothrips rufus*	Aptruf
Hemiptera	*Bemisia tabaci*	Bemtab	**Hemiptera**	*Frankliniella occidentalis*	Fraocc
Hemiptera	*Cimex lectularius*	Cimlec	**Hemiptera**	*Thrips palmi*	Thrpal
Hemiptera	*Cinara cedri*	Cinced	**Thysanoptera**	*Acromyrmex echinatior*	Acrech
Hemiptera	*Diaphorina citri*	Diacit	**Thysanoptera**	*Adelomyrmex longinoi*	Adelon
Hemiptera	*Diuraphis noxia*	Diunox	**Thysanoptera**	*Andrena crataegi*	Andcra
Hemiptera	*Eriosoma lanigerum*	Erilan	**Hymenoptera**	*Andricus curvator*	Andcur
Hemiptera	*Ericerus pela*	Eripel	**Hymenoptera**	*Andrena dunningi*	Anddun
Hemiptera	*Euschistus heros*	Eusher	**Hymenoptera**	*Andrena erythronii*	Andery
Hemiptera	*Ferrisia virgata*	Fervir	**Hymenoptera**	*Andrena forbesii*	Andfor
Hemiptera	*Gerris buenoi*	Gerbue	**Hymenoptera**	*Andricus grossulariae*	Andgro
Hemiptera	*Halyomorpha halys*	Halhal	**Hymenoptera**	*Andrena hippotes*	Andhip
Hemiptera	*Maconellicoccus hirsutus*	Machir	**Hymenoptera**	*Andrena imitatrix*	Andimi
Hemiptera	*Melanaphis sacchari*	Melsac	**Hymenoptera**	*Andricus inflator*	Andinf
Hemiptera	*Myzus persicae*	Myzper	**Hymenoptera**	*Andrena nasonii*	Andnas

Hymenoptera	*Andrena pruni*	Andpru
Hymenoptera	*Andricus quercuscalicis*	Andque
Hymenoptera	*Andrena regularis*	Andreg
Hymenoptera	*Andrena rugosa*	Andrug
Hymenoptera	*Anopheles fontenillei*	Anofon
Hymenoptera	*Aphaenogaster ashmeadi*	Aphash
Hymenoptera	*Aphidius ervi*	Apherv
Hymenoptera	*Aphaenogaster floridana*	Aphflo
Hymenoptera	*Aphaenogaster fulva*	Aphful
Hymenoptera	*Aphidius gifuensis*	Aphgif
Hymenoptera	*Aphaenogaster miamiana*	Aphmia
Hymenoptera	*Aphaenogaster picea*	Aphpic
Hymenoptera	*Aphaenogaster rudis*	Aphrud
Hymenoptera	*Apis mellifera carnica*	Apicar
Hymenoptera	*Apis mellifera caucasica*	Apicau
Hymenoptera	*Apis cerana*	Apicer
Hymenoptera	*Apis dorsata*	Apidor
Hymenoptera	*Apis florea*	Apiflo
Hymenoptera	*Apis mellifera intermissa*	Apiint
Hymenoptera	*Apis cerana japonica*	Apijap
Hymenoptera	*Apis laboriosa*	Apilab
Hymenoptera	*Apis mellifera*	Apimel
Hymenoptera	*Athalia rosae*	Athros
Hymenoptera	*Atta cephalotes*	Attcep
Hymenoptera	*Atta colombica*	Attcol
Hymenoptera	*Atta texana*	Atttex
Hymenoptera	*Belonocnema treatae/Belonocnema kinseyi*	Beltre
Hymenoptera	*Bombus bifarius*	Bombif
Hymenoptera	*Bombus bimaculatus*	Bombim
Hymenoptera	*Bombus borealis*	Bombor
Hymenoptera	*Bombus breviceps*	Bombre
Hymenoptera	*Bombus consobrinus*	Bomcon
Hymenoptera	*Bombus cullumanus*	Bomcul
Hymenoptera	*Bombus difficillimus*	Bomdif
Hymenoptera	*Bombus fervidus*	Bomfer
Hymenoptera	*Bombus griseocollis*	Bomgri
Hymenoptera	*Bombus haemorrhoidalis*	Bomhae
Hymenoptera	*Bombus ignitus*	Bomign
Hymenoptera	*Bombus impatiens*	Bomimp
Hymenoptera	*Bombus vancouverensis nearcticus*	Bomnea
Hymenoptera	*Bombus opulentus*	Bomopu
Hymenoptera	*Bombus perplexus*	Bomper
Hymenoptera	*Bombus picipes*	Bompic
Hymenoptera	*Bombus polaris*	Bompol
Hymenoptera	*Bombus pyrosoma*	Bompyr
Hymenoptera	*Bombus sandersoni*	Bomsan
Hymenoptera	*Bombus sibiricus*	Bomsib
Hymenoptera	*Bombus skorikovi*	Bomsko
Hymenoptera	*Bombus soroeensis*	Bomsor
Hymenoptera	*Bombus superbus*	Bomsup
Hymenoptera	*Bombus terrestris*	Bomter
Hymenoptera	*Bombus turneri*	Bomtur
Hymenoptera	*Bombus vagans*	Bomvag
Hymenoptera	*Bombus vosnesenskii*	Bomvos
Hymenoptera	*Bombus waltoni*	Bomwal
Hymenoptera	*Calopteryx splendens*	Calsp.
Hymenoptera	*Camponotus floridanus*	Camflo
Hymenoptera	*Campoletis sonorensis*	Camson
Hymenoptera	*Cataglyphis hispanica*	Cathis
Hymenoptera	*Cataglyphis niger*	Catnig
Hymenoptera	*Cecidostiba fungosa*	Cecfun
Hymenoptera	*Cecidostiba semifascia*	Cecsem
Hymenoptera	*Cephus cinctus*	Cepcin
Hymenoptera	*Ceratina australensis*	Ceraus
Hymenoptera	*Ceratina calcarata*	Cercal
Hymenoptera	*Ceratina dupla*	Cerdup

Hymenoptera	*Ceratosolen solmsi marchali*	Cermar
Hymenoptera	*Ceratina sp.*	Cersp.
Hymenoptera	*Chelonus insularis*	Cheins
Hymenoptera	*Colletes gigas*	Colgig
Hymenoptera	*Colletes inaequalis*	Colina
Hymenoptera	*Copidosoma floridanum*	Copflo
Hymenoptera	*Cotesia congregata*	Cotcon
Hymenoptera	*Cotesia vestalis*	Cotves
Hymenoptera	*Crematogaster levior*	Crelev
Hymenoptera	*Crematogaster nigropilosa*	Crenig
Hymenoptera	*Cyphomyrmex costatus*	Cypcos
Hymenoptera	*Cyphomyrmex rimosus*	Cyprim
Hymenoptera	*Diachasma alloeum*	Diaall
Hymenoptera	*Diadromus collaris*	Diacol
Hymenoptera	*Diadegma semiclausum*	Diasem
Hymenoptera	*Dinoponera quadriceps*	Dinqua
Hymenoptera	*Dolichoderus lamellosus*	Dollam
Hymenoptera	*Dufourea novaeangliae*	Dufnov
Hymenoptera	*Eufriesea mexicana*	Eufmex
Hymenoptera	*Euglossa dilemma*	Eugdil
Hymenoptera	*Eupelmus annulatus*	Eupann
Hymenoptera	*Eupelmus urozonus*	Eupuro
Hymenoptera	*Eurytoma adleriae*	Euradl
Hymenoptera	*Eurytoma brunniventris*	Eurbru
Hymenoptera	*Eurhopalothrix gravis*	Eurgra
Hymenoptera	*Fopius arisanus*	Fopari
Hymenoptera	*Formica exsecta*	Forexs
Hymenoptera	*Formica selysi*	Forsel
Hymenoptera	*Frieseomelitta varia*	Frivar
Hymenoptera	*Fulakora orizabana*	Fulori
Hymenoptera	*Ganaspis brasiliensis*	Ganbra
Hymenoptera	*Ganaspis sp. Gsp50*	Gansp.
Hymenoptera	*Gauromyrmex bengakalisi*	Gauben
Hymenoptera	*Gnamptogenys minuta*	Gnamin
Hymenoptera	*Goniomma blanci*	Gonbla
Hymenoptera	*Goniozus legneri*	Gonleg
Hymenoptera	*Habropoda laboriosa*	Hablab
Hymenoptera	*Harpegnathos saltator*	Harsal
Hymenoptera	*Lasioglossum albipes*	Lasalb
Hymenoptera	*Lasius niger*	Lasnig
Hymenoptera	*Leptopilina boulardi*	Lepbou
Hymenoptera	*Leptopilina clavipes*	Lepcla
Hymenoptera	*Leptopilina heterotoma*	Lephet
Hymenoptera	*Lepidotrigona ventralis hoosana*	Lephoo
Hymenoptera	*Linepithema humile*	Linhum
Hymenoptera	*Lysiphlebus fabarum*	Lysfab
Hymenoptera	*Macrocentrus cingulum*	Maccin
Hymenoptera	*Manica hunteri*	Manhun
Hymenoptera	*Megastigmus dorsalis*	Megdor
Hymenoptera	*Megalopta genalis*	Meggen
Hymenoptera	*Megachile rotundata*	Megrot
Hymenoptera	*Megastigmus stigmatizans*	Megsti
Hymenoptera	*Melipona quadrifasciata*	Melqua
Hymenoptera	*Microplitis demolitor*	Micdem
Hymenoptera	*Monomorium pharaonis*	Monpha
Hymenoptera	*Myrmecia varians*	Myrvar
Hymenoptera	*Nasonia giraulti*	Nasgir
Hymenoptera	*Nasonia longicornis*	Naslon
Hymenoptera	*Nasonia vitripennis*	Nasvit
Hymenoptera	*Neodiprion lecontei*	Neolec
Hymenoptera	*Neodiprion pinetum*	Neopin
Hymenoptera	*Neuroterus quercusbaccarum*	Neuque
Hymenoptera	*Nomada cressonii*	Nomcre
Hymenoptera	*Nomada luteoloides*	Nomlut
Hymenoptera	*Nomada maculata*	Nommac
Hymenoptera	*Nomia melanderi*	Nommel
Hymenoptera	*Nomada ovata*	Nomova

Hymenoptera	*Nomada pygmaea*	Nompyg
Hymenoptera	*Nylanderia fulva*	Nylful
Hymenoptera	*Octostruma amrishi*	Octamr
Hymenoptera	*Odontomachus brunneus*	Odobru
Hymenoptera	*Ooceraea biroi*	Oocbir
Hymenoptera	*Ormyrus nitidulus*	Ormnit
Hymenoptera	*Ormyrus pomaceus*	Ormpom
Hymenoptera	*Orussus abietinus*	Oruabi
Hymenoptera	*Osmia bicornis bicornis*	Osmbic
Hymenoptera	*Osmia conjuncta*	Osmcon
Hymenoptera	*Osmia cornifrons*	Osmcor
Hymenoptera	*Osmia lignaria*	Osmlig
Hymenoptera	*Osmia taurus*	Osmtau
Hymenoptera	*Paraponera clavata*	Parcla
Hymenoptera	*Patagonomyrmex angustus*	Patang
Hymenoptera	*Perissomyrmex snyderi*	Persny
Hymenoptera	*Pheidole fimbriata*	Phefim
Hymenoptera	*Pleistodontes nigriventris*	Plenig
Hymenoptera	*Pogonomyrmex barbatus*	Pogbar
Hymenoptera	*Pogonomyrmex occidentalis*	Pogocc
Hymenoptera	*Polistes canadensis*	Polcan
Hymenoptera	*Polistes dominula*	Poldom
Hymenoptera	*Polistes dorsalis*	Poldor
Hymenoptera	*Polistes fuscatus*	Polfus
Hymenoptera	*Polistes metricus*	Polmet
Hymenoptera	*Ponera clavicornis*	Poncla
Hymenoptera	*Ponera exotica*	Ponexo
Hymenoptera	*Ponera leae*	Ponlea
Hymenoptera	*Pristomyrmex bicolor*	Pribic
Hymenoptera	*Proceratium mancum*	Proman
Hymenoptera	*Pseudomyrmex concolor*	Psecon
Hymenoptera	*Pseudomyrmex cubaensis*	Psecub
Hymenoptera	*Pseudomyrmex dendroicus*	Pseden
Hymenoptera	*Pseudomyrmex elongatus*	Pseelo
Hymenoptera	*Pseudomyrmex feralis*	Psefer
Hymenoptera	*Pseudomyrmex flavicornis*	Psefla
Hymenoptera	*Pseudomyrmex gracilis*	Psegra
Hymenoptera	*Pseudomyrmex nigrocinctus*	Psenig
Hymenoptera	*Pseudomyrmex pallidus*	Psepal
Hymenoptera	*Pseudoneuroterus saliens*	Psesal
Hymenoptera	*Pseudomyrmex sp. PSW-54*	Psesp.
Hymenoptera	*Pseudomyrmex veneficus*	Pseven
Hymenoptera	*Pteromalus puparum*	Ptepup
Hymenoptera	*Rhopalothrix isthmica*	Rhoist
Hymenoptera	*Rogeria belti*	Rogbel
Hymenoptera	*Solenopsis fugax*	Solfug
Hymenoptera	*Solenopsis invicta*	Solinv
Hymenoptera	*Stegomyrmex manni*	Steman
Hymenoptera	*Synergus gifuensis*	Syngif
Hymenoptera	*Synergus itoensis*	Synito
Hymenoptera	*Synergus japonicus*	Synjap
Hymenoptera	*Synergus umbraculus*	Synumb
Hymenoptera	*Temnothorax curvispinosus*	Temcur
Hymenoptera	*Temnothorax longispinosus*	Temlon
Hymenoptera	*Tetramorium bicarinatum*	Tetbic
Hymenoptera	*Tetragonula carbonaria*	Tetcar
Hymenoptera	*Tetragonula clypearis*	Tetcly
Hymenoptera	*Tetragonula davenporti*	Tetdav
Hymenoptera	*Tetragonula hockingsi*	Tethoc
Hymenoptera	*Tetramorium immigrans*	Tetimm
Hymenoptera	*Tetragonula mellipes*	Tetmel
Hymenoptera	*Tetramorium parvispinum*	Tetpar
Hymenoptera	*Tetramorium simillimum*	Tetsim
Hymenoptera	*Tetheamyrma sp. CASENT0633220*	Tetsp.

Hymenoptera	*Torymus auratus*	Toraur
Hymenoptera	*Torymus geranii*	Torger
Hymenoptera	*Trachymyrmex cornetzi*	Tracor
Hymenoptera	*Trachymyrmex septentrionalis*	Trasep
Hymenoptera	*Trachymyrmex zeteki*	Trazet
Hymenoptera	*Trichogramma brassicae*	Tribra
Hymenoptera	*Trichogramma evanescens*	Trieva
Hymenoptera	*Trichogramma pretiosum*	Tripre
Hymenoptera	*Trichomalopsis sarcophagae*	Trisar
Hymenoptera	*Vespula germanica*	Vesger
Hymenoptera	*Vespa mandarinia*	Vesman
Hymenoptera	*Vespula pensylvanica*	Vespen
Hymenoptera	*Vespula vulgaris*	Vesvul
Hymenoptera	*Vollenhovia emeryi*	Voleme
Hymenoptera	*Wasmannia auropunctata*	Wasaur
Hymenoptera	*Xylocopa virginica*	Xylvir
Hymenoptera		
Hymenoptera	*Abscondita terminalis*	Abster
Hymenoptera	*Adranes taylori*	Adrtay
Hymenoptera	*Aenictocupidus jacobsonorum*	Aenjac
Coleoptera	*Aethina tumida*	Aettum
Coleoptera	*Agrilus planipennis*	Agrpla
Coleoptera	*Aleochara bilineata*	Alebil
Coleoptera	*Anoplophora glabripennis*	Anogla
Coleoptera	*Asbolus verrucosus*	Asbver
Coleoptera	*Callosobruchus maculatus*	Calmac
Coleoptera	*Chrysina resplendens*	Chrres
Coleoptera	*Coccinella septempunctata*	Cocsep
Coleoptera	*Cryptolaemus montrouzieri*	Crymon
Coleoptera	*Dendroctonus ponderosae*	Denpon
Coleoptera	*Diabrotica virgifera virgifera*	Diavir
Coleoptera	*Elaeidobius kamerunicus*	Elakam
Coleoptera	*Harmonia axyridis*	Haraxy
Coleoptera	*Hycleus cichorii*	Hyccic
Coleoptera	*Hycleus phaleratus*	Hycpha
Coleoptera	*Hypothenemus hampei*	Hypham
Coleoptera	*Ignelater luminosus*	Ignlum
Coleoptera	*Lamprigera yunnana*	Lamyun
Coleoptera	*Leptinotarsa decemlineata*	Lepdec
Coleoptera	*Leptinotarsa defecta*	Lepdef
Coleoptera	*Leptinotarsa haldemani*	Lephal
Coleoptera	*Leptinotarsa juncta*	Lepjun
Coleoptera	*Leptinotarsa lineolata*	Leplin
Coleoptera	*Leptinotarsa peninsularis*	Leppen
Coleoptera	*Leptinotarsa rubiginosa*	Leprub
Coleoptera	*Leptinotarsa texana*	Leptex
Coleoptera	*Leptinotarsa tumamoca*	Leptum
Coleoptera	*Leptinotarsa undecemlineata*	Lepund
Coleoptera	*Limonius californicus*	Limcal
Coleoptera	*Listronotus bonariensis*	Lisbon
Coleoptera	*Marronus borbonicus*	Marbor
Coleoptera	*Nicrophorus vespilloides*	Nicves
Coleoptera	*Onthophagus taurus*	Onttau
Coleoptera	*Ophraella communa*	Ophcom
Coleoptera	*Oryctes borbonicus*	Orybor
Coleoptera	*Oryzaephilus surinamensis*	Orysur
Coleoptera	*Photinus pyralis*	Phopyr
Coleoptera	*Pogonus chalceus*	Pogcha
Coleoptera	*Popillia japonica*	Popjap
Coleoptera	*Protaetia brevitarsis*	Probre
Coleoptera	*Propylea japonica*	Projap
Coleoptera	*Rhynchophorus ferrugineus*	Rhyfer

Coleoptera	*Sitophilus oryzae*	Sitory
Coleoptera	*Tenebrio molitor*	Tenmol
Coleoptera	*Tribolium castaneum*	Tricas
Coleoptera	*Tribolium madens*	Trimad
Coleoptera	*Trypoxylus dichotomus*	Trydic
Coleoptera	*Aedes aegypti*	Aedaeg
Coleoptera	*Aedes albopictus*	Aedalb
Coleoptera	*Anopheles albimanus*	Anoalb
Diptera	*Anopheles aquasalis*	Anoaqu
Diptera	*Anopheles arabiensis*	Anoara
Diptera	*Anopheles atroparvus*	Anoatr
Diptera	*Anopheles bwambae*	Anobwa
Diptera	*Anopheles christyi*	Anochr
Diptera	*Anopheles coluzzii*	Anocol
Diptera	*Anopheles cracens*	Anocra
Diptera	*Anopheles culicifacies*	Anocul
Diptera	*Anopheles darlingi*	Anodar
Diptera	*Anopheles dirus*	Anodir
Diptera	*Anopheles epiroticus*	Anoepi
Diptera	*Anopheles farauti*	Anofar
Diptera	*Anopheles funestus*	Anofun
Diptera	*Anopheles gambiae*	Anogam
Diptera	*Anopheles koliensis*	Anokol
Diptera	*Anopheles longipalpis*	Anolon
Diptera	*Anopheles maculatus*	Anomac
Diptera	*Anopheles melas*	Anomel
Diptera	*Anopheles merus*	Anomer
Diptera	*Anopheles minimus*	Anomin
Diptera	*Anopheles nili*	Anonil
Diptera	*Anopheles parensis*	Anopar
Diptera	*Anopheles punctulatus*	Anopun
Diptera	*Anopheles quadriannulatus*	Anoqua
Diptera	*Anopheles sinensis*	Anosin
Diptera	*Anopheles sp. NFL-2015*	Anosp.
Diptera	*Anopheles stephensi*	Anoste
Diptera	*Anopheles gambiae*	Anostr
Diptera	*Anopheles vaneedeni*	Anovan
Diptera	*Bactrocera dorsalis*	Bacdor
Diptera	*Bactrocera latifrons*	Baclat
Diptera	*Bactrocera oleae*	Bacole
Diptera	*Bactrocera tryoni*	Bactry
Diptera	*Belgica antarctica*	Belant
Diptera	*Bolitophila cinerea*	Bolcin
Diptera	*Bradysia coprophila*	Bracop
Diptera	*Calliphora vicina*	Calvic
Diptera	*Catotricha subobsoleta*	Catsub
Diptera	*Ceratitis capitata*	Cercap
Diptera	*Chaoborus trivitattus*	Chatri
Diptera	*Chironomus riparius*	Chirip
Diptera	*Chironomus tentans*	Chiten
Diptera	*Chrysomya rufifacies*	Chrruf
Diptera	*Cirrula hians*	Cirhia
Diptera	*Clogmia albipunctata*	Cloalb
Diptera	*Clunio marinus*	Clumar
Diptera	*Coboldia fuscipes*	Cobfus
Diptera	*Cochliomyia hominivorax*	Cochom
Diptera	*Contarinia nasturtii*	Connas
Diptera	*Condylostylus patibulatus*	Conpat
Diptera	*Culex quinquefasciatus*	Culqui
Diptera	*Culicoides sonorensis*	Culson
Diptera	*Dasypogon diadema*	Dasdia
Diptera	*Drosophila albomicans*	Droalb
Diptera	*Drosophila americana*	Droame
Diptera	*Drosophila ananassae*	Droana
Diptera	*Drosophila arizonae*	Droari
Diptera	*Drosophila asahinai*	Droasa
Diptera	*Drosophila athabasca*	Droath
Diptera	*Drosophila auraria*	Droaur

Diptera	*Drosophila azteca*	Droazt	**Diptera**	*Drosophila nasuta*	Dronas
Diptera	*Drosophila bakoue*	Drobak	**Diptera**	*Drosophila navojoa*	Dronav
Diptera	*Drosophila biarmipes*	Drobia	**Diptera**	*Drosophila neonasuta*	Droneo
Diptera	*Drosophila bifasciata*	Drobif	**Diptera**	*Drosophila nigromelanica*	Dronig
Diptera	*Drosophila bipectinata*	Drobip	**Diptera**	*Drosophila nikananu*	Dronik
Diptera	*Drosophila birchii*	Drobir	**Diptera**	*Drosophila novamexicana*	Dronov
Diptera	*Drosophila bocki*	Droboc	**Diptera**	*Drosophila obscura*	Droobs
Diptera	*Drosophila bunnanda*	Drobun	**Diptera**	*Drosophila orena*	Droore
Diptera	*Drosophila burlai*	Drobur	**Diptera**	*Drosophila pectinifera*	Dropec
Diptera	*Drosophila busckii*	Drobus	**Diptera**	*Drosophila persimilis*	Droper
Diptera	*Drosophila elegans*	Droele	**Diptera**	*Drosophila pseudoobscura*	Dropse
Diptera	*Drosophila erecta*	Droere	**Diptera**	*Drosophila punjabiensis*	Dropun
Diptera	*Drosophila eugracilis*	Droeug	**Diptera**	*Drosophila rhopaloa*	Drorho
Diptera	*Drosophila ficusphila*	Drofic	**Diptera**	*Drosophila robusta*	Drorob
Diptera	*Drosophila grimshawi*	Drogri	**Diptera**	*Drosophila rufa*	Droruf
Diptera	*Drosophila guanche*	Drogua	**Diptera**	*Drosophila sechellia*	Drosec
Diptera	*Drosophila gunungcola*	Drogun	**Diptera**	*Drosophila seguyi*	Droseg
Diptera	*Drosophila hydei*	Drohyd	**Diptera**	*Drosophila serrata*	Droser
Diptera	*Drosophila innubila*	Droinn	**Diptera**	*Drosophila simulans*	Drosim
Diptera	*Drosophila jambulina*	Drojam	**Diptera**	*Drosophila subobscura*	Droura
Diptera	*Drosophila kanapiae*	Drokan	**Diptera**	*Drosophila subpulchrella*	Drosub
Diptera	*Drosophila kikkawai*	Drokik	**Diptera**	*Drosophila suzukii*	Drosuz
Diptera	*Drosophila lacertosa*	Droosa	**Diptera**	*Drosophila takahashii*	Drotak
Diptera	*Drosophila lacteicornis*	Drolac	**Diptera**	*Drosophila tani*	Drotan
Diptera	*Drosophila leontia*	Droleo	**Diptera**	*Drosophila triauraria*	Drotri
Diptera	*Drosophila lowei*	Drolow	**Diptera**	*Drosophila truncata*	Drotru
Diptera	*Drosophila mauritiana*	Dromau	**Diptera**	*Drosophila virilis*	Drovir
Diptera	*Drosophila mayri*	Dromay	**Diptera**	*Drosophila vulcana*	Drovul
Diptera	*Drosophila melanica*	droica	**Diptera**	*Drosophila watanabei*	Drowat
Diptera	*Drosophila melanogaster*	dromel	**Diptera**	*Drosophila willistoni*	Drowil
Diptera	*Drosophila micromelanica*	Dromic	**Diptera**	*Drosophila yakuba*	Droyak
Diptera	*Drosophila miranda*	Dromir	**Diptera**	*Ephydra gracilis*	Ephgra
Diptera	*Drosophila mojavensis*	Dromoj	**Diptera**	*Eristalis dimidiata*	Eridim
Diptera	*Drosophila montana*	Dromon	**Diptera**	*Eutreta diana*	Eutdia

Diptera	*Glossina austeni*	Gloaus
Diptera	*Glossina brevipalpis*	Globre
Diptera	*Glossina fuscipes*	Glofus
Diptera	*Glossina palpalis gambiensis*	Glogam
Diptera	*Glossina morsitans*	Glomor
Diptera	*Glossina pallidipes*	Glopal
Diptera	*Haematobia irritans*	Haeirr
Diptera	*Hermetia illucens*	Herill
Diptera	*Holcocephala fusca*	Holfus
Diptera	*Liriomyza trifolii*	Lirtri
Diptera	*Lucilia cuprina*	Luccup
Diptera	*Lucilia sericata*	Lucser
Diptera	*Lutzomyia longipalpis*	Lutlon
Diptera	*Mayetiola destructor*	Maydes
Diptera	*Megaselia abdita*	Megabd
Diptera	*Megaselia scalaris*	Megsca
Diptera	*Mochlonyx cinctipes*	Moccin
Diptera	*Musca domestica*	Musdom
Diptera	*Paykullia maculata*	Paymac
Diptera	*Phlebotomus papatasi*	Phlpap
Diptera	*Phormia regina*	Phoreg
Diptera	*Phortica variegata*	Phovar
Diptera	*Polypedilum pembai*	Polpem
Diptera	*Proctacanthus coquilletti*	Procoq
Diptera	*Rhagoletis pomonella*	Rhapom
Diptera	*Rhagoletis zephyria*	Rhazep
Diptera	*Sarcophaga bullata*	Sarbul
Diptera	*Sarcophaga peregrina*	Sarper
Diptera	*Sarcophagidae sp. BV-2014*	Sarsp.
Diptera	*Scaptomyza flava*	Scafla
Diptera	*Scaptodrosophila lebanonensis*	Scaleb
Diptera	*Sitodiplosis mosellana*	Sitmos
Diptera	*Sphyracephala brevicornis*	Sphbre
Diptera	*Stomoxys calcitrans*	Stocal
Diptera	*Teleopsis dalmanni*	Teldal
Diptera	*Tephritis californica*	Tepcal
Diptera	*Themira minor*	Themin
Diptera	*Tipula oleracea*	Tipole
Diptera	*Trupanea jonesi*	Trujon
Diptera	*Zaprionus indianus*	Zapind
Diptera	*Zeugodacus cucurbitae*	Zeucuc
Diptera	*Glossosoma conforme*	Glocon
Diptera	*Hydropsyche tenuis*	Hydten
Diptera	*Limnephilus lunatus*	Limlun
Trichoptera	*Plectrocnemia conspersa*	Plecon
Trichoptera	*Sericostoma sp. HW-2014*	Sersp.
Trichoptera	*Stenopsyche tienmushanensis*	Stetie
Trichoptera		
Trichoptera	*Actias luna*	Actlun
Trichoptera	*Adela reaumurella*	Aderea
Trichoptera	*Adoxophyes honmai*	Adohon
Lepidoptera	*Agrotis ipsilon*	Agrips
Lepidoptera	*Amyelois transitella*	Amytra
Lepidoptera	*Antheraea mylitta*	Antmyl
Lepidoptera	*Antheraea pernyi*	Antper
Lepidoptera	*Arctia plantaginis*	Arcpla
Lepidoptera	*Bicyclus anynana*	Bicany
Lepidoptera	*Bombyx huttoni*	Bomhut
Lepidoptera	*Bombyx mori*	Bommor
Lepidoptera	*Busseola fusca*	Busfus
Lepidoptera	*Calycopis cecrops*	Calcec
Lepidoptera	*Calephelis nemesis*	Calnem
Lepidoptera	*Calephelis virginiensis*	Calvir
Lepidoptera	*Cecropterus lyciades*	Ceclyc
Lepidoptera	*Chilo suppressalis*	Chisup
Lepidoptera	*Cnaphalocrocis medinalis*	Cnamed
Lepidoptera	*Colias croceus*	Colcro

Lepidoptera	*Conopomorpha cramerella*	Concra
Lepidoptera	*Cydia pomonella*	Cydpom
Lepidoptera	*Danaus chrysippus*	Danchr
Lepidoptera	*Danaus melanippus*	Danmel
Lepidoptera	*Danaus plexippus plexippus*	Danple
Lepidoptera	*Delias pasithoe*	Delpas
Lepidoptera	*Dendrolimus punctatus*	Denpun
Lepidoptera	*Diatraea saccharalis*	Diasac
Lepidoptera	*Dione vanillae vanillae*	Diovan
Lepidoptera	*Drepana arcuata*	Drearc
Lepidoptera	*Elymnias hypermnestra*	Elyhyp
Lepidoptera	*Epargyreus clarus clarus*	Epacla
Lepidoptera	*Eumeta japonica*	Eumjap
Lepidoptera	*Galleria mellonella*	Galmel
Lepidoptera	*Graphium doson*	Grados
Lepidoptera	*Heliconius ethilla aerotome*	Helaer
Lepidoptera	*Heliconius melpomene aglaope*	Helagl
Lepidoptera	*Heliconius cydno alithea*	Helali
Lepidoptera	*Heliconius melpomene amaryllis*	Helama
Lepidoptera	*Helicoverpa armigera*	Helarm
Lepidoptera	*Heliconius elevatus bari*	Helbar
Lepidoptera	*Heliconius melpomene bellula*	Helbel
Lepidoptera	*Heliconius numata bicoloratus*	Helbic
Lepidoptera	*Heliconius cydno chioneus*	Helchi
Lepidoptera	*Heliconius cydno cordula*	Helcor
Lepidoptera	*Heliconius cydno cydnides*	Helcyd
Lepidoptera	*Heliconius melpomene cythera*	Helcyt
Lepidoptera	*Heliconius melpomene ecuadorensis*	Helecu
Lepidoptera	*Heliconius elevatus*	Helele
Lepidoptera	*Heliconius ethilla*	Heleth
Lepidoptera	*Heliconius hecale felix*	Helfel
Lepidoptera	*Heliconius hecuba flava*	Helfla
Lepidoptera	*Heliconius hecale*	Helhec
Lepidoptera	*Heliconius hermathena*	Helher
Lepidoptera	*Heliconius heurippa*	Helheu
Lepidoptera	*Heliconius hierax*	Helhie
Lepidoptera	*Heliconius erato x Heliconius himera*	Helhim
Lepidoptera	*Heliconius ismenius*	Helism
Lepidoptera	*Heliconius melpomene malleti*	Helmal
Lepidoptera	*Heliconius melpomene melpomene*	Helmel
Lepidoptera	*Heliconius melpomene meriana*	Helmer
Lepidoptera	*Heliconius nattereri*	Helnat
Lepidoptera	*Heliconius pachinus*	Helpac
Lepidoptera	*Heliconius pardalinus*	Helpar
Lepidoptera	*Heliconius melpomene plesseni*	Helple
Lepidoptera	*Heliconius melpomene rosina*	Helros
Lepidoptera	*Heliconius pardalinus sergestus*	Helser
Lepidoptera	*Heliconius pardalinus ssp. n. KMK-2015*	Helssp
Lepidoptera	*Heliconius melpomene thelxiopeia*	Helthe
Lepidoptera	*Heliconius timareta timareta*	Heltim
Lepidoptera	*Heliothis virescens*	Helvir
Lepidoptera	*Heliconius melpomene vulcanus*	Helvul
Lepidoptera	*Heliconius wallacei*	Helwal
Lepidoptera	*Heliconius cydno weymeri*	Helwey
Lepidoptera	*Heliconius xanthocles*	Helxan
Lepidoptera	*Helicoverpa zea*	Helzea
Lepidoptera	*Heliconius cydno zelinde*	Helzel

Lepidoptera	*Hyles vespertilio*	Hylves
Lepidoptera	*Hyphantria cunea*	Hypcun
Lepidoptera	*Hyposmocoma kahamanoa*	Hypkah
Lepidoptera	*Hypolimnas misippus*	Hypmis
Lepidoptera	*Laparus doris delila*	Lapdel
Lepidoptera	*Laparus doris*	Lapdor
Lepidoptera	*Leptidea sinapis*	Lepsin
Lepidoptera	*Lerema accius*	Leracc
Lepidoptera	*Lymantria dispar*	Lymdis
Lepidoptera	*Mamestra configurata*	Mamcon
Lepidoptera	*Maniola hyperantus*	Manhyp
Lepidoptera	*Maniola jurtina*	Manjur
Lepidoptera	*Manduca sexta*	Mansex
Lepidoptera	*Megathymus violae*	Megvio
Lepidoptera	*Melitaea cinxia*	Melcin
Lepidoptera	*Neruda aoede/Heliconius aoede*	Neraoe
Lepidoptera	*Operophtera brumata*	Opebru
Lepidoptera	*Ostrinia furnacalis*	Ostfur
Lepidoptera	*Ostrinia nubilalis*	Ostnub
Lepidoptera	*Papilio aristodemus*	Papari
Lepidoptera	*Papilio bianor*	Papbia
Lepidoptera	*Papilio glaucus*	Papgla
Lepidoptera	*Papilio machaon*	Papmac
Lepidoptera	*Papilio memnon*	Papmem
Lepidoptera	*Papilio polytes*	Pappol
Lepidoptera	*Papilio dardanus tibullus*	Paptib
Lepidoptera	*Papilio xuthus*	Papxut
Lepidoptera	*Pararge aegeria*	Paraeg
Lepidoptera	*Phoebis sennae*	Phosen
Lepidoptera	*Pieris rapae*	Pierap
Lepidoptera	*Plodia interpunctella*	Ploint
Lepidoptera	*Plutella xylostella*	Pluxyl
Lepidoptera	*Samia ricini*	Samric
Lepidoptera	*Spodoptera exigua*	Spoexi
Lepidoptera	*Spodoptera frugiperda*	Spofru
Lepidoptera	*Spodoptera litura*	Spolit
Lepidoptera	*Taenaris catops*	Taecat
Lepidoptera	*Trichoplusia ni*	TriniL
Lepidoptera	*Tuta absoluta*	Tutabs
Lepidoptera	*Vanessa tameamea*	Vantam
Lepidoptera	*Zerene cesonia*	Zerces

6.2 Tree file for the collapsed phylogenetic tree shown in Figure 61

Tree file (text) is for the tree that is shown collapsed in Figure 61; Species are indicated by four or six letter abbreviations, for full names see Appendix 6.1; Zebrafish (*danio rerio*) is indicated by Dr; Plants are indicated by Gm, At, St, Nt, and Os.

(St_GPA1:0.01325772,Nt_GA2:0.03186387,(At_GPA1:0.09538256,(Gm_GPA1:0.12265400,(Os_RGA1:0.20536733,(((((((Icho4HG12:0.19893675,(Icho2G12:0.00000001,Icho2.HG12:0.04074455)0.999850:0.22666398)0.994424:0.10736603,(Spar7G12:0.00851684,Sphsir7HG1:0.00000001)0.999850:0.34144097)0.999850:0.31957600,(Icho1HG12:0.70355333,(Caow2G12:0.49859655,((Mobr2G12:0.33224733,Saro1G12:0.65828529)0.999850:0.50463240,((Tric1G12:0.46128989,Tric2G12:0.64725598)0.999850:0.16849751,((Epmu11G12:0.30027340,Amqu11SG12:0.49903634)0.999850:0.37369227,((Befo10G12:0.10344834,(Mnle13G12:0.03942396,Plba5G12:0.15011412)0.700724:0.05432781)0.999850:0.45282739,((Dr_12:0.48590844,((HypdujTG12:0.82945995,(NaupoPOG12:0.23956444,((HapmaPOG12:0.00000001,(OctbiPOG12:0.00000001,(OctvuPOG12:0.01806590,OctsiPOG12:0.00310296)0.999850:0.04467515)0.989938:0.02143018)0.999850:0.19045894,(EupscPOG12:0.03471238,(SepphPOG12:0.08710072,WatscPOG12:0.08558877)0.557862:0.02469726)0.999850:0.06976085)0.999850:0.16037413)0.999850:0.09619777)0.999850:0.09818809,((Ptfl11G12:0.08733388,(Ptfl12G12:0.09021375,Sako7G12:0.09360358)0.304738:0.01498256)0.999850:0.08360477,((Stpu10G12:0.08510154,Lyva53G12:0.04422052)0.999850:0.14595748,(Acpl3G12:0.06444542,(Asru5G12:0.10206675,Asru9G12:0.21462397)0.999850:0.07778035)0.999850:0.16979698)0.999850:0.15978363)0.999850:0.07158949)0.750407:0.03615785)0.963638:0.04597580,((Auco21G12:0.17391029,(Hyvu18G12:0.14141129,(Auco22G12:0.11894314,Alal7G12:0.04859065)0.349002:0.00960903)0.791497:0.05921096)0.584655:0.06883312,(Exai4G12:0.13055749,(Acmi18G12:0.03209626,Styl4G12:0.13776981)0.999850:0.09026197)0.999850:0.07731850)0.985399:0.06255824)0.999850:0.08338313)0.864526:0.03720329)0.666196:0.06069757)0.999850:0.40461054)0.898521:0.06695397)0.999850:0.07303163)0.703818:0.01469441)0.345814:0.01632562,(Caow4Gq:0.61795483,(((Amqu13SGq:0.33318565,(Epmu3SGq:0.46257957,Epmu2SGq:0.25933924)0.999850:0.21806354)0.999850:0.22250767,((Epmu4SGq:0.12670849,((Plba13Gq:0.18368273,Plba7Gq:0.14933078)0.832509:0.05002142,(Befo15Gq:0.00000001,Mnle1Gq:0.10320126)0.980135:0.07275590)0.983049:0.36051684)0.936406:0.23729530,((Plba16Gq:0.31633415,(Plba6Gq:0.25638836,(Befo3Gq:0.11441652,Befo4Gq:0.00000001)0.999850:0.27781091)0.986305:0.43514899)0.619735:0.22384743,(Mobr4Gq:1.54670308,(Mnle18Gq:0.58617421,Plba11Gq:0.51616947)0.791701:0.23071542)0.430653:0.22158

141)0.998808:0.20637142)0.994258:0.20235108)0.506146:0.02705675,(Amqu12SGq:0.40417642,(Tric8Gq:0.24459099,(Hyvu16Gq:0.19670496,(Hyvu17Gq:0.31734109,(Exai1Gq:0.00000001,((Styl6Gq:0.00000001,Acmi12Gq:0.01289422)0.999850:0.02685262,(Dr_q:0.11797266,(Dr_14a:0.33030152,(Ptfl18Gq:0.00000001,(((NaupoPOGq:0.06811003,(HypdujTGq2:0.13564839,(HypdujTGq1:0.00794142,RamvarTaGq:0.05151763)0.999850:0.15409440)0.996220:0.02580053)0.991676:0.02526186,((Stpu3Gq:0.00000001,Lyva17Gq:0.06072275)0.954393:0.03435750,(Asru7Gq:0.00000001,Acpl1Gq:0.00000004)0.363453:0.07936771)0.999850:0.09736265)0.990222:0.01804696,(Sako13Gq:0.00000001,((SepphPOGq:0.24915247,(EupscPOGq1:0.14288091,(ArcduPOGq1:0.00000001,ArcduPOGq:0.00000001)0.999850:0.01459079)0.193051:0.00407299)0.991587:0.01447795,(((OctvuPOGq1:0.00000001,(OctsiPOGq1:0.00000001,OctbiPOGq:0.00000001)0.000000:0.00000001)0.000000:0.00000001,(HapmaPOGq1:0.00000001,(EupscPOGq:0.23691888,(OctvuPOGq:0.01457140,HapmaPOGq:0.00000001)0.424048:0.00222046)0.999850:0.06611899)0.974863:0.00344417)0.638970:0.01753288,(Mnle4Gq:0.02882618,Plba10Gq:0.12894786)0.999850:0.33517358)0.222912:0.00637693)0.999850:0.11772354)0.963417:0.02094243)0.227991:0.00395206)0.592771:0.01035120)0.000000:0.01042703)0.998768:0.01800747)0.942924:0.02639841)0.597596:0.02915490)0.299997:0.01799103)0.615053:0.01848195)0.999850:0.13928617)0.986992:0.04577624)0.999850:0.16123446)0.997342:0.11554306)0.999850:0.18021523,(Coli1Gs-f:1.58517625,((((Alal1Gs:0.23546146,Auco14Gs:0.41249669)0.999850:0.15024986,(((HypdujTGs3:0.09742516,RamvarTaGs:0.09513638)0.999850:0.15903976,((Dr_olf1:0.09707784,Dr_s:0.07934689)0.999850:0.12024582,(Lyva3Gs:0.12868635,(NaupoPOGs:0.06381173,((OctbiPOGs:0.00000001,(HapmaPOGs:0.00480048,(OctvuPOGs:0.00823890,OctsiPOGs:0.00000001)0.978166:0.00680698)0.988909:0.00678924)0.994292:0.02235539,(EupscPOGs:0.00079280,(ArcduPOGs:0.00000001,SepphPOGs:0.00000001)0.999850:0.08947612)0.998709:0.02699558)0.999850:0.07574781)0.999850:0.06594828)0.985866:0.03676545)0.998759:0.05633802)0.999850:0.08570867,(AmquGs:0.34230075,(EpmueGs:0.00434849,Epmu5SGs:0.10815891)0.999850:0.37450444)0.999850:0.22712364)0.993084:0.04393256)0.999850:0.26364156,(NaupoPOGf:0.47125503,(dromelDGf2:0.10864320,dromelDGf1:0.13196308)0.999850:0.86961074)0.999850:0.32737445)0.999850:0.28482135,(Caow5Gs-

f:0.38454643,(SphJP65HGs:0.62312509,Icho3HGs-
f:0.35969764)0.976434:0.10614180)0.553784:0.07783593)0.869487:0.09453466)0.
999850:0.26744076)0.999850:0.11703164,((Icho4HGv:0.36637381,(Caow6Gv:0.79
909642,(Pige4HGv:1.37428274,(SphJP64HGv:0.11587084,Sphsir2HGv:0.12420268
)0.999850:1.31167424)0.994444:0.57892938)0.995084:0.27449660)0.989721:0.165
28190,((Saro2Gv:0.10349822,Mobr3Gv:0.17846608)0.999850:0.24230700,((Amqu5
Gv:0.20454028,(Epmu1Gv:0.70126735,(Epmu10Gv:0.06571417,Epmu9Gv:0.30548
143)0.743841:0.04374420)0.999850:0.12569738)0.999850:0.19519693,((Befo13Gv:
0.05379607,Mnle16Gv:0.15283087)0.999850:0.37679793,(Dr_v1:0.33055070,((Ptfl1
3Gv:0.13884955,((Acpl18Gv:0.03429016,Asru1Gv:0.04554989)0.999850:0.1392857
2,(Lyva65Gv:0.02409980,Stpu7Gv:0.03004524)0.999850:0.18736649)0.999850:0.1
9060262)0.999850:0.06689156,(NaupoPOGv:0.11976917,((SepphPOGv:0.0282571
9,ArcduPOGv:0.01466839)0.999850:0.06610796,(HapmaPOGv:0.00445974,(OctsiP
OGv3:0.06903992,(OctsiPOGv1:0.00000001,(OctbiPOGv:0.04594186,OctvuPOGv:0
.00000001)0.000000:0.00000001)0.576507:0.00375012)0.999850:0.02493352)0.99
9850:0.16352521)0.999850:0.06653310)0.999850:0.19396054)0.384538:0.0299884
2)0.999850:0.20672706)0.971066:0.08266397)0.999850:0.27383150)0.999850:0.12
645288)0.995829:0.13565965)0.697655:0.03722942,(Caow3Gi:0.34998451,(((Icho1
HGi:0.69901496,(Icho7HGi:0.20721868,(Spar3Gi:0.05943001,(Sphsir9HGi:0.00000
001,Spar5HGi:0.00313151)0.000000:0.00000001)0.999850:0.28075867)0.695301:0
.02583662)0.999850:0.14090267,(Coli4Gi:0.29129684,(Pige14Gi:0.65774923,((Sph
sir5HGi:0.05203747,Spar1HGi:0.00000001)0.999850:0.68995854,((Spar3.Hgi:0.010
03739,Sphsir8HGi:0.01924455)0.999850:0.82408071,(Icho3HGi:0.39236564,Icho6
Gi:0.34988096)0.818283:0.05604224)0.774052:0.03737998)0.334154:0.00000008)
0.285307:0.02638186)0.999850:0.08107647)0.994126:0.04538911,((Ptfl8Gi:0.0943
4138,Sako16Gi:0.13210291)0.999850:0.41623136,(((((Hyvu19Gi:0.00000001,Hyvu4
Gi:0.04787742)0.999850:0.11701205,(Auco11Gi:0.47460588,(Alal19Gi:0.01188329,
Auco10Gi:0.09659339)0.000000:0.00821813)0.923286:0.02719499)0.999850:0.090
55211,(((Dr_i1a:0.08594861,(NaupoPOGi:0.05218158,((HapmaPOGi:0.00000001,(
OctsiPOGi1:0.00000001,(OctbiPOGi:0.00000001,(OctsiPOGi:0.00000001,OctbiPOG
i1:0.04454590)0.999850:0.07923345)0.000000:0.00000001)0.995143:0.01938397)0
.999850:0.09283119,(SepphPOGi:0.01909827,EupscPOGi:0.01676609)0.782651:0.

02406702)0.999850:0.11783195)0.999850:0.10197504)0.999850:0.03455025,(Lyva21Gi:0.02453727,(Sako17Gi:0.00000001,Ptfl10Gi:0.00727897)0.999850:0.04970518)0.995695:0.02105729)0.985190:0.02155105,(((Exai2Gi:0.00946633,(Styl2Gi:0.01003658,Acmi7Gi:0.03922917)0.905613:0.01171303)0.999850:0.03667343,((RamvarTaGi:0.13131569,HypdujTaGi:0.12989414)0.999850:0.29986398,(Tric6Gi:0.04100113,Tric5Gi:0.26352867)0.533365:0.03847395)0.742874:0.01428449)0.995582:0.01002423,((Lyva51Gi:0.00000001,Stpu4Gi:0.00000001)0.999850:0.08911801,(Acpl13Gi:0.00962080,Asru5Gi:0.00000001)0.000000:0.00000001)0.999850:0.08786640)0.000000:0.00000001)0.939444:0.03191445)0.998884:0.04943057,((Plba21Gi:0.10060216,(Befo5Gi:0.04678551,Mnle11Gi:0.02885988)0.982188:0.03346864)0.999850:0.19960408,(Plba2Gi:0.17615721,(Befo6Gi:0.05771037,Mnle12Gi:0.02454506)0.999850:0.10383451)0.999850:0.15248731)0.999850:0.19904330)0.983905:0.04557266,((Hyvu14Go:0.55223889,(Epmu7SGo:0.61079605,(Exai5Go:0.22243963,(Styl1Go:0.10140975,(Acmi4Go:0.00000001,Acmi13Go:0.01950830)0.954153:0.04578953)0.680006:0.06062050)0.730572:0.13351022)0.000000:0.00000007)0.994978:0.15309257,((((HypdujTaGo:0.01426909,(RamvarTGo1:0.12489305,RamvarTGo2:0.01698518)0.927977:0.01153913)0.999850:0.07680855,(NaupoPOGo:0.02873398,(EupscPOGo:0.00000001,(SepphPOGo:0.05485459,(ArcduPOGo:0.03926228,(OctsiPOGo:0.00000001,(OctbiPOGo:0.00000001,((OctvuPOGo1:0.00000001,HapmaPOGo:0.18513067)0.000000:0.00000001,(NaupoPOGo1:0.08649337,(OctvuPOGo:0.00831705,OctbiPOGo1:0.01038356)0.979675:0.03681443)0.999850:0.08723312)0.148097:0.00000001)0.000000:0.00000001)0.999850:0.01824682)0.000000:0.00000001)0.999850:0.00591500)0.987079:0.01132557)0.999850:0.02745202)0.999850:0.08619605,(Tric4Go:0.77235923,Dr_o1:0.06660919)0.968383:0.05197700)0.982115:0.02812687,(((Stpu5Go:0.02753643,Lyva47Go:0.03963047)0.981363:0.05939103,(Asru3Go:0.01575288,Acpl16Go:0.11658798)0.993936:0.06438352)0.999850:0.15979331,(Sako19Go:0.07720559,Ptfl17Go:0.02786743)0.956854:0.03979708)0.995447:0.06106119)0.978725:0.02919283)0.999850:0.12684183)0.999850:0.07734182)0.999850:0.09009924)0.999850:0.06299729)0.999850:0.14605907)0.999850:1.64159983)0.939784:0.05963704)0.258027:0.00470066)0.999850:0.07439004);

6.3 Tree file for the collapsed phylogenetic tree shown in Figure 62

Tree file (text) is for the tree that is shown collapsed in Figure 62; Species are indicated by four or six letter abbreviation, for full names see Appendix 6.1; Zebrafish (*danio rerio*) is indicated by Dr; Plants are indicated by Gm, At, St, Nt, and Os.

(Gm_GPA1:0.12236291,At_GPA1:0.09110469,((St_GPA1:0.00745853,Nt_GA2:0.03488724)0.999850:0.07565737,(Os_RGA1:0.19704201,((Caow6Gv:1.17661064,((Saro2Gv:0.11620005,Mobr3Gv:0.17220271)0.999850:0.22016597,(Epmu10Gv:0.41634277,(Befo13Gv:0.40487694,(Dr_v1:0.31227237,((Ptfl13Gv:0.15475745,Lyva65Gv:0.37849653)0.999011:0.05812406,((NaupoPOGv:0.11893439,(OctsiPOGv1:0.00532028,OctsiPOGv3:0.06677328)0.999850:0.23911774)0.999850:0.14582963,(((PlapelArGw:0.08226690,(OppnovAGw2:0.15596246,(AchcolArGw:0.16502837,(OppnovAGw1:0.01496114,MedsubArGw:0.13194853)0.999850:0.09452672)0.194207:0.02049265)0.570964:0.05253931)0.999850:0.08724125,(((GradosLeGw:0.00000001,RhianiPGw:0.00000001)0.999850:0.20156059,(ActlunLeGw:0.00857641,(TyrputArGw:0.00000001,RhazepDGw:0.00000001)0.508797:0.00695096)0.999850:0.19208809)0.942358:0.04518921,(SarscaArGw:0.21410438,(DerpteArGw:0.01760368,(DerfarArGw:0.04730447,EurmayArGw:0.05249915)0.981025:0.00793591)0.999850:0.09115180)0.999850:0.09316525)0.999850:0.22016778)0.999850:0.29461332,(TricorMyGv:0.11830170,(CataquDpGv:0.27500501,(AndmauArGv:0.18562642,(EphdanPGv1:0.21654999,(RhianiPGv2:0.05401278,(DiacitHeGv:0.23262332,((CrysecBlGv:0.00956953,CopforBlGv:0.01680036)0.999850:0.03200357,((TricasCOGv:0.01125205,(ProjapCGv2:0.06062613,AsbverCOGv:0.00287489)0.000000:0.00341411)0.999850:0.11402988,(PleconTrGv:0.06100404,StetieTrGv:0.03450025)0.999850:0.24813106)0.999850:0.06091142)0.998794:0.02791598)0.706779:0.00954241)0.999850:0.03124567)0.999850:0.02938944)0.337683:0.02926077)0.999850:0.07886362)0.999850:0.05015428)0.997724:0.05148023)0.965264:0.04613081)0.944059:0.04198842)0.999850:0.20731731)0.991147:0.11953111)0.999850:0.26152622)0.979466:0.10329624

)0.860708:0.11488407,(((Sphsir8HGi:0.01860743,Spar3.Hgi:0.00869449)0.999850:0.76367524,((Spar1HGi:0.64263517,(Coli4Gi:0.29024820,((Icho1HGi:0.69041914,(Icho7HGi:0.20506743,(Sphsir9HGi:0.00318835,(Spar5HGi:0.00000001,Spar3Gi:0.05501830)0.000000:0.00000001)0.999850:0.27314469)0.603274:0.01604005)0.999850:0.14652171,(Caow3Gi:0.36165446,(Ptfl8Gi:0.42832870,(((Hyvu14Go:0.52707087,(Epmu7SGo:0.53237867,Exai5Go:0.26586280)0.214250:0.00975430)0.989866:0.14278601,(((Lyva47Go:0.20992021,Ptfl17Go:0.03549605)0.799742:0.04032492,((OctsiPOGo:0.01900564,NaupoPOGo:0.04308349)0.999850:0.01975832,((Mesbel25NG:0.09711758,(HypdujTaGo:0.00951842,(RamvarTGo2:0.01762294,RamvarTGo1:0.12512883)0.999850:0.01373989)0.999850:0.04984756)0.998810:0.02610470,((((DerfarArGo:0.00593125,(DerpteArGo:0.00292202,(SarscaArGo:0.00888328,EurmayArGo:0.00871125)0.000000:0.00000001)0.961334:0.00285515)0.999850:0.01461636,(TyrputArGo:0.00001427,ActlunLGo2:0.00000002)0.000000:0.00727185)0.999850:0.00774988,(AndmauAGo1:0.04189726,(RhianiPaGo:0.01329673,(AchcolAGo1:0.24084484,(MedsubArGo:0.00000001,OppnovArGo:0.02035651)0.900530:0.01434909)0.610311:0.00710757)0.986206:0.00372051)0.824663:0.00362623)0.999850:0.01881030,(PsepalCBGo:0.00000002,(PenvanCrGo:0.01852074,((PsespiCBGo:0.03648335,(StrmarMyGo:0.01360071,HolduoCBGo:0.00651895)0.000000:0.00000001)0.995409:0.00691300,(TricorMyGo:0.03230537,(ParhawCrGo:0.10947260,((dromelDiGo:0.00363285,DrovulDiGo:0.00000001)0.999850:0.05312946,(CataquDpGo:0.02590635,(((ProjapCOGo:0.03650479,CimlecHeGo:0.01969588)0.978296:0.00958862,(ThrpalThGo:0.04979805,(CopforBlGo:0.00000001,(CrysecBlGo:0.02497912,StetieTrGo:0.09903150)0.999850:0.03768039)0.973608:0.00385867)0.961718:0.00432627)0.986963:0.00506395,(AcypisHeGo:0.05199415,AndmauAGo2:0.04511653)0.884961:0.00547643)0.934650:0.00439069)0.000000:0.00000002)0.994197:0.00754869)0.642930:0.00272775)0.955788:0.00550525)0.000000:0.00000001)0.000000:0.00000002)0.431935:0.01072735)0.998895:0.02247668)0.999850:0.02791886)0.999850:0.05028227)0.933759:0.04898263,(Dr_o1:0.06554290,Tric4Go:0.67181778)0.938826:0.03807416)0.999850:0.04870966)0.999850:0.10224287,((Befo6Gi:0.25499895,Befo5Gi:0.23113726)0.999850:0.22155762,(((Dr_i1a:0.10329005,(Ptfl10Gi:0.05346320,Lyva21Gi:0.02409516)0.999850:0.03405630)0.992662:0.02530938,(((((ParhawCrGi:0.16253481,PenvanCrGi:0.04150763)0.999850:0.04107808,((EphdanPaGi:0.0166674

9,(RhianiPGi3:0.03819450,((CimlecHeGi:0.04313766,(AcypisHeGi:0.09297324,(Thr
palThGi:0.03753325,FraoccThGi:0.00601761)0.999850:0.03626392)0.814849:0.009
29958)0.769217:0.00289358,((DiacitHeGi:0.11120277,(((PlenigHGi2:0.07063953,Pl
enigHGi1:0.06691652)0.999850:0.08926118,(PseeloHyGi:0.00000009,PlapelAGi3:0
.80209038)0.196986:0.03696557)0.999850:0.03799414,((AnofonHyGi:0.06437023,(
RhazepDGi2:0.08013286,(dromelDGi1:0.02377834,(DrovulDiGi:0.00000001,dromel
DGi2:0.00867788)0.617088:0.00945112)0.999850:0.06171756)0.999850:0.0486502
8)0.999850:0.02691819,(ProjapCOGi:0.06557218,(TricasCOGi:0.00880080,Asbver
COGi:0.02281072)0.796901:0.00641989)0.999850:0.01647313)0.999850:0.026341
04)0.000000:0.00039000)0.965778:0.01169326,(CrysecBlGi:0.01544780,CopforBlGi
:0.00834014)0.999850:0.01343433)0.925640:0.00496010)0.999850:0.01096823)0.8
98045:0.00464891)0.999850:0.01766613,(CataquDpGi:0.11640711,(HolduoCBGi:0.
17776873,(PsespiCGi2:0.08950381,PsespiCGi1:0.43836132)0.999850:0.07064172)
0.994655:0.04527728)0.950248:0.01171206)0.999850:0.01604778)0.879113:0.008
13135,(StrmarMyGi:0.07836482,(TricorMyGi:0.05438987,Tric5Gi:0.26888238)0.996
323:0.04754420)0.904380:0.01245624)0.999850:0.03521591,((RamvarTaGi:0.1236
3518,HypdujTaGi:0.13255024)0.999850:0.29287460,((PlapelAGi2:0.02431708,(Plap
elAGi1:0.04768279,(OppnovArGi:0.02839653,MedsubArGi:0.03247477)0.942928:0.
01903949)0.671649:0.03591942)0.685812:0.04842727,((TyrputArGi:0.00965012,Ac
tlunLGi:0.02392174)0.999850:0.23199555,((DerfarArGi:0.02026103,(EurmayArGi:0.
04850172,DerpteArGi:0.01600076)0.928667:0.01126985)0.999850:0.12075394,(Sar
scaArGi:0.24223558,RhianiPGi1:0.10707645)0.981972:0.04181502)0.693892:0.035
14901)0.999850:0.23921409)0.442963:0.01797706)0.985600:0.03987687)0.999850
:0.04682487,(Exai2Gi:0.03612359,(Mesbel18NG:0.23073195,(NaupoPOGi:0.05729
331,OctsiPOGi:0.19635167)0.999850:0.06063540)0.978771:0.01390978)0.000000:
0.00000001)0.999850:0.02579797)0.987108:0.03786544,(Auco11Gi:0.48416553,Hy
vu19Gi:0.11908331)0.999850:0.05931635)0.998902:0.05966677)0.995219:0.06015
822)0.993711:0.06355834)0.994756:0.08157838)0.988037:0.06334195)0.998938:0.
06464659)0.986269:0.04977998)0.537831:0.01364487,(Icho6Gi:0.35976799,Icho3
HGi:0.41535339)0.963177:0.07704697)0.989854:0.05786503)0.999850:0.20283252
,((Coli1Gs-f:1.60492748,(Caow5Gs-
f:0.45838380,(((Auco14Gs:0.48521486,((AsbverCGs3:1.42815613,(ProjapCGs2:0.3

3739613,(AsbverCGs2:0.07422358,TricasCGs2:0.17155253)0.999850:0.13314828)0.893914:0.07666890)0.999850:0.20726854,(EpmueGs:0.53625827,(Lyva3Gs:0.13100974,((Dr_olf1:0.09854711,Dr_s:0.07154187)0.999850:0.12928255,(((((PenvanCrGs:0.05114681,ParhawCrGs:0.11400665)0.999850:0.05214016,(CataquDpGs:0.03681687,(((ThrpalThGs:0.00260141,FraoccThGs:0.00506970)0.999850:0.02906793,(EphdanPaGs:0.05459874,((CopforBlGs:0.00259473,CrysecBlGs:0.00000001)0.999850:0.03128969,(CimlecHeGs:0.06466591,(DiacitHeGs:0.06234230,AcypisHeGs:0.05564257)0.997831:0.01202720)0.990923:0.01145652)0.999850:0.01195771)0.942977:0.00699921)0.999850:0.02217186,((HolduoCBGs:0.01038453,PsepalCGs1:0.00000001)0.999850:0.02527288,(PsespiCBGs:0.05649497,PsepalCGs2:0.27017390)0.000000:0.01236667)0.999850:0.04003199)0.983080:0.01367556)0.733548:0.01226516)0.931786:0.01821938,((Mesbel23NG:0.17294507,(RamvarTaGs:0.08994842,HypdujTGs3:0.07723415)0.999850:0.14363614)0.999850:0.05532661,(PseeloHyGs:0.07314780,(((TricasCOGs:0.04014847,(AsbverCOGs:0.04761722,ProjapCOGs:0.04952263)0.322111:0.00597480)0.999850:0.07830444,((PleconTrGs:0.01803189,StetieTrGs:0.01520371)0.999850:0.03451722,(DelpasLeGs:0.03627251,ActlunLeGs:0.02587154)0.984543:0.01057205)0.999850:0.05212328)0.995773:0.02409643,(AnofonHyGs:0.01671735,(RhazepDiGs:0.01606515,(dromelDiGs:0.00507401,DrovulDiGs:0.00000001)0.999850:0.02362197)0.999850:0.03600794)0.999850:0.03656980)0.988875:0.03215219)0.927262:0.02148425)0.999036:0.03371041)0.963812:0.01855030,((SarscaArGs:0.00886828,(DerfarArGs:0.00000001,(DerpteArGs:0.00000001,(EurmayAGs2:0.00527855,EurmayAGs1:0.00000002)0.000000:0.00000001)0.996234:0.00578939)0.589932:0.00401170)0.999850:0.04065397,(PlapelAGs2:0.00609973,(PlapelAGs1:0.00000001,(RhianiPGs:0.02264078,((AchcolArGs:0.00435500,(MedsubArGs:0.01010706,OppnovArGs:0.00267977)0.999850:0.01080998)0.999850:0.01816422,(ActlunLGs2:0.00188500,(TyrputArGs:0.00000001,RhazepDGs2:0.00000001)0.999850:0.01421651)0.999850:0.02513038)0.858425:0.00313717)0.000000:0.00000001)0.422226:0.00418370)0.956562:0.00864726)0.999850:0.05692380)0.995995:0.03000690,(NaupoPOGs:0.05498578,OctsiPOGs:0.09979762)0.999850:0.07349991)0.999850:0.04789813)0.741206:0.02144986)0.999850:0.11891994)0.998637:0.07526784)0.000000:0.01830095)0.999850:0.18073464,(NaupoPOGf:0.41382195,(((PenvanCrGf:0.45462223,(CataquDpGf:0.56694680,(PsepalCBGf:0.12448403,Holduo

CBGf:0.22005268)0.999850:0.44924428)0.000000:0.05194641)0.995676:0.095681
05,((PlapelArGf:0.08947534,(AchcolArGf:0.14046954,(OppnovArGf:0.04623682,Me
dsubArGf:0.05528195)0.999850:0.07245237)0.974698:0.04289912)0.999850:0.116
61492,((SarscaArGf:0.17912077,(ActlunLGf2:0.02679716,(RhazepDGf:0.00000001,
TyrputArGf:0.01472055)0.892648:0.01721634)0.999850:0.10908356)0.998657:0.03
105960,(DerfarArGf:0.23570049,(EurmayArGf:0.27008946,DerpteArGf:0.06444593)
0.512985:0.00965817)0.999850:0.27362782)0.999850:0.14243396)0.999850:0.170
36230)0.997155:0.12629349,((ProjapCOGf:0.36816963,(TricasCOGf:0.19969271,A
sbverCOGf:0.20312725)0.999850:0.34159990)0.999850:0.42266047,(RhianiPGf2:0.
75255627,(PleconTrGf:0.63115909,(DelpasLeGf:0.36542730,(AnofonHyGf:0.63360
252,(dromelDGf1:0.11368195,(DrovulDiGf:0.03021664,dromelDGf2:0.03389491)0.9
99850:0.09967253)0.999850:0.45566286)0.999850:0.22775392)0.955303:0.094394
13)0.963649:0.09318569)0.529713:0.08551768)0.984279:0.09984864)0.999850:0.1
9366190)0.999850:0.29992788)0.999850:0.28272666,(Icho3HGs-
f:0.36768697,SphJP65HGs:0.63494621)0.993912:0.11316589)0.000000:0.0374595
0)0.998519:0.11722280)0.999850:0.24715531,(Icho1HG12:0.68976317,((Caow2G1
2:0.51931989,((Sphsir7HG1:0.00000001,Spar7G12:0.00816661)0.999850:0.363181
20,(Icho4HG12:0.19018487,Icho2G12:0.21244531)0.884988:0.07251061)0.999850:
0.23717457)0.997067:0.09396393,(((Epmu11G12:0.71545064,(Befo10G12:0.50180
070,((Lyva53G12:0.32415741,((Ptfl11G12:0.14955053,(Dr_12:0.43827317,(Exai4G1
2:0.21535511,(Hyvu18G12:0.12479662,Auco22G12:0.10941535)0.999850:0.10649
174)0.999850:0.08634195)0.990959:0.03867681)0.870958:0.02647695,(HypdujTG1
2:0.77810122,(NaupoPOG12:0.28548648,OctsiPOG12:0.44230405)0.981416:0.080
22564)0.993942:0.06145963)0.704394:0.03036353)0.999850:0.06003592,(((Plapel
G12b:0.42115535,((RhazepDiG1:0.00000001,(ActlunLG12:0.06722945,TyrputAG12:
0.00000002)0.000000:0.00000001)0.999850:0.30062678,(RhianiPG12:0.28206052,(
SarscaAG12:0.22293227,(DerpteAG12:0.12141107,(EurmayAG12:0.04383410,Derf
arAG12:0.07289843)0.996816:0.05001248)0.999850:0.16232109)0.999850:0.21649
834)0.999850:0.12056243)0.999850:0.39691359)0.965850:0.08319310,(PlapelG12
c:0.21682627,(PlapelG12a:0.06266700,(MedsubAG12:0.16321996,OppnovAG12:0.
15428407)0.432093:0.01211713)0.909664:0.09011255)0.846186:0.06796098)0.999
850:0.17652996,((Andmau12b:0.21741778,(StrmarMG12:0.26678499,TricorMG12:0

.23756899)0.553551:0.05426457)0.661803:0.02359626,((((CataquDpG1:0.3410283
4,RhianiPaG1:0.28933569)0.988670:0.08942217,(Mesbel9NG1:0.83397757,(Psespi
CBG1:0.25385965,(PsepalCBG1:0.06202846,HolduoCBG1:0.00412770)0.999850:0.
48961283)0.999850:0.16743537)0.999850:0.15021996)0.999850:0.07292632,((Eph
danPGi2:0.22036433,((FraoccTG12:0.00683966,ThrpalTG12:0.12941197)0.999850:
0.13919933,(CimlecHeG1:0.15154136,(AcypisHeG1:0.17656220,DiacitHeG1:0.3124
1155)0.984434:0.02659432)0.987689:0.03088842)0.824385:0.01980666)0.922471:
0.02804372,((CrysecBG12:0.03004445,CopforBG12:0.08270627)0.999850:0.07901
349,((Andmau12c:0.11912089,(PseeloHyG1:0.07695840,PlenigHyG1:0.03655515)0
.681998:0.01856848)0.999850:0.06404018,(((DelpasLeG1:0.08811680,(PleconTG1
2:0.11966181,StetieTG12:0.05710567)0.999165:0.09049464)0.999850:0.08912032,
(AnofonHyG1:0.10401398,(RhazepDG12:0.08576967,(dromelDG12:0.02690706,(Dr
ovulDiG1:0.06522398,dromelDiG1:0.06038171)0.000000:0.01125066)0.997612:0.0
4594046)0.999850:0.14413087)0.999850:0.14979295)0.999850:0.12014841,((Proja
pG12c:0.09603529,ProjapG12e:0.00000001)0.999850:0.16119514,(AsbverCG12:0.
01922100,TricasCOG1:0.04239949)0.999850:0.14261460)0.999850:0.09776065)0.
999850:0.04569798)0.999850:0.03726478)0.992994:0.03061712)0.999850:0.07123
648)0.999850:0.06326929,(PenvanCrG1:0.13586868,ParhawCrG1:0.48464660)0.99
9850:0.23797475)0.999850:0.07162102)0.999850:0.12911364)0.999850:0.0675527
4)0.947866:0.03148472)0.864153:0.07465103)0.647990:0.03132204,(Tric1G12:0.4
4802987,Tric2G12:0.64819704)0.999850:0.18611809)0.999850:0.51153628,((Derfa
rAGr2:1.00516151,Caow4Gq:0.62253618)0.584997:0.03446355,((Epmu2SGq:0.664
12102,((Mobr2Gq:0.40323237,Saro1Gq:0.61889775)0.999850:0.37167249,(Mobr4G
q:1.31264512,Befo4Gq:0.70650784)0.912236:0.20532115)0.329689:0.15033501)0.
993953:0.10749805,(Hyvu16Gq:0.12786847,(Tric8Gq:0.23560480,(Exai1Gq:0.0299
9571,(Dr_q:0.11696313,(OctsiPOGq1:0.13342129,((HypdujTGq2:0.12306449,(Ptfl1
8Gq:0.00000001,(RamvarTaGq:0.04989442,HypdujTGq1:0.00770783)0.999850:0.1
5697789)0.000000:0.00000002)0.996182:0.02822831,(NaupoPOGq:0.06313282,(((
PenvanCGq2:0.09946178,((ParhawCGq2:0.04301861,ParhawCGq3:0.29744689)0.
999850:0.10458457,(((RhazepDGq1:0.01929531,(EphdanPGq2:0.21466752,dromel
DGq2:0.04966833)0.999850:0.12091112)0.999850:0.03235455,((dromelDGq1:0.00
296339,DrovulDiGq:0.00296966)0.964755:0.00605052,(dromelDGq5:0.34689944,dr

omelDGq3:0.14253410)0.999850:0.15843993)0.994939:0.01389362)0.999850:0.05911652,((PlenigHyGq:0.03936709,(((AnofonHyGq:0.01525903,(PleconTrGq:0.00618038,StetieTrGq:0.00862790)0.999850:0.02279151)0.957865:0.00385261,(CimlecHeGq:0.02011139,AcypisHeGq:0.03343791)0.999850:0.01032833)0.999850:0.00687645,((EphdanPaGq:0.03919059,(PseeloHyGq:0.01353002,(FraoccThGq:0.01067490,ThrpalThGq:0.00733097)0.999850:0.06370882)0.814632:0.00654721)0.956474:0.00620547,((AsbverCOGq:0.01656933,TricasCOGq:0.00740837)0.999850:0.06136502,(CopforBGq2:0.00589711,CrysecBGq1:0.00000002)0.703959:0.00264541)0.986256:0.00589520)0.904830:0.00313120)0.989373:0.00615320)0.999850:0.02902187,(PsespiCGq1:0.01631187,(PsepalCBGq:0.00390060,HolduoCBGq:0.00800464)0.9999850:0.06616171)0.999850:0.01939992)0.984126:0.01938592)0.231529:0.00778679)0.982709:0.01474900)0.694677:0.00850991,((PsespiCGq2:0.36340319,Dr_14a:0.27732900)0.000000:0.02853882,(Mesbel17NG:0.08027000,Lyva17Gq:0.21892472)0.999850:0.04075773)0.999850:0.03525378)0.428534:0.01627193,(((TyrputAGq1:0.00000001,(RhazepDGq2:0.00000001,ActlunLeGq:0.00000001)0.000000:0.00000001)0.999850:0.03164904,(OppnovAGq5:0.00000001,(AchcolArGq:0.00293012,(MedsubAGq5:0.00604805,PlapelAGq2:0.01305880)0.998816:0.00297869)0.000000:0.00000001)0.999850:0.01115117)0.999850:0.01085439,((StrmarMyGq:0.07782801,(TricorMGq1:0.05102010,(TricorMGq3:0.44046725,(((PlapelAGq1:0.14887043,((AchcolAGq4:0.22542468,(MedsubAGq1:0.10155808,OppnovAGq1:0.09144047)0.999850:0.12685376)0.999850:0.17639270,(((ActlunLGq2:0.01617560,TyrputAGq2:0.02506027)0.999850:0.21314289,(DerfarAGq4:0.07345897,(DerpteArGq:0.07725517,EurmayAGq1:0.03983891)0.288828:0.00569640)0.999850:0.36912707)0.176253:0.04705038,RhianiPaGq:0.40049474)0.999850:0.55716043)0.999250:0.09731709)0.998508:0.07004209,(CrysecBGq2:0.08881028,CopforBGq1:0.11267725)0.999850:0.56197465)0.000000:0.01599180,(TricorMGq2:0.40559222,((MedsubAGq2:0.10814586,(OppnovAGq4:0.00327629,OppnovAGq2:0.03948066)0.999850:0.05334775)0.999850:0.13570356,(OppnovAGq3:0.05812763,(MedsubAGq3:0.02423200,MedsubAGq4:0.00000002)0.999850:0.32848052)0.999850:0.05428249)0.999850:0.29794469)0.999850:0.12205816)0.455723:0.02177137)0.999850:0.04541268)0.988904:0.01202829)0.993012:0.01236652,(RhianiPGq1:0.02668758,(SarscaAGq1:0.00577340,(DerpteAGq1:0.00000001,(DerfarAGq3:0.00000001,(EurmayAGq2:0.02290225,DerfarAGq2

:0.00896330)0.942592:0.01136953)0.999043:0.00588892)0.999850:0.00973135)0.999850:0.01069614)0.999850:0.01103925)0.572444:0.00455903)0.999850:0.02186998)0.999850:0.03067260)0.966158:0.01875674)0.999151:0.03468046)0.928064:0.02860463)0.965291:0.02735008)0.999850:0.04183633)0.939450:0.03753502)0.999850:0.16797905)0.999850:0.12517601)0.984888:0.06467383)0.998585:0.06118570)0.758153:0.02545462)0.999850:0.14071724)0.932020:0.06456794)0.838016:0.08690395)0.999850:1.42674173)0.991584:0.06122736)0.000000:0.00241733);

6.4 Publications

Manuscripts in preparation:

Abu Obaid A., Ivandic I., Korsching S. Unraveling the function of the fifth G protein family: Gv has many actions but potentially a single role, in cationic homeostasis.

Abu Obaid A. and Korsching S. The fifth G protein family is born in the common ancestor of animals and unicellular Holozoa and is sculpt by rare gene gains and a fractal pattern of gene losses.